2003

High-speed National Research Network and its New Applications

research activity annual report

Identification code:	MSM 000000001
Carrier:	CESNET, z. s. p. o.
Carrier representative:	Ing. Josef Kubíček managing board chairman Prof. RNDr. Milan Mareš, DrSc. managing board vice chairman
Research leader:	Ing. Jan Gruntorád, CSc.
Contact address:	CESNET, z. s. p. o. Zikova 4 160 00 Praha 6 Czech Republic phone: +420 224 355 207 fax: +420 224 320 269 E-mail: info@cesnet.cz

GEANT

Editors:
Pavel Satrapa, Ladislav Lhotka, Pavel Vachek

Authors of individual parts:
1 Jan Gruntorád
2 Pavel Satrapa and others
3 Václav Novák, Tomáš Košňar
4 Milan Sova
5 Stanislav Šíma, Lada Altmanová, Jan Radil, Václav Fanta, Martin Míchal, František Potužník, Leoš Boháč, Miroslav Karásek, Sven Ubik, Miloš Wimmer, Miloš Lokajíček, Jiří Novotný
6 Ladislav Lhotka, Martin Pustka, Pavel Satrapa
7 Ladislav Lhotka
8 Ivo Hulínský and others
9 Luděk Matyska
10 Miroslav Vozňák, Michal Neuman, Jan Růžička
11 Sven Ubik
12 Helmut Sverenyák
13 Luděk Matyska
14 Vladimír Smotlacha, Sven Ubik
15 Ladislav Lhotka
16 Tomáš Zeman, Jaromír Hrad
17 Miroslav Vozňák, Radek Holý
18 Jan Nejman
19 Jan Haluza, Filip Staněk
20 Pavel Satrapa
21 Pavel Vachek, Miroslav Indra, Martin Pustka
22 Vladimír Smotlacha
23 Michal Krsek
24 Otto Dostál, Michal Javorník, Karel Slavíček
25 Stanislav Šíma, Helmut Sverenyák

ISBN 80-239-2176-2

Contents

1 Introduction

The report presented herein describes the progress in solving the research plan titled *High-speed National Research Network and its New Applications* together with results achieved in 2003. This was the fifth and last year of this research plan.

Completion of the research plan is a reason for looking back at the past and for evaluation. Research and development activities carried out within the framework of this research plan contributed significantly to the fact that the National Research and Education Network of the Czech Republic belongs to the foremost European academic networks.

Backbone parameters, capacity of international lines and quality of member connections improved significantly during last five years. In the beginning of 1999, most parts of the CESNET backbone consisted of 34 Mbps lines, no backup lines existed and the international links were heavily overloaded most of the time. By the end of 2003, the backbone capacity of 2.5 Gbps offers full redundancy and sufficient capacity of international lines. Typical capacity of the most significant member institution circuits has reached 1 Gbps. It means that the backbone capacity as well as that of the international links increased more than 70× during those five years. Both the performance and reliability of the whole networking environment improved.

Other services follow this trend, especially in the reliability field. In addition, the portfolio of network services has been extended significantly: the IP telephony, various multimedia services, native IPv6 backbone, AAA services, Grid services and others have been added. This quantitative and qualitative improvement was achieved using an almost constant financing from the Ministry of Education and CESNET members.

Another positive fact is the growth of user groups using the high-speed features of the CESNET2 network. The IT professionals and users from the natural science fields, especially the high energy physics, have always been very active and demanding. During the last years, use of the network for medical applications has been evolving successfully – especially thanks to the long-term cooperation between the IT institutes and faculty hospitals.

Regardless of the almost traditional problems concerning the research plan financing, solving the research projects – parts of the whole research plan – continued successfully. In addition to the internal projects, our research team participated successfully in the *GÉANT* project research activities titled TF-NGN (Task Force – Next Generation Networking). Besides the GÉANT project, we participated in three other international projects of the 5th EU Framework Programme: *DataGrid*, *SCAMPI*, and *6NET*.

During 2003 we worked also on proposing a new research plan for the years 2004–2010 titled *Optical National Research Network and its New Applications*. We also participated in formulating proposals for the 6th EU Framework Programme projects – in particular, the *Multi-Gigabit European Academic Network (GN2)*, *Enabling Grids for E-Science and industry in Europe (EGEE)*, *Global Advanced Research Development Enviroment and Network (GARDEN)*, and *Grid Aware Network Developments in Europe (GRANDE)*.

The following chapters contain detailed information about the progress of individual projects solved within the research plan. Because of the wide extent of these projects, different authors wrote their separate chapters; therefore, this document may be regarded as a collection of papers connected with a common theme.

2 Summary

The research plan *High-speed National Research Network and Its New Applications* pursues the following fundamental objectives:

- operate a high-speed national research network, CESNET2
- ensure its further development, according to the needs of users and current status of the technology
- participate in analogous projects at European and global levels
- carry out original research in the area of networking technologies and their applications
- actively seek, adapt and develop corresponding applications.

The specific aspect of this research plan is that it has, to a great extent, the character of a service. Most of the investments are spent on the operation and development of a communication infrastructure for science, research and education. A number of other projects and activities benefit from the project, even if not directly associated with this research plan or CESNET, since an adequate communication infrastructure is often an implicit assumption.

Due to a broad range of the research plan, all activities were divided into thematically defined projects, classified further into three categories: strategic, international and other. The remainder of this chapter includes a brief summary of the activities and results of individual projects. For a more detailed information see the following chapters.

2.1 CESNET2 Backbone and its Services

2.2 Strategic Projects

The main objective of the project *Optical networks and their development* has been the development of optical technologies and their application in national research networks. In 2003 we essentially finished transition of CESNET2 network to dark fibres. We have also been engaged in the analysis of the possibilities for international dark fiber interconnection. We started building an experimental network named CzechLight, which is dedicated to the research in optical transmission technologies.

We achieved very good results in the domain of fibres with no equipment along the route (Nothing In Line, NIL). We managed to increase the distance reachable by this technology to 235 km (Brno–Ostrava). We increased significantly the number of NIL lines in the CESNET2 network – more than one half of the backbone lines are currently of the NIL type. We simulated long-haul data transmissions using 1, 2.5 and 10 Gbps.

Single-fibre equipment is another interesting technology we dealt with. After testing it on an experimental line connecting the National Library to CESNET, we deployed five single-fibre lines with transmission speeds of 100 Mbps. These connections are suitable especially for connecting smaller nodes to the backbone (Cheb, Opava, Karviná, etc.).

Implementation of IPv6 in CESNET2 network has been focusing on the development and deployment of the new version of IP. In 2003 we advanced significantly by removing previously used tunnels and starting a native IPv6 service over MPLS transport (using the so-called 6PE technology). IPv4 and IPv6 protocols are now equal in the backbone. We tried to treat both protocols equally also in access networks, but the delay in shipment of new control engines for access routers hindered our effort. We nevertheless expect to be able to configure our access routers for an IPv4/IPv6 dual-stack operation within first months of 2004. We also established IPv6 peerings with six inland networks.

In October, we organised a successful workshop titled *IPv6 – development and deployment*. The event met with very positive responses from networking professionals.

The *Liberouter* project continues the development of a PC-based accelerated IPv6 router. During 2003 we designed and created daughter cards for the COMBO6 motherboard, which implements hardware packet forwarding. Because of the flexibility of the programmable hardware it is based on, several other projects got interested in using it for their special purposes. We advanced significantly with the software development and expect to deliver the first prototype router in 2004.

The project *Multimedia transmissions* focuses on the transport of multimedia contents over data networks. In collaboration with MU Brno we developed the first AccessGrid network node allowing high-quality videoconferencing and creation of a virtual working environment. Workspace for 3D videoconferencing is included in the AccessGrid node, which also enables research in this field.

Most activities that were started in the previous years continued in 2003 – development of a unicast videoconferencing mirror, building the H.323 infrastructure, support of pilot groups, and collaboration with Czech Radio on high-quality network broadcasts of its programs.

The *MetaCentre* project has been developing the national grid – a distributed platform for demanding computations. It was upgraded by adding new 64 Intel Pentium IV Xeon 2.4 GHz processors. In addition to this, disk, memory and networking capacities of the existing grid nodes have also been significantly expanded.

We made important changes to the information and authentication services. Along with increasing compatibility with common standards, we reworked the system Perun for managing user accounts so that it is now ready to support large-scale homogeneous grids. We continued testing the behaviour of various task types in a highly distributed computing environment.

The project *Voice services in CESNET2* has been working further on our IP telephony platform. Its core is essentially operating in a production mode – 20 association members are connected to our IP telephony network and use it rather heavily. The overall traffic volume increased by a factor of more than six compared to the end of 2002. We also have several new foreign partners reachable over IP telephony.

In the research area we focused on experiments with advanced features of IP phones and different setups of the infrastructure. We continued our research of the SIP protocol that is supposed to become the successor of H.323.

The *End-to-end performance* project investigates methods of bandwidth provisioning, throughput monitoring and other related topics. We concentrate in particular on end-host tuning, which can significantly improve the transmission performance. We have also contributed to international initiatives, such as PERT, whose mission is to establish a support body for helping users solve their performance-related problems.

2.3 International Projects

The goal of the *GÉANT* project has been to create a high-speed backbone connecting the National Research and Education Networks (NRENs) in European countries. In 2003 we took part in the operation and development of this network (including IPv6 deployment), including the workgroups of TF-NGN, which represent the research part of the project. Our contribution was mainly in the fields of optical technologies, quality of service, IPv6, and network monitoring.

In the *DataGrid* project, which focuses on large-scale grids for processing huge data volumes, we have been responsible for the development of a logging service and security protocols. In 2003 we concentrated on the deployment of version 2 that was started in the previous year. Along with the development of version 2.1 we have also been preparing the forthcoming version 3. We created an interface of the logging service with R-GMS (grid monitoring architecture). However, since we found out that the R-GMS implementation suffers from many flaws, we started the development of our own variant of R-GMS at the end of 2003.

We participated on finalising the proposal for the *EGEE (Enabling Grids for E-science and industry in Europe)* project, which can be considered a successor to *DataGrid*. In this project we will continue the development of the grid middleware. We are the only participant from Central Europe who obtained direct financial support. The project proposal was accepted and will start on April 1st, 2004.

The *SCAMPI* project focuses on high-speed network monitoring. Our participation was originally aimed at equipment evaluation, testing and measurements. However, after one of the SCAMPI partners failed to deliver the promised hardware measurement device, we were invited to fill this gap with the COMBO6 card, namely to adapt it for the purposes of high-speed network monitoring. As a consequence, CESNET increased its person-month capacity in the project and Masaryk University was accepted as an additional project partner.

In the first stage we intend to use the current version of COMBO6 for monitoring using the existing Gigabit Ethernet daughter cards. Later we plan to create a daughter card capable of monitoring 10 Gbps lines – this is the target speed of the SCAMPI project.

Finally, the fourth project of 5th EU Framework Programme with CESNET participation is *6NET*. Its goal is to build a large-scale experimental IPv6-only network and gain practical experience with its operation. The primary contribution of CESNET to this project is the development of an IPv6 router based on the COMBO6 card. Apart from that, we also participate in other activities, for example in building and operating the 6NET backbone and in experiments with IPv4/IPv6 coexistence.

2.4 Other Projects

The project *Infrastructure and Technologies for On-Line Education*, being a pilot project of the Department of Telecommunication Engineering (Czech Technical University in Prague, Faculty of Electrical Engineering), aims at developing and evaluating tools for an electronic support of education (eSupport). In 2003 we finished the WWW portal containing educational materials, live transmissions and recordings of lectures. We have also been preparing lectures to be shared by remote universities.

The project *Distributed Call Centre* represents an advanced IP telephony application. It integrates many technologies and offers various options for the communication between a user and an operator – from the common phone call to sharing applications and remote guidance. In 2003 we realised some configuration changes and developed scripts for controlling the components of the call centre.

The main result of the *Intelligent NetFlow Analyzer* project is the distribution version of the *NetFlow Monitor* program for network traffic evaluation. The program was downloaded by more than 1000 organisations in 60 countries since its release (1st quarter of 2003). During 2003 we have been improving and enhancing the program in many ways – five new versions were released till October.

The project *Storage over IP* has been testing HyperSCSI – a protocol allowing transport of SCSI commands over the network. We tested its software implementations for Linux and MS Windows operating systems. Especially the Linux implementation was found to be functional and usable for building cheap local storage networks. So far, the only supported protocol is Ethernet, which disables the use of this method in wide area networks.

The *Presentation* project has been working on a framework for further development of our WWW server. This year we implemented a significant technological upgrade and face-lift by converting the code to the combination of strict XHTML 1.0 and CSS. We released 30 technical reports, two thirds of them being written in English. We also organised two successful workshops – *New Ways in Development of High-speed Networks and their Applications*, and *IPv6 – development and implementation*.

Network security is the subject of another project named *Security of local CESNET2 networks*. The ever increasing number of attacks targeting both commercial and academic networks calls for new countermeasures that sometimes utilise rather unconventional methods. The goal of the project team is thus to mediate information about freely available systems for detecting and preventing intrusions, develop our own extensions and, last but not least, implement these tools in our networks in order to make them more secure.

The objective of the project *NTP server controlled by the national time etalon* is to provide a time server offering an “official time”. We designed and implemented the architecture of the server already in 2002, this year we continued with various tests and improvements (for example, we developed a new version of the control software). The achieved accuracy of 500 ns surpassed our original expectations.

Finally, the project *Platforms for video transmission and production* focuses on the technologies for video transmission and broadcasting. Beside realising many live broadcasts and enhancing the video archive, we have been engaged in an international collaboration within the TF-Netcast group. Our most important contribution to the activities of this group was an improved announcement portal *prenosy.cesnet.cz*. In collaboration with the Jyxo company we developed a system for searching metadata of multimedia files. In the second half of 2003 we created the portal *streaming.cesnet.cz* dedicated fully to multimedia transmissions.

Part I

Backbone and Services

3 CESNET2 Backbone Network – Operations and Development

CESNET2 backbone network was entirely reconstructed during the year 2003 both with regard to the logic topology, technology, stability and general availability of the operated services. We succeeded in fulfilling most of the stated goals; however, in the meantime it became obvious that in some cases interim solutions must be used. The final planned stage should be reached in the beginning of 2004. The problems were partly caused by a delayed delivery of new hardware (Supervisor Engine Sup720 with HW support for IPv6 ordered for access routers OSR7609 in the network core). CESNET2 network is the first client of Cisco Systems to deploy these new high-performance control cards in a production network.

The development and changes of the backbone network can be briefly summarised as follows:

- Deployment of new Cisco OSR7609 access routers (temporarily with the Supervisor2/PFC2 Engine) and an entire reconfiguration of all GigaPoPs (March–May).
- Launch of the direct peering with the Slovak academic network SANET (April).
- Fully redundant network core and setup of back-up connections for all GigaPoPs (May).
- Back-up of the connections to Telia International Carrier (our provider of international connectivity) and to NIX (June).
- Deployment of IPv6 in the production network and migration to a dual-stack IPv4/IPv6 operation on the peering with the GÉANT network (June).
- General increase of the network's stability and availability through the use of Cisco NSA services.
- Migration from leased data circuits (services) to leased optical fibres terminated with our own technology.

3.1 Topology of the backbone network

The basic logical topology of the backbone network is formed by ten nodes (GigaPoPs) interconnected over data circuits with a transmission capacity of at least 1 Gbps (see Figure 3.1). The backbone circuits use GE (Gigabit Ethernet) and POS STM-16/OC-48 having transmission capacity of 2.5 Gbps. Most gigabit nodes are connected to the backbone network via two data circuits for redundancy. The optical circuits are operated in several configurations:

GBIC-CWDM-1550: CWDM GBIC can be used only for GE on shorter distances. We use a version for the 1550 nm wavelength, which has minimal signal attenuation (according to its producer's statement, a typical operating distance is around 100 km with the attenuation limit of 32 dBm). These are used for example on the links Prague–Pardubice, Prague–Ústí nad Labem, Ústí nad Labem–Liberec and Olomouc–Zlín.

These GBICs are also deployed on the international link Brno–Bratislava where we had to use the Catalyst 3524 switch as a signal regenerator due to the total distance that had to be spanned.

Optical EDFA amplifiers: Following the extensive tests performed by the project *Optical Networks and their Development* (see below), we started deploying EDFA amplifiers on the majority of our long-haul optical links (Prague–Pilsen, Pilsen–České Budějovice, České Budějovice–Brno, Brno–Ostrava etc.). In all these cases, active elements are located only at the ends of the routes, hence the name "Nothing in Line" (NIL). Furthermore, the optical routes are protocol transparent.

Optical SDH regenerators STM-16/OC-48: We used Cisco ONS15104 optical regenerators for the first optical routes of the backbone network (for example Prague-Brno). The regenerators perform a complete opto-electro-optic regeneration according to ITU SONET/SDH standards (with a signal delay of approximately 20 μs) on the line and section layer levels. They use 192 kbps SDCC channel for in-band management (telnet/ssh access and SNMP). The disadvantage of the regenerators is their relatively small operating distance (approximately 80 km) so they must be deployed in the middle of an optical route (for example we used three of them on a 320 km-long route Prague–Brno). Another drawback is their protocol dependence (support only SONET/SDH).

From the operational point of view the regenerators are perfectly suitable for a production environment because of their administration possibilities, flawless function and reliability – we have not experienced any problem with them since they were deployed (in 1999). We expect to construct a new route Prague–Brno with the use of DWDM which will replace the current technology.

The transition of all backbone links to the dark optical fibres will be finished by the end of January 2004. We were sometimes forced to adapt the network

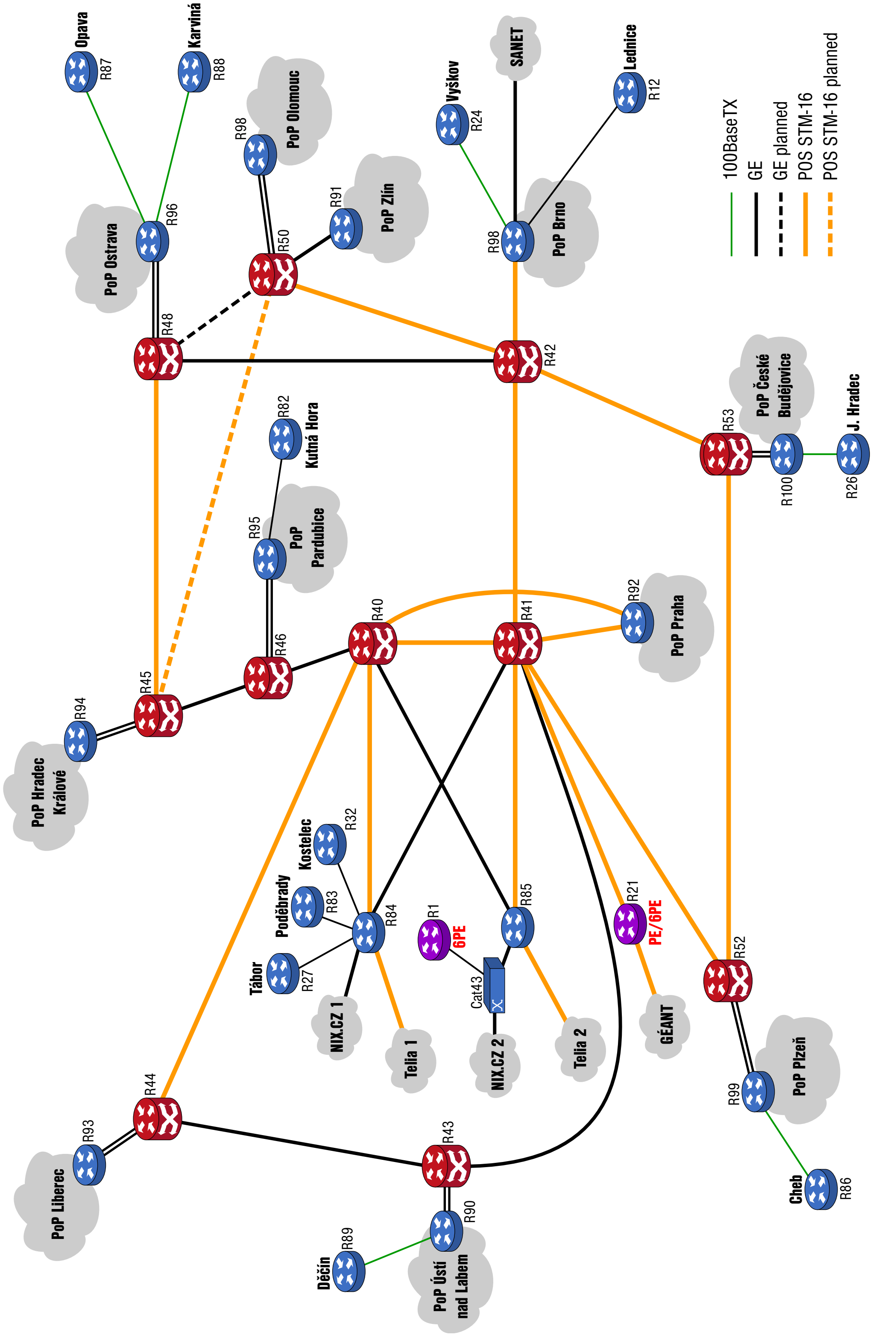


Figure 3.1: Current topology of CESNET2 backbone network

topology to the availability of the optical routes between the individual nodes. Some planned routes could not be implemented because of prohibitive cost (due to long distances), and in some cases it turned out that a particular optical route passes through the location of one of our nodes and thus it made more sense to split the route into two (for example, the route from Ústí nad Labem to Pilsen crosses Prague and we thus split it up into two routes, both terminated in our Prague PoP).

The new routes to be implemented are Hradec Králové–Olomouc and Olomouc–Ostrava. The latter will substitute the current route Hradec Králové–Ostrava. In order to complete the transition to the leased optical fibres, it still remains to solve some issues like delayed deliveries of optical amplifiers or technical problems with certain equipment.

The basic transmission protocol of the backbone network is IP/MPLS. We use OSPFv2 as the IGP protocol of the MPLS network. The logic topology of the network is divided into two functional levels to which the topology of individual GigaPoPs is adapted:

Network Core (Core Backbone Area): The network core itself is formed by GSR12016 routers (red color in the Figure 3.1), on which all backbone data circuits are terminated. They perform only MPLS functions (except for IP multicast) and are transparent from the viewpoint of unicast IP. The IOS version used (12.0(21)S7) supports only TDP, not LDP. The switch-over to backup circuits is controlled by the OSPF protocol (we reduced the timing parameters towards a faster convergence, for example OSPF Hello to 1 s). The type of HW interface we use and the transition to GE will not allow to use the more efficient and faster re-routing mechanisms (Fast Re-routing or DPT) with re-routing time around 50 ms.

Access network part (GigaPoP Area): On the edge of our backbone – as MPLS Provider Edge (PE) devices – we use Cisco OSR7609 access routers (with Supervisor2/PFC2 Engine), except for smaller nodes (Ústí nad Labem, Zlín) where Cisco 7206-VXR access routers with NPE-G1 are used. These routers perform all functions and transport services of the backbone network (MPLS, MPLS VPN, QoS, IPv4 routing, IPv4 multicast, export of NetFlow statistics, access filters) for the connected sites. While academic metropolitan networks (PASNET, BAPS,..) and university networks are connected over gigabit Ethernet, other participants and remote work places are often connected with smaller capacities (Ethernet, Fast Ethernet). The access routers and network core routers are always connected through two interfaces GE or POS STM-16/OC-48 with load distribution at the OSPF level.

> The smaller network nodes (PoP) are connected to these access routers by 100 Mbps optical circuits (a single fiber connection, green lines in the Figure 3.1), 34 or 10 Mbps microwave links (we use the conversion to 100BASE-T at both ends of the 34 Mbps circuits) and leased fixed circuits. Smaller routers with a limited functionality are used in these nodes as MPLS Customer Edge (CE) devices (Cisco 2621, C2651-XM or C2691).

A typical GigaPoP topology is shown in the Figure 3.2 – the architecture of the Liberec node serving as an example. Due to the limited hardware support of MPLS in the OSR7609 router (with the current Supervisor2/PFC2 engines, MPLS is supported only on OSM modules), the connection to the core router is accomplished by two interfaces of the OSM-4GE-WAN module. Unfortunately, MPLS is not supported on WS-X6516-GBIC and WS-X6548-RJ45 access modules. We had to create an MPLS Intra-PoP external connection for connecting other routers with MPLS support (a port of OSM-4GE-WAN 3/3 module and GE1/2 on the supervisor) and distributing MPLS signalisation to the adequate LAN port over 802.1Q VLANs.

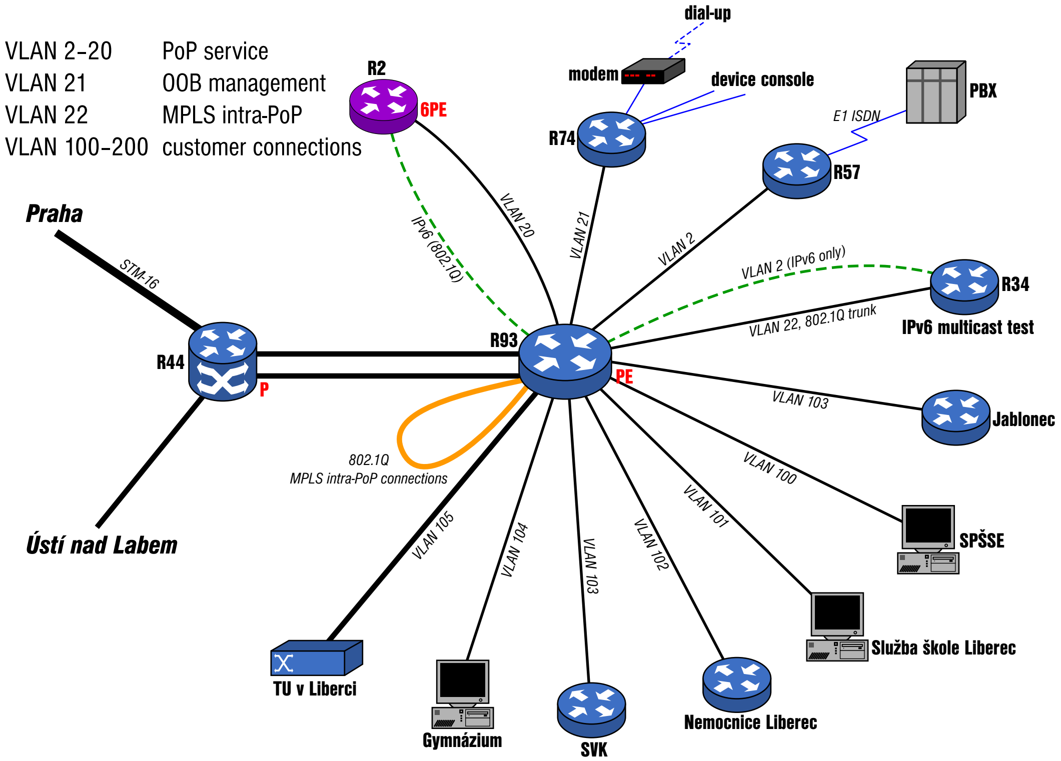


Figure 3.2: GigaPoP topology example (Liberec)

Each node's functionality is divided into three parts: MPLS support, connection of service servers and connection of end users. Cisco 7500 routers serve as 6PE routers (see below) or as test PE/6PE routers, e.g., for IPv6 multicast experiments. The service segments include OOB access, VoIP gateway, UPS and other devices and servers. End users are connected over 802.1Q virtual LANs. If MPLS VLANs are needed, we use the 802.1Q trunk and take care of mapping MPLS VLANs into 802.1Q.

3.2 IPv4 Routing

We use internal BGPv4 (iBGP) between PE routers as the internal routing protocol conveying information about network prefixes reachable in individual nodes (see Figure 3.3). The external routers R84, R85 and R21 are (temporarily) configured as redundant route reflectors RR1, RR2 and RR3 (iBGP neighbours are shown in Figure 3.3 only for RR1). The other PE routers run as route reflector clients. The use of route reflectors reduces the number of neighbor adjacencies in the backbone network.

The GigaPoP access routers propagate only static aggregated prefix blocks to the backbone, i.e., no redistribution from the inner routing protocols takes place. Metropolitan networks use OSPFv2 internally (with an OSPF process identifier different from the one used on the backbone network) and smaller networks often use static routes. Two notable exceptions are the networks of PASNET and ČVUT that act as pseudo-autonomous systems (using private AS numbers 65001 and 65002) and must thus be routed using external BGP.

3.3 IPv4 Multicast Routing

We use internal MBGP for the exchange of multicast routing information. The configuration of iMBGP is similar to the iBGP configuration of PE routers with three route reflectors on R84, R85 and (temporarily) R21. However, for multicast the route reflector clients must be configured on *all* P and PE routers (see Figure 3.4), or otherwise the RPF (Reverse Path Forwarding) checks would fail on the core routers and all multicast packets would be discarded.

We use PIMv2-SM for multicast routing. The backbone network is divided into separate PIM domains (see Figure 3.5), each having its own rendezvous point (RP). A multicast data source first registers with the nearest RP, which results in creating the so-called SPT (Shortest-path Tree) from RP towards the source. Starting from the last hop designated router, which connects a network where one or more recipients exist, all routers along the way to the RP send their register messages (hop by hop). This procedure results in the so-called shared tree (*,G) where the asterisk is a wildcard indicating any source. The initial

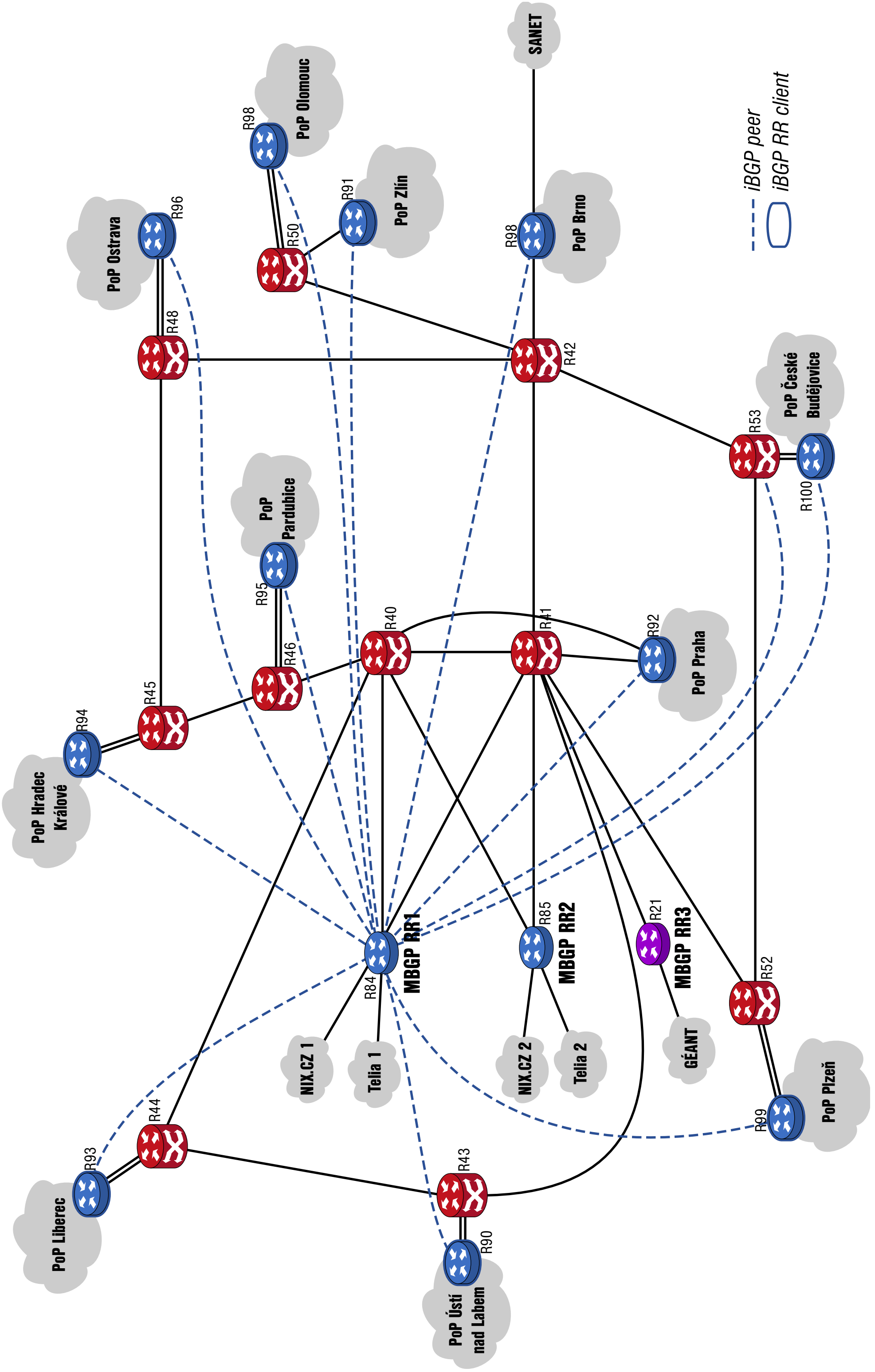


Figure 3.3: Internal unicast routing (route reflector RR1 on R84)

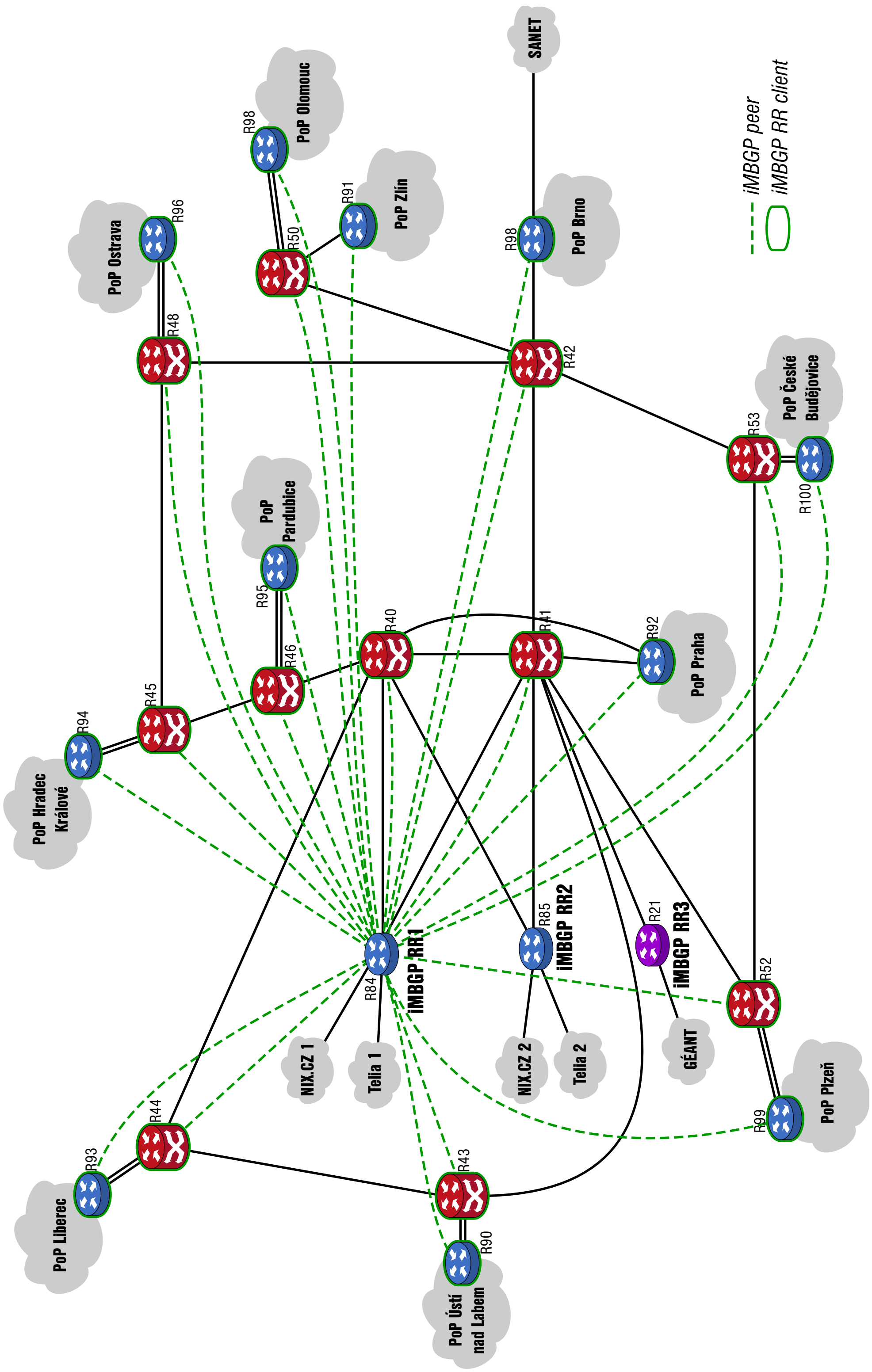


Figure 3.4: Multicast Routing (iBGP RR1 on R84)

data transfer proceeds towards RP via SPT and then towards the recipient via the shared tree. However, the data transfer using the shared tree is often not optimal and so a path optimization can take place – the last hop router switches over to the optimal shortest path tree.

The RP in each of the multicast domains is used for all connected networks (with an exception of PASNET and ČVUT). An RP is elected dynamically through the use of the Auto-RP protocol (Cisco proprietary protocol), or the BSR (Bootstrap Router) protocol. The latter is an IETF standard implemented in the routers of other producers. While most participants are already connected over a PIMv2-SM interface, some participants are unfortunately still connected over PIMv2-DM, mostly because of the limitations imposed by the technology they use (for example, Extreme Networks switches). As the SM/DM boundary creates a number of problems, we ultimately aim at using PIMv2-SM exclusively.

We use MSDP protocol in the full-mesh group configuration between all nodes for propagating information about active sources of multicast data. This configuration is rather complicated and difficult to administer. The use of the mesh group reduces the volume of reports that have to be exchanged between the MSDP peers and enables the exchange of SA (Source Active) reports between all iMSDP routers regardless of the RPF check mechanism.

RPF check failures in relation to MSDP, resulting in SA messages being dropped, was the main source of problems with multicast distribution in the backbone network. The reason behind is that, logically, the unicast and multicast topologies are not congruent since unicast is encapsulated in MPLS whereas multicast is transported as plain IP packets.

Following the recommendations from Cisco Systems and the GÉANT network, we have also set access lists for MSDP reports filtering out certain reported active sources, thus eliminating the unwanted traffic resulting from misconfigured protocols and applications like Novell NDS, ImageCast and others.

At present, we prepare a new backbone setup for IPv4 multicast in cooperation with Cisco NSA. The expected implementation of the Anycast RP mechanism should lead to a more reliable multicast operation – its advantage is the possibility of load-balancing and setting up redundant RPs. In this case, the RP configuration is entirely static, which should make troubleshooting easier and the network more resistant against flooding.

Unfortunately, the implementation of the Anycast RP technology will have a considerable impact on the configuration of multicast in the off-backbone nodes and further in the networks of the connected organisations, due to the necessity of a manual RP configuration and the use of MSDP protocol on their upstream interfaces. We also plan to install the HP OpenView NNM Multicast Monitor on our second monitoring station for monitoring the backbone multicast operation.

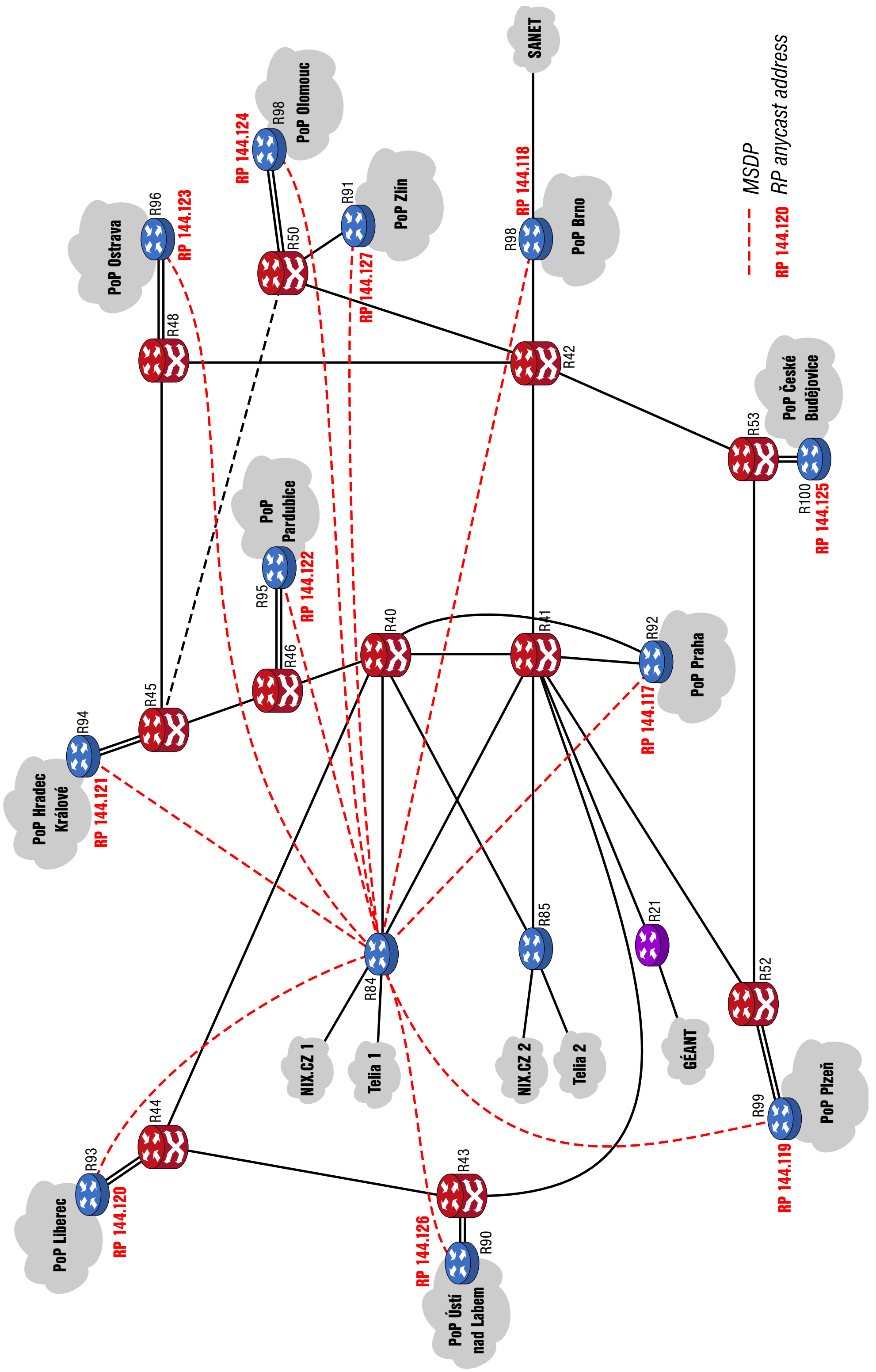


Figure 3.5: CESNET2 multicast topology

The statistical evaluation of multicast traffic requires that the routers support both SNMP and NetFlow v9.

While multicast distribution on the level of the network backbone is essentially flawless, many problems still persist on the side of the connected participants (diverse technologies, incompatibilities etc.). Therefore, we cannot yet claim to have a reliable end-to-end multicast service and potential users of multicast applications (e.g., videoconferences) thus often resort to alternative unicast-based technologies.

3.4 IPv6 Implementation

In 2003, one of the important goals was to integrate the experimental IPv6 backbone, which was based mainly on PC routers connected by tunnels, into the standard production environment of the backbone network. Although hardware support for IPv6 forwarding was one of the tender requirements already in mid 2002, by the end of 2003 the selected OSR7609 platform still lacks this functionality, mainly due to the Supervisor 720 engine being seriously delayed. Consequently, we had to look for alternative solutions.

A beta version of IOS with IPv6 software support, which was offered as an interim solution for the current Supervisor2/Engine2, turned out to be unusable in the production network and, as a matter of fact, the producer stopped its development. A reasonable solution thus seemed to be to use the existing Cisco 7500 routers, which had been replaced by the new OSR7609 routers during the backbone upgrade. The 7500 routers, configured as so-called 6PE devices, encapsulate IPv6 in MPLS packets and and forward them to the OSR7609 routers via a VLAN – see Figure 3.6.

The advantages of this solution are:

- easy implementation on dedicated routers
- no need for reconfiguring or upgrading the backbone network
- combination with tunnels still remains an option
- a good router performance due to CEF6

During the deployment of 6PE we encountered a number of various bugs and IOS-related problems. At present, the routers run an engineering version derived from IOS 12.3.

IPv6 multicast has not been supported yet on the backbone. The current architecture requires that the external neighbours be treated as CE routers connected to our 6PE routers. This turned out to be impossible for the existing IPv4 peering with the GÉANT network where we had an OSR7609 router (R85) with a

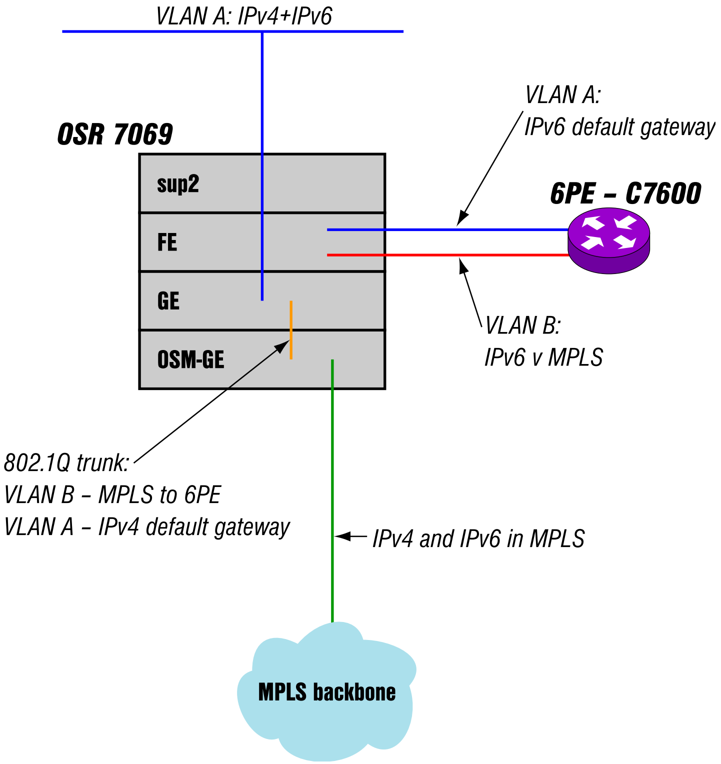


Figure 3.6: Description of the interim IPv4/IPv6 implementation

POS STM-16/OC-48 interface. The only chance for a native IPv6 peering was to use temporarily a Cisco GSR12008 router (R21) with IOS 12.0(25)S1 supporting IPv4/IPv6 dual-stack. The current IPv4/IPv6 topology is shown in Figure 3.7.

We use iBGP as an internal routing protocol for IPv6 (as well as for IPv4) with two route reflectors on the external routers R1 and R62. The R1 router is also used for IPv6 peerings towards NIX. IPv6 traffic is merged with IPv4 on an L2 switch (Cat43), where the second GE circuit is also terminated. Dual-stack is configured, apart from R21, on the routers in Ústí nad Labem and Zlín (R90 and R91) – both are Cisco 7206-VXR with NPE-G1 and IOS version 12.(3)1. Connections of end sites are implemented either by tunnels (if the end site is not IPv6-ready), or natively (e.g., over a dedicated VLAN) using static routing.

The connection of smaller PoPs outside of the MPLS core network, will be also implemented natively and OSPFv3 will be used as the internal gateway protocol. At present, we use this configuration between GigaPoP Ostrava and PoPs Karviná and Opava. However, IOS version 12.2(15)T is needed, which requires at least 128 MB memory and is not available for older routers like Cisco 2600 that are still servicing a number of smaller PoPs. Therefore, it will be necessary to upgrade these routers in order to further extend native IPv6 backbone operation.

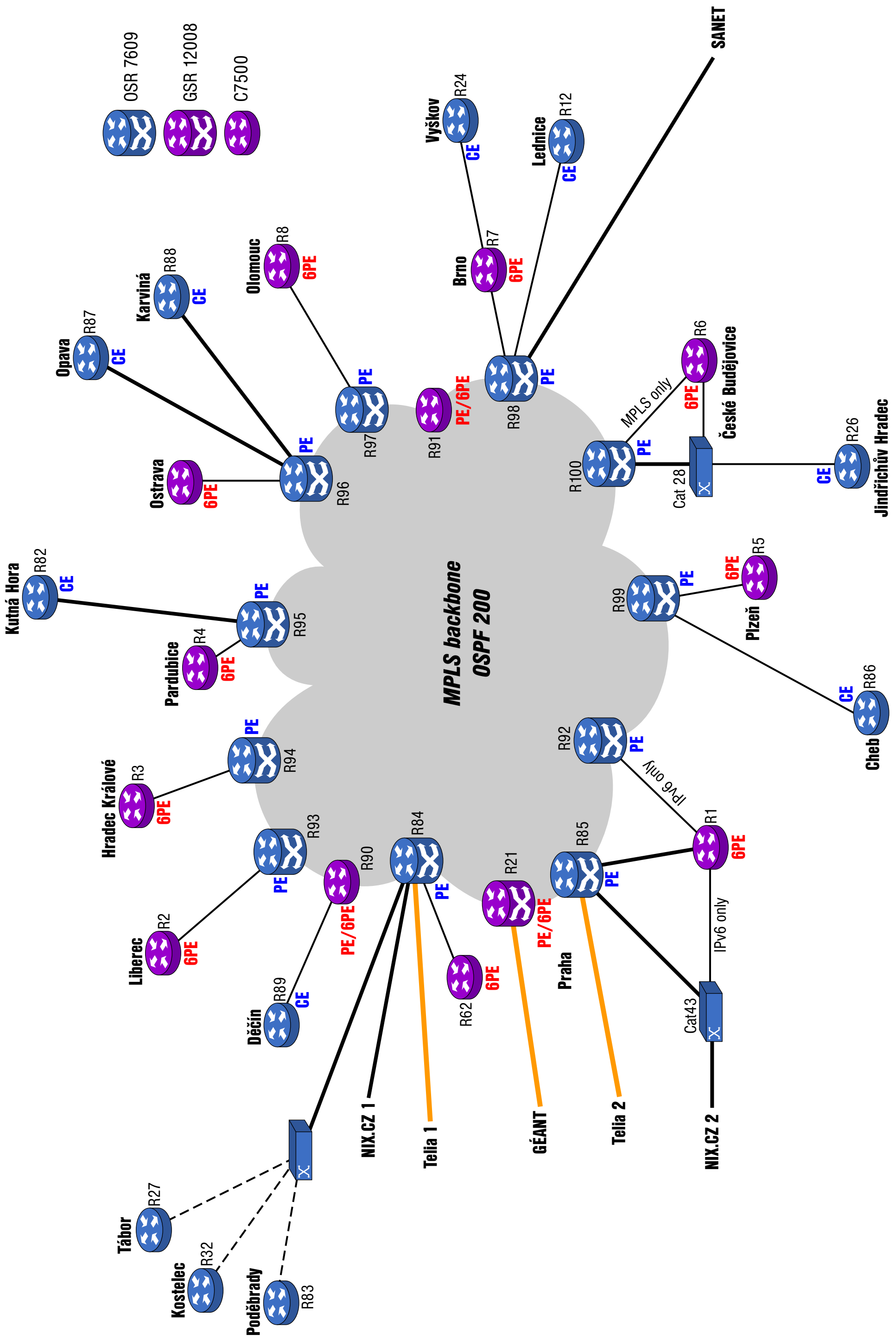


Figure 3.7: IPv4/IPv6 MPLS backbone topology

The necessary condition for the transfer of the entire backbone network to the dual-stack IPv4/IPv6 operation is the upgrade of all Cisco 7609 routers to the new supervisor engines Sup720 together with a stable IOS supporting 6PE. Sup720 has higher demands on both the input power and cooling compared to the current Supervisor2/PFC2 engine. The existing OSR7609 chassis cannot cope with these requirements and have thus to be exchanged as well.

A stable IOS release supporting Sup720 is not available either, but is planned for early 2004 under the code name 12.2S Tetons-2. The vendor provided us with a beta version of IOS, which we currently test on the standby devices. We have been communicating the problems discovered during this early Field test (EFT) directly to the developers so that they can be corrected. The aim of our participation in this EFT is first to get acquainted with the features of the new IOS, and second to help debug and finalise the awaited production version. We are ready to deploy it in our backbone as soon as it proves to be stable.

During the upgrade, we will have to rearrange all modules in the chassis because the existing Supervisor2/PFC2 occupies positions 1 and 2 while Sup720 must be placed in positions 5 and 6, originally reserved for the switching matrix. Configurations will also have to be appropriately changed.

After the upgrade, the OSR7609 routers will assume the role of the existing 6PE routers Cisco 7500, though only for the unicast IPv6. IPv6 multicast support is currently unclear and we plan to utilise the old Cisco 7500 routers again, this time as tunnel end-points (this solution is currently being verified).

3.5 External Connectivity

CESNET2 network uses the following external connections and peerings (see Figure 3.8):

3.5.1 Foreign Connectivity (Commodity Internet)

Telia International Carrier provides the global foreign connectivity. The connection capacity is 800 Mbps (software-limited) on a POS STM-16/OC-48 circuit terminated at the R84 router. The backup connection is realised by the analog circuit from the Telia node in Prague to the R85 router.

Telia provides the unicast connectivity for IPv4 (IPv4 multicast is planned) and IPv6 (a tunnel connection to London).

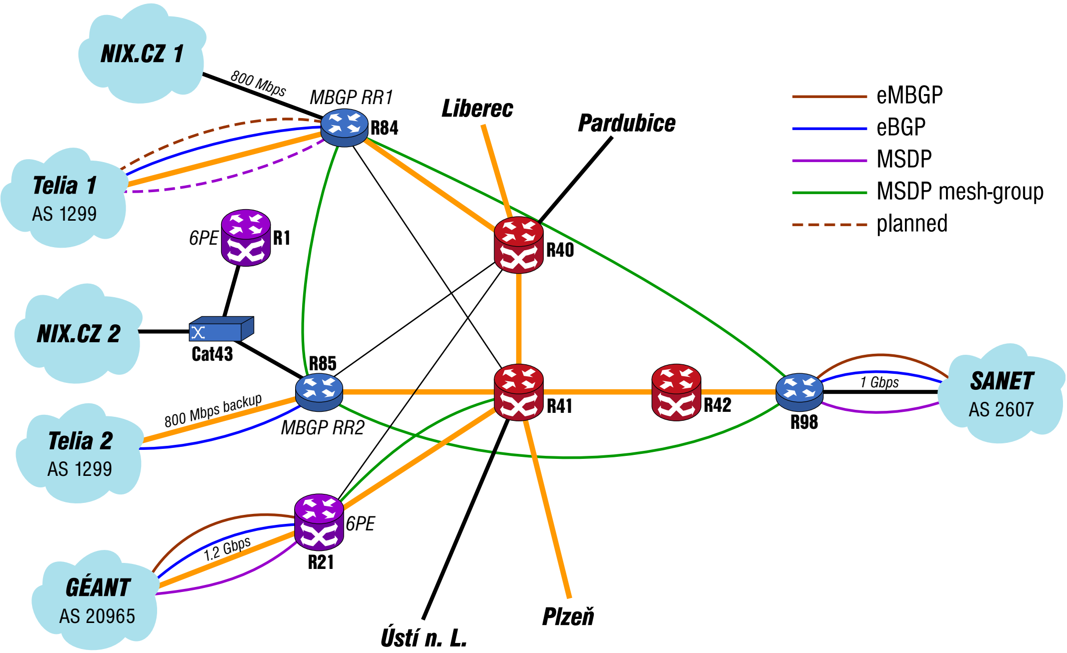


Figure 3.8: External connectivity of CESNET2 network

3.5.2 GÉANT Network Connection

Our connection to the GÉANT network is realised by a POS STM-16/OC-48 circuit to the collocated GÉANT node. The capacity of this connection is limited to 1.2 Gbps. On our side it is temporarily terminated at the dual-stack IPv4/IPv6 router R21 (GSR12008) in order to allow the native IPv6 connection. IPv4 multicast is routed through this connection as well.

The GÉANT node in Prague has a 10 Gbps connection to Germany (Frankfurt am Main) and two POS STM-16/OC-48 circuits to Poland (Poznań) and Slovakia (Bratislava).

The link to GÉANT is mainly useful for a communication with national research and education networks (NRENs), but generally not with other pan-European Internet providers and peering centres. The detailed information on the GÉANT network topology can be found at *www.geant.net*.

3.5.3 National Peering in NIX.CZ

NIX.CZ access is realised by two GE circuits terminated at two border routers R84 and R85 (for back-up purposes). The circuits are set up on leased optical fibres using corresponding GBICs (GBIC-LX/LH and CWDM-GBIC on the longer backup line). Native IPv6 peering is implemented via the 6PE router R1.

3.5.4 Peering with SANET

The connection to Slovak academic network SANET has uses leased optical fibres Brno–Bratislava. The link is equipped with CWDM-GBIC-1550 and a Catalyst 3524 switch is used instead of a repeater.

We use 802.1Q VLAN for the device administration and point-to-point links for connecting the border routers in Brno (R89) and on the SANET side. Through this peering we also provide SANET access to NIX.CZ and SANET provides CESNET2 reciprocally access to the Slovak peering centre SIX, saving both networks some capacity on their international links. Slovak networks are advertised to NIX.CZ and Czech ones to SIX tagged with appropriate BGP communities.

3.6 Backbone network administration

The backbone network is being continuously supervised by the CESNET NOC (Network Operating Centre) on a 24/7 basis. Network problems that require a detailed diagnosis and/or reconfiguration of active elements are reported to the administrator in service by the NOC operators. CESNET has enhanced hardware and software service agreements covering all network devices with a guaranteed time for resolving the defects (for key devices it is 4 hours). Cisco NSA (Advanced Services) support is used for escalating the pending issues if necessary. As part of this support, two TAC engineers have been assigned to the CESNET2 network with a detailed insight into the network topology, services used and a direct access to network devices.

We use the following tools for the backbone network administration:

Management of the whole backbone network: The real-time network status is monitored and predefined events observed by means of an *HP OpenView NNM 6.31* monitoring station (UltraSparc 420R, Solaris 2.8) equipped with the ET (Extended Topology) license. Backup station for network administration is planned to be used for multicast monitoring. To this end, an installation of the *HP OpenView NNM Multicast monitor* is currently under preparation.

Management of active network elements (routers, switches,. . .): We use *CiscoWorks 2000 RWAN 1.3* over the ssh protocol for secure access to network devices. The most used functions from the CiscoWorks portfolio are the following: RME (Resource Management Essential) for creating device surveys and automatic configuration backup, and Syslog Analyzer for evaluating specific records and logs about the operation of network devices.

Network Services Monitoring: We also use the *Nagios v1.0* program (*www.nagios.org*), which is a follow-up of the previously used *NetSaint*. Our installation primarily monitors availability of the network servers (mail, DNS, WWW). The Nagios system server is also used for monitoring the IPv6 network. The advantage of Nagios system is the fact that it is an open-source system and we can easily add our own changes and extensions.

Statistics about devices and network traffic: *GTDMS* is a system developed in-house. Its main purpose is to gather different statistics about the operation and performance of network devices. We also added a number of alarms that fire when certain router- and circuit-related thresholds are exceeded (CPU overload, free memory size, feeding sources, interior temperature, congestion and error rate on circuits). In addition, we use our own system called *NetFlow Monitor* for processing and viewing the statistics (see below). It is intended first for assessing the traffic volumes recorded for individual member and non-member institutions, and second for tracking security incidents (evaluation of the observed flows according to predefined conditions). All backbone PE routers export NetFlow data. At present, we are able to collect only IPv4 operation statistics because the implementation of NetFlow v9, which includes statistics about IPv6 and multicast operation, is not available yet for the production versions of IOS.

Request Tracker (RT): The role of this trouble ticket system is to process various requests (their creation, monitoring and archiving) pertaining to almost all areas of network operation and administration. A detailed description of the RT software can be found at *www.fsck.com*. Messages from each of the specialised queues (noc, trouble, admin,...) are distributed to a designated group of receivers (network administrators, NOC or users).

Out-of-Band management (OOB): Remote access to all active network elements in all backbone nodes is guaranteed even in the case of network failures.

3.7 Statistical Traffic Analysis

3.7.1 Average Long-term Utilization of Backbone Network Core

At the first sight, the core of the CESNET2 backbone seems to have enough free capacity. However, we cannot simply characterise it as an "over-provisioned" network. The gigabit backbone lines have a long-term average utilization below 15 percent. Nevertheless, in 2003 we have been observing much more frequent traffic peaks that significantly exceed the average level. On certain lines these peaks cannot be considered a transient and random phenomenon anymore. In the 2002 report we illustrated the dependence of the measured line utilisation on the time step of the measurement. It turned out the actual line load over short periods highly exceeds the values obtained from the standard operational measurement, whose time steps are in the range of several minutes. During 2003 we recorded cases where the average load during the whole (usually 5–10 minutes) measurement step highly exceeded the long-term average and the sustained load on the corresponding lines was more than 30 % of the nominal line rate.

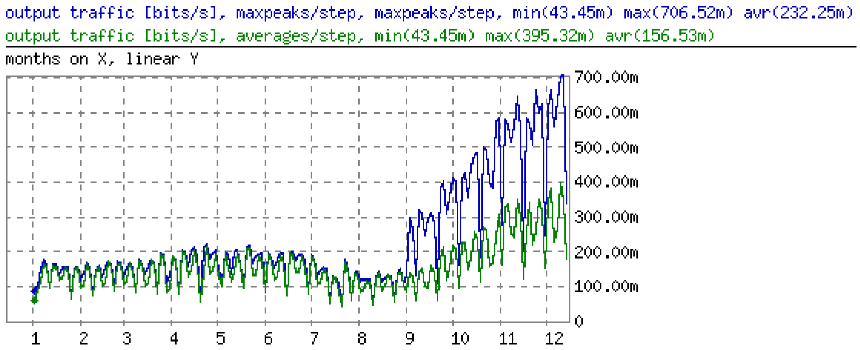


Figure 3.9: Long-time traffic peaks appearance on the Prague–Brno line 2.5 Gbps in 2003, direction to Brno

In the graph above, the noticeably less frequent appearance of traffic peaks before September 2003 is mainly due to a different aggregation strategy of aged data that was in place at that time.

3.7.2 Traffic trends in 2003

Compared to year 2002 we observed very few changes in the general trends of network utilisation. Specifically, the overall traffic volume was slowly but steadily growing with the typical plateau during the summer months. In the last three months of 2003 we can see a dynamical increase, especially on the lines that aggregate several traffic tributaries.

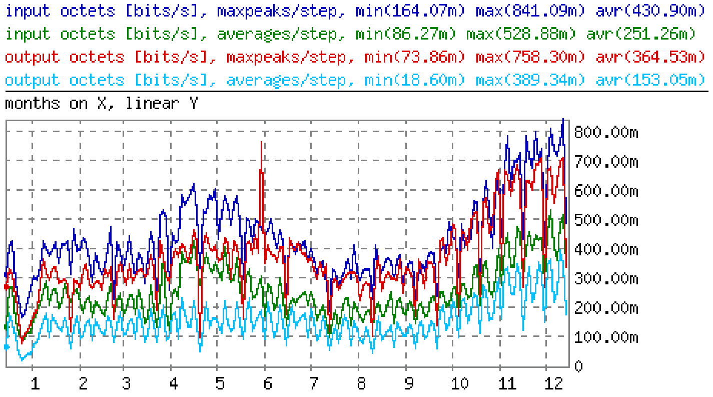


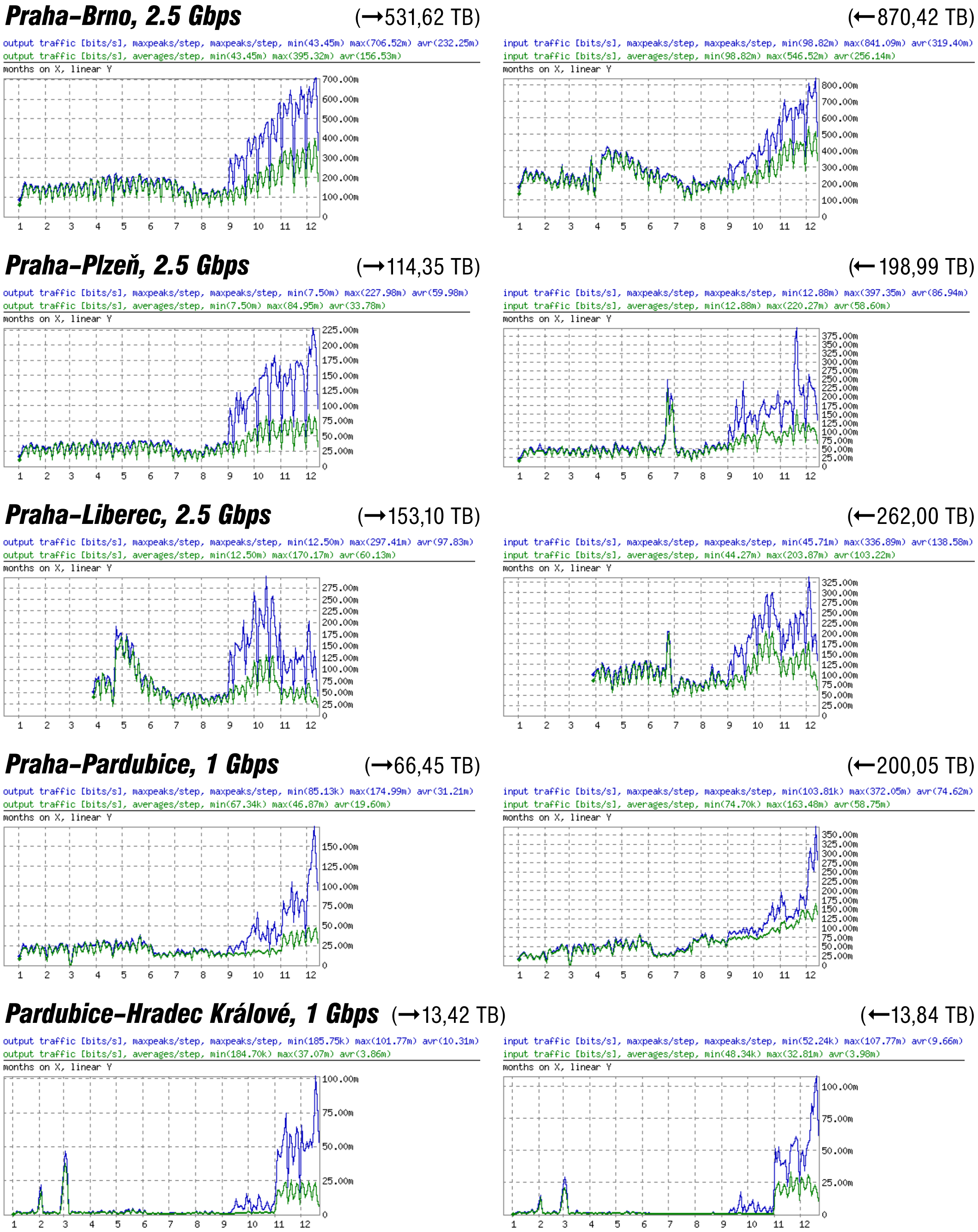
Figure 3.10: Traffic growth dynamic significant from September to November on Prague–Brno line 2.5 Gbps in 2003

3.7.3 CESNET2 network utilization in 2003

Some backbone core lines experienced a noticeable growth of the total traffic volume. For the sake of clarity we plotted volumes for both traffic directions separately including, whenever possible, the total amount of transferred data. The missing parts of some graphs are caused by changes in network architecture that took place during the year (topology, line changes) or by changes in the measurement system configuration. The low frequency of peaks before September 2003 is again an artefact induced by a change in the strategy of aged data aggregation in the measurement system.

3.7.4 External lines

All external lines have enough free capacity, the only exception being our connection to the global Internet in the outgoing direction. This line is rate-limited by software means, hence some traffic peaks can exceed the configured maximum. As before, the gaps in the graphs are caused by problems in the measurement system rather than traffic outages. The data, including the total volumes of transferred data, are again plotted separately for each traffic direction.



Praha–Brno, 2.5 Gbps (→531,62 TB) (←870,42 TB)

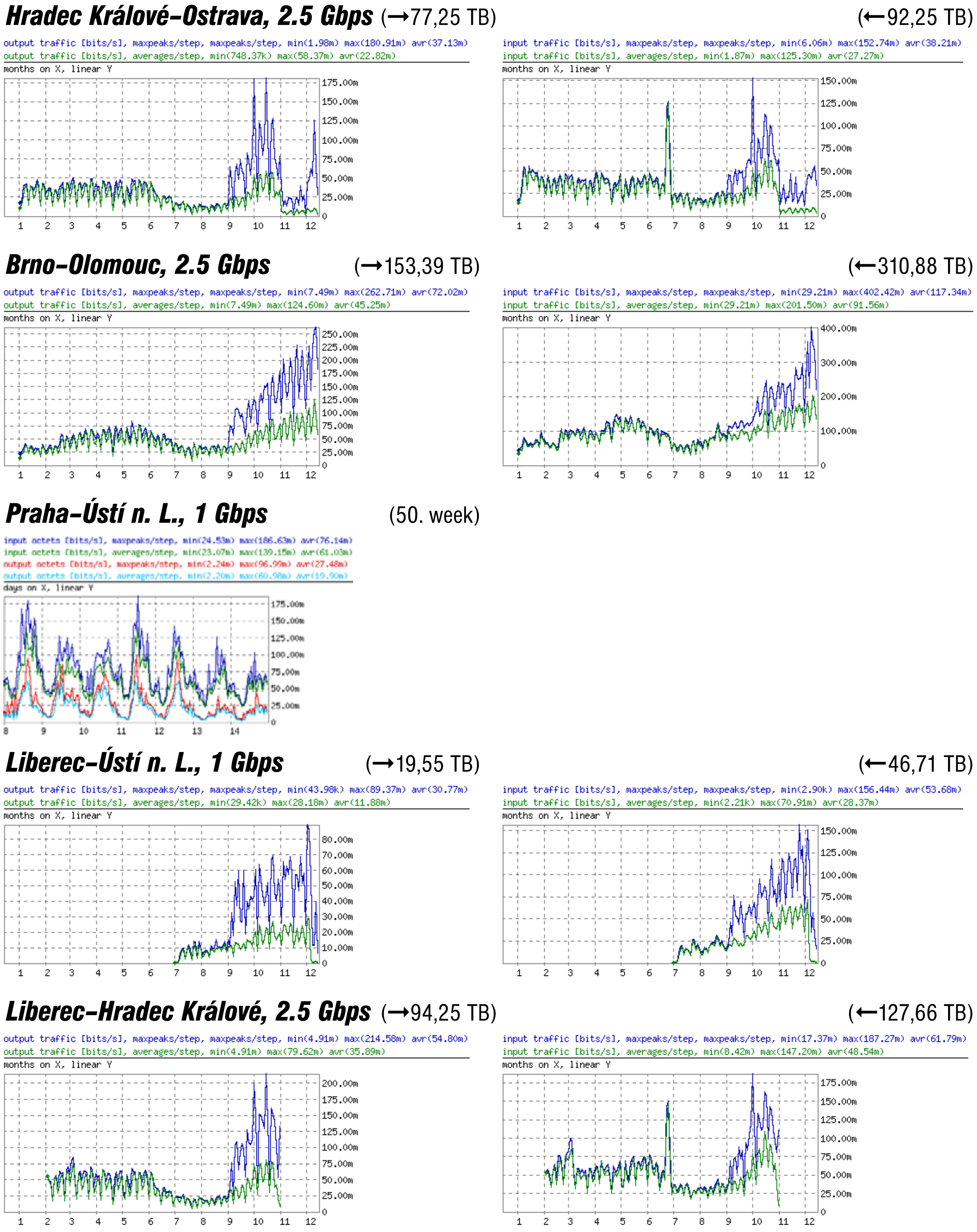
Praha–Plzeň, 2.5 Gbps (→114,35 TB) (←198,99 TB)

Praha–Liberec, 2.5 Gbps (→153,10 TB) (←262,00 TB)

Praha–Pardubice, 1 Gbps (→66,45 TB) (←200,05 TB)

Pardubice–Hradec Králové, 1 Gbps (→13,42 TB) (←13,84 TB)

Figure 3.11: Load on backbone lines in 2003



Hradec Králové-Ostrava, 2.5 Gbps (→77,25 TB) (←92,25 TB)

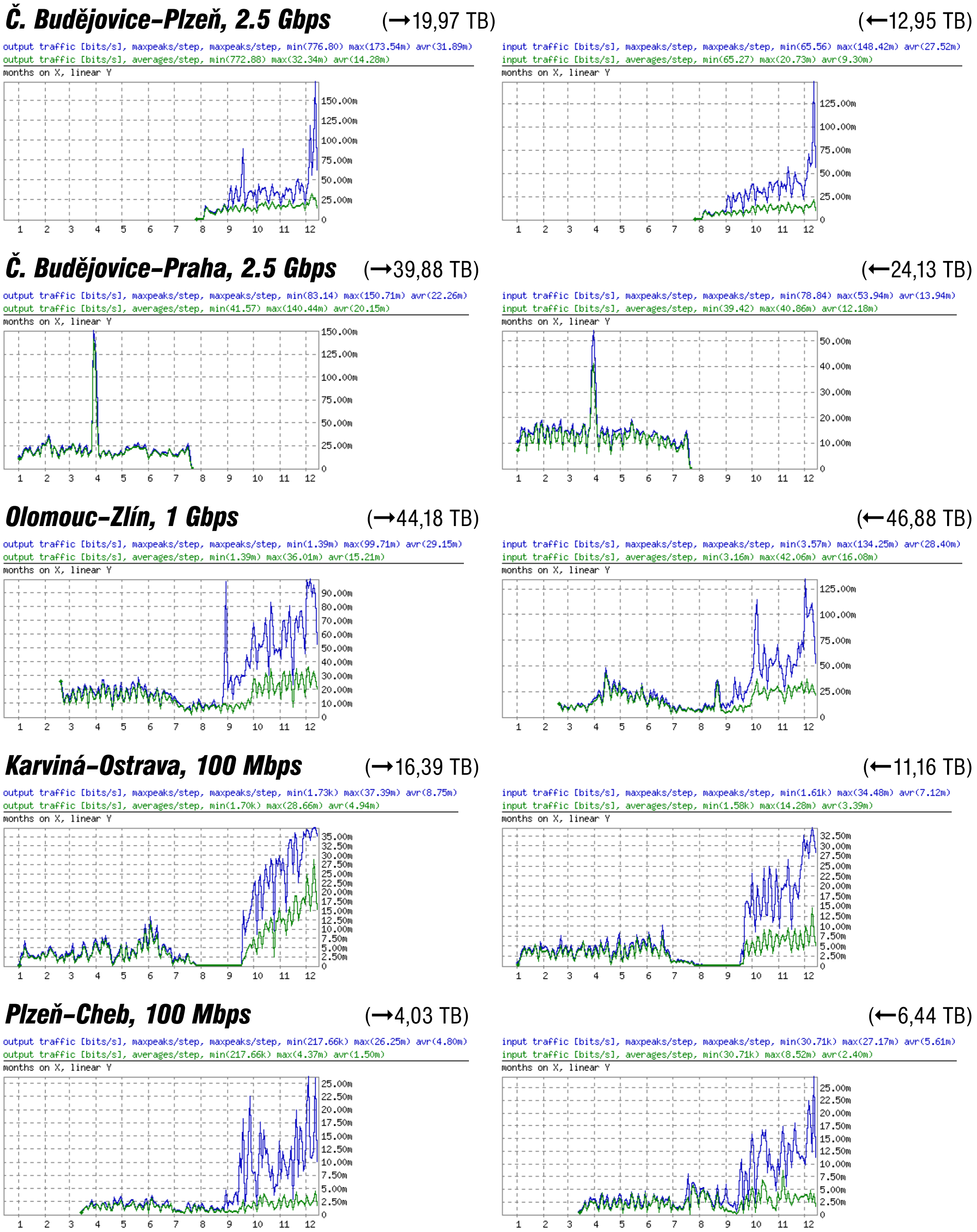
Brno-Olomouc, 2.5 Gbps (→153,39 TB) (←310,88 TB)

Praha-Ústí n. L., 1 Gbps (50. week)

Liberec-Ústí n. L., 1 Gbps (→19,55 TB) (←46,71 TB)

Liberec-Hradec Králové, 2.5 Gbps (→94,25 TB) (←127,66 TB)

Figure 3.12: Load on backbone lines in 2003



Č. Budějovice-Plzeň, 2.5 Gbps (→19,97 TB) (←12,95 TB)

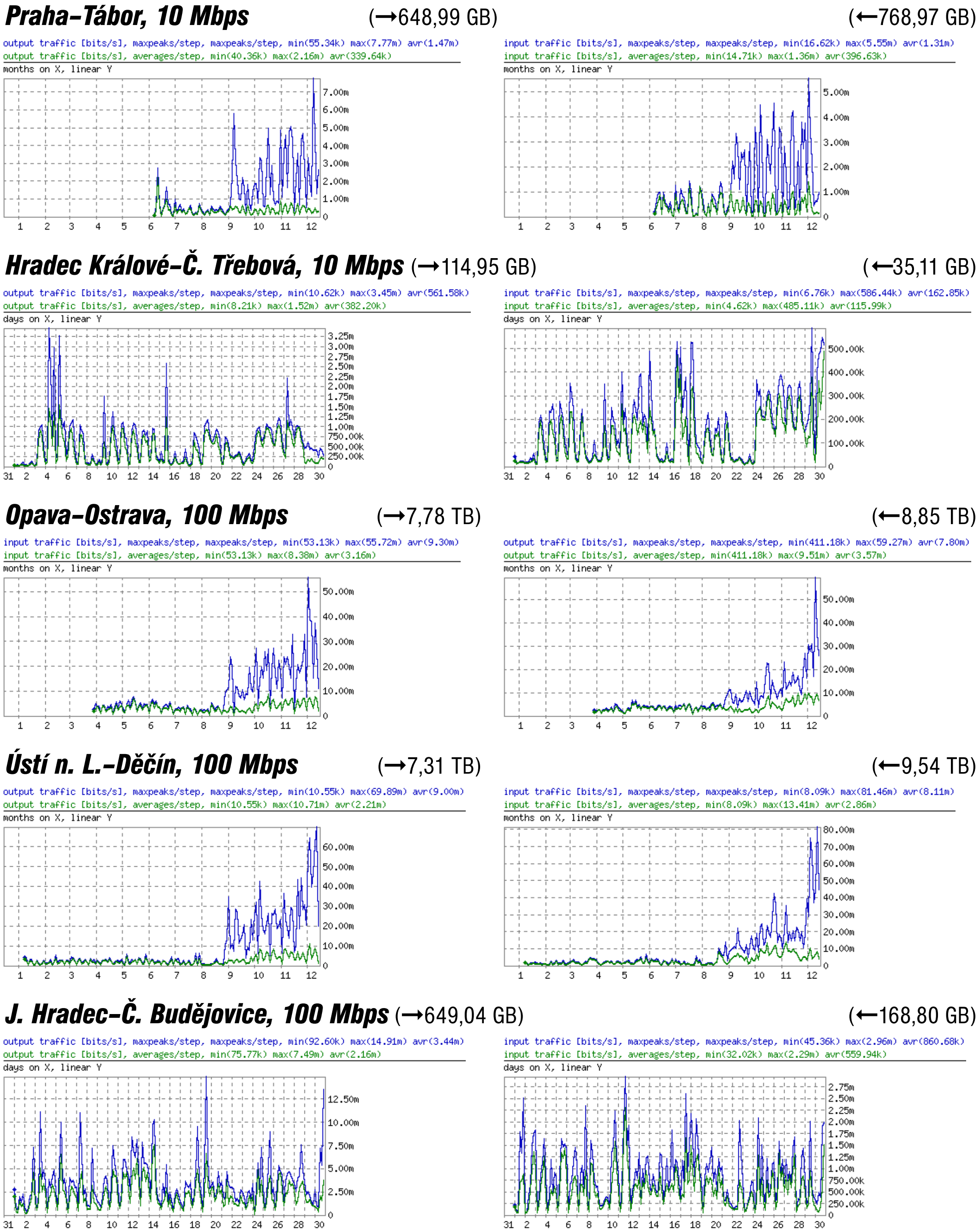
Č. Budějovice-Praha, 2.5 Gbps (→39,88 TB) (←24,13 TB)

Olomouc-Zlín, 1 Gbps (→44,18 TB) (←46,88 TB)

Karviná-Ostrava, 100 Mbps (→16,39 TB) (←11,16 TB)

Plzeň-Cheb, 100 Mbps (→4,03 TB) (←6,44 TB)

Figure 3.13: Load on backbone lines in 2003



Praha-Tábor, 10 Mbps (→648,99 GB) (←768,97 GB)

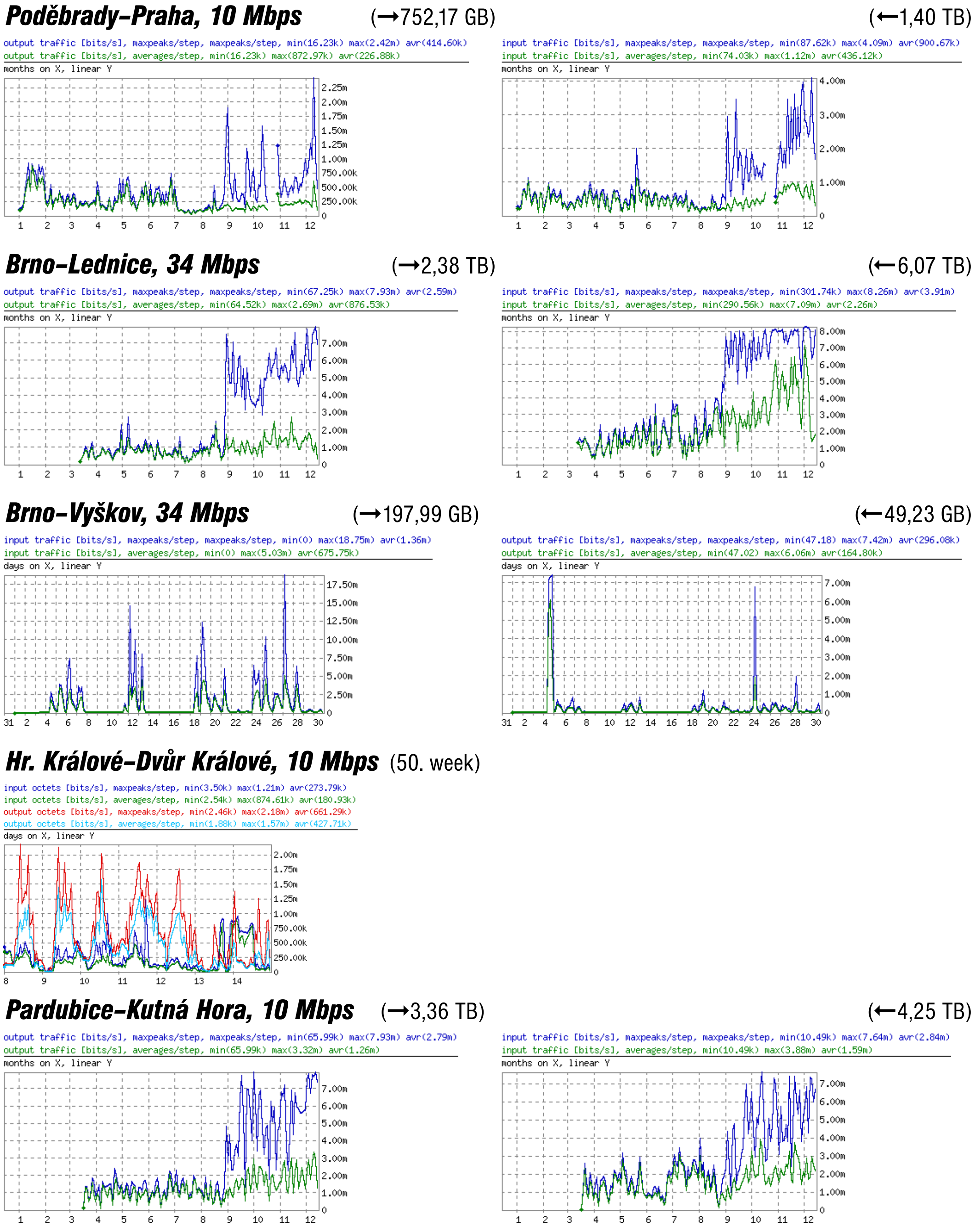
Hradec Králové-Č. Třebová, 10 Mbps (→114,95 GB) (←35,11 GB)

Opava-Ostrava, 100 Mbps (→7,78 TB) (←8,85 TB)

Ústí n. L.-Děčín, 100 Mbps (→7,31 TB) (←9,54 TB)

J. Hradec-Č. Budějovice, 100 Mbps (→649,04 GB) (←168,80 GB)

Figure 3.14: Load on backbone lines in 2003



Poděbrady–Praha, 10 Mbps (→752,17 GB) (←1,40 TB)

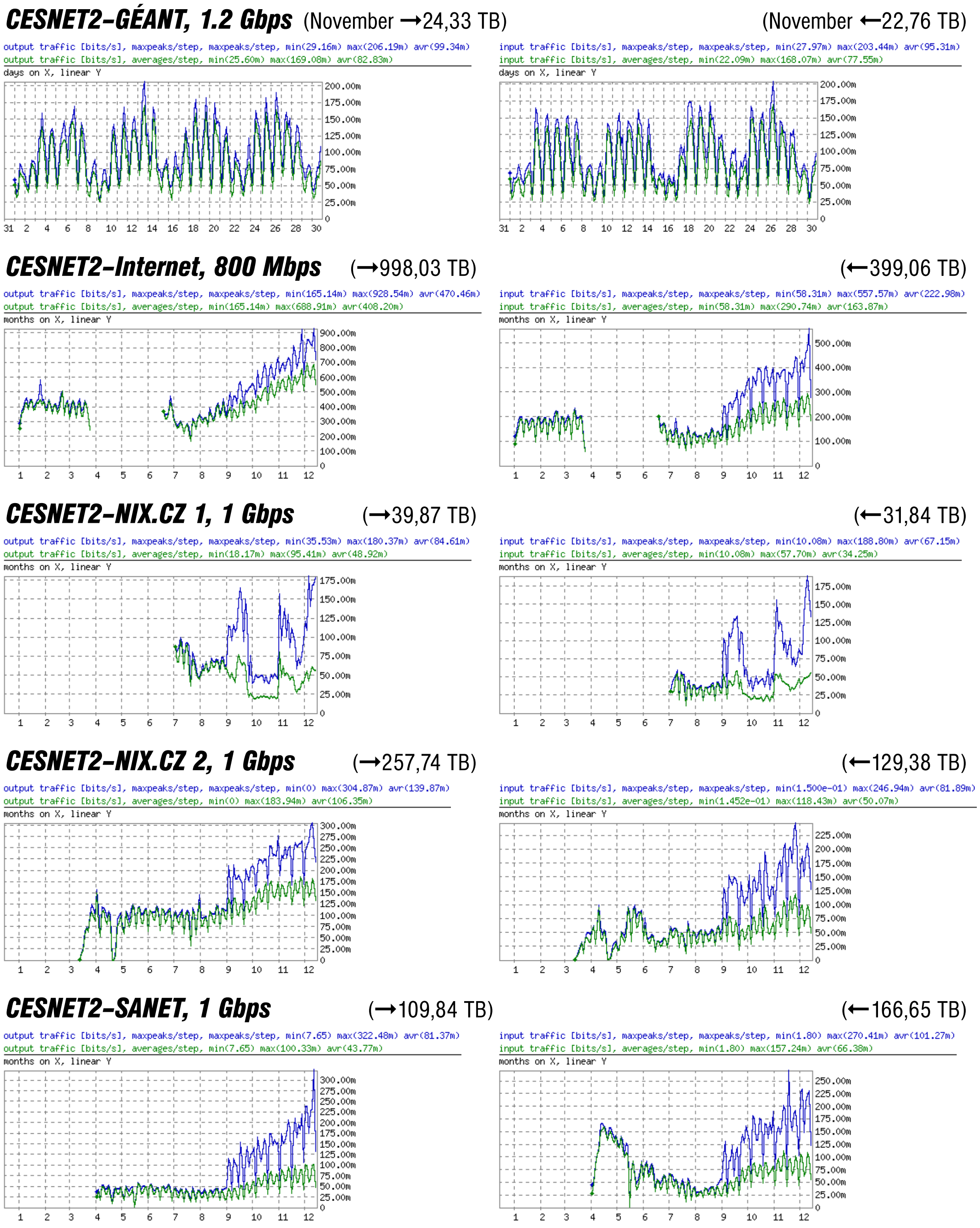
Brno–Lednice, 34 Mbps (→2,38 TB) (←6,07 TB)

Brno–Vyškov, 34 Mbps (→197,99 GB) (←49,23 GB)

Hr. Králové–Dvůr Králové, 10 Mbps (50. week)

Pardubice–Kutná Hora, 10 Mbps (→3,36 TB) (←4,25 TB)

Figure 3.15: Load on backbone lines in 2003


CESNET2-GÉANT, 1.2 Gbps (November →24,33 TB) (November ←22,76 TB)

CESNET2-Internet, 800 Mbps (→998,03 TB) (←399,06 TB)

CESNET2-NIX.CZ 1, 1 Gbps (→39,87 TB) (←31,84 TB)

CESNET2-NIX.CZ 2, 1 Gbps (→257,74 TB) (←129,38 TB)

CESNET2-SANET, 1 Gbps (→109,84 TB) (←166,65 TB)

Figure 3.16: Load on external lines in 2003

3.8 Future plans for the backbone network development

In the first quarter of 2004 we plan to finish the upgrade of the backbone Cisco 7609 routers to the Supervisor 720 engine and, consequently, reconfigure the PE routers for full IPv4/IPv6 dual-stack operation. Another important change will be the new architecture of multicast operation and administration. Regarding IPv6, we will continue the deployment of native IPv6 unicast in the remaining PoPs and also implement the interim solution for IPv6 multicast using Cisco 7500 routers and tunnels.

We can provide other services that may be demanded either by projects or connected institutions, such as MPLS VPN, Ethernet over MPLS (see [MTV01]) or Quality of Service. As concerns the latter, so far there has been very little motivation for implementing QoS due to quite sufficient available bandwidth.

Further development of the optical routes involves a gradual transition to 10GE on the route Prague–Brno, including a necessary upgrade of the routers in the network core and external routers. Other technological changes will be mostly driven by specific needs of both research projects and connected institutions.

4 Authentication and Authorization Services

4.1 Authentication and Authorization Services

The development of the CESNET2 network and its applications posed new requirements on the authentication and authorization services.

4.1.1 The Central Authentication and Authorization System

The Central Authentication and Authorization System (*CAAS*) was implemented in 2001 to support integrated administration of users, resources and services, and access rights within the research projects. Since then, the CAAS provides for authentication and authorization services to WWW applications and access control for the backbone routers and other equipment. In 2003 the system has been extended to support control of terminal access to host computers and access to some centrally managed services (e.g., e-mail).

Architecture

The *CAAS* data are stored in several LDAP servers. In 2003 we upgraded the existing installations from *iPlanet Directory Server* version 4.1 to *Sun ONE Directory Server* version 5.2. The master server is used for active data management through the WWW interface. Data modifications are replicated on-line to replicas that serve authentication and authorization requests sent by individual network services.

Main replicas are operated on hosts *ldap1.cesnet.cz*, *ldap2.cesnet.cz*, and *ldap3.cesnet.cz* that are located in important network nodes (Prague, Brno, Ostrava). *TACACS+* servers running on these hosts provide access control services to backbone nodes (*NAS*). Remaining network services that use CAAS (HTTP servers, CESNET IMAP server, shell access to operating systems. . .) communicate with the replicas via the secured LDAPS protocol.

WWW interface for CAAS administration

WWW interface for CAAS management is built upon Perl libraries *perlOpenLDAP* (the LDAP protocol), *myPerlLDAP* (object oriented interface to LDAP data

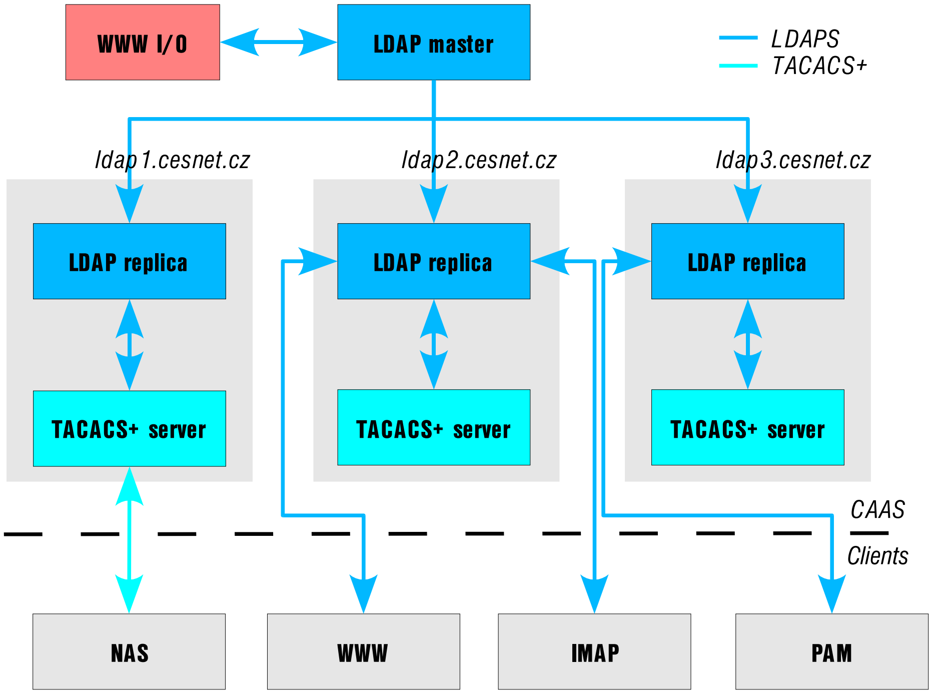


Figure 4.1: CAAS Architecture

objects), and *myAppFramework* and *myForms* (for WWW applications programming). The presentation level, appearance, and localization is controlled by XSLT stylesheets. Most of the applications are localised to both Czech and English languages. The run-time environment for the applications is provided by the Apache HTTP server with the *mod_perl* module operated on a dedicated host (see [Sov02]). Secured LDAPS protocol is used for the communication between the application and the master LDAP server.

Two new applications have been developed in 2003: the group management application enabling group managers to modify their groups' membership lists. The user management application has been created for the Human Resources Department personnel.

CAAS Clients

The CAAS was deployed to serve various access control clients. Backbone nodes enforce access control based on the communication with *TACACS* servers. The access control rules and authentication data are stored in the CAAS LDAP database. WWW applications are served by our *auth_ldap* and *auth_ldap_x509* Apache modules. WWW users can be authenticated either by their personal X.509 certificate or by their user name and password. Both authentication methods map the user to a corresponding LDAP entry which is then used for an authorization decision. Shell/terminal access to operation system services is controlled by the standard *PAM-LDAP* module.

4.1.2 CAAS Reconfiguration

After three years of continuous operation, the CAAS was substantially reconfigured in 2003. We upgraded the LDAP servers, increased the number of replicas and dedicated the master server exclusively to data modification. The reconfiguration process was also an opportunity to change the data information tree suffix. We have moved from the original "classic" suffix *o=ten.cz* based on ITU X.500 recommendation to the "modern" directory component-based format – *dc=cesnet,dc=cz*. The new suffix should facilitate integration of our directory services into the national and international contexts. At the same time, all of the LDAP servers were moved from the *ten.cz* DNS zone to the *cesnet.cz* zone.

The reconfiguration was challenging not only for technical reasons but mainly for its organisational demands. Uninterrupted operation of all CAAS services during the whole period of reconfiguration was an absolute requirement.

We accomplished the reconfiguration in the following steps:

1. New master server installation on a new server.
2. Creation of DNS aliases for all LDAP servers in the *ten.cz* zone pointing to new names in the *cesnet.cz* zone.
3. Sequential re-installation of both existing LDAP servers so that all the time at least one of them was providing CAAS services. The servers were initialised with the new DNS names.
4. The original master database was write-locked and exported.
5. The exported data was modified to reflect the suffix change.
6. The modified data was imported to the new master database.
7. The replicas were initialised for the new suffix.
8. New master database was write-locked. At this point of time the CAAS servers were accessible both under the new and old DNS names. The data was stored under both suffixes.
9. The administrators of all services were asked to reconfigure their systems so that they use the new DNS names and the new suffix.
10. After the reconfiguration of the dependent services the old suffixes were read-locked. Thanks to the excellent cooperation of all parties this step was finished in two days.

11. The new master server was opened for data modification.

12. Installation and initialisation of the third replica.

4.2 Certificate Authority CESNET CA

Experience from the deployment of Public Key Infrastructure (*PKI*) gained by the support of Czech users and researchers of the *DataGrid*[1] project and the pilot PKI project for CESNET members led to a new remarkable achievement. The *CESNET Certification Authority (CESNET CA)* was officially opened for the Czech academic community on April 1st, 2003.

The extension of the user community mandated the creation of a new certification policy – *CESNET CA Basic Level Certificate Policy*[2] and modifications to the *CESNET CA Certificate Practice Statement version 1.2*[3]. These documents reflect the requirements specified within the *DataGrid* project as well as the results of the pilot project that provided PKI services to selected universities.

Our experience gained from operating a certification authority open to a wide academic community positively influenced the work of *DataGrid*'s *WP6-CA Managers* workgroup. We are so far the only member of this group running such an open CA – all others are providing CA services just within their grid projects. It was only by the end of 2003 that the possibility of incorporating other open certification authorities started to be discussed. We believe our active cooperation within the group will help this process to continue without major obstacles.

The *DataGrid* project was finished by the end of 2003. The original *WP6-CA Managers* workgroup will continue its work as the *Policy Management Authority*, being part of the new *EGEE* project. We intend to continue the active cooperation on this new platform aiming towards the best services to our users.

As a consequence of opening the CESNET CA to the users outside the grid projects, we had to increase the availability of the *Registration Authority (RA)* – the workplace intermediately serving the users. We prepared a new release of the RA software that enables concurrent access of several RAs to a shared database. The data are managed in an LDAP server. The main reasons for choosing LDAP instead of a relational database are:

- native access control provided by LDAP server up to the level of individual data items
- standardized and secured network interface
- the possibility of integrating RA data with other LDAP services we operate.

[1] *http://www.eu-datagrid.org/*

[2] *http://www.cesnet.cz/pki/CP/Basic/1.1/CESNET_CA_Basic_CP_1.1.pdf*

[3] *http://www.cesnet.cz/pki/CPS/1.2/CPS.pdf*

Three registration authorities operate for the CESNET CA by the end of 2003. All of them are located in CESNET premises in Prague. As the number of certification requests from users outside Prague increases, we intend to establish new RAs operated by CESNET member institutions in their localities. We plan to establish first two RAs outside Prague in the first half of 2004.

4.2.1 CESNET CA Services Utilization

By the end of November 2003, 12 Czech academic institutions were registered by the CESNET CA (an institution is considered to be registered when its first user sends in a certificate request). At the same time there were 175 valid certificates in operation, 51 of them being personal and 124 server certificates.

The statistics above show that the certificates are mainly used for server authentication under the SSL/TLS protocols. User certificates are used by GRID users and server administrators for authenticating the server certificate requests.

Such a situation is rather typical for open CAs operated by European NRENs. One of the main reasons is seen in the lack of applications using PKI for user authentication. To help overcome this we plan a massive deployment of the *auth_ldap_x509* Apache module in the majority of CESNET WWW services. This way, we intend to make the user certificates more attractive. When the number of user certificates exceeds the "critical mass", they can be used in other applications, e.g., e-mail.

One of the long-term problems preventing a wide PKI penetration to academic networks is the management of user certificates. Students typically share computers in public computer rooms and thus cannot safely store their private keys on local hard disks. A new technology that may help to overcome this obstacle emerged during 2003 – "smart" chip cards able not only to store the private keys but also to perform cryptographic operations directly in the card's processor. The private key then never leaves the card and can be securely used even on a shared computer. These cards are a promising candidate technology for the new generation of ID cards to be issued by universities for their students and employees.

However, the new ID card technology deployment has considerable demands on both investments and organisational preparation. In a close cooperation with the *ID-Cards* workgroup associating issuers of university ID cards, we set as a strategic target to prepare the transition so that the PKI-enabled ID cards can be deployed within few next years.

First experiences were shared at the workshop *Chip technologies for applications requiring identification and electronic signature* held in Prague on December 4, 2003. The lecturers from universities, CESNET and CoProSys presented their views on the requirements and possible solutions of the issue. Some of the universities started to test SPK2.5DI cards with the aim of verifying their compatibility with the existing applications. The results of the tests, expected during 2004, will be used in the process of preparing the strategy of ID cards technology transition.

Part II

Strategic Projects

5 Optical Networks and their Development

Researchers have examined mainly the development of national research and education networks (NREN) and experimental research networks, such as the global TransLight and National LambdaRail networks in the USA, and the development of optical transmission systems for these networks. The acquired knowledge was applied to work on the international project SERENATE, to the development of CESNET 2 network, to the preparation of the proposal of new research intent, to the preparation of international projects GN2, GRANDE and GARDEN and to cooperation with other NRENs and to the development of CzechLight and TransLight lambda networks.

We have applied mainly the following knowledge:

1. The transfer to fibres has strategic significance. The ownership or the right to use the fibres connecting network points of presence enables the designers to choose from a much wider scale of network solutions when they are creating one, than in the case of purchase of telecommunication service. The result should be an optimal network solution which will meet the requests of the users much more sufficiently. The number of lambdas and the speed of transmission is not restricted and charged by telecommunication operators. The transmission system is not chosen by the telecommunication operator and so a better technology which is tailor-made to the network requirements can be used.

2. The use of advanced transmission technology, which in addition has configurations and parameters chosen according to the NREN users' requirements, enables much lower costs of the network construction and operation. The use of GE and 10GE transmissions instead of SDH transmissions also lowers the costs of interface circuits and requirements of telecommunication knowledge and thus decreases the costs of both the project and the operation.

3. Applied research in the field of optical transmissions is among NREN activities unusual, but it has started bringing results making it possible to build better networks. The results are immediately applicable to the transfer of NREN to fibres.

4. The method of realisation of optical lines without in-line equipment – if it is possible – which has been introduced at TERENA conference in Limerick in 2002, has become known under the name "NIL approach" (Nothing In Line) and has turned out as very suitable for NREN. This is caused mainly

by relatively short distances between university workplaces, their ability to locate the equipment in their premises and to provide local assistance in necessary cases, while the equipment is remotely monitored and set from the network centre.

5. We expect that in near future the use of PCs with optical transceivers GE and 10GE of small size (XFP MSA) and long reach (80–130 km) supported by programmable hardware (such as COMBO6 card) for higher throughput and, if need be, completed with the equipment for optical amplification (prolonging the optical reach to double distance) or coloured transmissions (WDM), will be significant for the construction of NRENs.

6. We expect the rise of open hardware systems. Open software systems (e.g. Linux, Globus) successfully compete with products of giant companies (e.g. Microsoft and IBM) because they follow open worldwide research and development cooperation of organizations and citizens on their perfection and because they provide users with wide possibilities of adapting the systems effectively so that they meet the users' requirements.But at the same time, the supply of routers, switches and transmission systems for the Internet and also the supply of devices and instruments for many other branches is very much monopolized and there is no competition in open solutions for them yet. This makes the prices very high, restricts the possibilities of development and discourages from the applications with special requirements of safety, security of personal data, price, etc.

The transfer of CESNET2 production network to fibres was in the main completed in 2003. In January 2004, NREN in the Czech Republic is going to have 2,354 km of operating double-fibre lines and 360 km of single-fibre lines, which is the total of 2,714 km to approximately 10.2 mil people and 78,866 km^2 of area. For comparison we note that the most advanced countries in applying fibres in Europe are Poland, Slovakia and Switzerland. For NREN, Poland has its 2,600 km of fibre to 38.7 mil people and 312,683 km^2 of area, in Slovakia NREN has its 1,370 km of fibre operating to 5.4 mil people and 49,035 km^2 of area, in Switzerland it is 1,200 km of NREN's operating fibre to 7.3 mil people and 41,293 km^2 of area. Next fibres are in most cases being prepared for use or they are booked. In other European countries, the use of fibres for NREN is lower, in some of them the transition to fibres is still being considered. In GÉANT network fibres are not used. In the USA, the dedicated fibres for the national production network (Abilene2) are not used but the use of fibres for the national production networks of individual states of the union is very frequent and 2,600 miles of fibres are booked for the national experimental network National LambdaRail with the possibility of the network's widening (only a small part is now in operation).

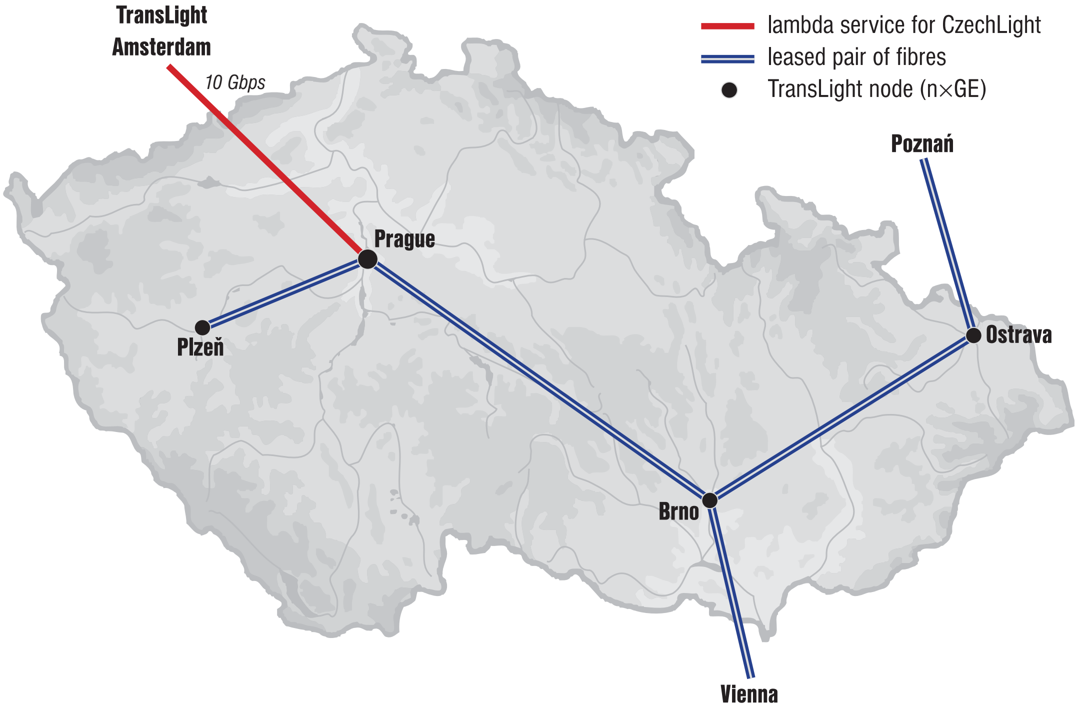


Figure 5.1: CzechLight 2004 (expected topology)

In the Czech Republic, we have preferred transformation of the CESNET2 production network on fibres in consequence of price level. We have started preparation of intercity fibre connection for CzechLight in end of year 2003. For year 2004, international fibre lines between Prague and Vienna and between Prague and Poznań (see Figure 5.1) is prepared; realisation depends on cooperation and support from our partners, Cisco Systems an EU.

5.1 Intention of the Research and Development of Optical Networks and the Main Results

The researchers have focused on the following activities and they have reached the following main results:

1. The transition of the main lines of CESNET2 network to leased optical fibres which enables the development of transmission parameters of the network independently of the providers of telecommunication services. The use of the fibres gradually puts its way through in the world as the way

of NREN perfection and the Czech Republic belongs to the first countries engaged. The results in the reconstruction of CESNET2 network have been reached thanks to the cooperation with the network development and operation team.

2. The lease of the fibres including the transmission facilities is for NREN in the Czech Republic much more advantageous than the purchase of gigabit telecommunication service. Moreover, a favourable situation and well-specified tenders for the fibre leasing lead to price decrease by 50–75 %.

3. The decrease in prices made it possible to change the topology of the network so that the access to the nodes is better (backbone nodes are accessible by means of at least two gigabit circuits) and the reliability of the provided service is much higher.

4. Gaining the option of the realization of the fibre first mile to the members' workplaces to order (which is now among NREN unusual); physical access of the most important nodes of CESNET2 network by fibres of 3–4 different owners.

5. Economically effective connection of the members in Děčín , Cheb, Jindřichův Hradec, Karviná and Opava by means of single fibre lines. At the same time, it is a step forward to decrease the difference between the access to information service and the chances of attending contemporary projects of the research and development among Czech regions.

6. Creating the first NRENs connection in Europe by means of dark fibre (Brno–Bratislava); gaining similar offers from Vienna, Munich, Frankfurt a. M., Nurnberg, Dresden, Berlin, Poznań and Bielsko-Biala. Some of them will probably be used for the international projects GN2 JRA4 and GARDEN.

7. Lowering the prices of the lease of the fibres also enabled the start of the construction of CzechLight network (which is of experimental character) physically separately from CESNET2 network providing production service – similarly to TransLight and National LambdaRail, which are physically separated from Abilene, GÉANT and NREN. This has opened significant new chances of research in this area for the researchers. It is known that the chances of research of new transmission systems and services on production networks are very much limited.

8. Setting up lambda service Prague–Amsterdam 2.5 Gbps, installation and testing CzechLight, finding the errors of the equipment and complaint of them. Lambda services CzechLight, NetherLight and CERN have been

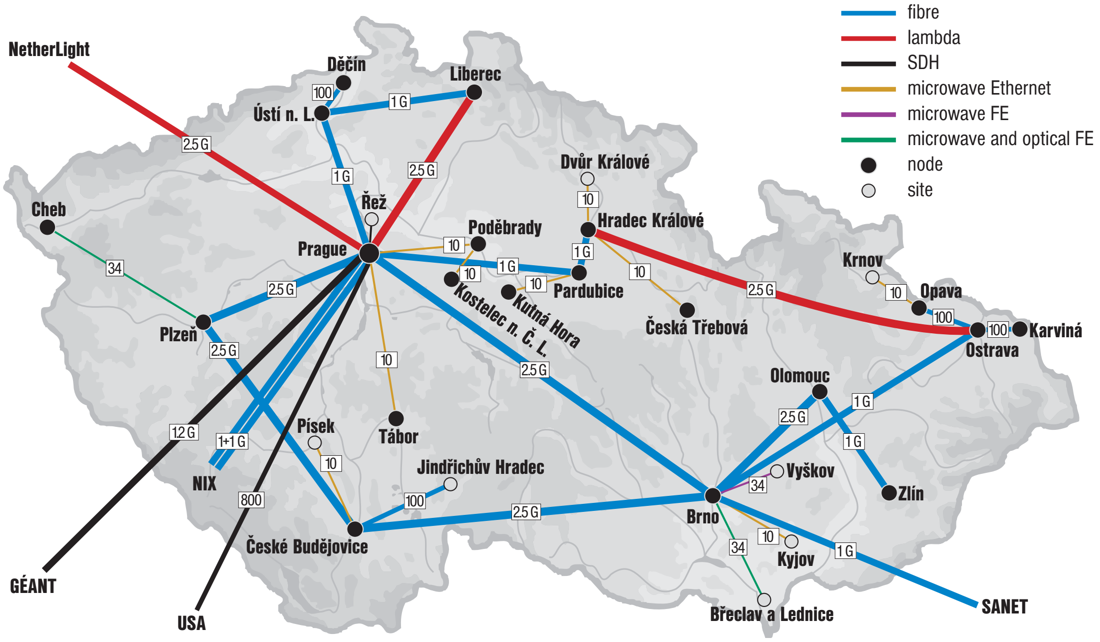


Figure 5.2: CESNET2 topology (December 2003)

experimentally used for the transmission of data between CERN and ASCR Institute of Physics (IoP) at Mazanka, Prague. Also Prague Academic Network (PASNET), ASCR (Academy of Science of the Czech Republic) – network of Na Mazance area – and IoP local network have been involved in realization of this end-to-end service.

9. Gaining the possibility of financially reasonable upgrade of the circuit Prague–Amsterdam from 2.5 Gbps to 10 Gbps and thus gaining more interest or acceptability for experiments shared with institutions in Europe, the USA, Canada and for presentation of the results.

10. Simulations and testing of the setting of optical amplifiers EDFA for GE and 2.5 Gbps transmissions without in-line equipment up to approximately 250 km; putting it on the operational line of 235 km (which is probably the world's primacy in production network).

11. Simulation and testing of the setting of optical amplifiers EDFA for 10GE transmissions without in-line equipment up to 250 km long line (the reach supposed by 10GE standard is up to 40 km). Testing of optical amplifying of WDM transmissions. Cisco company has been interested in cooperation in optical transmissions development.

12. Testing of Raman amplifiers for prolonging the distances noted above.

13. The presentation of the results in setting up dark fibres and testing the possibilities of transmission systems, participation in the preparation of GN2, GRANDE and GARDEN projects.
14. The proposal of setting the fibre line Prague–Frankfurt with optical amplifiers as the base for DANTE, simulation of running of the signals, consultation to “Invitation to Tender for Network Element for the GN2 Network”.
15. Evaluation of the situation in the area of wireless microwave and optical transmissions (mainly 802.11a and 802.11h) from the view of their usability for the first mile of inter-city transmission circuits. The preparation of the prototype of the equipment for free space optical transmissions 100 Mbps, the first result may be expected by the end of the year 2003.
16. The testing of the possibility of building the regenerator for optical fibre transmissions by programmable hardware is running through.
17. The work on replacement of expensive transmission equipment or interface cards and independent optical amplifiers by equipment with FPGA, XPF, transceivers, circuits for optical amplification, ASIC and hybrid IO is running through. The advantage should be Open hardware and remarkably lower price.

When solving the project, we have reached the results recognized abroad (see the list of publications and presentations).

5.2 Cooperation of NRENs on Setting Dark Fibre

We presented the results of CESNET2 network design and operation at international seminars TF-NGN in February 2003 in Rome and in September 2003 in Cambridge. These presentations and following discussions helped us to make contact with experts from individual NRENs which deal with fibre lines and transmission systems procurement and implementation. By the end of the year 2003, for active participants we have started international mailing list *CEF-Networks* for support of dark fibre implementation in European NRENs. Reciprocal consultations about dark fibre implementation ran through mainly with our colleagues in Ireland, the Netherlands, Poland, Portugal, Slovakia, Slovenia, Serbia and Switzerland. The aim is above all to gain and widen information about the ways of acquirement of fibre lines and the solutions of transmission systems which have been proved on experimental or operational NREN lines. It’s turned out recently that this effort will even be supported by international GN2 project (Joint Research Activity 4 – Testbed) of the FP6 program of the EU.

5.3 National Fibre Footprint

In the USA, the project called National LambdaRail (NLR) rose in 2003. This project contractly uses some of the fibres of Level 3 company (see Figure 5.3) and cals them National Fibre Footprint. This infrastructure extends the fibres used in individual states and regions in research and education networks in the USA. This has significantly strengthened the focus on the use of fibres in research and education networks, which was not so much used before. GN2 project in Europe now is trying to find out to what extend fibres are suitable for the pan-European network.

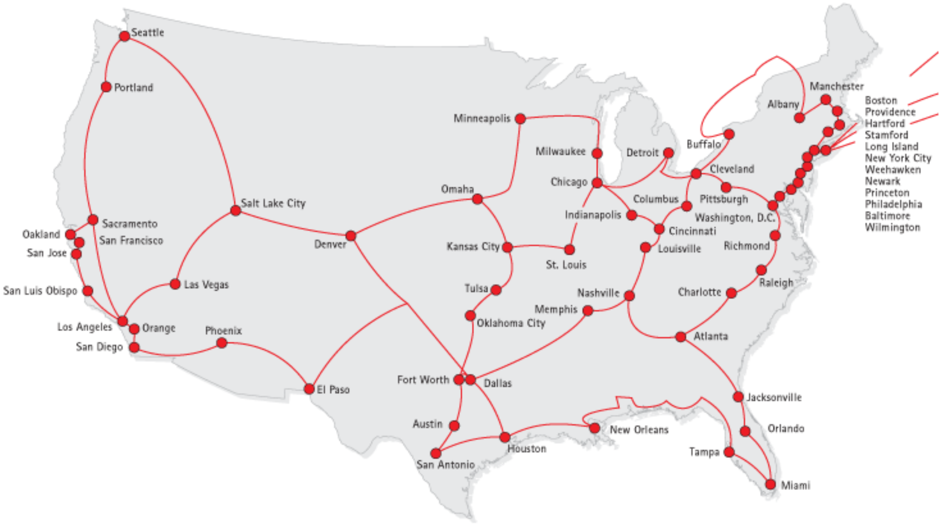


Figure 5.3: National LambdaRail (NLR)

In the Czech Republic, the national fibre footprint of the CESNET2 production network is already in operation.

At the end of 2002, we announced a tender for setting up and lease of fibres for seven backbone circuits. Only the routes Plzeň–České Budějovice and Ústí nad Labem–Liberec were selected for realization. Other circuits were not recommended for realization, mainly due to too long fibres or unacceptable requirements of the tenderers. The circuits were much cheaper than the service used till then. We also gained the possibility to use free of charge testing loops of fibres of various length for testing transmission systems in the laboratory. In March 2003, we announced a tender for setting up and lease of fibres for nine routes, out of which seven were backbone lines. The aim of this tender was not setting up new lines but replacement of the existing ones – with better conditions, not only from financial but also from technical point of view. This means acquiring fibres with better parameters (e.g., shorter lines with lower attenuation) and lines which use at least partly G.655 fibre.

The lease of the fibres was offered by eight competitors, two of them had been refusing fibre lease before. The prices of leased fibres for NREN are lower than 1.25 CZK/m/pair/month (about 0.5 EUR/pair/m/year), in some cases they are much lower. There has mainly been decrease in prices of routes between main regional cities where the fibres of the providers remain unused for a longer time. When evaluating the offers, we were searching for an optimal solution with respect not only to the price but also to the length of fibres and necessary HW for transmission. On longer routes, the purchase of amplifiers is necessary and thus the technical and operational requirements of the route grow.

What has been an essential knowledge gained during the solution is how to specify tender for fibres and what criteria to choose so that the realized network is of top parameters and keeps open possibility of further development for attainable financial resources. CESNET2 network now has the fibre footprint longer than 2,000 km, low expenses on its lease and wide scale of possibilities of increasing transmission speeds and capacities (see Figure 5.4).

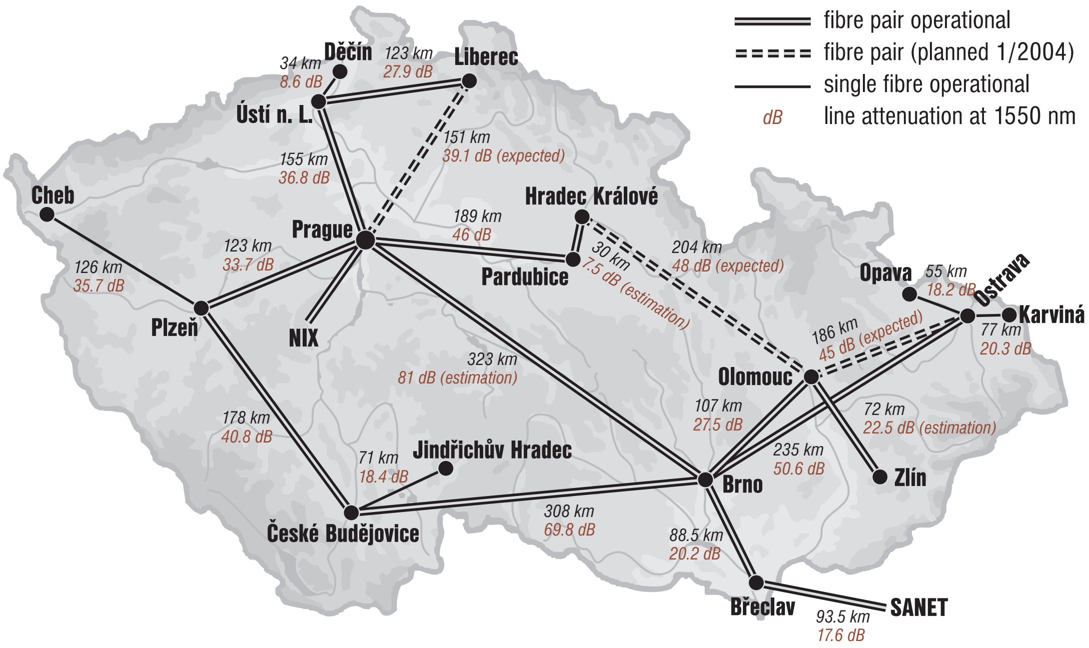


Figure 5.4: Fibres of CESNET2 (December 2003)

5.4 Optical Transmissions in Customer Empowered Fibre Networks

Customer Empowered Fibre networks (CEF networks) are networks where (some of) the users of the network own the fibres or have the right to use the fibres and to decide the way of the network design (mainly the transmission system) and network operation.

Gaining the fibres is followed by the question how to solve the transmission system best. Of course, offers of transmission systems for telecommunication operators may be used but these are usually very expensive. It also turns out that NRENs have different requirements and so it is often most acceptable to build a “tailor-made” transmission system and upgrade it to higher transmission speed or higher number of lambdas at the time when it’s really necessary.

It is also important to figure that the prices of fibres and equipment change (now are decreasing very significantly) and so it is economically very risky “to build for future needs”. Moreover, the uncertainty connected with future needs often results in the selection of more universal solutions which are usually more expensive. Sometimes it is more profitable to rent another pair of fibres than to invest into the transmission equipment. According to what is feasible for us, we have dealt with the analysis of this situation and evaluation of the individual methods of solutions and we have proved some of them in laboratory or in CESNET2 production network.

5.4.1 Transmission Routes without In-Line Equipment

One of the significant methods of CEF network setup is NIL which is trying to solve the topology of the network so that it is not necessary to use in-line optical amplifiers or regenerators. This means that all devices are located in PoPs of the network. Figure 5.5 shows to what extent this method is applied in CESNET2 network. Dark blue routes are fibres without in-line equipment, i.e., NIL fibres. The transition of the remaining light blue and red routes (except for the route Prague–Brno) to this way of transmission should be carried out by the end of January 2004.

5.4.2 Single Fibre Transmission

CESNET has used new types of converters which implement bi-directional transmission over one optical fibre for economically effective connection of sites of members or some of the customers.

We opted for new single-fibre lines using tested twisted pair–fibre converters made by MRV (e.g., *http://www.mrv.com/product/MRV-FD-FS*). This way is economic and technically attractive. It is possible to implement the transmission for Ethernet, Fast Ethernet (FE) and Gigabit Ethernet. The converters are made for transmission over single fibre and pair of fibres. Equipment for transmission over single fibre has shorter reach. We opted for two-way transmission over

	Fibre				*Transmission*	
Line	*length*	*Atten.*	*Band-*	*Operation*	*equipment*	*Note*
	[km]	*[dB]*	*width*	*since*	*deployed*	
Pardubice–H. Králové	30	7.5*	1 GE	15. 1. 02		* estimated value
Olomouc–Zlín	72	22.5*	1 GE	19. 2. 03		
Plzeň–Č. Budějovice	178	40.8	2.5 G	25. 7. 03	2×24 dBm	
Ústí n. L.–Liberec	123	27.9	1 GE	30. 6. 03		
Brno–Ostrava	235	50.6	1 GE	6. 6. 03	2×27 dBm + 2×10 dBm	
Č. Budějovice–Brno	308	69.8	2.5 G	15. 9. 03	2×10 dBm + 2×27 dBm	+ 2×Raman since 12/2003
Praha–Ústí n. L.	155	36.8	1 GE	8. 12. 03	Catalyst (temporary)	2×21 dBm, delivery 01/2004
Praha–Brno	323	81.0*	2.5 G	10. 1. 00	3×ONS15104	changed in 2004
Praha–Liberec	151	39.1*	2.5 G	1. 2. 04*	2×21 dBm	
Praha–Plzeň	123	33.7	2.5 G	8. 9. 03	2×10 dBm	both amplifiers in Prague
Praha–Pardubice	189	46.0	1 GE	17. 5. 02	2×21 dBm	
H. Králové–Olomouc	204	48.0*	2.5 G	1. 2. 04*	2×27 dBm + 2×10 dBm	
Brno–Olomouc	107	27.5	1 GE	25. 11. 03		
Olomouc–Ostrava	186	45.0	2.5 G	1. 2. 04*	2×24 dBm + 2×10 dBm	
Total	*2,385*					

Table 5.1: Fibre pairs in CESNET2 network

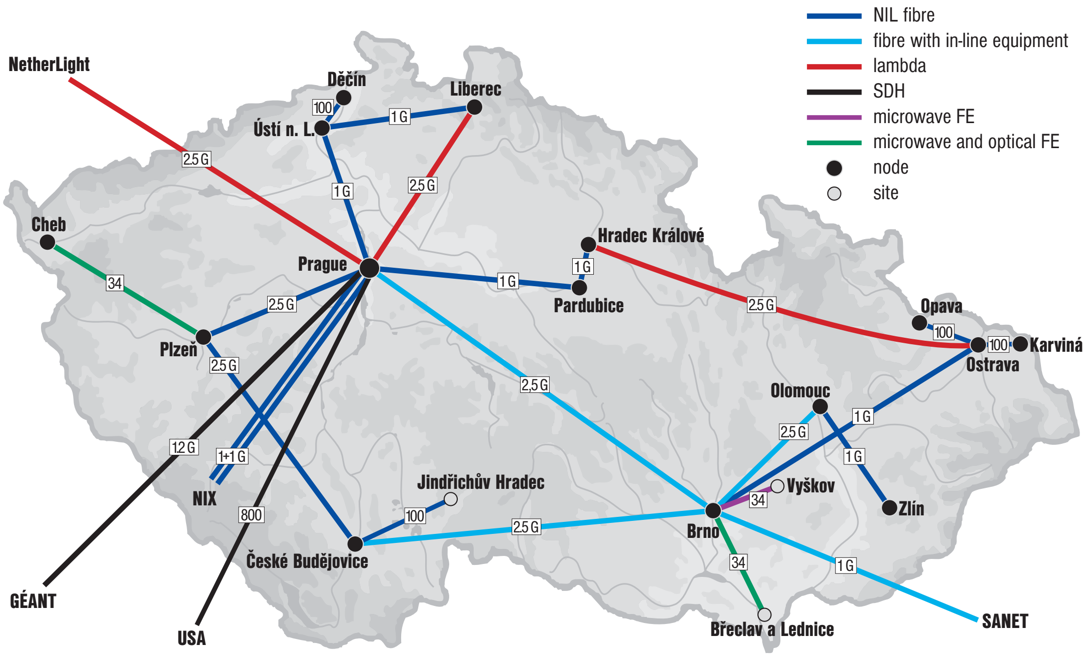


Figure 5.5: NIL fibres in CESNET2 (December 2003)

a single fibre, using different wavelengths 1520 nm and 1560 nm for different directions. According to the available information, this transmission system is more reliable than the system that uses of the same wavelength for transmission in both directions.

During the year 2003 we deployed five single-fibre lines with FE converters. We obtained leasing of one fibre under 60 % of price of leasing pair of fibres for five-year contract from one provider. We didn't manage to obtain the leasing of one fibre from other providers because the demand for these fibres is still quite low and the second fibre from the pair remains unused. This situation might change when one fibre transmission in both directions becomes better known.

At the end of March 2003 we launched the first operational long-distance line Ostrava–Opava. The distance is 55 km and we used MRV converters EM316 WFC/S4 & MRV EM316 WFT/S4 (formerly Nbase-Xyplex) for Fast Ethernet for both ends of line. The operation on this line is error-free. The independence of the converters on the Cisco Systems equipment, software and support is a great advantage.

We used the same model with S4 converters for two other lines, Ostrava–Karviná and České Budějovice–Jindřichův Hradec. The pair of S3 converters with the reach 20–50 km sufficed for the shortest line Ústí nad Labem–Děčín. We deployed the S5 converters for the longest line Plzeň–Cheb with its distance of 126.4 km and attenuation 35.7 dB. After deployment the line was working with-

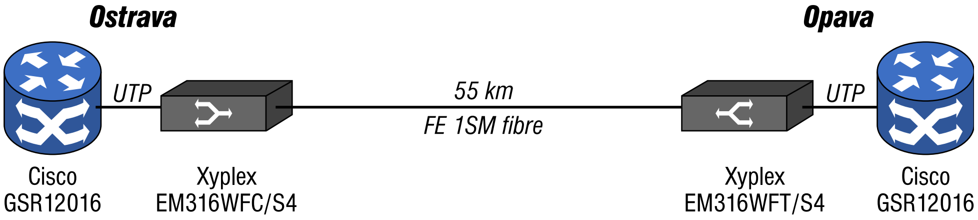


Figure 5.6: Single-fibre line Ostrava–Opava using MRV converters

out errors for one month although the parameters of the line are worse then those specified by the converter provider. This result we consider interesting in spite of the fact that it originally occurred due to low financial resources for purchase of device. After one month the line error rate raised and it was necessary to switch the traffic to the old line. The measurement of this line demonstrated that the parameters of the fibre had not worsened. We found out during the test that this pair of converters had higher error rate on shorter line too. This means that there was error on the equipment. The S4 converters on the line Ostrava–Karviná had a breakdown too. The result is the complaint of the fault of the device, purchase of the pair of S3 converters for elongating of reach on the line Plzeň–Cheb and purchase of spare converters for the case of the fault on some single-fibre line. Despite these problems with reliability of some MRV equipment, single-fibre lines seem to be the most advantageous solution in a specific situation.

		Atten.	*Operation*		*Reach*
Line	*km*	*[dBm]*	*since*	*Equipment*	*[km]*
Ostrava–Opava	55	18.2	1. 3. 03	S4	40–100
Ostrava–Karviná	77	20.3	1. 7. 03	S4	40–100
Plzeň–Cheb	126	35.7	1. 7. 03	S5	40–125
Ústí n. L.–Děčín	34	8.6	1. 7. 03	S3	20–50
Č. Budějovice–J. Hradec	68	18.4	1. 9. 03	S4	40–100

Table 5.2: Parameters of single-fibre lines

Fast Ethernet is operational on all single-fibre lines. These lines are suitable for transmission speed 100 Mbps. It is also possible to equip them with the converters of Gigabit Ethernet, but it does not seem necessary in consideration of the workload now. On the whole, single-fibre lines are financially comparable to microwave circuits but they are faster and more reliable. The difference in reliability is more significant on longer lines. Long microwave lines are sensitive to atmosphere troubles. The exception which is financially not very favourable is when it is necessary to lay long segment of fibre (principally first mile). The payoff of the investment to MRV converters is very good. The longer the line, the more favourable is the investment.

Line	*km*	*Payoff [mon]*
Ústí n. L.–Děčín	34	18.5
Ostrava–Karviná	77	9.5
Plzeň–Cheb	126	6.9

Table 5.3: Payoff of the investment to MRV converters

Deployment of the 10GE technology on a single fibre: At our request, PASNET (Prague Academic and Scientific network) has experimentally built and tested the interconnection between the ASCR Institute of Physics at Mazanka (IoP) and the office of CESNET Association at Zikova.

For the high-speed transmission to the remote sites of their cooperating partners outside the Czech Republic, the ASCR Institute of Physics uses the international lambda services provided by CESNET from the Zikova location in Prague. The link between the Institute's premises at Mazanka and the CESNET site at Zikova is provided by the PASNET.

IoP's computers at Mazanka are connected to a Gigabit switch from Cisco Systems. This switch is connected to the backbone device Catalyst 6509 at Ovocný trh location with a single single-mode fibre (7 km length), using the 1000BASE-SX GBIC module (product designation WS-G5484) and converters from MRV. From Ovocný trh, the service continues as a specific VLAN through PASNET's 10 Gbps backbone link to Catalyst 6506 at Zikova location with 802.1Q protocol. The link terminates on the ONS 15450 device at the CESNET centre.

Based on economic considerations, we have opted for the solution with optical converters to minimize operational costs. We deployed MRV's protocol-independent converters, product designation EM316WGC-T. These converters use different wavelengths (1310/1550 nm) for either direction of communication on a single fibre. According to the documentation, this type can be used for the distances up to 25 km.

It is worth mentioning the fact that the backbone line of the Prague academic network (an SM segment of 5 km) between locations Ovocný trh and Zikova was also operating with passive splitters on a single fibre for the period of three months. The technology was 10GE between Catalyst switches (modules WS-X6502-10GE fitted with WS-G6488-10GBASE-LE, i.e., using the wavelength of 1310 nm). Converters EM316SC3S were used as suitable for this wavelength. There was neither deterioration of service nor increase in error rate resulting from the use of optical splitters.

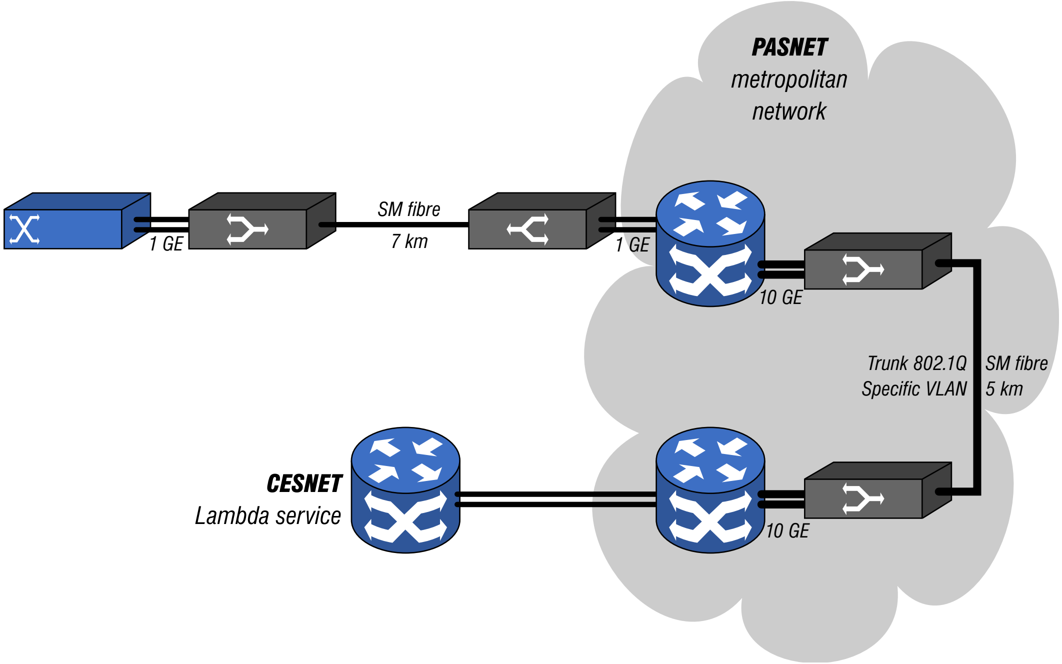


Figure 5.7: Experimental link Mazanka–Zikova

There was no unscheduled outage on the above described link IoP Mazanka–Zikova throughout the whole period of six months when the service was used.

We were looking for the best solution in realization of single-fibre inter-city circuits, but we detected among other NRENs only one. Swiss network SWITCH (The Swiss Education & Research Network) is using an interesting solution for the cheap single-fibre long-distance connections, which has been used for more than one year. For bidirectional Gigabit Ethernet, the Cisco CWDM GBICs are used in the routers and the POCs – Passive Optical Couplers – on both ends of single fibre. The equipment POC consists of the splitter for two wavelengths (1530 and 1550 nm) and OADM-1 Channel Optical Add/Drop Multiplexer. OADM takes out the reflections due to bad connectors and fibre impurities.

Without another amplifiers the distances up to 100 km were achieved. With the help of EDFA amplifiers (16 dBm) – between the output of GBIC and POC – the distances up to 150 km are reachable.

SWITCH has the leased pairs of dark fibres in use, but after the tender and selection procedure it is using the transfer equipment Sorento (for single fibre) for the backbone lines. The second fibre is verywell used for connecting other institutions along the lines. The topology of SWITCH is shown in Figure 5.8.

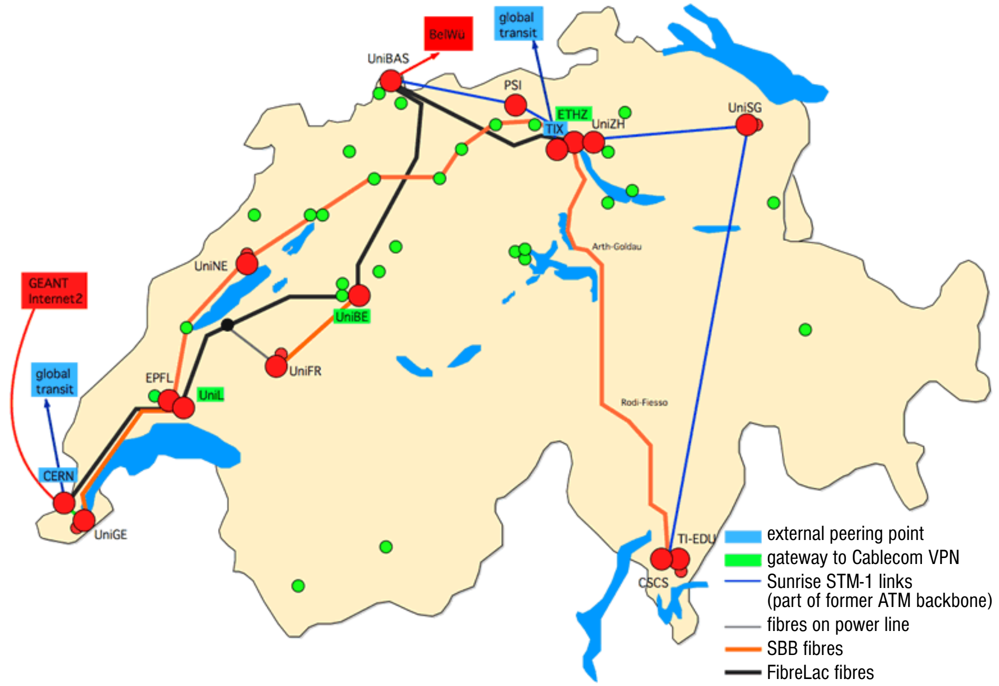


Figure 5.8: Topology of Swiss network SWITCH

5.4.3 Simulations, Testing and Deployment of GE NIL Long-Haul Transmissions

In the area of optical signal amplification we have taken up with our knowledge made back in 2002, when line Praha–Pardubice with the length of 189 km was put into service. This year, another NIL line Brno–Ostrava with the length of 235 km was put into service. With the use of Gigabit Ethernet (GE) as transport protocol, it was possible to deploy high-power erbium doped fibre amplifiers (EDFA) only. We were able to test 2.5 Gbps PoS, too. In this case it was necessary to deploy additional EDFA preamplifiers (the reason is lower sensitivity of PoS line cards). A migration to PoS technology did not occur yet and the line works error-free. As far as we know, no line with similar parameters is operated in any other network (neither research nor production).

With standard CWDM GBIC it is possible to overcome distance up to 250 km, which we verified in lab environment on G.652 spools kindly loaned by OFS, Denmark (former Lucent). Detailed schemes of tested configurations are to be found in Annual Research Report 2002.

5.4.4 Simulations, Testing and Deployment of 2.5 Gbps NIL Long-Haul Transmissions

Nowadays, optical amplifiers are used on the following backbone lines: Praha–Plzeň, Plzeň–České Budějovice and Brno–České Budějovice. On the first line, two 10 dBm EDFA amplifiers are used in non-typical configuration: one amplifier is used as a preamplifier and the other one is used as a booster. This brings the advantage of placing both of them in Praha, which means that deployment and maintenance of the amplifiers is easier. The second line with the length of 178 km is equipped with boosters only.

The line Brno–České Budějovice has length 308 km and to keep NIL approach, EDFAs are not sufficient and one has to utilize an amplification effect based on the principle of stimulated Raman backscattering (RFA). Some problems occurred within delivery of Raman amplifiers and we tested the amplifiers in our lab together with G.652 spools from OFS at the beginning of December. In agreement with the results of simulations and experiments, it is possible to equip the line by means of NIL method before the end of 2003 (for every direction, it is necessary to deploy one EDFA booster and one EDFA preamplifier, Raman fibre laser and optical filter to suppress the noise). For some time, due to these reasons, the line is equipped with EDFA amplifiers only, one of them works as booster and the other one works as in-line amplifier. The line works error-free in this configuration. 2.5 Gbps technology was tested in the lab in configuration with in-line amplifiers up to distance 350 km, and no compensation of chromatic dispersion or the use of optical filters to suppress the noise was necessary.

5.4.5 Field Trial with 10 Gigabit Ethernet Adapters for PC

We have acquired two 10 Gigabit Ethernet adapters for PC (Intel PRO/10GbE) for measurement and experimental work in the SCAMPI project and for evaluation of issues in end-to-end 10 Gigabit Ethernet communication. As a first step we tested achievable throughput with Linux PC and estimated the maximum possible distance between the sender and receiver. We described our observations in CESNET technical report 10/2003. We summarize here some of our findings.

Test Setup

The test setup is shown in Figure 5.9. Each adapter was installed in a PCI-X 64-bit 133 MHz slot of a Dell 2650 server. This slot has its own PCI bus, that is no other device shared the same PCI bus. The server was equipped with one Intel Xeon 2.4 GHz processor and 1 GB RAM. The two adapters were connected back-to-back with a short optical patch cable. Both machines ran Debian Linux with 2.4.22 kernel.

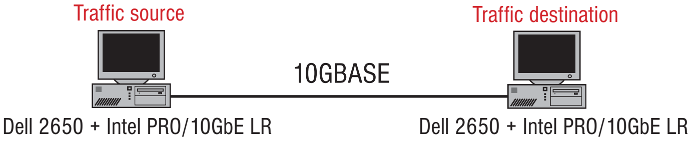

Traffic source
Traffic destination
10GBASE
Dell 2650 + Intel PRO/10GbE LR
Dell 2650 + Intel PRO/10GbE LR

Figure 5.9: Test setup

Optical Power Budget

The Intel PRO/10GbE LR adapter uses a 1310 nm laser with the specified reach of 10 km. Unfortunately, the transceiver is of the "300-pin" type and cannot be replaced with another transceiver (such as one using a 1550 nm laser, which would be an attractive option), which is possible with XFP or XPAK type transceivers.

The output power measured by the Expo FOT-90A fibreoptic power meter was –3.2 dBm on one adapter and –5.0 dBm on the other adapter. We inserted the Expo FVA-60B variable attenuator between the two adapters to find the maximum acceptable power budget between the sender and receiver. We found that the maximum acceptable attenuation with no packet losses was 7.85 dB in one direction and 8.30 dB in the other direction. Taking the lower value and considering attenuation of 0.35 dB/km at 1310 nm on standard optical fibres, we can estimate the maximum possible distance between the sender and receiver to approximately 22 km.

Interconnection with 1550 nm Devices

High-speed optical ports on routers and switches use 1550 nm lasers more frequently than 1310 nm lasers. The reason is most optical fibres have lower attenuation at 1550 nm and the signal at this wavelength is also easier to amplify, which allows to span longer distances. It is important for high-speed ports, which are usually used for long-distance backbone circuits. On the other hand, PC adapters are usually designed for use in local networks and are therefore often equipped with cheaper 1310 nm lasers. In this case, for end-to-end communication over a wide-area network we need to resolve how to connect 1550 nm and 1310 nm devices together.

Optical wavelength converters are not yet commercially available. We can expect that if such a device becomes available, it will be expensive. We can however take advantage of broadband sensitivity of most optical receivers. We tried to connect the Intel PRO 10GbE LR adapter directly to the Cisco Catalyst 6500 switch with the WS-X6502-10GE 1-port 10 Gigabit Ethernet adapter, which uses an extended reach 1550 nm laser. The 1-port adapter has a fixed non-interchangeable transceiver, unlike the more expensive 2-port and 4-port adapters, which use XENPAK type interchangeable transceivers.

A 5 dB attenuator was inserted in each direction to protect the receiver from a possible damage by a high power level from the sender (it was actually needed only in the direction from the extended reach laser). Communication worked without any problem and we successfully sent and received 6 millions of 1500-byte packets, that is 3.6×10^{11} bits without any packet loss or damage. This shows that the Intel PRO 10GbE LR adapter can be reliably used to connect to the Catalyst's 1550 nm transceiver, thus enabling an end-to-end 10 Gigabit Ethernet communication over a wide-area network.

Throughput

We used *iperf* to measure TCP throughput. Both sender and receiver socket buffers were set to 1 MB. Transmission interface queue (txqueue) was set to 10,000 packets. This value was computed using a rule of thumb that it should cover transmission at the physical interface rate (10 Gbps) for the duration of the operating system scheduling timer (10 ms). PCI-X burst transfer size was increased from the default value of 512 bytes to 4096 bytes. We used a standard MTU of 1500 bytes and the maximum MTU of 16,114 bytes supported by the adapters. We also monitored CPU load with the *top* command and the number of generated interrupts by reading the */proc/interrupts* file before and after the test.

We have measured throughput of 1.3 Gbps for 1500-byte packets and 2.6 Gbps for 16,114-byte packets. In the latter case, the sender CPU load was almost 100 %. The receiver CPU load as well as the load on both sides with 1500-byte packets were lower and did not limit the achieved throughput. The CPU load increased probably as a consequence of significant increase in number of interrupts. It appears that interrupt coalescing did not work properly.

Conclusion

We achieved the maximum TCP throughput 2.6 Gbps with 16,114-byte packets on Dell 2650 servers. If the problem with interrupt coalescing is resolved, the achieved throughput could probably be higher. However, the use of jumbo packets (larger than 1500 bytes) is required to achieve throughput significantly higher than 1 Gbps.

The maximum acceptable power budget should allow communication up to approximately 22 km. Interconnection with 1550 nm transceiver on Cisco Catalyst switch worked without problems. This enables an end-to-end 10 Gigabit Ethernet communication over a wide-area network (although not at the full speed with current PCs).

The adapters could be also used to build a relatively inexpensive 10 Gigabit Ethernet router or switch, with multiple adapters in one PC running proper routing or switching software. With current PCs, the throughput would be however much lower than the full line rate. Another possible use is for emulating fast long-distance networks with *NIST Net* network emulator.

5.4.6 Simulations and Testing of 10 Gbps NIL Long-Haul Transmissions

In the first half of the year, we were evaluating performance limits of NRZ data transmission at 10 Gbps over standard single mode fibre (SMF, G.652) without the deployment of in-line EDFAs with the aim to find out the limits of the maximum transmitter–receiver distance and to keep the bit-error-ratio (BER) below 10^{-12}. For these numerical simulations we used commercial software *OptiSystem 2.0* from OPTIWAVE and software *OptSim* from ARTIS.

The effect of input optical powers to SMF and to the dispersion compensating fibre (DCF), degree of GVD compensation of the SMF, and the effect of group velocity dispersion (GVD) compensation schemes – post-compensation and pre-compensation, when DCF is placed at the beginning or at the end of line – has been investigated. We have performed a comprehensive analysis and we have found that optimal degree of compensation is approximately 85 % for post-compensation scheme and 90 % for pre-compensation scheme. Assuming optimistically the attenuation of G.652 fibre to be 0.22 dB/km, it follows from analysis that maximum transmitter–receiver distance is 270 km for post-compensation scheme and 230 km for pre-compensation scheme with $\text{BER}<10^{-12}$.

At the same time, we began with practical verification of simulations. We bought Cisco Catalyst 6503 with 10GE 1550 nm line cards with maximum distance of 40 km and we used field-tested Keopsys optical amplifiers. Cisco systems routers and switches are deployed in CESNET2 production network and therefore the research results can be utilized for partial upgrade to 10GE.

The results can be summarized as follows: for distances up to 100 km one can use boosters only without compensation of chromatic dispersion. When booster and preamplifier are used, it is possible to cover the distances up to 200 km, for these distances it is necessary to compensate the effect of chromatic dispersion. With the use of 3-W Keopsys Raman fibre laser we were able to cover a record-breaking NIL distance of 250 km. During these experiments, the best results were accomplished with the combination of both post and pre-compensation schemes.

It has to be kept in mind that transceivers and receivers of 10GE line cards have parameters according to standards, when high-quality DWDM lasers are used, it is possible to expect additional increase of maximum distances with NIL approach. In that case, it is possible to cover the distances among all major cities in the Czech Republic and the results are employable for most of the European countries. Further research plans were framed in application form for grant entitled "Optimization of data transmission at 10 Gbps over G.652 fibres without the deployment of in-line EDFAs with respect to maximum transmission distance", the grant was awarded by the Grant Agency of the Czech Republic. Duration of the project is 3 years, beginning in 2004.

Results of these practical experiments were presented at Terena conference in 2003 *Optically Amplified Multigigabit Links in the CESNET2 Network*. Another contribution was presented at ConTel 2003 conference in Zagreb. Both presentations were positively accepted and some of NRENs were interested in 10GE deployment with NIL approach.

Former results of simulations and experiments entitled *Optical networking in CESNET2 gigabit network* were accepted for publication in journal Annales of Telecommunications. Brand new theoretical and experimental results were summarized in paper *Optimization of NRZ data transmission at 10 Gbps over G.652 without in-line EDFAs* to be published in Fiber and Integrated Optics journal.

5.4.7 Gain Stabilization in All-Optical Gain-Clamped Lumped Raman Fibre Amplifier

In optical networks with WDM channel addition/removal, fibre amplifiers with fast gain stabilisation must be used to suppress transition effects as a result of changes in the number of transmitted channels. For these reasons, we started to investigate an application of all-optical gain-clamped lumped Raman fibre amplifiers. We have developed our own simulation software for numerical analysis of transition effects in RFA and theoretical results were verified experimentally. In experiments, DCF module was used as RFA and channel addition/removal was simulated by transmitting signals of two lasers, light of one of the lasers was square-wave modulated at 500 Hz, power fluctuations of the other laser caused by cross-gain modulation of the RFA were monitored at the output of the amplifier with a digital oscilloscope with and without all-optical feedback. With optical feedback applied, we achieved suppression of transition effects by 10 dB.

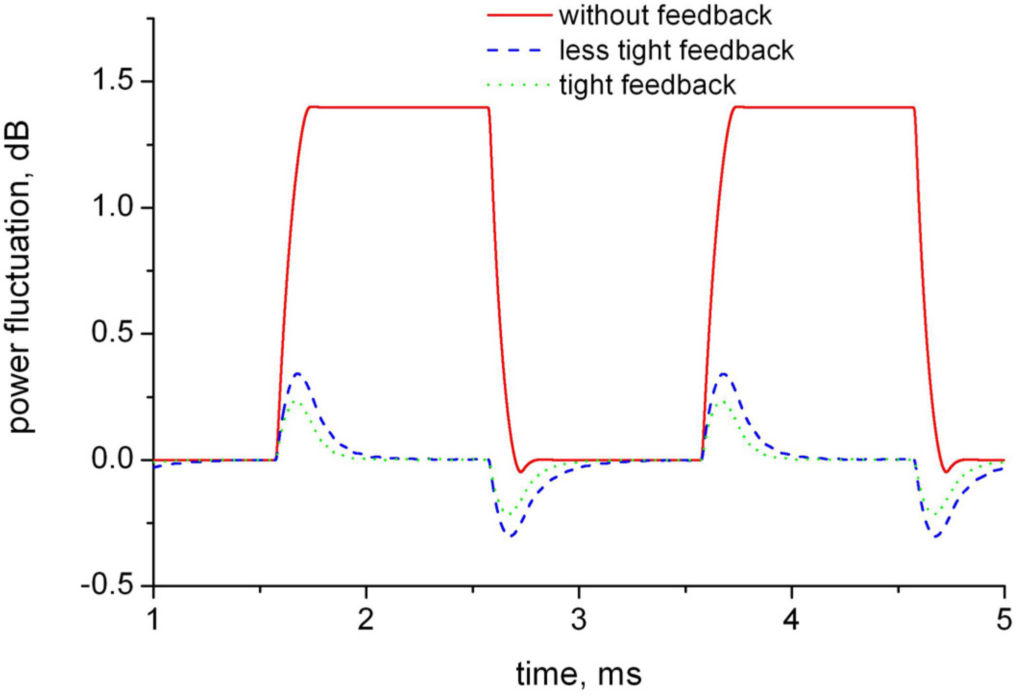


Figure 5.10: Output power fluctuation in fibre: numerical simulations

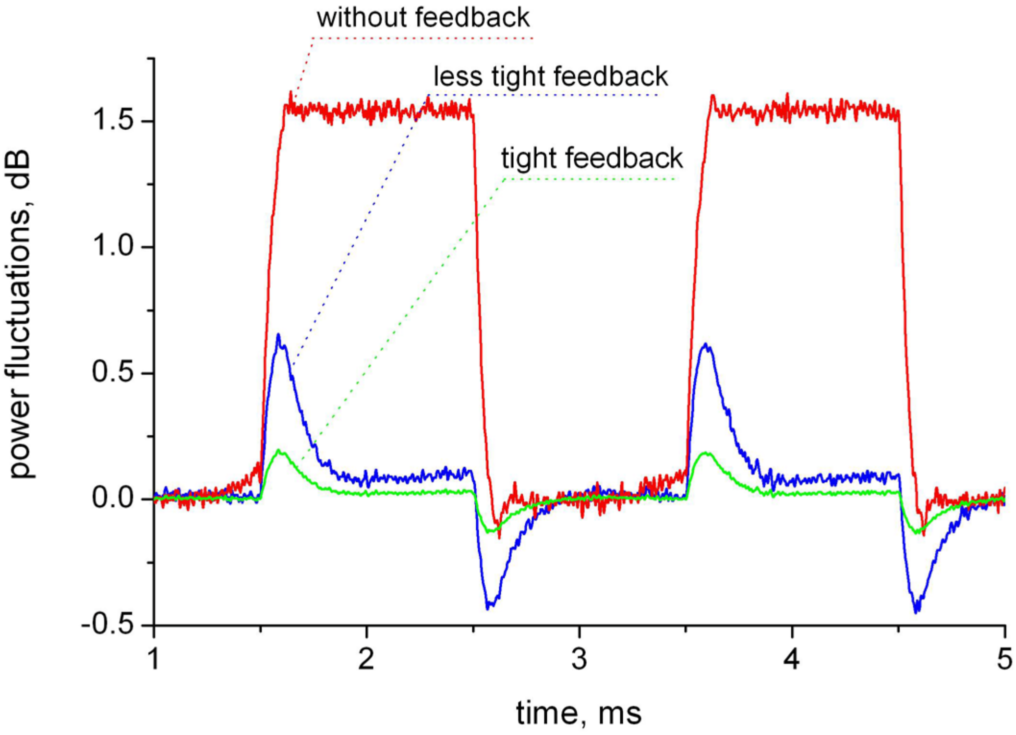


Figure 5.11: Output power fluctuation in fibre: experimental verification

Theoretical and experimental results were summarized in the following papers:

- M. Karásek, J. Kanka, P. Honzátko, J. Radil: *Channel Addition/Removal Response in All-Optical Gain-Clamped Lumped Raman Fiber Amplifier* sent for publication to IEEE Photonics Technology Letters
- M. Karásek, J. Kanka, P. Honzátko, J. Radil: *Protection of Surviving Channels in All-Optical Gain-Clamped Lumped Raman Fibre Amplifier: Modelling and Experimentation*, sent for publication to Optics Communications
- M. Karásek, J. Kanka, P. Honzátko, J. Radil: *All-optical gain-controlled lumped Raman fibre amplifier* was accepted for oral presentation on Optical Network Design and Modelling conference, ONDM 2004, Ghent, Belgium

5.4.8 Experimental 10G WDM Transmission System

As another step we have decided to assemble and to test in lab environment experimental WDM transmission system designated for deployment with dark fibres. We primarily tested multiple NIL transmissions of Gigabit (GE) and 10 Gigabit (10GE) Ethernet by means of NIL method.

Signals were combined via conventional 1×N directional couplers and the same couplers are used to split signals at the end of fibre. Subsequently, tunable optical filters must be used (DWDM or CWDM multiplexers can not be used because routers line cards are "of grey wave length", i.e., they are not tuned to ITU grid). It means the overall line length is smaller due to higher insertion loss of couplers in comparison to DWDM multiplexers. Achieved results are very interesting. We were able to transmit 2×GE and 2×10GE over 200 km standard G.652 fibre with EDFA amplifiers and Raman laser. Measured bit error rate for 10GE channels was better than 10^{-13}.

In another experiment, two 10GE signals were successfully transmitted over 250 km without use of in-line EDFAs. In the experiment, standard directional couplers and optical filters were used to combine and split individual channels, 3 EDFAs, 3 DCF and transmission fibre was pumped by Raman laser.

These results are encouraging both for NREN operators and especially for experimental networks like CzechLight, which is now projected, because standard equipment from major vendors (Cisco Systems, Nortel Networks, Lucent Technologies) is not always suited for specific requirements of these networks. Another important factor can be the price of such experimental WDM system and possibility for quick reconfiguration according to the needs of administrators or even end users.

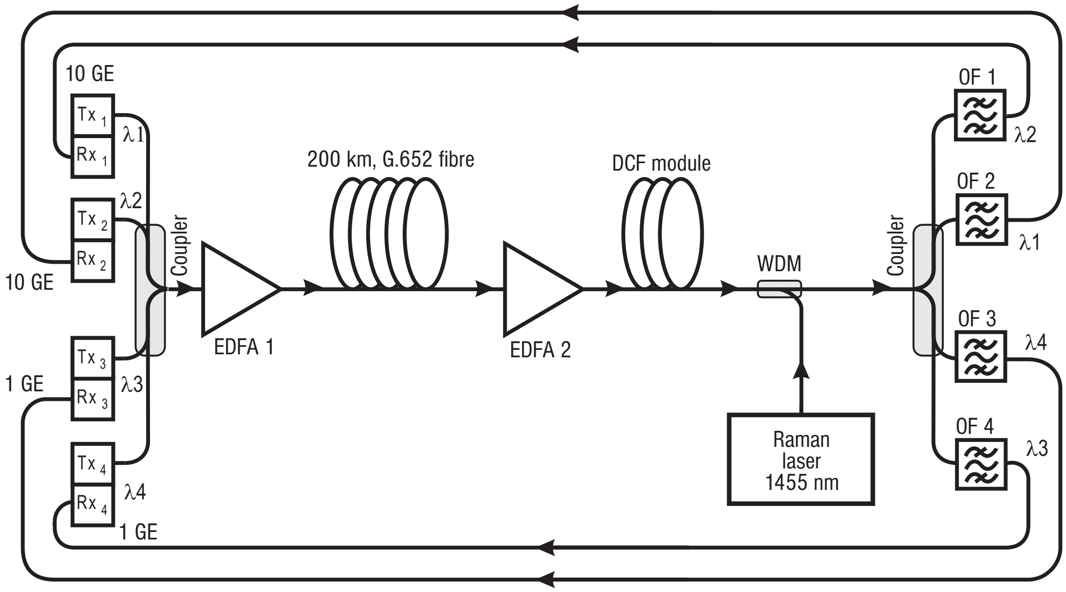


Figure 5.12: Experimental WDM system for transmission of 2×GE and 2×10GE over 200 km

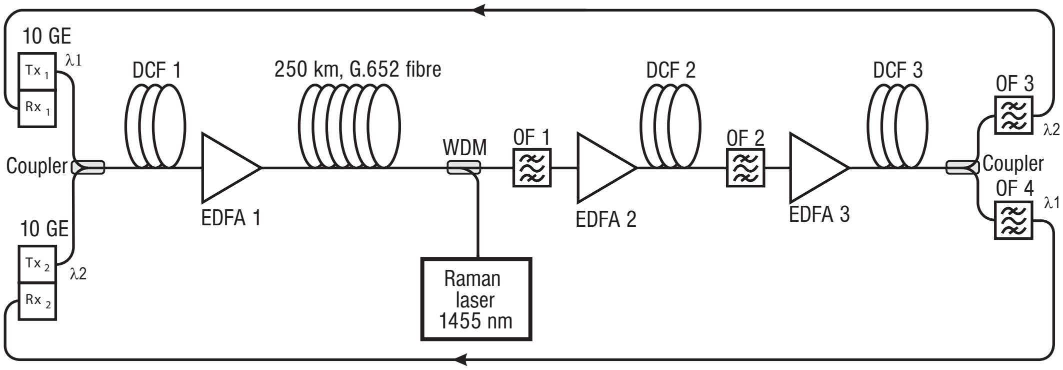


Figure 5.13: Experimental WDM system for transmission of 2×10GE over 250 km

5.5 Dark Fibre International Connection of NRENs

We checked out leasing possibilities of dark fibres over border in connection with preparation of participation of CESNET Association in international projects GARDEN and GRANDE and for other international activities.

During the year 2003 we preliminary inquired the circuits to Slovakia, Poland, Germany and Austria from inland providers of fibres. At first we demanded lines from Prague to Bratislava, Poznań, Frankfurt a. M., Munich, Berlin and Vienna. Based on acquired knowledge we requested the possibilities of connection for university towns or NREN PoPs near border. As a reply, we obtained proposals of international connections Brno–Bratislava, Brno–Vienna, Plzeň–Munich, Plzeň–Nurnberg, Ústí nad Labem–Dresden and Ostrava–Bielsko Biala. The table 5.4 shows rough overview of possibilities. We point out that the prices for real procurement would be different (probably lower) and that the connection to Poznań would probably be agreed with Polish NREN, which is the owner of fibres in Poland.

We have already used the experience from negotiations about leasing of fibres in the international projects, principally SERENATE and GN2. The expenses on optical transmission systems for a neighbouring border NREN PoPs are lower (for example 50 %) than connection of NREN centres (located usually in capitals). We called this method NoB (Near over Border). Overlaying European infrastructure, like GÉANT, would be able to have lower number of circuits (e.g. only circuits longer than 1000 km) for lower expenses and better transmission of high data volume for long distances. Real possibilities of using this method will appear during procurement for data lines and fibres for project GN2.

Connection of			*Connection of*		
centres	*km*	*Eur/mon*	*PoPs near to border*	*km*	*Eur/mon*
Praha–Poznań	970	40,400	Ostrava–Bielsko Biala	150	6,300
Praha–Frankfurt	665	18,900	Plzeň–Munich	330	13,800
			Plzeň–Nurnberg	227	9,500
Praha–Wien	520	11,500	Brno–Wien	191	7,500
			Č. Budějovice–Linz	200	8,300
Praha–Bratislava	432	11,600	Brno–Bratislava	182	7,600
Praha–Berlin	530	22,100	Ústí n. L.–Dresden	230	9,600

Table 5.4: Contemporary possibilities of dark fibre connection abroad

The first international fibre connection of NRENs Brno–Bratislava by Gigabit Ethernet was realised in April 2003 thanks to the initiative of our colleagues from SANET (Slovak Academic Network) and laying shorter first mile fibre line to CESNET PoP at Masaryk University in Brno. It was found out that the traffic on this line is busy[1] (more than 150 Mbps), despite of the connection between CESNET and SANET from Praha to Bratislava. The traffic was even increased later by using other applications, especially experimental streaming of television and radio broadcast from SANET.

[1] see *http://www.cesnet.cz/provoz/zatizeni/*

5.5.1 Equipment for Planned Dark Fibre Line Praha–Frankfurt a. M.

In connection with the possibility of dark fibre leasing between Praha and Frankfurt a. M. terminated in GEANT PoP, we investigated the optimisation of 10G single channel transmission system. We took into account the possibility to lease G.652 as well as G.655 fibres. *OptiSystem* software by Optiwave was used for numerical simulations and analysis, the total length of the line was 665 km. Required BER was better than 10^{-12}.

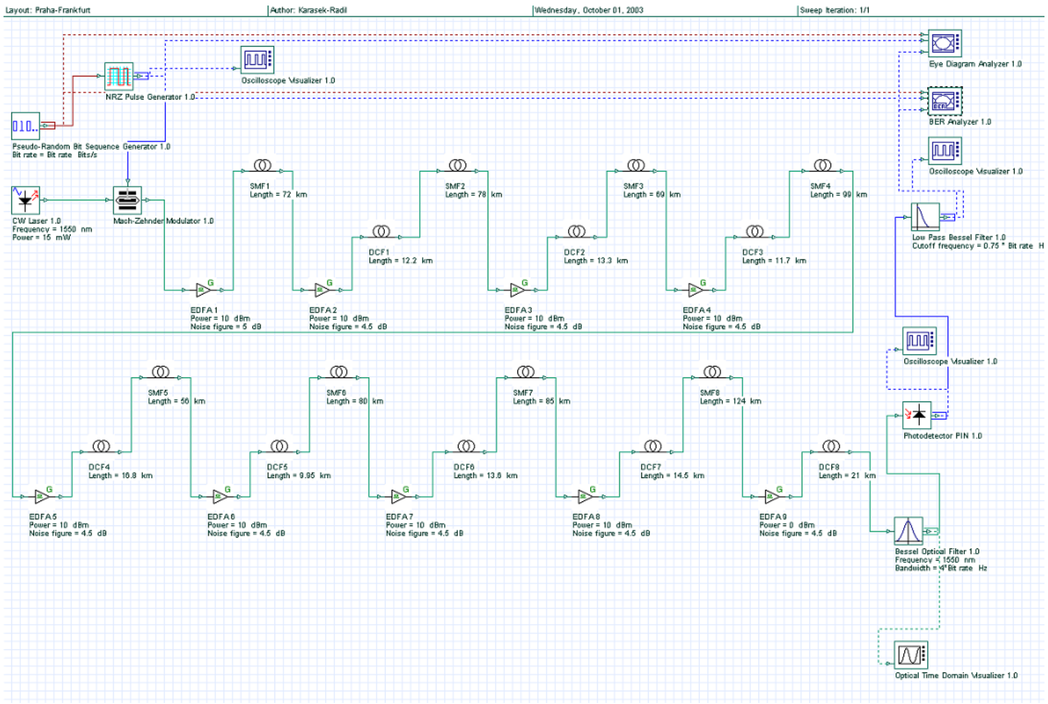
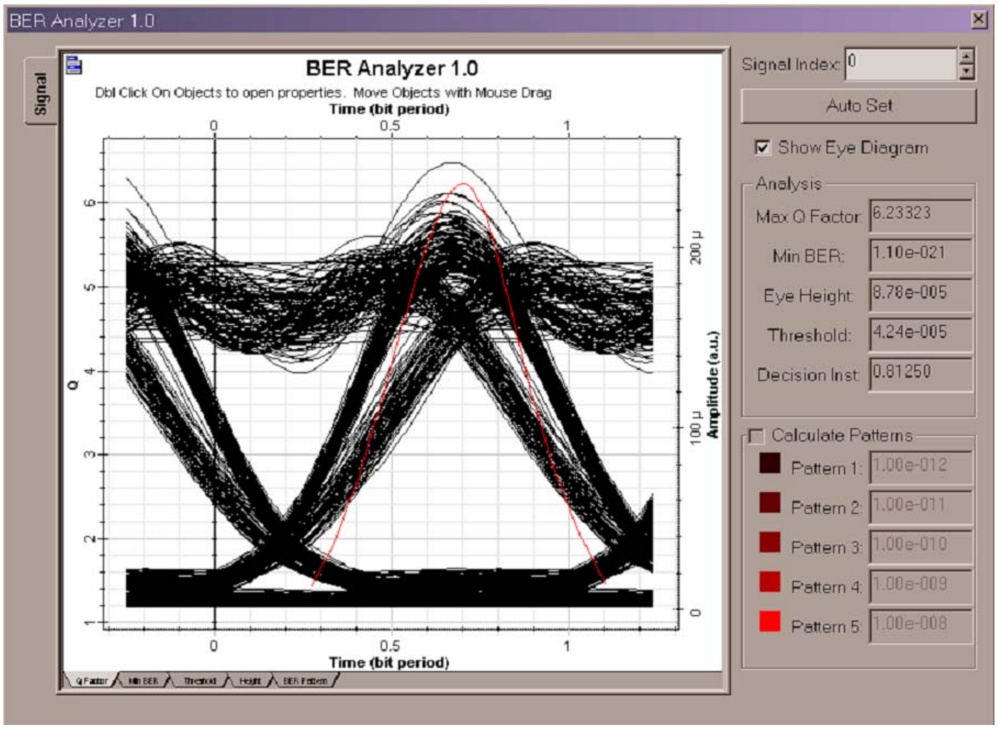


Figure 5.14: Parameters of optical amplifiers and optical fibres, simulation in OptiSystem software

In case of G.652 fibre, 9 EDFAs and 8 DCF modules for compensation of chromatic dispersion have to be deployed for both transmission directions. If G.655 is leased, the number of EDFAs is reduced to 7 and only one DCF module is required. This reduction of both active and passive elements along fibre is very significant from financial point of view, even if the price for leasing of G.655 is usually higher than for standard G.652 fibre. It is not easy to perform exact economical analysis because prices of fibre leasing, EDFAs and DCF modules are vendor and intention dependent. In agreement with available background for our case study, the leasing of dark fibre and deployment of our own equipment was profitable for 25 % in comparison with purchasing of 10 Gbps lambda from a Telco operator.



Signal Index: 0

Auto Set

☑ Show Eye Diagram

Analysis

Max Q Factor: 6.23323

Min BER: 1.10e-021

Eye Height: 8.78e-005

Threshold: 4.24e-005

Decision Inst: 0.81250

□ Calculate Patterns

■ Pattern 1: 1.00e-012

■ Pattern 2: 1.00e-011

■ Pattern 3: 1.00e-010

■ Pattern 4: 1.00e-009

🟥 Pattern 5: 1.00e-008

Figure 5.15: Q factor and BER for 9 EDFAs

Another step will be simulations of DWDM transmission system for the same dark fibre line.

5.6 Microwave-Based First Mile Connection for NREN Data Circuits

The goal of the project is to analyze the suitability of new first-mile technologies concerning the needs of the CESNET2 gigabit network. Furthermore, the project aims to propose and assess new solutions based on new standards regarding high-speed microwave devices (IEEE 802.11a/g/h).

5.6.1 Testing the Equipment Needed for 802.11g Transmissions in the Czech Republic

In June, the IEEE agreed upon the final version of the 802.11g standard concerning the 54 Mbps transmissions inside the 2.4 GHz band. There is reverse compatibility between 802.11g-compliant devices and older 802.11b-compliant ones. Mutual communication may be carried out at 11 Mbps.

The increase in speed was made possible by substituting the original CCK (Complementary Key Coding) modulation method used with the 802.11b DSSS (Direct Sequence Spread Spectrum) by the OFDM (Orthogonal Frequency Division Multiplexing) method, which is being utilized by the 802.11a Standard (5 GHz) as well. But 802.11g and 802.11a devices are not mutually compatible.

All 802.11b-compliant devices use the DSSS method and provide for maximum theoretical transmission speed of 11 Mbps (in reality, the speed reaches approx. 5.5 Mbps). Some companies have implemented the PBCC (Packet Binary Convolutional Coding) modulation method in their products. PBCC may double the speed, however the receiver has to be more sensitive and the signal-to-noise ratio has to be higher. This proprietary technology is often referred to as 802.11b+.

When employing the OFDM modulation used by the 802.11g/a standard, the band is divided into many narrow channels. These channels are used to transmit data at a relatively slow speed providing for more robust transmission than in the case of PBCC. The total data flow equals to the sum of the flow at all the channels and may reach 54 Mbps.

The theoretical transmission speed of 802.11g is 54 Mbps and it should work at the same range as 802.11b. However, the speed decreases rapidly with growing distance, while in the case of 802.11b, the speed should remain constant even close to the limiting distance. Another advantage of 802.11g is a simple migration of 802.11b to higher speeds.

802.11b and 802.11g may coexist as they both utilize the 2.4 GHz band. Thanks to that, existing 802.11b adapters are able to communicate with 802.11g access points. Naturally, the communication will be carried out at only 11 Mbps and the access point has to be configured to allow 802.11b communication (by decreasing the maximum transmission speed). 802.11a networks cannot be gradually upgraded in the above mentioned manner, as they utilize 5 GHz transmission band.

802.11g devices use modern hardware capable of encrypting the data being transmitted with just a slight drop in transmission speed (a few percent). On the other hand, encrypting data being transmitted by a 802.11b device may slow down the communication by up to 30 %.

Many 802.11g devices are currently available at the market.

- ORiNOCO – AP-2000 Access Point, AP-600
  http://www.proxim.com/products/wifi/11bg/
- D-Link – AirPlus Xtreme G
  http://www.dlink.com/
- Buffalo Technology – AirStation G54 Broadband Router Access Point (AP)
  http://www.buffalotech.com/wireless/index.php

- Linksys – Wireless-G Access Point WAP54G
  http://www.linksys.com/products/
- and others . . .

Manufacturers offer a broad scale of products ranging from PCMCIA cards and USB devices to access points combined with simple switches or firewalls.

A PCMCIA card costs approx. CZK 3,000, while the cost of access points varies from CZK 8,000 to CZK 15,000 and sometimes even to CZK 20,000. Cheaper access points are not usually equipped with monitoring tools and that is why choosing the best position of the antenna and finding a free channel may be somewhat difficult. Also, it is not usually possible to find out how fast the devices communicate.

We have performed several tests at the turn of September and October. We have tested an 802.11g Buffalo AirStation-G54 purchased by the University of West Bohemia for the purpose of connecting student dormitories. We have completed indoor tests at ranges of 1 to 30 meters as well as outdoor tests at ranges varying from 1 to 6 km.

Buffalo AirStation-G54 has been designed to be used mainly indoors. It supports both 802.11b and 802.11g and may work either as a multi-client access point or as a point-to-point bridge. It features an internal antenna used for indoor operation. For outdoor operation, it is possible to connect an external antenna (the connector is of the type used by ORiNOCO as well). Besides that, the access point is equipped with four Ethernet ports and may be used as a switch.

The device is configured by means of an embedded web server (and we do not consider it as very well-considered). It lacked the means to monitor signal intensity (signal/noise ratio) or the transmission speed used to communicate with the partner device. The device supports encryption and access lists.

Transmission speed tests have been carried out using two identical Buffalo AirStation-G54 access points working in the point-to-point mode. We tried to transfer files of different sizes, i.e., megabytes, tens or hundreds of megabytes (the size did not affect the speed) via HTTP and FTP. The access points were allowed to use 802.11g only and we tried several power output values ranging from 8 to 22 mW.

Indoor use, 1 m, internal antenna, 22 mW power output: The actual transmission speed maintained throughout the test was 21.6 Mbps. In our opinion, this is the maximum the device is capable of. After enabling data encryption, the transmission dropped slightly to 20.8 Mbps.

Indoor use, 30 m, internal antenna, 22 mW power output: With direct visibility, we managed to reach the transmission speed of 21.6 Mbps. However, every obstacle did cause a significant drop. Maintaining communication through eight walls is extremely difficult and the communication becomes unstable. To build a good-quality internal infrastructure, it would be necessary to use better internal antennae.

Outdoor use, 1 km, ext. antenna (17 dB gain), 8 to 22 mW power output: Noise generated by the city interferes with the transmission and it is difficult to find a free channel. Under the same conditions, the ORiNOCO 802.11b device communicates at 3–4 Mbps with data encryption disabled. Buffalo can communicate at approx. 10 Mbps and encryption causes just a slight decrease to the speed. The transmission speed maintained the same rate when used to connect two computers as it did when set up to provide communication for virtually hundreds of computers found in the student dormitories.

During a short-term test at 22 mW of power output (this value exceeds limits set by the standards), the transmission speed alternates from 7 to 11 Mbps depending on the level of interference. At 8 mW (complying with the standards), the transmission speed alternates from 6 to 10 Mbps.

Outdoor use, 6 km, external antenna (17 dB gain): No noise, all channels are free. The ORiNOCO 802.11b device maintains transmission speed of 5.5 Mbps (maximum speed achievable with this kind of equipment regardless of the distance). After enabling encryption, the speed drops down to 3.7 Mbps.

Buffalo communicates at 1 Mbps. At this range, it is possible to perceive a dramatic reduction of the transmission speed. After switching on the 802.11b mode, the speed increases to 4 Mbps regardless of the encryption setting.

The Buffalo AirStation-G54 rates among cheaper 802.11g/b access points – its price is approx. CZK 8,000. As well as in the case of 802.11b devices, the suitability of the device for an actual solution depends on distances and local conditions. Thanks to the technology, the device may communicate faster with encryption on (recommended) and mode set to 802.11b (which may be done, should the need occur). In a city generating a lot of noise, it may reach 10 Mbps at 1 km. We were not able to test the following situation but it is possible to assume that in a noiseless environment, it could reach communication speeds of 10 to 20 Mbps at 1 to 2 km.

It is possible to use a circulator to boost the signal without exceeding legal limits. The circulator is positioned between the wireless card and the antenna and separates signal emitted by the transmitter from that emitted by the receiver. Basically, this means replacing one antenna with two antennae. The power output of the transmitting antenna may be low (lower than the limits), while the gain of the other antenna may be relatively high (for example 24 dB).

When acquiring new devices for new wireless links, it is preferable to buy 802.11g devices, as the prices of 802.11g and 802.11b equipment are almost the same. Transmission speeds should be higher and in case it proves necessary, the device may be configured to work in the 802.11b mode. For primary connections, it is more suitable to use higher-class devices (such as ORiNOCO AP-2000, AP-600), as they feature better throughout configuration system and monitoring tools.

5.6.2 802.11a and 802.11h Standards and Devices in the Czech Republic

After a long time, the Czech Telecommunication Office has issued an information on The Operation of RLAN Wireless Devices[2] in the 5 GHz band (dated July 2003). According to this document, operation of 802.11a is prohibited. On the other hand, the 2.4 GHz band is public (covered by a general license) and generally available 802.11b (11 Mbps) a 802.11g (54 Mbps) devices may be used freely (a fact that causes a significant overload inside this band). The 5 GHz band is reserved for devices capable of Dynamic Frequency Selection (DFS) and Transmit Power Control (TPC). Contemporary 802.11a (54 Mbps) devices do not meet this requirement.

The situation is very similar in other European countries. However, in July 2003, the World Radiocommunication Conference WRC-03 proposed changes to the Radiocommunication Code valid since January 1st 2005. According to the new regulation, individual countries will be allowed to implement their own general licenses concerning the 5 GHz band. The Czech Telecommunication Office is expected to follow this ruling.

In October 2003, the Czech Telecommunication Office proposed a Provisional General Licence[3] allowing the 5 GHz band to be used freely to build wireless networks. However, the DFS and TPC conditions are not met by 802.11a devices. They are met by the 802.11h standard but no 802.11h devices are available at the moment. Theoretically, it should be possible to make a purely-software upgrade from 802.11a to 802.11h. However, this can hardly be guaranteed for all 802.11a devices available at present.

In future, it is possible to expect a new standard (802.11n), which will provide for transmission speeds exceeding 100 Mbps.

[2] *http://www.ctu.cz/art.php?iSearch=&iArt=292*
[3] *http://www.ctu.cz/art.php?iSearch=&iArt=323*

In conclusion, it is not possible to recommend investments into any infrastructure based on 802.11a. At present, the 802.11g seems to be the most suitable technology for legal high-speed connections.

5.7 CzechLight and TransLight

Within the frame of international projects on lambda services, one Cisco ONS 15454 box was purchased at the beginning of 2003 and it was connected to international exchange point NetherLight in Amsterdam via 2.5 Gbps circuit. Cisco ONS 15454 became the core of prepared network called CzechLight.

When some initial problems were solved (the very first box of its kind in the Czech Republic), we started to test and verify gigabit connectivity (GE channels) to Amsterdam and Geneva. During the first few months, some problems were discovered and they were not resolved until help from TAC Cisco. At the present time, CzechLight is used for connection between Institute of Physics ASCR Mazanka and CERN, Geneva and other lambdas are being prepared for international IPv6 connectivity and for experiments between CESNET and DataTag (Geneva and Chicago).

It was turned out during year that biggest problem is mutual agreement of end users and sometimes the problem is the lack of both of free ports and/or bandwidth (especially abroad) for requirements of various experiments. This kind of problems is now beginning to be solved within the framework of other international projects, especially for networks not under single administration policy (bandwidth on demand in multidomain environments).

CzechLight is an experimental network and it is possible to use it for potentially disruptive experiments. It is completely independent of production network CESNET2, which can't be used for this kind of experiments. As a consequence, some parts of this network are down and can't be used at all. For example, 10GE long-haul tests were performed on international NetherLight line between Amsterdam and Geneva during the summer and connectivity to CERN was completely lost.

Nowadays, CzechLight is a part of international experimental network TransLight which associates experimental networks of Canada, the USA, the Netherlands, Great Britain and northern European countries. The aim of TransLight is to verify the possibility to transfer large amount of data among relatively small number of users on international scale. With the help of this approach it is possible to eliminate expensive IP routers and to deploy new and cheaper equipment like TDM (time division multiplexers) or all-optical switches.

Experimental status of these networks allows to perform experiments even on the lowest layers i.e. directly on optical layer. Combined together with good availability of dark fibres (even international ones today), it affords opportunity to expand CzechLight into other cities in form of an experimental WDM system and to take advantage of our knowledge of NIL long-haul solutions. It offers brand-new opportunities to experimental and research networks.

5.7.1 Experimental Line Praha–CERN for Elementary Particle Physics Research

The full GE capacity link between Prague and the European laboratory CERN in Geneva was established for the use of the Institute of Physics AS CR (FZU) at the Mazanka campus in Prague. The link was used for the large-scale computer simulations of the detectors for the LHC (Large Hadron Collider) experiments (planned startup in 2007). FZU collaborates on the two LHC experiments – ATLAS and ALICE. We have taken part in both experiments Data Challenges – mass detector simulations – with the help of local computing farm Golias. One simulation task lasts typically one day and several hundreds simulations tasks are submitted during one simulation campaign. The results of one task are several output data sets, one of them with size of 200 MB is copied to CERN. During the tests in 2003 some 2 TB of data were transferred to CERN.

The amount of simulation tasks and volume of data transfers will grow in 2004 as a result of improvement in automation of task submission and monitoring of the simulation tasks what is one of the results of the CERN LCG (LHC Computing Grid) project. In the same time, Institute of Physics AS CR runs large-scale detector simulations for the project D0 of the Fermi National Accelerator Laboratory (FNAL or Fermilab) in Batavia, Il. Here we pay our contribution to the detector maintenance and operations by the delivery of the computing services. Further increase of the link capacity over the CzechLight and NetherLight to StarLight is desirable (realization of the Fermilab connection to the StarLight is expected soon).

5.8 Programmable Equipment for Long-Haul Transmissions – Perspective and Possibilities

In the present time we have begun the work on modular system that includes different optical and electronic units. Our goal is to integrate electronic units (based on the COMBO system in particular) together with optical transceivers and optical amplifiers (preamplifiers, in-line amplifiers, boosters) in 19” chassis suitable for placing in professional racks.

Unified access via SNMP will be used for monitoring of all modules. This conception allows us relatively simple creation of complicated functional blocks for both research and development plans and tasks and for use in production environment.

One of the most interesting applications, which uses both programmable logical arrays and optical components, is the design of electro-optical switch for speeds from 1 Gbps to 10 Gbps. We’d like to begin work on this task at the beginning of 2004.

In the first phase, we will focus on research and development of optical amplifiers from standard components (optical EDFA and RFA modules, powers, industrial PC). From these components it will be possible to design and assembly optical amplifiers suited for specific requirements of experimental and research networks (for example CzechLight) and for considerably lower prices than one can expect for commercially available and comparable equipment.

5.8.1 Design of programmable gigabit repeater

The design of the gigabit repeater is based on the card COMBO-4SFP (CESNET Technical Report 12/2003). The COMBO-4SFP is equipped with four SFP cages, two XILINX VIRTEX II FPGA, two SRAM and 3 serial flash EEPROM for the necessary configuration information. One EEPROM is used to keep info about the board (type of VIRTEX’s, ID number of the board, etc.), the other two others are used for configuration of SFP’s.

The COMBO-4SFP has been developed in the frame of *Liberouter*[4] project as interface card for the COMBO6 (PCI hardware accelerator). Both power supply and download of firmware for the COMBO-4SFP are supported by host PC computer through COMBO6.

The desired functionality of the gigabit repeater is not as complex as the necessary functionality needed in the *Liberouter* project. All functions can be provided on the COMBO-4SFP with support of low end processor, boot engine and power supply.

We have built the low cost card COMBO-BOOT (CESNET Technical Report 14/2003) as the inexpensive replacement of the COMBO6 and the host computer. The COMBO-BOOT has two power supplies (input 12–20 V), flash memory for

[4] *http://www.liberouter.org/*

the firmware of both FPGA's, RS 232 interface and Texas Instruments MCU. Free development tools[5] (including C language) for the MCU are available.

The most important goal for the MCU is to download firmware into flash EEPROM (through RS232) and boot FPGA's after power up. The MCU could be also used for the configuration of firmware (either through RS 232 or on board switches) and allowed to add SNMP functionality for the monitoring and control of power supply and link quality.

We suppose to place repeater into 19" 1U chassis with one or two means of power supply (any combination of 220 V and 48 V).

On the COMBO-4SFP we use ser/des chip VSC7145[6] which is able to work with four speeds:

- 1.25Gb/s Gigabit Ethernet
- 2.5Gb/s Infiniband
- 1.062GB/s Fiber channel
- 2.12Gb/s OC48

The repeater will be able to work with all these speeds.

As the COMBO-4SFP can work with all standard SFP transceivers the gigabit repeater is also able to provide functionality of media converter and could also be extended with optical amplifiers.

The design of the 2-port 10GbE card (COMBO-2XFP) for the SCAMPI project is being designed now. When the development is finished, we will use this card for the design of 10GE repeater.

[5] *http://mspgcc.sourceforge.net/*

[6] *http://www.vitesse.com/products/briefs/VSC7145_PB_v3.pdf*

6 Implementation of IPv6 in the CESNET2 network

In 2003 we saw a relatively massive penetration of the IPv6 protocol into the research and education networks. The pan-European GÉANT network started a production IPv6 service in the dual-stack mode and most connected NRENs have also activated IPv6 on their access links to GÉANT. As far as the CESNET2 backbone is concerned, we implemented a new technology instead of the previously used tunnelling, namely IPv6 transport over MPLS (Multiprotocol Label Switching). In the new setup, IPv6 is in a position that is theoretically almost equivalent to IPv4. I reality, though, the situation is considerably worse since

- IPv6 support in routers and other devices is still weak compared to IPv4 in terms of performance, features and stability,
- global IPv6 routing is far from being optimal, especially due to a huge number of tunnels that distort the view of the global topology that is used for BGP routing decisions,
- attractive servers and services, which definitely represent the most important aspect for end users, still rather rarely support IPv6.

In order reach the critical mass of users and through it the ultimate worldwide expansion, IPv6 surely needs a certain degree of a “political” support, primarily from the side of institutions that finance research and development projects. Recently, however, we have been witnessing an unfortunate shift where IPv6 is no more a pragmatic solution to the objective problems of the Internet but rather as a totem or trump card in the technological competition (especially) between Europe and North America. This trend is behind the attempts to accelerate the deployment of IPv6 in end-user networks or even notions like “migration” to IPv6. However, such an approach may potentially lead to adverse effects. Users accustomed to the comfort of IPv4 cannot be expected to take into account “higher interests” and will reject IPv6, unless it provides a comparable quality.

In our view, the best strategy of IPv6 development is its gradual and (as much as possible) non-invasive deployment, starting from backbone networks through campus and local networks down to the end users. In the ideal case the distinction between IPv4 and IPv6 will be completely blurred and either of the two protocols will be used as needed without any intervention or perhaps even knowledge on the side of the user. Therefore, CESNET recognises its role namely in

- an improved support for IPv6 in the backbone network,
- gradual deployment of production IPv6 in the nodes,
- providing the services and applications on top of IPv6 (along with IPv4),
- dissemination of know-how to both users and network administrators.

6.1 Building the IPv6 backbone network

In 2003 we realised the transition from tunnelling IPv6 in IPv4 to the MPLS technology that is used for IPv4 transport as well. It means that IPv6 is now handled almost identically as IPv4. This fact resulted in an improved stability of the IPv6 backbone. As 6PE routers we currently use the Cisco 7500 routers that are still available in most backbone PoPs. So far we were not able to install IPv6 directly on the Cisco 7600 routers that serve as PE devices for IPv4. The reason was a seriously delayed delivery of crucial hardware and software components (in particular the Supervisor 720 engine). Previously used PC/Unix routers are not suitable for this backbone architecture since they lack support for MPLS.

In the course of 2003 IPv6 has been enabled in all gigabit nodes of the CESNET2 network: Praha, Ústí nad Labem, Hradec Králové, Pardubice, Ostrava, Zlín, Olomouc, Brno, Plzeň and České Budějovice. Two smaller nodes (Karviná and Opava) have also been connected to the IPv6 backbone. The configuration of the entire backbone (including international lines) is shown in Figure 6.1.

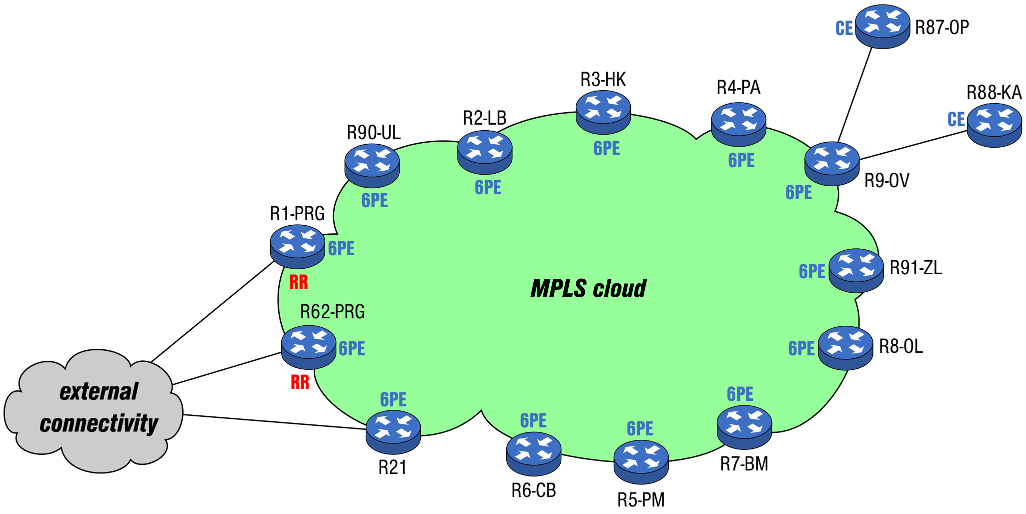


Figure 6.1: Topology of the IPv6 backbone

For IPv6 routing we use the internal BGP protocol configured with two route reflectors. The choice of iBGP was essentially mandated by MPLS. For testing purposes we also use OSPFv3 among the nodes Ostrava, Karviná and Opava. The consolidation of the router platform enabled us to phase out the RIPng routing protocol that was still in use in 2002.

The connectivity and basic services of IPv6 are being monitored using the standard monitoring system *saint.cesnet.cz*. We use either the modules available for the *Nagios*[1] program, or modules developed by ourselves. The monitoring system has a L2 connection to the IPv6 network. We monitor the following areas:

- reachability of backbone routers,
- reachability of border routers of our peering neighbours,
- status of the important IPv6 services,
- BGP routing data.

The migration to MPLS has had an unpleasant consequence for network statistics: we are not able to distinguish IPv6 from IPv4 any more. For this reason we have temporarily suppressed the specific measurements of IPv6 traffic volumes on backbone links and thus have only aggregate volumes for both protocols.

6.2 IPv6 in the networks of CESNET member institutions

Table 6.1 gives an overview of institutions that have IPv6 deployed in at least one network. It is subdivided according to the individual backbone nodes and also the type of connection.

Node	*Native*	*Tunnel*
Praha	CESNET, ČVUT Dejvice	FzÚ AVČR, VŠE
Brno	MU, VUT	–
Ostrava	VŠB, SLU Karviná, FPF	SLU Opava
Plzeň	ZČU, Služba škole	–
Hradec Králové	FAF	–
České Budějovice	JU	–
Liberec	TU	–
Olomouc	–	–
Pardubice	–	–
Ústí nad Labem	–	–
Zlín	–	–

Table 6.1: IPv6 availability in CESNET member institutions

The PASNET network in Prague has so far no support for IPv6 and so the schools and institutes of the Czech Academy of Sciences outside of the Dejvice campus can connect only via tunnels. Nevertheless, CESNET NIC has already included PASNET in the addressing scheme and the address allocations follow common scheme so that after migrating to a native connection all addresses should remain valid.

[1] *http://www.nagios.org/*

The addressing scheme is still the same as described in the technical report 3/2001. The fact that our original prefix 2001:718::/35 was shortened to /32 has not been projected into the allocation policy yet and we leave the extra address space open for future purposes.

6.3 IPv6 peering in the Czech Republic

Other autonomous systems in the Czech Republic can peer with the CESNET2 network only through the national exchange point NIX.CZ. On the CESNET side, the border router towards NIX.CZ is R1.

By the end of 2003 CESNET has the peering agreements with the following subjects:

- GTS CZECH (AS 2819)
- Tiscali (AS 3257)
- IPEX/GIN (AS 9080)
- Pragonet (AS 12767)
- Casablanca (AS 15685)

Peering with the XS26 network (AS 25336) is currently implemented using an IPv6-over-IPV4 tunnel terminated at the R62 router.

6.4 Foreign connectivity

In 2003 we realised native IPv6 interconnections with the pan-European research network GÉANT and with the experimental network of the 6NET project. The same access line to GÉANT in now used for both IPv4 and IPv6 protocols. The router on the CESNET side (R21, Cisco GSR 12008) operates in the dual-stack mode. In contrast, the 6NET access link carries only IPv6 since the 6NET network is IPv6-only.

Transit IPv6 connectivity is provided to CESNET by Telia International Carrier (AS 1299).

National and international external IPv6 connectivity is summarised in Figure 6.2.

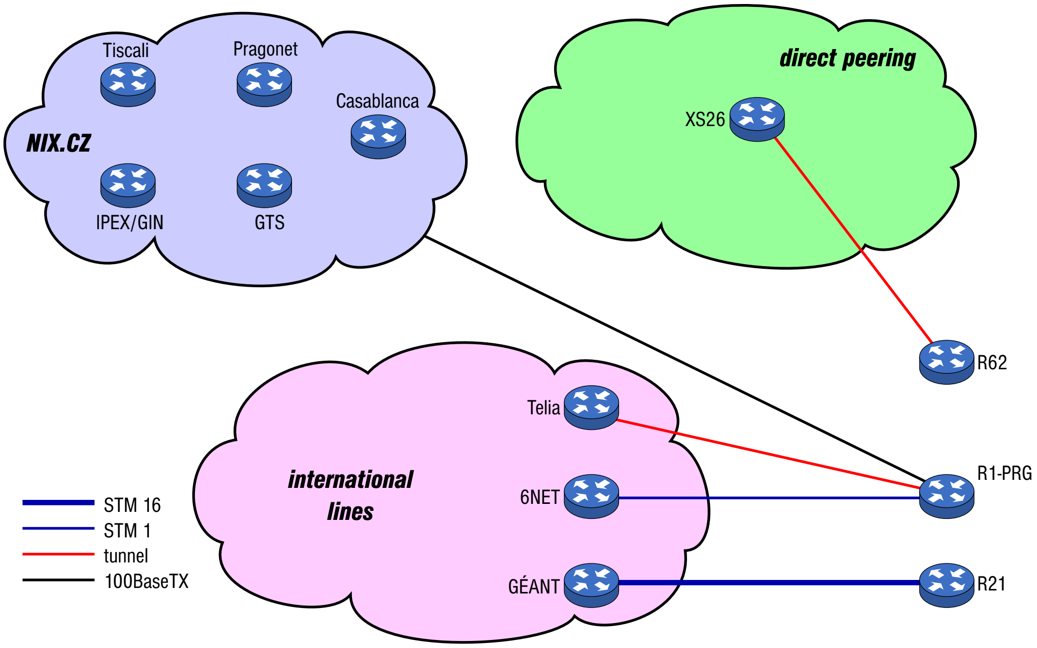


Figure 6.2: External IPv6 connectivity of CESNET2 network

6.5 Project presentations and publications

6.5.1 Seminars

An important event aimed both at a general IPv6 technology dissemination and presentation of our results in this area was the seminar *IPv6 – development and implementation* on October 22nd, 2003. We presented four lectures:

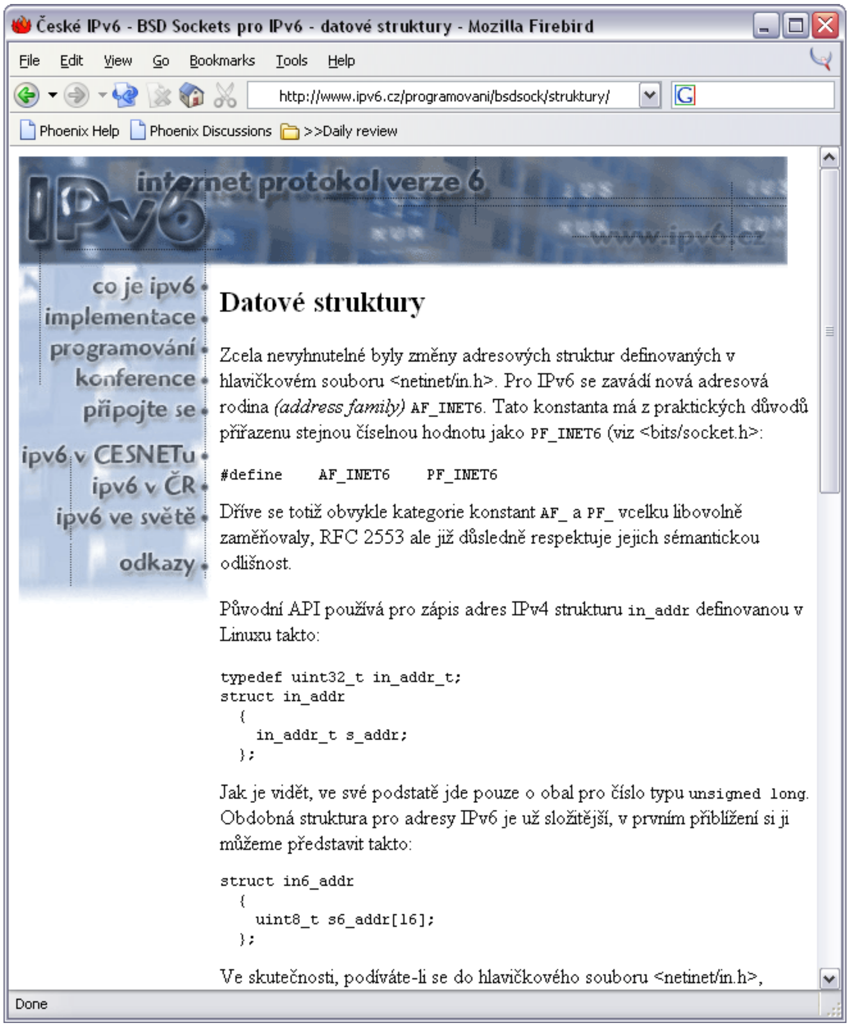
- Recent status of IPv6 standards by Pavel Satrapa
- Current worldwide IPv6 infrastructure by Ladislav Lhotka
- Experiences with IPv6 service in the CESNET2 network by Martin Pustka
- Liberouter project by Jiří Novotný

The seminar was received very positively and the page with the presentations and video recordings of the lectures, accessible from *www.cesnet.cz*, has already had many visitors.

Apart from this event we have been regularly reporting about our work at the meetings of EU projects TF-NGN and 6NET and recently we also took part in the preparation of new projects for the 6th Framework Programme of the EU.

6.5.2 Web

We have been continually updating the WWW pages of the IPv6 project on *www.cesnet.cz*, which contains diverse information about the current topology and status of our network, addressing and so on.



České IPv6 - BSD Sockets pro IPv6 - datové struktury - Mozilla Firebird
File Edit View Go Bookmarks Tools Help
http://www.ipv6.cz/programovani/bsdsock/struktury/
Phoenix Help
Phoenix Discussions
>>Daily review
IPv6
internet protokol verze 6
www.ipv6.cz
co je ipv6
implementace
programování
konference
připojte se
ipv6 v CESNETu
ipv6 v ČR
ipv6 ve světě
odkazy
Done

Datové struktury

Zcela nevyhnutelné byly změny adresových struktur definovaných v hlavičkovém souboru <netinet/in.h>. Pro IPv6 se zavádí nová adresová rodina *(address family)* `AF_INET6`. Tato konstanta má z praktických důvodů přiřazenu stejnou číselnou hodnotu jako `PF_INET6` (viz <bits/socket.h>:

```
#define    AF_INET6    PF_INET6
```

Dříve se totiž obvykle kategorie konstant `AF_` a `PF_` vcelku libovolně zaměňovaly, RFC 2553 ale již důsledně respektuje jejich sémantickou odlišnost.

Původní API používá pro zápis adres IPv4 strukturu `in_addr` definovanou v Linuxu takto:

```
typedef uint32_t in_addr_t;
struct in_addr
  {
    in_addr_t s_addr;
  };
```

Jak je vidět, ve své podstatě jde pouze o obal pro číslo typu `unsigned long`. Obdobná struktura pro adresy IPv6 je už složitější, v prvním přiblížení si ji můžeme představit takto:

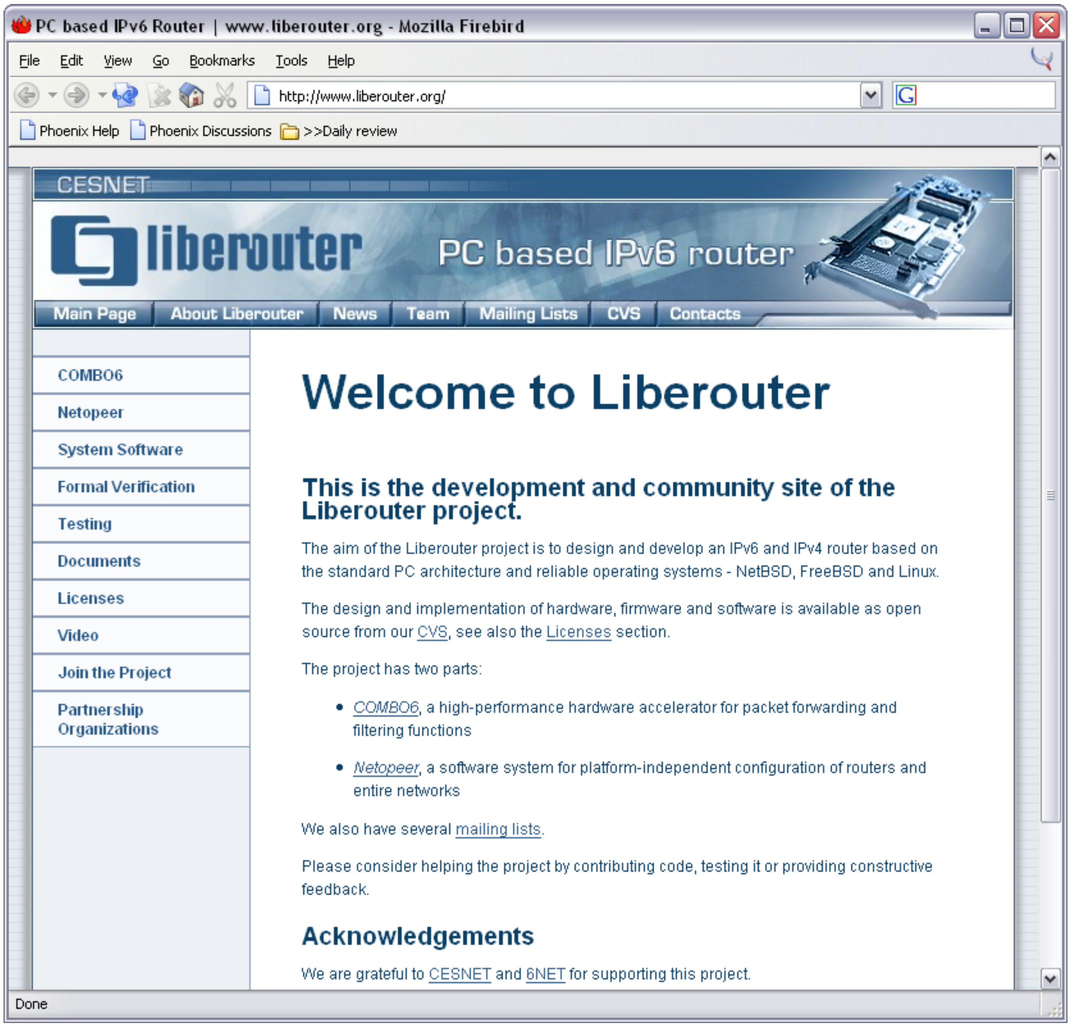
```
struct in6_addr
  {
    uint8_t s6_addr[16];
  };
```

Ve skutečnosti, podíváte-li se do hlavičkového souboru <netinet/in.h>,

Figure 6.3: Server *www.ipv6.cz*

Another server administered by the project group is *www.ipv6.cz* that is intended as a source of general information about IPv6. Apart from texts describing various aspect of the IPv6 technology and implementations, this portal offers a number of IPv6-related mailing lists. There have been no substantial changes in 2003, only the existing contents were updated.

Finally, the *www.liberouter.org* server is devoted to the presentation of the *Liberouter* project. The contents are entirely in English, as this project also has foreign participants. During 2003 the server was renamed (from the original



Welcome to Liberouter

This is the development and community site of the Liberouter project.

The aim of the Liberouter project is to design and develop an IPv6 and IPv4 router based on the standard PC architecture and reliable operating systems - NetBSD, FreeBSD and Linux.

The design and implementation of hardware, firmware and software is available as open source from our CVS, see also the Licenses section.

The project has two parts:

- *COMBO6*, a high-performance hardware accelerator for packet forwarding and filtering functions
- *Netopeer*, a software system for platform-independent configuration of routers and entire networks

We also have several mailing lists.

Please consider helping the project by contributing code, testing it or providing constructive feedback.

Acknowledgements

We are grateful to CESNET and 6NET for supporting this project.

Figure 6.4: Server *www.liberouter.org*

name *www.openrouter.org*), its structure and design was significantly improved and a lot of new information added.

6.5.3 Other publications

In 2003 we also published six contributions in the on-line magazine *Lupa* dealing with diverse IPv6-related topics like deployment, evolution of standards and technological trends.

7 Liberouter

Recent experiences indicate CESNET can play a very useful coordination role in research projects with many participating institutions beside CESNET. This model has been successful especially for complex projects like *DataGrid* (see Chapter 13) or in the last two years *Liberouter*. An effective and targeted cooperation of diverse research organisations seems to be a general problem, evidence of which is for example the effort invested by the EU into support for "Networks of Excellence" in the 6th Framework Programme. CESNET can serve in such cases as a needed neutral platform, assume overall responsibility for the research results and, last but not least, provide such distributed teams with all facilities that are necessary for a successful cooperation – videoconferencing equipment, Web portals and other services.

The basic aim of the *Liberouter* project is to develop a PC-based IPv6 and IPv4 router. At present approximately 50 persons collaborate on this project, about half of that number being students of five universities (MU and VUT in Brno, ČVUT and VŠE in Prague and ZČU in Plzeň).

The Liberouter should provide the following benefits beyond a software-only PC router:

1. Higher throughput in the range of 5–10 Gbps, i.e., roughly ten times the throughput of a PC with standard hardware.
2. Uniform configuration interface comparable to commercial routers.

The first goal is addressed by a hardware accelerator based on the technology of Field-Programmable Gate Arrays (FPGA). In the case of the second goal we decided to pursue a more general solution and create a generic configuration system, which could be used for all common router platforms and also for an effective configuration of entire networks.

The results of the project are publicly available including source code. We use appropriate open-source licenses for different parts of the project, namely:

- GNU General Public License (GPL).
- BSD-type license.
- OpenIPCore Hardware General Public License[1] – this is an analogy of the GPL license tailored for the specific aspects of hardware and firmware development.
- GNU Lesser General Public License (LGPL), suitable primarily for software libraries.

We were also involved in the preparation of CESNET licensing policies and directives for intellectual property protection.

[1] *http://www.opencores.org/OIPC/OHGPL.shtml*

7.1 Hardware accelerator

The architecture of the hardware accelerator as described in the last year's report [Gru03] remains valid. Its leading idea is the so-called hardware-software co-design, which tries to find an optimum division of functions between hardware and software. In the case of an IP router such a division is quite straightforward:

- *Control plane* (routing protocols, configuration and administration subsystems etc.) is implemented in the software of the host computer.
- *Data plane*, which realises packet forwarding between network interfaces, is implemented using programmable hardware, in our case gate arrays Virtex II from Xilinx, Inc.

An important principle allowing us to reuse most of the existing software is a full integration of the specialised hardware into the operating system environment. In other words, the operating system deals with our hardware accelerator in exactly the same way as if it was a common multiport Ethernet card.

We started the year 2003 already with the first prototype of the main board named COMBO6. During January and February the board had been activated and tested. We discovered two minor design flaws that do not affect card's functions but do not allow to achieve the planned performance. We nevertheless decided to manufacture four additional cards using the same design, which have been used by our developers. Then in the second half of 2003 we designed a new corrected revision of the card.

The COMBO6 card is itself not equipped by any communication interfaces. These are supposed to be added by means of an add-on daughter interface card. Two such cards were designed and manufactured in 2003:

1. COMBO-4MTX with four metallic ports of Gigabit Ethernet, see Figure 7.1.
2. COMBO-4SFP with four ports of Gigabit Ethernet in the form of SFP cages. SFP GBIC adapters are available for a variety of media (single- and multimode fibre, CWDM and even metallic). See Figure 7.2.

Figure 7.3 shows the resulting "sandwich", where COMBO6 is connected with the COMBO-4SFP interface card through special connectors.

A card with two ports of 10GE is also being designed. However, we face certain difficulties due to a limited availability of 10GE chipsets. The first 10GE card, to be finished during the first quarter of 2004, therefore internally splits the clean 10GE channel into four channels, each with 1 Gbps. For the next spin we will then either use a suitable 10GE chipset, if it is available by then, or implement all necessary functions ourselves in the FPGA.

For a more detailed description of the hardware design see the paper [NAF03] and technical reports 12/2003 and 13/2003.

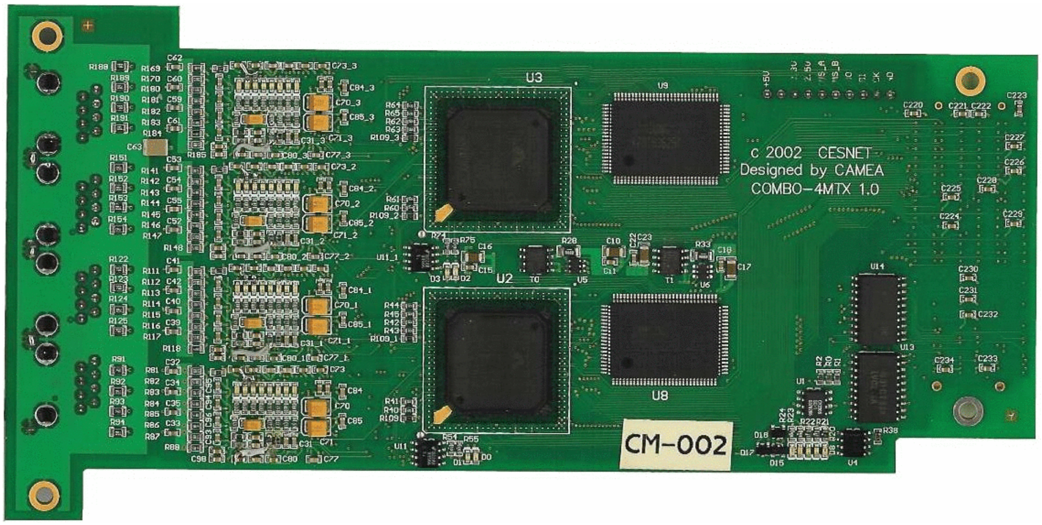


Figure 7.1: COMBO-4MTX card

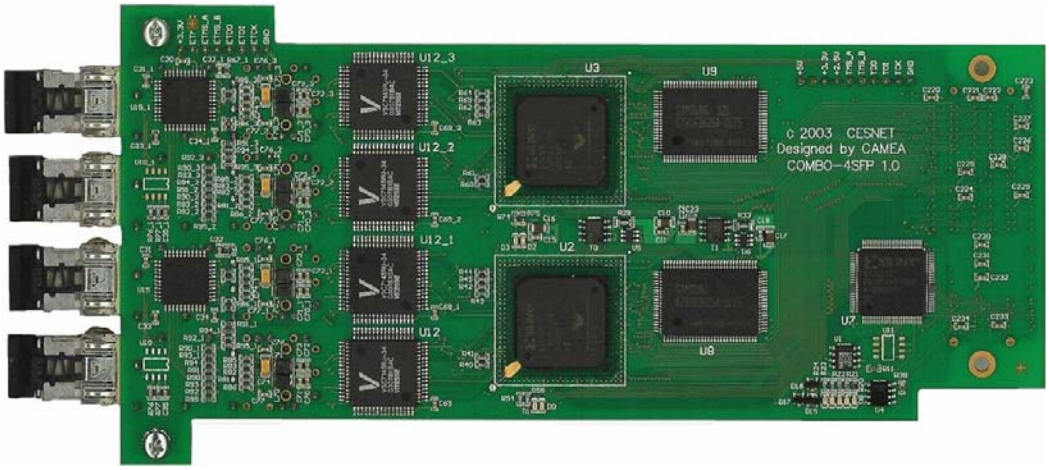


Figure 7.2: COMBO-4SFP card

7.2 Firmware

The most difficult development task, unfinished so far, is undoubtedly the firmware, i.e., special programs for the gate array that handle packet reception from an input network interface, processing the header data, decision about the appropriate action and, finally sending it to the right interface.

Figure 7.4 shows an overall diagram of the firmware, which consists of the following blocks:

HFE (Header Field Extractor): The task of this module is to extract all necessary information from the link and network layer headers (L2 and L3

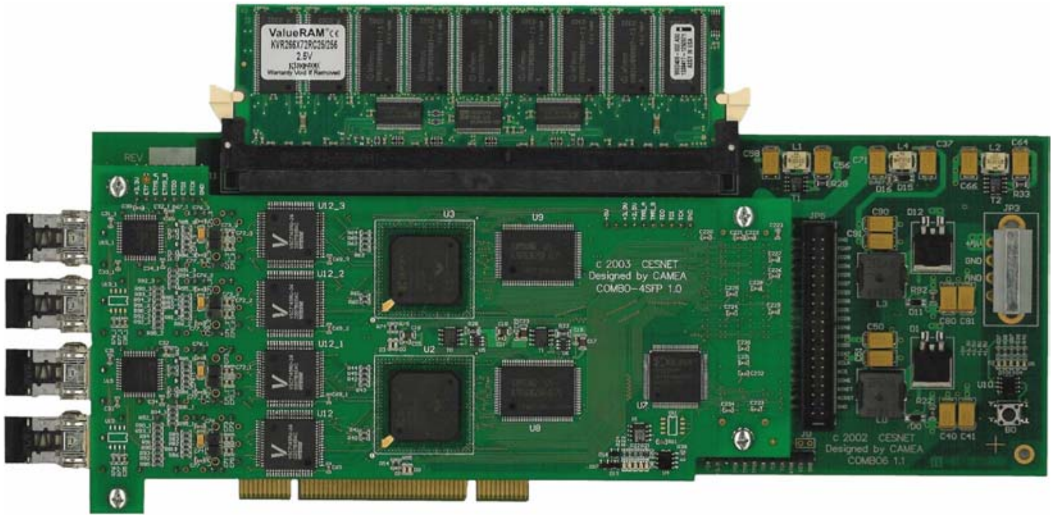


Figure 7.3: COMBO-4SFP card attached to COMBO6

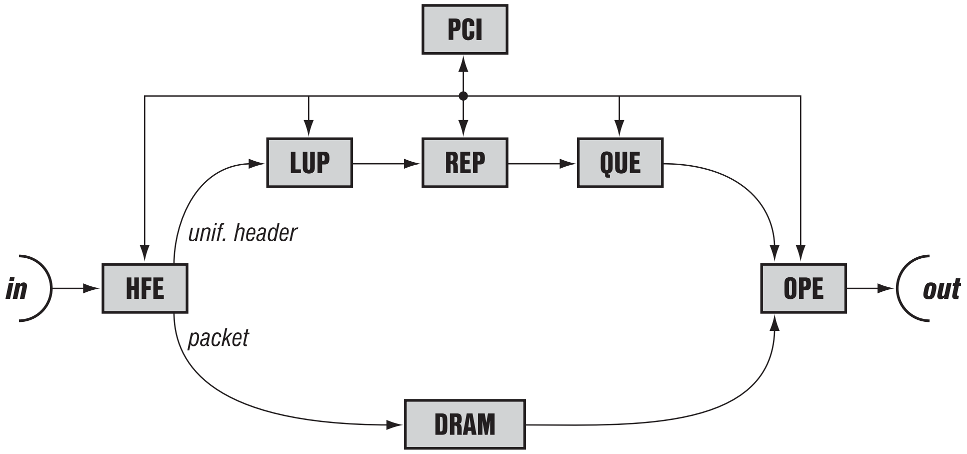


Figure 7.4: Block diagram of the COMBO6 firmware

status registers, source and destination MAC and IP addresses, source and destination port, VLAN tag etc.) and store them in a structure called *Unified Header* (UH). It is 596 bits long and contains the data items at fixed offsets.

DRAM (Dynamic Random Access Memory): During header processing the entire packet waits in the dynamic memory of the COMBO6 card. The other modules deal only with its DRAM address and only at the end of the process the OPE module picks the packet contents from the DRAM and assembles the outgoing packet.

LUP (Look-up Processor): The most complex module, which has to decide about the fate of every single packet, in the first place whether it should be forwarded at all, and if so, how it has to be modified and what are the proper outgoing interfaces. One can imagine the look-up process as a search tree that reads, step by step, the fields in the Unified Header and according to their contents decides about further actions. Our LUP implementation addresses this problem in two phases: First one is a fast search in the associative memory (CAM), which can cover at most one half of the UH bits though and, moreover, is not suitable for an effective matching of ranges (e.g., port ranges). The remaining UH bits are thus handled by an algorithm that is implemented directly in the FPGA. The design of the LUP module is described in [ICN04].

REP (Replicator): This module replicates packets as necessary, or rather just pointers to packets. This is useful primarily for multicast but also for supporting functions like *tcpdump*.

QUE (Output Queue): A packet that is ready for sending is inserted into one of the output queues. In the current firmware version, the set of these queues is managed in the mode of priority queues.

OPE (Output Packet Editor): The task of OPE is to reassemble the entire packet and modify the header data according to both the requirements of a particular communication protocol (for example the TTL field in the IPv4 header) and the instructions issued by the LUP module.

Most of the firmware modules described above are present in four instances – one for each input or output port. The number of output queues (QUE) is configurable in the firmware design.

Figure 7.4 also suggests that all modules can communicate with the PCI bus. This data path is used e.g., for the handling of exceptions that cannot be processed in the hardware (yet).

All modules mentioned above are now under simultaneous development. The developers also already started the tedious process of integration of the entire firmware design and its placement in the gate array.

7.2.1 Methods and tools of firmware development

Programs for the gate arrays are being developed in the VHDL language (Very High Speed Integrated Circuits Hardware Description Language). The firmware developers have access to a professional VHDL development environment consisting of the programs *Leonardo Spectrum*, *FPGA Advantage*, *HDL Designer* and *Modelsim*. In some ways, however, these programs do not support the style of work of a distributed team:

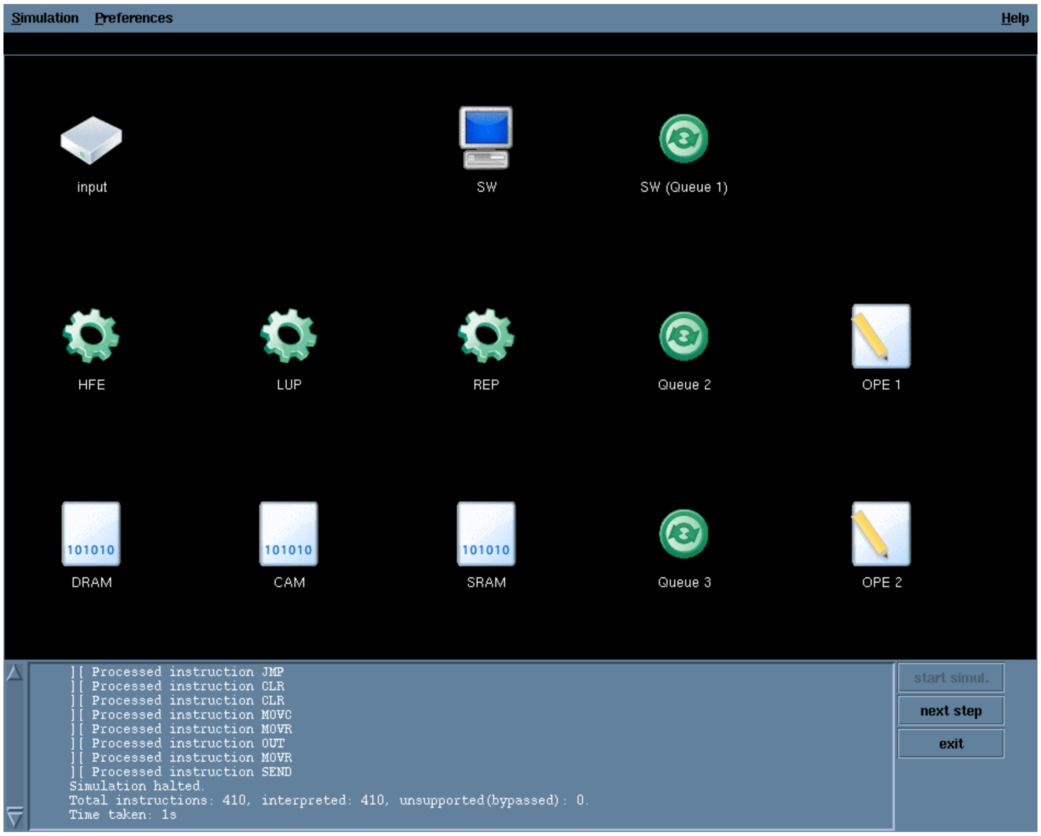
1. The default user interface is graphical, which makes the shared and/or remote access difficult and also does not allow the development tasks and sequences to be effectively automated and documented.
2. Open source development relies to a large extent on contributions from independent volunteers that help with debugging and in some cases also with the development proper. As long as there is no free development kit, a real VHDL programming and debugging is available only to those with access to the expensive professional environment.

Solving the first issue turned out to be quite easy, thanks to the fact that all development steps from the VHDL source code to the microcode of a particular chip are performed by programs with a command line interface anyway. We thus created an alternative development environment that is based on the traditional *make* command. The development cycle can now be fully controlled from a line terminal and, moreover, the *Makefile*s help document the individual compilation procedures and make them repeatable.

The ultimate solution to the problem of opening the VHDL development to a wider community – a free development environment – is not in sight and so we resorted to a partial solution, namely adding another abstraction layer to the firmware design, so-called *nanoprocessors*. These are processors implemented inside the FPGA with a very small but highly specialised instruction set. Hence, the firmware modules described above are implemented in two stages: First a nanoprocessor is designed and then a "nanoprogram" is written for it using the nanoprocessor's instruction set. Whereas the first stage (implementation of a nanoprocessor) required the full-fledged VHDL development kit, nanoprograms can be written or modified relatively easily.

A direct interaction with the machine code is not really enjoyable even for nanoprocessors and so we tried to create a more comfortable development environment enabling the use of assembler-like language for nanoprocessor programming. Filip Höfer solved this issue very elegantly in his Bc. thesis [Hof03]. The resulting program *nsim* is a *generic* compiler, simulator and debugger for *any* nanoprocessor – the instruction set and semantics is passed to the *nsim* program by means of a simple declarative language. The only missing link in the nanoprocessor development tool chain is a disassembler, but it is now also being developed.

In order to be able to perform combined simulations of several nanoprocessors, and also for presentation purposes, we created a graphical front-end to *nsim* named *xnsim*, see Figure 7.5. This program is able to simulate and visualise individual steps of data processing in the COMBO6 firmware. The upper part of the window displays the firmware modules and dynamical exchange of information between them, while the lower parts shows the nanoprocessor instructions being executed and their output. The simulation operates on a sequence of packets that are given to the *xnsim* program as input in the *tcpdump* format.



```
    ][ Processed instruction JMP
    ][ Processed instruction CLR
    ][ Processed instruction CLR
    ][ Processed instruction MOVC
    ][ Processed instruction MOVR
    ][ Processed instruction OUT
    ][ Processed instruction MOVR
    ][ Processed instruction SEND
    Simulation halted.
    Total instructions: 410, interpreted: 410, unsupported(bypassed): 0.
    Time taken: 1s
```

Figure 7.5: Screenshot of the *xnsim* program

7.3 System software

Communication between the COMBO6 card and the host operating system has two main forms:

1. COMBO6 card together with a daughter interface card works as a plain multiport Ethernet adapter.
2. For some purposes (firmware download, debugging) special mechanisms are needed.

We are developing software support for both forms of communication under NetBSD and Linux operating systems.

The first form of communication is enabled by a thin layer of drivers and other system software that hides the peculiarities of the COMBO6 card. Consequently, configuration, administration and monitoring of network interfaces can be accomplished by the usual operating system commands like *ifconfig*, *tcpdump* etc.

The look-up processor (LUP), which is an important part of the COMBO6 firmware, manages its own data structures that govern packet classification and the decision about the fate of each packet. These structures must nevertheless reflect analogical information that is present in the operating system kernel, especially in the routing and filtering tables. For a correct function of the hardware forwarding accelerator, the LUP data must agree with those in the kernel. This synchronisation is taken care of by the *combod* daemon, which is notified about the changes in the kernel tables, e.g., through the *netlink* socket.

A direct access to the COMBO6 and interface cards is enabled by a low-level driver and realised through two character device files */dev/combosix0* and */dev/combosix1*. We have developed utilities for the initialisation of both cards and firmware download. Another interesting program is *comboctl*, which allows the developers to easily perform simple operations like reading and writing from/to specific registers or memory location, and even write simple scripts in the TCL language.

7.4 Formal verification

Formal verification methods, which are also known in the domain of software programming, are needed even more for proving correctness of hardware designs, since errors discovered only after a design has been realised tend to be rather expensive. However, famous flaws in products of even the biggest companies (like the Pentium processor by Intel) indicate that a complete formal verification of designs of very complex components is still just a dream.

The *Liberouter* project also involves students from a group of formal verification at the Faculty of Informatics MU in Brno. Our aim is to apply the formal verification approaches to smaller hardware and firmware modules, but also use the methodology for imposing some discipline on the programmers, for instance by requiring them to intersperse their VHDL code with certain logical formulas (assertions).

The method we use for formal verification is so-called *model checking* that attempts to prove the correctness of a specification of the given system on a certain level of abstraction that is determined by a system's model. The validity of the specification is verified by going through the entire state space of possible inputs and comparing their corresponding output against the specification. If they do not match, one immediately obtains a counterexample.

Automated formal verification is done using the program *NuSMV*[2], which uses a special SMV language for model description. Our formal verification group also developed a set of scripts for transforming VHDL programs into SMV.

Formal verification was applied, for example, to a simplified model of the look-up processor for proving that the CAM and SRAM memories are correctly shared

[2] *http://nusmv.irst.itc.it/*

by all four channels. This application and other experiences are described in the technical report 17/2003.

7.5 Netopeer configuration system

One of the most distinguishing features of commercial routers compared to their PC-based counterparts is the way in which these devices are configured and administered. Router vendors like Cisco Systems or Juniper Networks offer for these purposes very comfortable and consistent command line interface, which covers configuration and monitoring of all hardware and software subsystems. In contrast, PC/Unix router configuration typically boils down to editing diverse configuration files and init scripts.

The goal of the *Netopeer* software is to create both a configuration interface comparable to commercial routers and a platform-independent system for managing configurations of routers and entire networks. The basic idea, and a hard problem at the same time, is to define a neutral configuration format from which and into which platform-specific configurations can be translated. We have chosen XML as the language for the neutral configuration format, primarily because its tree structure is a good match for the structure of typical configuration data, and also because a number of well-established tools for XML processing are available.

Configurations expressed in XML are stored in the configuration repository, which, beside other functions, also supports version control. Software modules that interface with the user are called *front-ends*; their main role is to take a configuration expressed in a user-friendly way and transform it into the XML form. We are working on both *batch* front-ends that do only this transformation and *interactive* ones that also aid the user with preparing the configuration and store it into the repository. On the other side, the interface between the Netopeer system and concrete routers is a realm of so-called *back-ends*. They handle the reverse transformation from XML into the platform-specific configuration language, and may also handle other tasks like installing and activating the configuration in the target router.

As most of the functions are common to several (or all) front- and back-ends, we decided to implement them in the form of two specialised libraries:

- *Netopeer-XML* library provides the front-ends with an application programming interface (API) for dealing with XML files, validation against XML schemas, and accessing element and attribute values by means of XPath expressions. The library also creates a useful software layer, which hides the details of the particular underlying XML parser.

- *Repository library* provides an API for interacting with the configuration repository – storing and retrieving configurations, listing the contents of the repository, displaying differences between versions etc. The library also makes a potential migration to another versioning system relatively painless.

Figure 7.6 shows the relations of back-ends and both types of front-ends to these libraries.

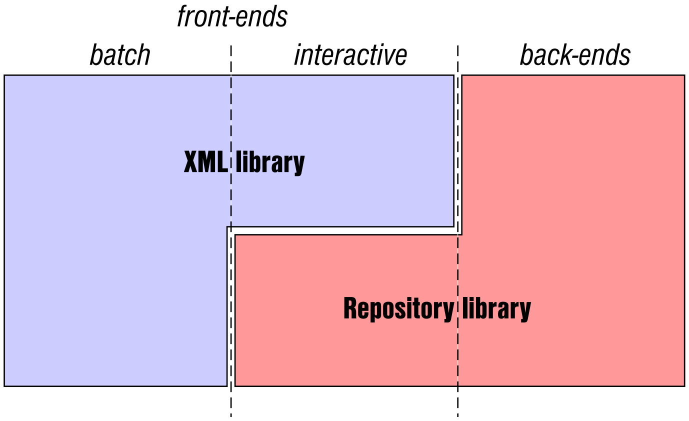


Figure 7.6: Utilisation of Netopeer libraries by front- and back-ends

In the following subsections we will describe the current status of individual software components.

7.5.1 XML schema of the configuration data

Significant part of a typical router configuration can usually be translated from one router configuration language into another. Nevertheless, configurations of certain subsystems (e.g., packet and route filters that are not covered by any IETF standards) are often vendor- or even model-specific and thus hard to translate between configuration languages. Hence, defining an universal and neutral configuration language compatible with at least the most popular router platforms seems to be next to impossible. We thus decided to provide, in an otherwise generic XML schema, a possibility for verbatim sections of platform-specific configuration commands.

Nonetheless, a full coverage of the "convertible" subset is still far from straightforward. In order to enable the development of the Netopeer modules, we decided to define an interim XML schema covering network interface configuration (including tunnels and VLANs), static routing, RIP and RIPng routing protocols, and finally packet and route filters. This schema is described in the technical report 2/2003.

7.5.2 XML tools

We originally intended to base our XML subsystem on the *Xerces*[3] parser (C++ version), which by then seemed to have the best support for W3C XML standards. However, our practical experiences were not satisfactory: Xerces is a rather bloated and slow piece of software with annoying deficiencies (for example in error handling) and incomplete documentation. Fortunately, during the year 2003 another XML library – *libxml2*[4] – reached a reasonable level of maturity. In contrast to Xerces, libxml2 is fast, modular and well-documented. For these reasons we migrated our XML library to libxml2. The current version of the Netopeer-XML library is described in the technical report 22/2003.

7.5.3 Repository

The most popular system for version control, *CVS*[5], which is otherwise used by all members of the Liberouter team for their development work, is not really appropriate as a foundation of the Netopeer repository. From our point of view, the most serious drawback of CVS is the absence of an application programming interface. This turned out to be a problem for several reasons, for instance error handling is very tricky and fragile. We thus adopted another versioning system, *Subversion*[6], which is comparable to CVS in terms of functionality but also provides a decent API. On top of Subversion we built a specific Netopeer repository library. Its installation and API is documented in the technical report 21/2003.

7.5.4 Front-ends and back-ends

At present, all front- and back-ends that are either finished or under development work with the above-mentioned interim XML schema. Completely finished is the batch front-end for Cisco IOS. A similar front-end for JUNOS is also functional but not fully conformant to the interim schema yet. Interactive front-ends are still being developed.

On the back-end side we have two functioning modules, both based on XSL transformations: one for Cisco IOS and another for Linux.

Contrary to our original plan, we decided to stop the development of a SNMP front- and back-end. As it turned out, in the configuration area SNMP is relatively useless and does not cover all necessary parts of router configuration.

[3] *http://xml.apache.org/xerces-c/index.html*

[4] *http://xmlsoft.org/*

[5] *http://www.cvshome.org/*

[6] *http://subversion.tigris.org/*

7.5.5 Metaconfiguration

An interesting application, which will cooperate with Netopeer in the future, is the so-called *metaconfiguration*. It should allow to configure entire networks on a higher level of abstraction and assist during the network design process as some kind of an expert system. Metaconfiguration will also support templates that may be reused many times with varying parameters. The network design and parametrisation will be described again in XML, but using its own schema (i.e., not the one of Netopeer). From this description, configurations of individual routers will be generated automatically, but this time already according to the Netopeer schema.

Metaconfiguration is the topic of the PhD. thesis of Miroslav Matuška, and currently it is in the stage of architectural design.

7.6 Testing

A systematic and independent testing of all hardware and software components is considered an important part of the Liberouter project. In 2003 we set up two test laboratories (in Prague at FEL ČVUT and in Brno at ÚVT MU) and furnished them with the necessary equipment – apart from PCs we also purchased a powerful network analyser, Spirent AX4000.

The testing team elaborated a methodology for IPv6 router conformance testing based on the *TAHI Conformance Suite*[7] and also evaluated available software tools for router performance testing. The results are summarised in the technical report 18/2003.

7.7 Project management, publications and presentations

The number of people directly participating on the Liberouter project is currently around 50, half of it being students. The task of coordinating such a large team, difficult by itself, is further complicated by the distributed character of the team. In 2003 we introduced a management hierarchy with delegated competences. The following workgroups have been formed:

1. VHDL – 17 members, leader Jan Kořenek (VUT Brno)
2. System software – 11 members, leader David Antoš (MU Brno)
3. Formal verification – 5 members, leader David Šafránek (MU Brno)

[7] *http://www.tahi.org/*

4. Testing – 4 members, leader David Rohleder (MU Brno)
5. Netopeer – 10 members, leader Ladislav Lhotka (CESNET)

The main instruments for team coordination have been

- regular weekly videoconferencing meetings of the whole team,
- videoconferences of individual workgroups, mostly also with a weekly period,
- short weekly reports that are obligatory for all team members,
- 7 mailing lists,
- outdoor seminars, organised roughly every three months.

Various technologies and services support team collaboration. Apart from videoconferences and mailing lists we have been successfully using the version control system CVS, where the team members are required to store all their results and documentation. The contents of the CVS repository are open to the public. Another much-needed service is a bug and request tracking system, for example *Bugzilla*[8]. The test operation is already underway and we plan to deploy it in February 2004.

The main project web server is *www.liberouter.org*. It also serves as a portal for accessing other information sources (CVS via a web interface, archives and administrative pages of the mailing lists etc.). The large extent of hardware and software development also means a lot of diverse documentation, that could be immediately presented on the web and discussed both within and outside the project team. Despite considerable efforts and extensive tests of systems for contents management (e.g., *Plone*[9]), we have not arrived at a satisfactory solution yet.

We presented our results at three international conferences (TNC 2003 in Zagreb, FPL 2003 in Lisbon, ICETA 2003 in Košice). Five additional papers were presented at national conferences. An entire session at XIIIth EurOpen conference at Strážnice was devoted to *Liberouter*. We also contributed to the success of the seminar *IPv6 – development and implementation* organised by CESNET on October 22, 2003. Liberouter team members authored or co-authored 11 CESNET technical reports in 2003.

In the area of presentation we also cooperated with a designer on a project logo. The result is shown in Figure 7.7. Consequently, we also ordered and purchased T-shirts, glasses and cups with this logo.

[8] *http://www.bugzilla.org/*

[9] *http://www.plone.org*

Figure 7.7: Liberouter project logo

7.8 Conclusions

Although the *Liberouter* project has not been finished yet, some of its preliminary results – in particular the COMBO6 card – were already used in other projects, for example *SCAMPI* (see Chapter 14) and *CzechLight* (Chapter 5.7), and more applications are expected to be included in forthcoming projects of the EU 6th Framework Programme, GN2 in the first place.

The Liberouter project clearly documents the opportunity for CESNET to act as a coordinator of large-scale research and development projects involving universities and institutes of the Academy of Sciences as partners. As a matter of fact, such a platform is often hard to find otherwise.

Beside the visible results, the Liberouter project has had a number of other positive side-effects, especially in the area of human resources. The students involved benefit greatly from their participation: apart from the technological know-how they also learn how to cooperate in such a diverse team and how to present their results. So far, two Bc. and three M.A. theses stemmed from the project and all got excellent marks. Jan Kořenek's M.A. thesis was nominated for the Prize of the Minister of Education.

8 Multimedia Transmissions

8.1 Objectives, Strategy, and Structure of the Project

In the last two years, the *Multimedia transmissions* project integrated several activities in the area of multimedia data processing and transmission. CESNET is well aware of the importance of supporting advanced applications and this project is thus naturally one of the strategic projects of the research plan.

Our goal has been to create a system for remote collaboration using multimedia applications with a range of audio and video transmission demands. We have aimed to cover different areas of collaboration – from collaboration of individuals to interconnections of specialised centres – by further developing videoconferencing systems, video tools and tools for workspace sharing.

During the last two years project researchers have been focusing on videoconferencing tools suitable for a routine use, systems for implementation of on-demand audio and video transmission and specific applications in the area of asynchronous communication of individuals and groups. The development of network support for multimedia transmission has also been an integral part of this project.

Our effort has been targeted especially on delivering technologies developed by us to potential end users. To this end, we run portals providing basic information that should help end users to select the most appropriate technology and also recommendations and detailed instructions for particular products and possible ways of their utilization. Continuous support of pilot groups and consulting services on multimedia technologies have become an important part of our activities.

The field of multimedia transmissions is evolving quite dynamically and so even after two years of the project we can hardly consider it finished. Despite the multitude of resolved problems we have still many tasks to carry on and new tasks continue to emerge. An important part of our work is an involvement in international activities as well as establishing contacts with similar projects abroad. A close cooperation with researchers working in related areas pursued by CESNET is a commonplace.

The project underwent a large-scale reconstruction at the beginning of 2003. The internal organization was reshaped and we reduced a number of researchers to increase efficiency. The results achieved during the last year suggest this decision was correct. Project researchers published their results in 6 technical reports and 7 journal articles or conference papers. The results were also presented on several important conferences including 3 lectures and 2 posters.

8.2 Collaborative Environment Support

Working on large projects requires coordination and communication between people from the academic centres in the Czech Republic, Europe and often even across the continents. Travelling over long distances is often too expensive and/or time-consuming, while e-mail and telephone communication may not always suffice. A possible solution can then be a videoconferencing service providing real-time video, sound, and shared workspace (editor, whiteboard, . . .).

8.2.1 The reflector

Communication of researchers working on large projects often involves multipart sessions where each participant can simultaneously exchange information with a number of others. One possible implementation of multidirectional transmission is multicast. It can be described as a family of protocols designed to provide management of a communicating group (group creation, login and logout of group members, etc.) and routing of data packets. All the group members receive the data at the same multicast IP address. Data replication and its delivery to all the group members is handled by network nodes supporting the multicast protocols. In particular, the data replication occurs automatically inside the network so that at most one copy of the data is sent over any single line.

This communication scheme brings about increased demands on active components of the computer networks. Current implementations of routing protocols for multicast are often imperfect and cause a high instability of multicast operations, especially in large and heterogeneous networks. It is thus not always possible to deploy a reliable solution based on multicast network services and another more stable system is needed that is also less dependent on the underlying network.

As a consequence of the principles of collaborative sociology and human psychology, communication in a work group typically involves just a limited number of active members. Current high capacity network lines can transfer a reasonable number of data copies without any problem and we can therefore utilize simpler data distribution procedures. One of the possible approaches is data distribution among communicating partners through one central active component (server). Due to the data “reflection” functionality we call such an element a *reflector* or a *mirror*. When using a reflector, each of the group members needs just the ubiquitous unicast network connection and no special network services. Compared to multicast, this approach requires larger amounts of data to be transmitted by the network and so the number of participants is necessarily limited. This scalability issue can be seen as the main disadvantage of such an approach.

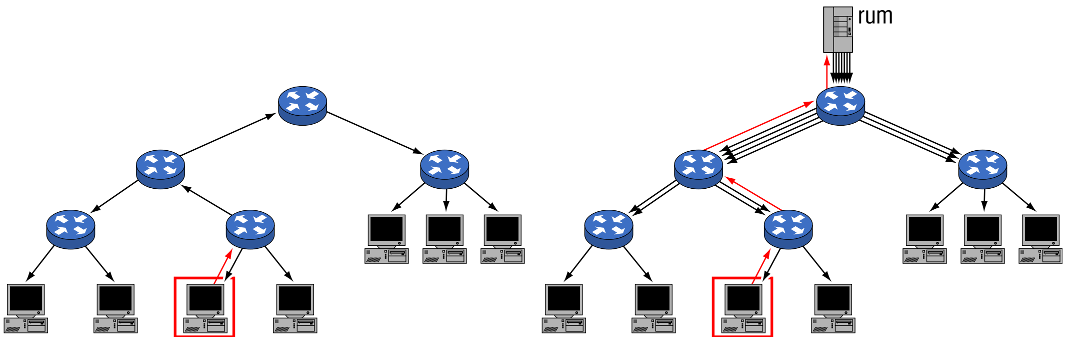


Figure 8.1: Comparison of multicast and unicast communication schemes

A reflector architecture must be flexible enough to allow for implementation of all required features and perhaps even capabilities not yet envisaged. We have utilised the following conceptual approaches to achieve our goal:

- active networks with direct and indirect programmable nodes
- overlay networks for implementing specific services over best effort networks
- ability of an end user to create and to manage the service

The result is a UDP reflector fully controlled by the end users. The basic reflector function – data reception and replication for a group of clients – is augmented by a direct management by the end users and a possibility of incorporating other user-provided modules to accomplish specific tasks.

The reflector consists of a number of components responsible for reflector management and data processing. The data processing is structured into modules for data receiving, shared memory, data classification, process scheduling, specific data processing and data sending. Each module for data receiving is connected to a single port. Received data are placed into an incoming queue, classified and checked against certain permission rules (defined by an AAA policy). The session management module maintains a list of client addresses to which the reflector should send the data. At the end of the processing, prepared packets are placed into an outgoing queue and injected into the network by the packet scheduler. The process scheduler takes care about running the modules for data processing while checking the limits of the allocated resources and providing statistical data. The administrative part of the reflector ensures the communication of the user with the reflector through messaging modules, a control module and an AAA module. The communication uses a specific RAP language.

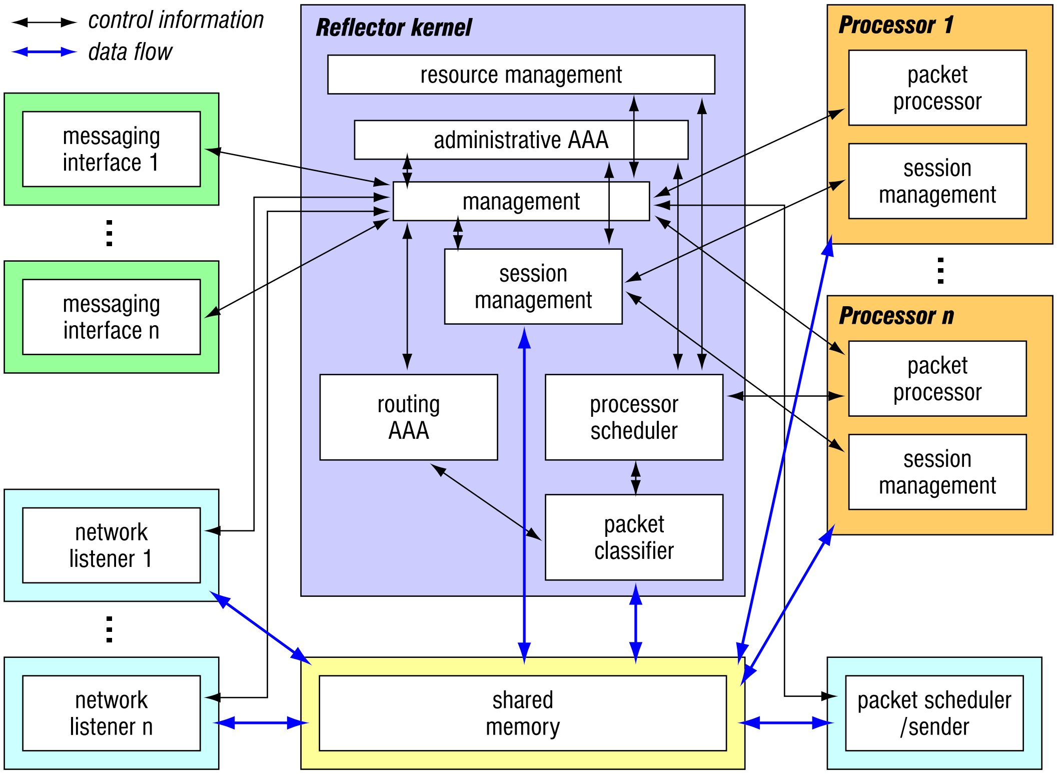


Figure 8.2: UDP reflector – RUM

To ensure better scalability, we can interconnect several reflectors using a mesh of tunnels. The reflector management also supports monitoring of relevant events. The actual processing of data passing through the reflector depends on the used modules. There are modules for recording, data transformation, synchronization of streams, or even for combining several streams into one stream to save bandwidth.

Due to the existence of a separate data copy for each client, this environment is also suitable for strong security. This can be achieved by establishing a secure connection for exchanging encryption keys between the reflector and each client. We can also utilize reflectors in a hostile network environment (e.g., for networks obstructed by firewalls). In this case it is possible to transfer data encapsulated in other protocol (either TCP or UDP), which is permitted by the firewall (e.g. HTTP) between two reflectors. The advantage of this solution is that no changes in the configuration of an existing network are necessary.

8.2.2 AGP, PIG, and the Laboratory of Advanced Networking Technologies

At the beginning of 2003, the Laboratory of Advanced Networking Technologies (ANT) was established at the Faculty of Informatics of the Masaryk University (MU) in Brno. The laboratory stemmed from the joint activities of the Faculty of Informatics and Institute of Computer Science at the MU and the CESNET association. It is a research laboratory specialised in advanced networking protocols and applications requiring high-speed networks. The laboratory provides space and facilities for students who work on projects related either to their curriculum or to their bachelor or master theses. The laboratory is also open to doctoral students working on their Ph.D. degree.

The laboratory is currently equipped with a high-end visualization technology, including one 3D and several 2D projection systems and high-end audio facilities. This equipment is hooked together via a programmable RGB and audio signal switch connected also directly to the IP network for remote steering. Equipping the laboratory with these facilities was a part of the effort to build the first Czech Access Grid node. All laboratory premises are covered by both gigabit wired and 802.11b wireless networks. The laboratory provides enough space for new facilities and enough workplaces for students, including two areas with built-in cubicles.

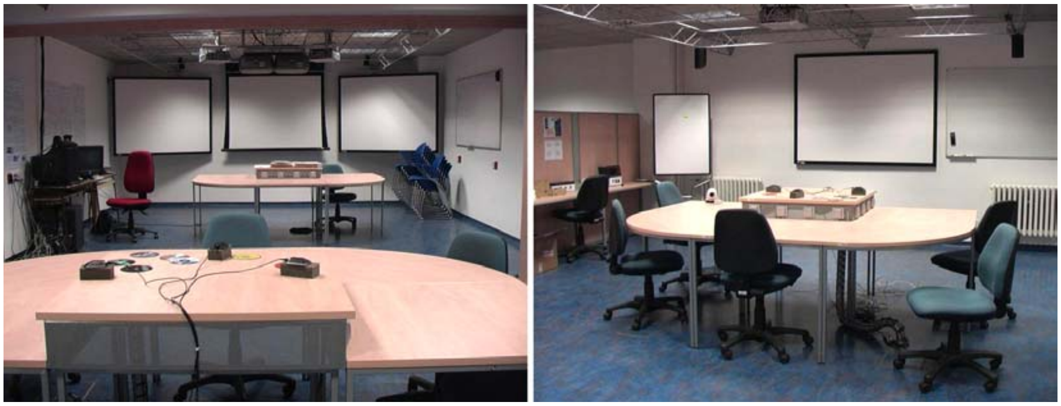


Figure 8.3: AGP Sitola

In the laboratory the very first Access Grid node in the Czech Republic has been build. In order to achieve the main purpose of the Access Grid node, i.e., to ensure high-quality communication with similar installations around the world, we need to satisfy a number of conditions. A recommended architecture, including required technical equipment, is available at *www.accessgrid.org*. Having in mind the fast development in the area of computer technology, multimedia, and computer networks, we decided to modify the recommended architecture so that it would better utilize the available performance and the possibilities of powerful computing servers.

We can see the resulting architecture in the figure 8.4. The main modification compared to the standard scheme is a reduction in the number of computers, compensated by a heavy increase of their performance:

- combination of computers for audio and video acquisition and encoding resulted in a single dual-processor computer with the FreeBSD 5.1 operating system instead of Linux,
- combination of the visualisation and control computers into a second dual-processor computer running Windows 2000.

In the opposite direction, the installation of a passive 3D projection (optional for a standard AG node) was an AG scheme extension. Due to the focus on computational chemistry applications and on the research of transmission protocols for synchronous multichannel transmission (multi-channel sound or stereoscopic 3D video), it became one of cornerstones of the laboratory equipment.

All the equipment of the AG node can be remotely controlled. In the next stage we expect to extend the current solution that uses touch panel with our own programmable system. Its web interface will allow to specify and program even more complex scenarios including video, sound, data routing through the network and other elements of the environment (e.g., lighting). The first objective is the development of the software for room control with the following features:

- easy addition and removal of controlled equipment,
- authentication,
- user access control based on a system of roles,
- pre-defined scenarios for particular types of videoconferences.

The software should be integrated with the present AG version 2.0, which is based on the OGSA technology. Another ongoing activity is a research on synchronization protocols for 3D projection and the proposal of system for 3D data storage and sharing.

8.2.3 3D Video Transmission

The research in the area of 3D video transmission in best-effort networks concentrated on the use of DV technology, which appears to be promising thanks to the high number of equipment available at reasonable cost. The advantage of this system is a possibility of implementing a solution without any license fees.

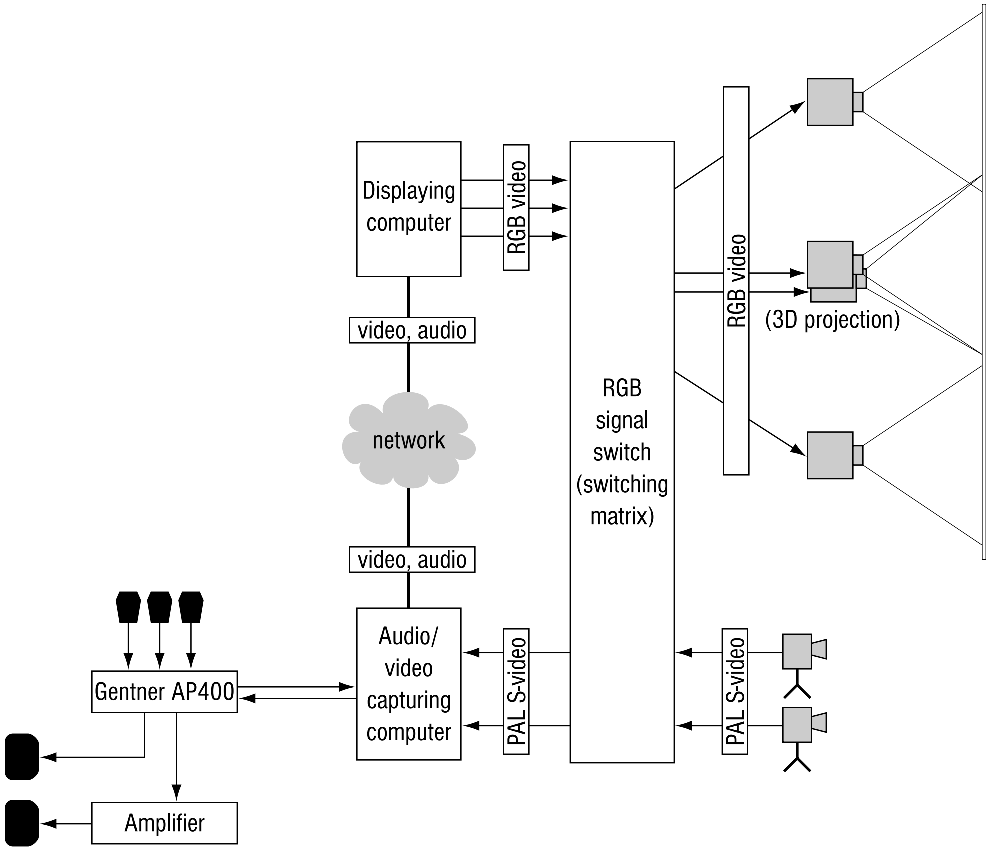


Figure 8.4: The AGP Laboratory scheme

A software solution has been implemented by the *DVTS* project, which also delivered two RFCs specifying transmission of DV through IP networks (RFC 3189 and RFC 3190). The implementations for the Linux and *BSD platforms comprise tools for reading the DV stream from the IEEE-1394 interface and sending this stream through the IP network. A separately maintained *xdvshow* program can then be used for viewing received data in X Windows. A Windows 2000/XP version of this display tool also exists and is based on the DirectShow technology.

Recently, the main developer of *DVTS* left the project and its further development has been stalled. We therefore decided to continue the development of the *xdvshow* tool which had the worst deficiencies among all *DVTS* tools. After implementing both direct imaging using the a X11 interface and imaging via the SDL library, we also added support for full-screen mode based on SDL. The new version of *xdvshow* uses multi-threaded architecture that results both in a more robust data reception from the network and in higher performance enabling the display of two simultaneous DV streams on an ordinary PC without any hardware acceleration. It is thus suitable for software-based display of a stereoscopic video using two independent DV sources.

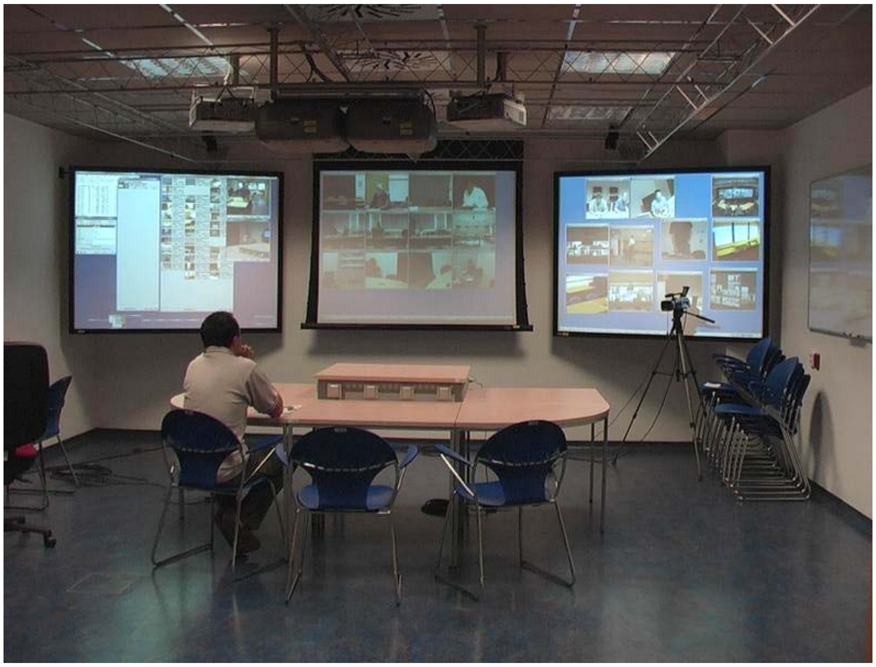


Figure 8.5: AGP Videoconferencing

The 3D scene reading is performed using two DV cameras placed on a stereoscopic tripod head. The data are transferred to the computer via an IEEE-1394 interface and sent over the network with an aggregate data stream of more than 50 Mbps. The display computer receives the data from the network and displays them using two projectors with polarisation filters with orthogonal planes of polarization. The projection uses a special non-depolarising screen, i.e., the polarised light reflected off the screen almost completely preserves its polarisation characteristics. The observer then wears glasses with orthogonal polarising filters that re-create the stereoscopic effect.

8.2.4 Distributed Environment for Video Editing and Encoding

The ever increasing number of requests on video archives in the academic community motivated us to create a pilot editing workplace at the Institute of Computer Science of MU and, specifically, develop an environment for distributed video encoding. The editing workplace is located at the ANT Laboratory premises and is based on technologies from AVID and Adobe. We mainly use the AVID Xpress Pro product, but Adobe Premiere Pro and Adobe AfterEffects are also available.

The output from the editing systems is encoded by a distributed processing environment, typically into the RealMedia format intended for streaming. For this

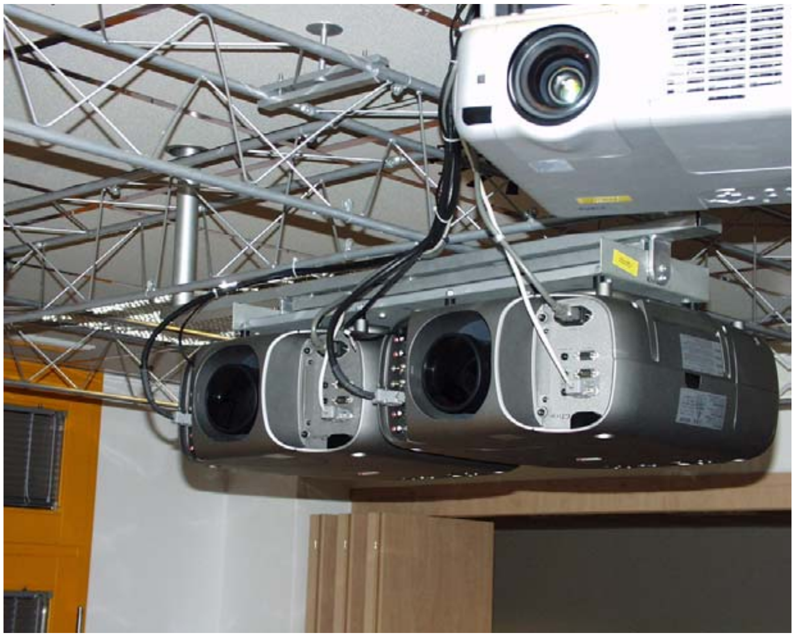


Figure 8.6: 3D display – Laboratory

purpose we use the extensive computing and storage capacity of the MetaCenter project. We have developed a prototype of a distributed encoding environment capable of parallel encoding using up to several hundred processors simultaneously. The stream designated for encoding is first stored to a distributed data storage based on the IBP protocol, then automatically split into smaller chunks, which are processed in parallel, and finally joined into the resulting file. Moreover, the available computing power allows complex image transformations (e.g., high-quality de-interlacing or resolution downsampling) to be performed on the fly.

8.2.5 H.323 Videoconferencing Infrastructure

During 2003 we continued to work on the stabilisation and development of the H.323 videoconferencing infrastructure. In particular, we aimed at improving the integration of the existing end stations into the current infrastructure. We distributed several videoconferencing kits *Polycom ViewStation* and *ViaVideo* to a number of sites in the Czech Republic. The sites with extensive videoconferencing demands and/or many users are equipped with the stations supporting conferences in the 4CIF resolution (full PAL) and have the functionality of a small MCU.

We cooperate with the manufacturer on solving errors in the current station software and intend to ask them to add new functions in future software versions.

Figure 8.7: Polycom FX and ViaVideo videoconferencing stations

As of late 2003, up-to-date versions of firmware providing increased resilience against various network attacks have been installed in most devices.

Currently we are evaluating the quality of the H.264 video codec implementation and the encryption support (AES) in these stations. We expect to deploy these new functions on the major MCUs (Prague and Brno) in early 2004.

An ongoing cooperation with researchers working on the strategic project *Voice services in the CESNET2 network* resulted in an alternative numbering plan that also incorporates videoconferencing stations. The adopted numbering plan enables stations to access the VideNet videoconferencing network, public telephony network, and CESNET IP telephony network. We are continuously carrying out interoperability tests of videoconferencing stations, IP phones and audio and videoconferencing software applications. We also participate in an inter-project activity that works towards the implementation of a directory services extensions defined by the H.350 standard.

8.2.6 Small multimedia platforms

In 2003 we started the development of our own multimedia platform. Our aim is to create mobile and multi-purpose systems for multimedia transmission focused especially on videoconferences, multimedia acquisition and presentation. Its benefit will be in a simple and uniform control, mobility, low operation costs and reduced noise. We are also developing new acceleration modules, which can make these systems attractive even in the area of special high-quality A/V transmission.

Regarding hardware design we have finished work on two prototypes of this platform for videoconferencing and streaming. Another set of devices is under development supporting the MPEG4@IP transmission using an accelerator card based on programmable hardware (FPGA).

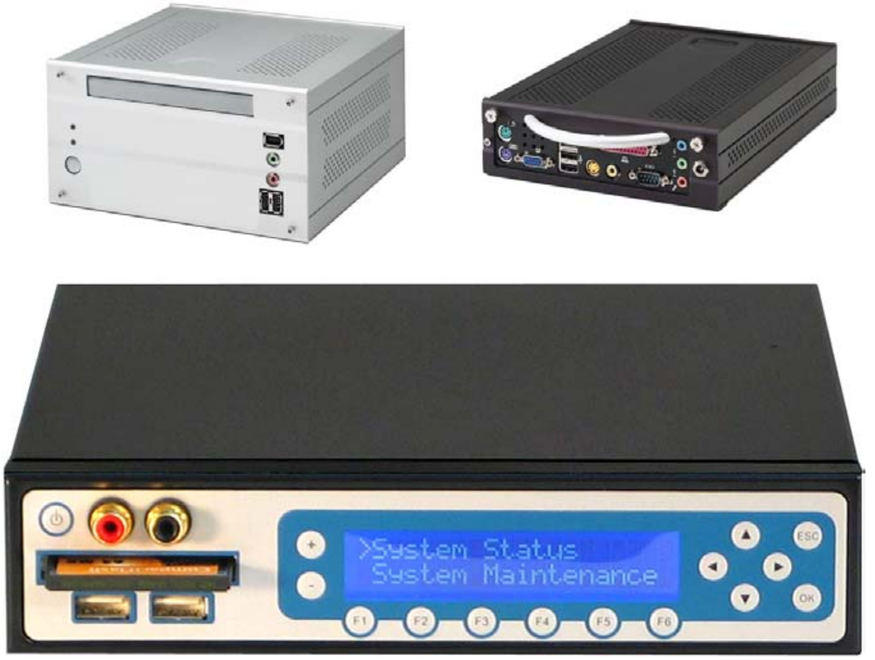


Figure 8.8: Examples of small multimedia platform prototypes

8.2.7 Support of Pilot Groups

We also continue our support to the projects that use videoconferencing facilities (e.g., traditionally *DataGrid* and *IPv6*). The experience gained by these groups provides us with important feedback, often indicating weak points in the services provided by our project.

We have created several new types of videoconferencing sets and delivered them to users.

8.3 Special Projects and Activities Support

8.3.1 Cooperation with AVC SH CVUT

Another interesting cooperation that we started in 2003 involves the student group AVC SH (Audiovisual Center Silicon Hill). The goal of this group is to establish a sophisticated semi-professional studio capable of producing and processing high quality digital video – both on-line live streaming and off-line processing of recordings. Their current efforts focus on activities and events organised by the Silicon Hill and the Student Union of CVUT, e.g., OpenWeekend,

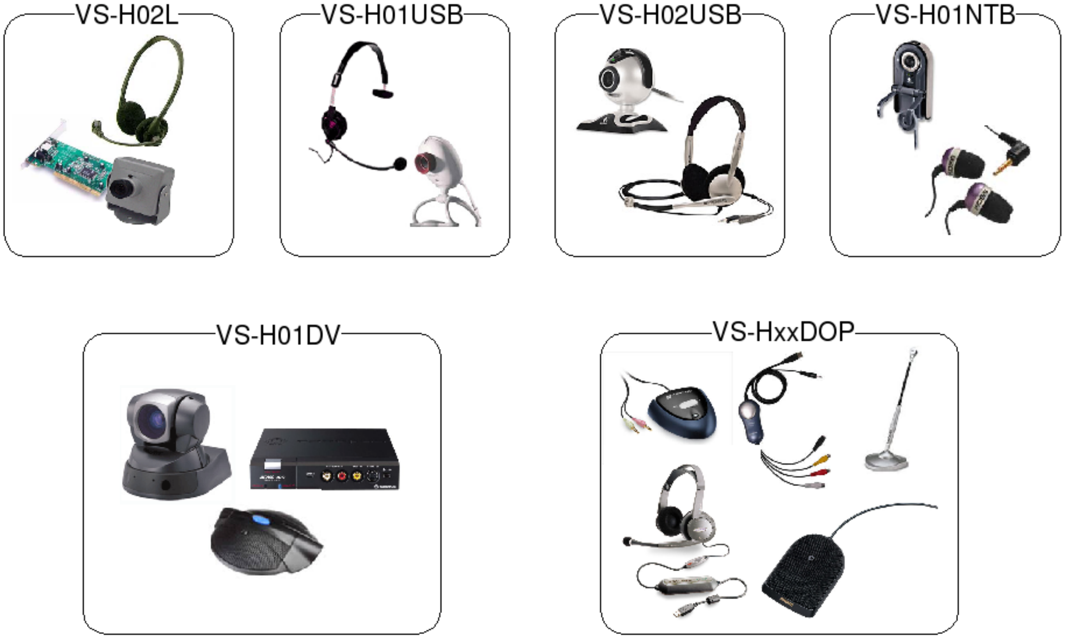


Figure 8.9: Examples of our most widely used videoconferencing sets

InstallFest, CryptoFest, and lectures arranged by the SU CVUT. The long-term objective of this group is to implement webcasting of lectures at CVUT. A number of such webcasted lectures have already been organized, most notably the Physics Thursdays, a unique cycle of lectures and seminars arranged by the Department of Physics at FEL CVUT and Programmer Evenings at FEL CVUT as well. All essential information about these activities is available at *avc.sh.cvut.cz*.

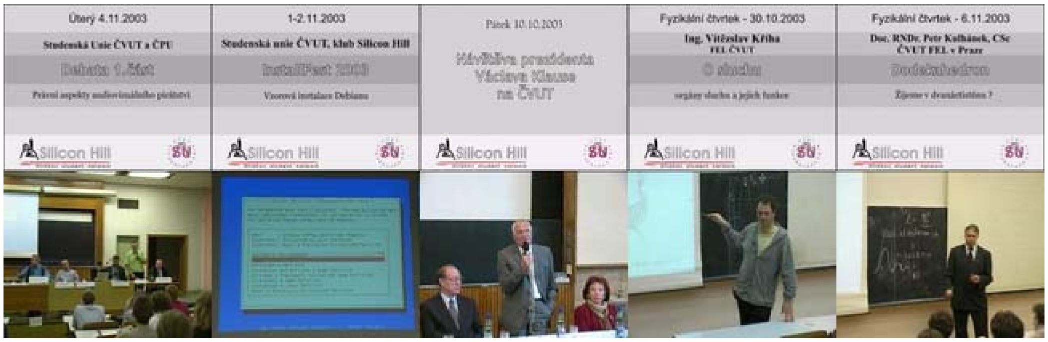

Úterý 4.11.2003
Studenská Unie ČVUT a ČPU
Debata 1.část
Právní aspekty audiovizuálního pirátství
1-2.11.2003
Studenská unie ČVUT, klub Silicon Hill
InstallFest 2003
Vzorová instalace Debianu
Pátek 10.10.2003
Návštěva prezidenta Václava Klause na ČVUT
Fyzikální čtvrtek - 30.10.2003
Ing. Vítězslav Kříha
FEL ČVUT
O sluchu
orgány sluchu a jejich funkce
Fyzikální čtvrtek - 6.11.2003
Doc. RNDr. Petr Kulhánek, CSc
ČVUT FEL v Praze
Dodekahedron
Žijeme v dvanáctistěnu ?
Silicon Hill
SU

Figure 8.10: Results of the efforts of the AVC SH student group

Our contribution to this activity lies mainly in consulting, helping with some scenario preparations, and lending A/V technology.

8.3.2 Recoding and streaming of lectures and other activities at FI MU

Lecture recordings consist of several parts: video, sound and written materials communicated to the students either using a data projector or traditional means like a blackboard. The video part is captured with an ordinary consumer-class tripod-mounted camera (e.g., one of the cameras from the Sony HandyCam series). The camera is connected to the S-video computer interface. We use a PC with an ATI TV Wonder card and the Linux operating system. A virtual driver – so-called video loopback – has been added to the Linux kernel. Inside the computer, the file for streaming and download is created using the program RealProducer 8.51 Basic by Real Networks. Sound is acquired by interconnecting the lecture hall sound system with a sound card in a PC. It is also possible to use sound captured by the camera, though its quality is lower.

Written materials can be captured using the camera. Actually, this is the only option if the teacher writes on the blackboard. Otherwise, if the presented materials are available in digital form, we can insert them directly into the recording and obtain much better image quality. We use the *multiplexor* software for adding the presentations.

The recordings are available from the web portal *video.fi.muni.cz* and immediately accessible from any computer in the *muni.cz* domain. Access from the rest of the Internet requires authentication. One can either directly play the recording by clicking an appropriate link, or download it to the local disk for later playback. The size of an average lecture recording is approximately 300 MB when using single stream with bit rate of 512 kbps. It is thus possible for students to burn the recordings on CDs if needed.

As a test of our abilities in video acquisition and processing, we tried to record on-stage a student performance of Shakespeare's Hamlet. We used several cameras and subsequent video editing. The entire performance was recorded using 3 different DV cameras (Canon XM-2, Canon XM-1, and Sony TRV-30E). This setup led to a subtle complication, since even in the case of prosumer cameras the recording speeds were not identical and we had to correct (re-synchronise) it during editing phase.

We used cassettes with 80-min. tracks in order to avoid changing the cassettes during the performance. The content of the cassettes was transferred through a FireWire interface (IEEE-1394) to a computer disk and stored in the AVI format. The video was saved directly in the DV format without any re-compression (thanks to the fact that same format was used on the cassettes). Minimum requirements for the acquisition computer are thus a sufficiently fast disk and a simple FireWire interface like the OHCI chip by Texas Instruments, which is now often directly integrated on commodity PC boards. As the acquisition

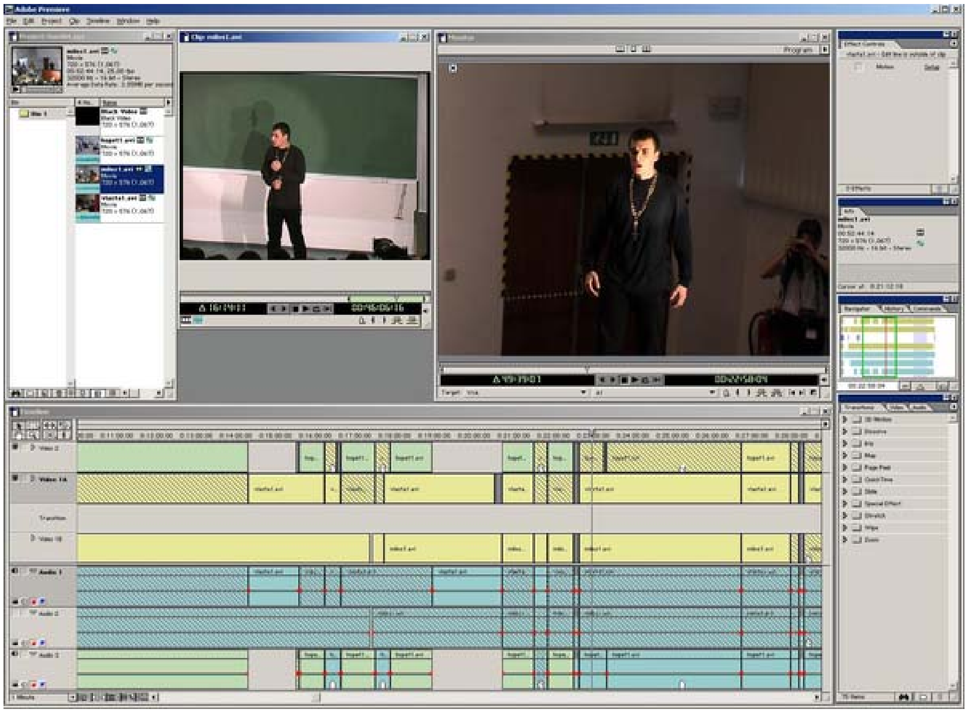


Figure 8.11: Hamlet – processing a theater performance recording

software we use Adobe Premiere (version 6.0 or higher). It is also important to have enough free disk capacity – one hour recording corresponds to a file of approximately 14 GB.

The editing has been performed using Adobe Premiere 6.0 as well. Since we are not professional video editors, we spent about 14 hours before becoming reasonably satisfied with the result. This is by no means a stunning efficiency for one hour of video – hopefully the new editing workplace at ICS/FI MU will help us in this respect. We asked the director of the theatre performance to evaluate the edited video. Following his comments, we added front and rear subtitles. The final result was transcoded to the RealMedia format for streaming from the CESNET streaming server in two versions: the first with a maximal quality, intended for powerful computers with a broadband network connection (data stream is around 3 Mbps), and the second one of a lower quality but also less demands on the connection capacity and performance of the receiving computer.

8.3.3 Virtual Participation at APAN Conference – A Telepresentation

Although the ANT Laboratory at the Faculty of Informatics of MU is primarily a research and development lab, the technology is also utilised for supporting regular videoconferences and special events, for example a virtual participation and lecture at *APAN* conference. This international conference, which is organized by universities and providers of high-speed academic computer networks in East Asia and Pacific areas, was held in Busan, Korea. The videoconferencing tools helped us to participate and even present a lecture without the expensive and time-consuming travelling otherwise needed for attending the conference in person. Of course, part of the time thus gained was amortised in the preparation and technical support of the videoconference.

Conference organisers proposed to use H.323 protocol for video and audio transmission combined with an ordinary presentation in PowerPoint format. As easy as it may seem, we had to cope with unexpected geographical timing issues. Our lecture was included in the session on so-called Logistical Networking but, as it turned out, none of the session lecturers was present in person and all lectures thus were to be presented from remote places. Two of the lecturers were actually in the same time zone as the conference venue (Singapore and Korea) but another was from the USA and yet another (us) from Europe. It was thus impossible to find a time slot comfortable to all participants. Finally, the session took place from 7:00 to 8:30 of local Korean time, corresponding to the interval 1:00–2:30 AM in Brno.

In order to test the videoconferencing setup and network connectivity between us and Korea, we organised two testing sessions in advance. Videoconferences based on the H.323 protocol don't have high bandwidth requirements (typically below 1 Mbps) and we thus supposed the intermediate academic networks would provide enough capacity for ensuring a good transmission quality. However, first test results were an unpleasant surprise for us – network failures and data losses in the network degraded the image and sound quality and rendered it practically unusable. Upon a closer investigation, the connectivity problems were localised on the Korean side and successfully eliminated. A consequent test then resulted in an excellent video and audio quality.

The lecture mentioned above demonstrated that with existing technologies and the capacity of academic computer networks similar activities can be organised with minimum risk. After a careful preparation and with an appropriate technical support (microphones, camera, proper teleconferencing facilities), one can transmit video and audio data through the network essentially to any place connected to the world-wide academic network. With a proper preparation

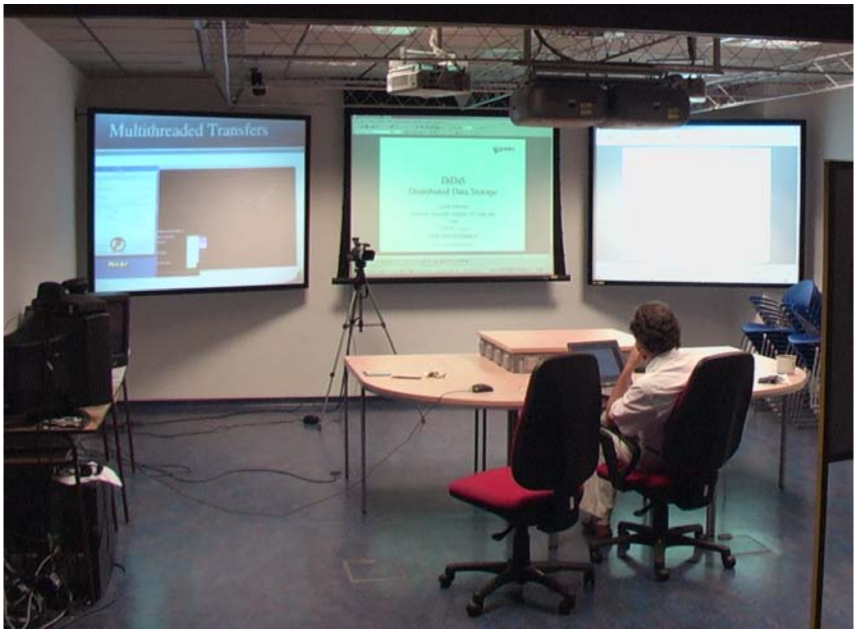


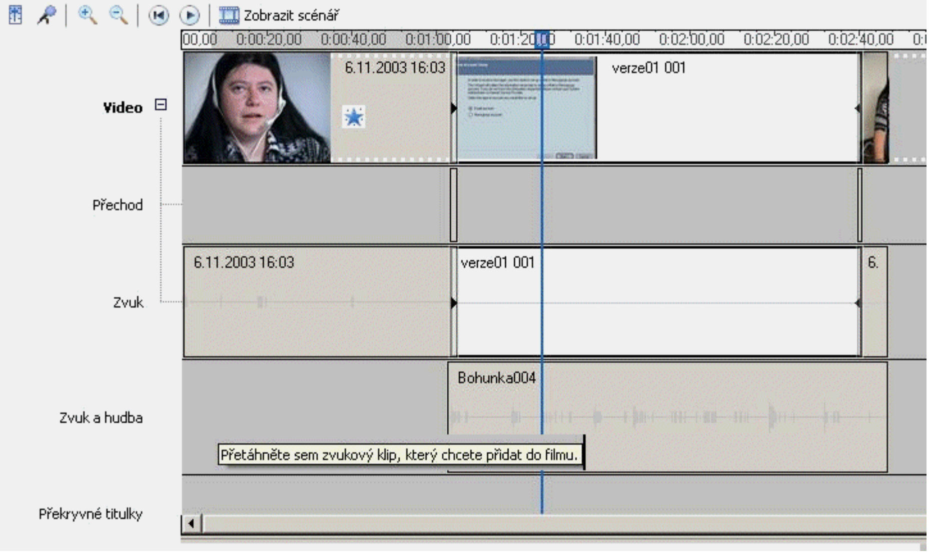
Figure 8.12: Virtual APAN participation – telepresentation

and some training of the lecturer, remote lectures may help to avoid excessive travelling of today's busy professionals.

8.3.4 Education Support at the University of Economics (VSE)

Activities aimed directly at the education support were an important part of the research effort. The creators of multimedia educational applications have to face the problem of a seamless integration of standard video contents with whatever happens on the computer screen. To help with these issues and disseminate knowledge about the solutions we tried, we published a CESNET technical report (11/2003) called "Use of Digital Non-Linear Editing Machinery for Creating Multimedia Education Lessons" (in Czech). In this report we presented the verified and recommended procedures. The proposed technology is based completely on free software thus minimising implementation costs.

Further activities focused on proposing and implementing a technological platform for recording and direct broadcasting of activities associated with education. Already in 2001 it was a technological tool chain for recording and direct broadcasting based on the analog technology. In cooperation with the Department of Philosophy at VSE, we tested the system on live lectures ("Philosophy and science methodology" lectured by doc. Pstružina for doctoral students).



Přetáhněte sem zvukový klip, který chcete přidat do filmu.

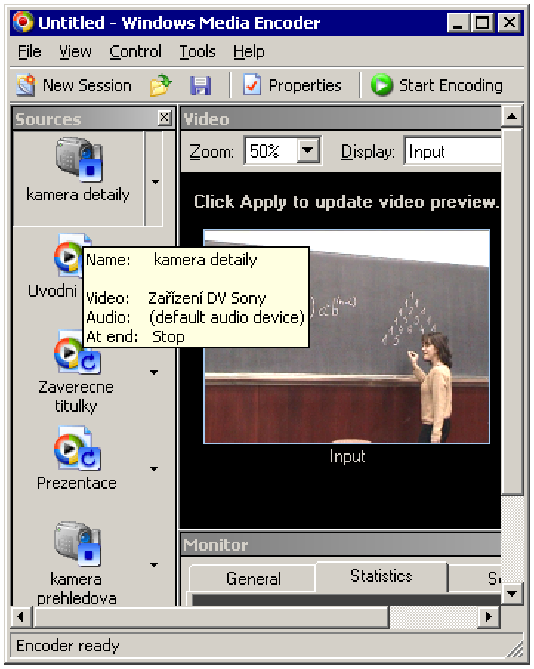
Figure 8.13: Digital editing workplace at VSE

In 2003, based on these experiments and experiences, we have proposed, implemented, and operationally verified a new technology based on digital video. Our results are described in the technical report 16/2003 “Small digital studio”. The solution is based on modern digital cameras connected using the IEEE 1394 (FireWire) interface directly to computers and *Windows Media Encoder 9* by Microsoft. This technology allows for a simple switching between two cameras with a possibility of inserting video sequences from data files containing e.g., introductory and final subtitles, presentations and optionally other audiovisual information.

The studio technology is mainly software-based, meaning that it uses a minimum of additional hardware components and largely depends on the processor performance. With high-end PCs (Pentium 4, Athlon XP) we can reach the standard PAL resolution in real time. The benefits of the new technology can be seen in the improved quality of the recorded and transmitted image and especially in the simplification of the post-production phase. The technological chain can be applied to recording or on-line broadcasting of lectures by important visitors, seminars, conferences and other similar activities.

8.3.5 Cooperation with Czech Radio

The cooperation with Czech Radio has several levels. This year, we supported a project called “Peregrine Falcons in the Heart of the City”, following the tradition of similar projects in the past (Peregrine Falcons in the Heart of the City 2001 and 2002, Millennium Cub, Kristyna Live, etc.). These projects are organized by the Czech Radio and they aim at allowing the general public to look at what



Zoom: 50%
Display: Input

Click Apply to update video preview.

Name: kamera detaily
Video: Zařízení DV Sony
Audio: (default audio device)
At end: Stop

Encoder ready

Figure 8.14: Small digital studio at VSE

is otherwise impossible to observe, and popularise the issue of endangered species protection at the same time. We have constructed a tool chain for audio and video presentation of the peregrine falcon family nesting in Pilsen. Due to problems with the nesting of this particular couple, an alternative locality and content has been selected – Zoo at Chomutov.

Another area of collaboration with Czech Radio is audio streaming. Based on successful tests of a system for permanent on-line broadcasting that we proposed and implemented in 2002, this year we started routine broadcasts of the Czech Radio stations to the Internet in high quality.

The system was initially based on the technology of audio signal transmission in MPEG (MP3) and Ogg compression formats with bit rate 128 kbps, later we added an Ogg stream with a 256 kbps variable bit rate (that means the bit rate is allowed to oscillate around the average value of 256 kbps depending on sound scene complexity). The decision to broadcast in the two formats followed from the fact that while MP3 format is the most widespread in the world and more-or-less automatically expected, the newer Ogg format is considered to be the format of the future. The latter has been developed in a way similar to open source software and is free from the patent and licensing restrictions that hamper free use of the MP3 format. Furthermore, we consider the Ogg format technically superior, able to provide comparable or better quality than MP3 with lower bandwidth requirements. The Ogg format also is supported by the most widespread audio players (e.g. xmms, WinAmp, zinf, qcd, etc.).

Currently we are streaming the following stations: CRo1-Radiožurnál, CRo2-Praha and CRo3-Vltava, all in MP3 at 128 kbps, Ogg at 128 kbps, and Ogg at 256 kbps formats. Since May 2003, we have added the broadcasting of a CVUT student radio – Radio Akropolis. In this case we offered just the capacity of the broadcasting server and the contents are produced by Radio Akropolis itself. They use data streams in the Ogg format with parameters of 44.1 KHz, 16 bits, 2 channels, 128 kbps and variable bit rate.

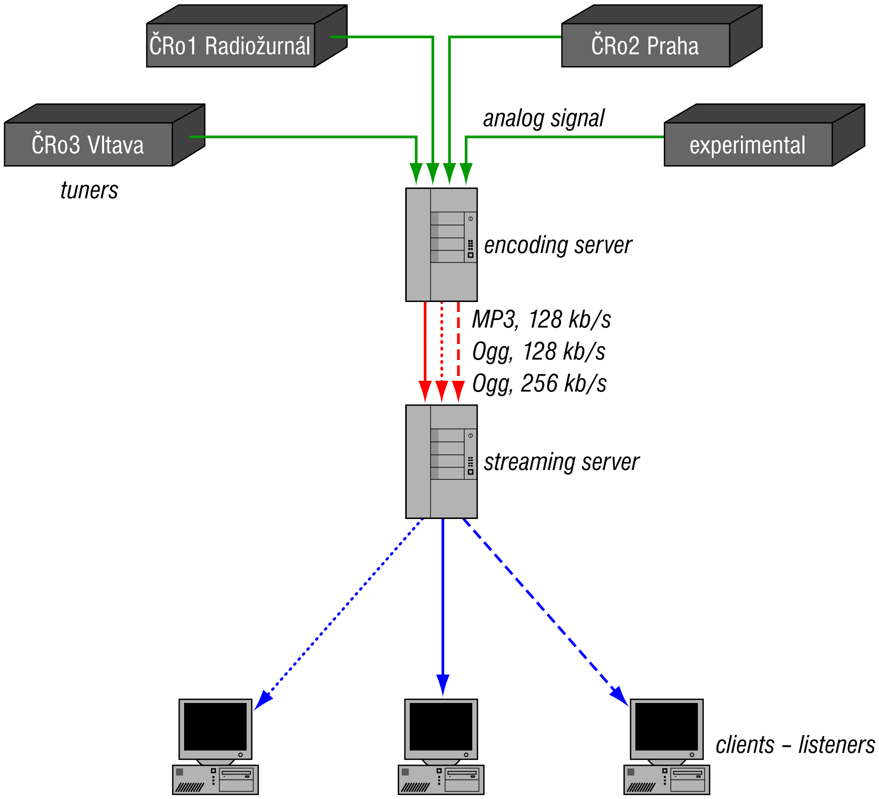


Figure 8.15: Block diagram of the audio broadcasting technological chain

The operation of the whole system is stable. The dual-processor encoding server works in real time under a constant load of 1.35 while encoding 9 concurrent streams—every encoder instance consumes approximately 13 % of the CPU resources of the server. The streaming server works under a variable load exceeding 0.5 only in case of 100 or more simultaneous clients.

We consider the sound quality to be very good. The highest quality is definitely provided by the Ogg stream with 256 kbps bit rate. With its clear sound scene it is close to the CD quality. In the case of a 128 kbps bit rate, the sound scene is slightly narrower and less balanced. These differences become evident especially with high quality equipment – ordinary computer speakers tend to flatten them. Differences between MP3 and Ogg at 128 kbps are rather minor. In our opinion, though, the Ogg format offers a somewhat higher quality, which can be manifested especially in compositions with a large dynamic range. At the bit rate of 64 kbps, the Ogg format is definitely better than MP3. Judging from the listeners' feedback, the quality of all the streams is perceived as high.

The time delay of the received stream is very small (1–3 seconds) and can be tuned by setting the buffer memory size on the audio client side. While setting a bigger buffer size eliminates the jitter and other irregularities in the received data that are frequent on lines with lower capacity, it obviously increases the delay. For a 256 kbps stream we recommend to increase the buffer size, because the default value corresponds to less then 1 second of play time and this is usually not sufficient.

The streaming servers support both IPv4 and IPv6. IPv4 streams can be selected at
http://radio.cesnet.cz:8000/stream_identification
(e.g., *http://radio.cesnet.cz:8000/cro3.ogg*)
and their IPv6 equivalents at
http://amp-ipv6.cesnet.cz:8006/stream_identification
(e.g., *http://amp-ipv6.cesnet.cz:8006/cro3.ogg*).

	CRo1	CRo2	CRo3	Akropolis
data volume per month [GB]	1135	402	434	238
average daily data volume [GB]	37	13	14	7
number of clients per month	25828	5570	6591	9027
number of MP3 clients included	21202	4255	4820	9027
number of Ogg clients included	4626	1315	1771	-
average number of clients per day	860	185	219	301
maximum number of clients per day	2201	302	291	556
avg. number of clients simultaneously	35	7	7	6
max. number of clients simultaneously	67	21	23	20
average connection duration [min]	49	80	73	29

Table 8.1: Statistical summary for November 2003

The following graphs display utilization statistics of the Internet streaming service and document an increasing popularity of this type of service.

As a result of an agreement with Czech Radio, we have access to daily exports of their program schedule in the XML format containing program names and timing data. This information is automatically extracted, processed and added to each stream as a supplementary text data (so-called metadata) describing the program just played.

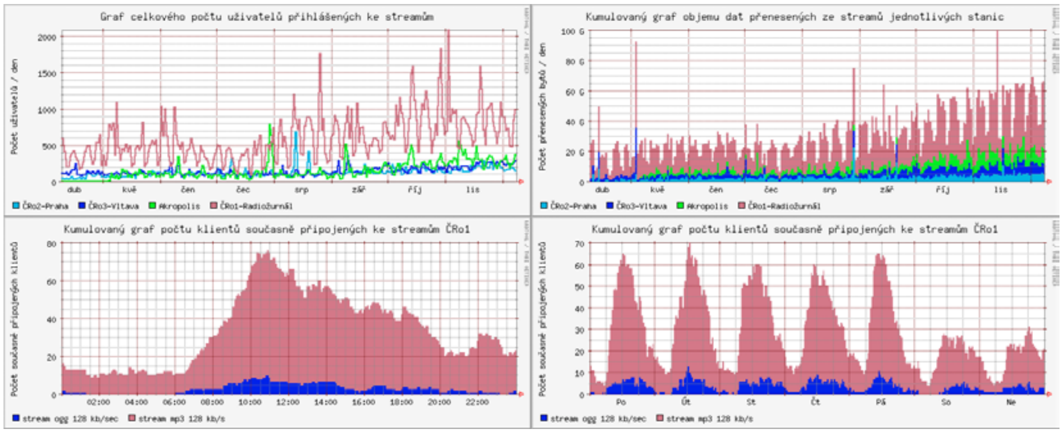


Figure 8.16: Graphic overview of on-line radio utilization

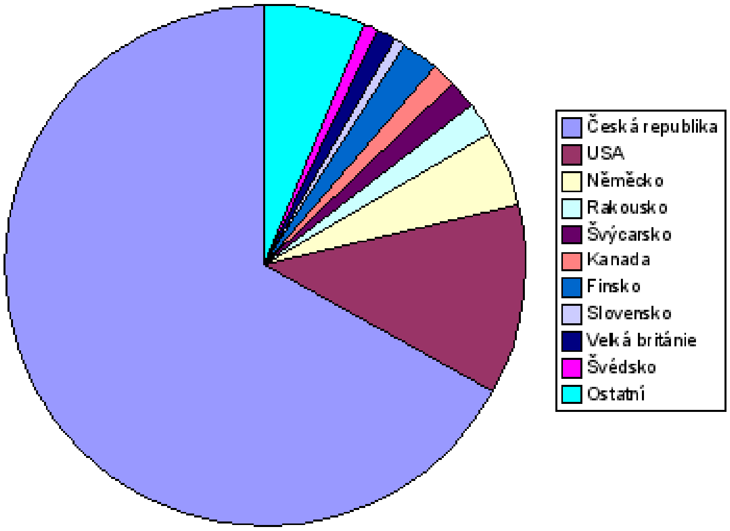


Figure 8.17: Distribution of online radio clients

The issues of selecting technology for such an application are discussed and the final solution described in CESNET technical report number 24/2003.

8.3.6 Piano recital – telepresentation

On November 6th, 2003 we organised a recital of Jakub Litoš, a young pianist and composer, at CESNET premises in Prague. A live stream of this performance was transmitted as a telepresentation over the Internet to audience at the Northern State University (Aberdeen, South Dakota, USA). From the technical point of view, the transmission was implemented as a two-point H.323 videoconference. We used the *Polycom ViewStation FX* and *TANDBERG 880* facilities as the end stations.

Figure 8.18: Telepresentation of the recital by Jakub Litoš, a pianist and composer

The reactions on this event were very positive. Listeners from the Northern State University described the image and especially sound quality as very good. We are thus ready and willing to provide technical support to similar activities again.

8.4 Future Work

Our future plans follow from the CESNET research strategy that also includes support for complex applications utilising the national research network. We want to extend our activities in the following areas:

- completely digital A/V tool chains
- HD capturing, transmission, and displaying
- IPv6 multicast applications
- 10 Gbps end points for A/V applications
- programmable hardware (FPGA, DSP) for A/V
- development of our own A/V platforms
- building new A/V laboratories and AGP nodes
- SIP for A/V and integration with H.323 infrastructure
- methodology for information dissemination, publishing usage scenarios.

9 MetaCentre

9.1 General information

The aim of the MetaCentre project is to provide academic users with a sophisticated computing environment, which fully exploits the possibilities of high-speed computer network and provides access to computing and disk resources of the biggest academic computing centres in the Czech Republic. The main emphasis is on the support of *operational infrastructure*, whose needs and requirements then determine the orientation of necessary research and development activities. Another important objective of the MetaCentre project is a close cooperation with analogical international projects in order to build a sound base of all necessary professional and technical know-how.

The result of the project is a *national Grid*, i.e., a distributed virtual computer that enables both simultaneous utilization of computing resources and the possibility of using individual nodes without knowing exactly the location and to a certain extent even the architecture of individual computers. While CESNET invests into own cluster computing capacities (based on the IA-32 architecture), the project also integrates external computing systems (Alpha, SGI), many terabytes of disk capacities in individual centres and a tape back-up library, purchased several years ago by CESNET.

The main activity of the project is development and operational maintenance of the Czech national grid. In the development area we pursued a closer cooperates with two international projects of the 5th EU Framework Program – *DataGrid* (see below) and *GridLab*, where Czech Republic is represented by the Masaryk University.

The following centres participated in the project in 2003:

- Institute of Computer Science, Masaryk University in Brno
- Institute of Computer Science, Charles University in Prague
- Centre for Information Technology, University of West Bohemia in Pilsen
- Computing Centre, Technical University of Ostrava

Except for the last one, all the centres contribute their computing capacities, particularly large computing systems from SGI and Compaq.

MetaCentre also directly administers or provides professional support to the managers of external clusters, in particular NCBR (National Centre for Biomolecular Research) at Masaryk University in Brno and NTC (New Technologies Research Centre) at University of West Bohemia in Pilsen. The total number of processors administered this way is about 165.

9.2 Operation

The basic structure of MetaCentre has been stable for several years. The computing resources are situated in four localities and three cities: Brno (ÚVT MU), Prague and Pilsen (CIV ZČU). The two workplaces in Prague are located at the premises of ÚVT UK and CESNET. All centres are directly connected to the high-speed backbone of the CESNET2 network with a 1 Gbps link. This capacity can be upgraded according to the needs.

The bulk of computing power is provided by clusters based on processors with the IA-32 architecture. Processor counts and general features of these clusters are as follows:

- Brno: 32×Pentium III 1 GHz, 64×Pentium 4 Xeon 2.4 GHz
- Prague: 64×Pentium III 700 MHz
- Pilsen: 32×Pentium III 1 GHz

In total 192 processors are instantaneously available in the dual-processor configuration (i.e., 96 nodes) with RAM capacity of 1 GB per unit. The disk size increases from 9 GB per node in the oldest (and least powerful) nodes through 18 GB per node up to 36 GB in the latest nodes.

All cluster nodes are equipped with a Fast Ethernet network adapter (100 Mbps). Cluster nodes in Pilsen and in Brno are interconnected with high-speed networks. In Pilsen it is the Myrinet network (1.2 Gbps full duplex), while in Brno each node includes two active Gigabit Ethernet interfaces (integrated in the cascade of ProCurve switches from HP) and, in addition, 16 nodes are equipped with a Myrinet network interface (transmission speed up to 2 Gbps full duplex). This distribution enables solving jobs with enormous demands on inter-node transmission throughput and also testing the influence of different interfaces on program scalability.

However, the computing capacities of the MetaCentre are not limited only to clusters. At the end of the year 2002, CESNET purchased an HP server with two Itanium II processors (1 GHz, IA-64 architecture), 6 GB of internal memory and 100 GB of local disk space. The primary role of this server was to provide a test environment for this type of architecture. We found out that the real maximum performance was around 140 % of the performance of a Pentium 4 Xeon processor with 2.4 GHz frequency. For the *sander* program computations (LES – equilibration), we were able to achieve a performance almost identical to that of a dual-processor Pentium 4 Xeon with MPI/shmem compiled with a PGI compiler. On the other hand, operations with big blocks of memory (routines *mem** in the *libc* library) are not really optimal – the achieved performance was often only half of the performance of the Pentium 4 Xeon processor. Compilers for the IA-64 architecture also have aggressive optimisations turned on by default (extensive reorganisations of floating point operations), which can cause computational errors.

The participating centres provide also their own computing capacities, consisting mostly from SGI computers with MIPS processors (ÚVT UK and ÚVT MU) and Compaq computers with Alpha processors (ZČU). In addition to the computing capacities, disk arrays are also available with a total capacity exceeding 5 TB.

All cluster nodes run the *Debian Linux* operating system, which was upgraded to version 3.0 during 2003. All computing resources of the MetaCentre are accessed through the *PBSpro* batch system, maintained by Masaryk University in Brno. The *PBSpro* system cooperates with native batch systems of big computers (e.g., NQE). Furthermore, we have installed the *globus* system (version 2.2.4) together with a gateway between this system and PBS. Moreover, in the second part of the year we started the testing of *globus* version 3.0.

Data backup is realised by a large-volume tape library at ÚVT MU using the NetWorker system by Legato. Although the backup covers all machines and disk capacities and the volumes of backup data steadily increase, the transmission capacity of the CESNET2 network is still sufficient. The capacity of the tape library is now fully utilised (with backups being archived for a period of at least 3 months).

Many programs and software systems are available to the users of the MetaCentre. Their summary can be found at the portal web site *meta.cesnet.cz*. The project budget is used for updates of the development environment running on clusters, namely compilers (Portland Group and Intel) and monitoring and debugging tools for parallel programs (*TotalView*, *Vampir*, and *VampirTrace*). Among application software a special mention deserves the maintenance of the *Matlab* system. Data access is managed in part by the distributed AFS file system. On clusters we currently use the free *OpenAFS* software.

9.2.1 Extension in 2003

Since the equipment ordered in 2002 was delivered as late as in February 2003, the budget allocated for the latter year was used for enlarging capacity of selected nodes instead of increasing the number of processors. In particular, additional 32 disks with the capacity of 36 GB were purchased for the newest cluster (with Xeon processors) and in the half of the nodes the RAM memory capacity was increased up to 2 GB. This extension of memory and disk capacities enables us to utilize more widely the advanced methods of job planning, especially displacing the running jobs with lower priority when a job with a higher priority arrives – the so-called preemptive planning. However, this method requires that both the old job (with lower priority) and the new one (with higher priority) fit into disk and memory at the same time. Preemptive planning enables a faster execution of short application jobs with high priority. In addition to the extension of computing nodes we upgraded the disk capacity of cluster front-ends and purchased another one (dual CPU Intel Pentium 4 Xeon 3 GHz).

In reaction to an increasing demand for processors with a 64-bit architecture, we purchased an IBM p615 computer with two Power4+ processors (1.2 GHz, 6 GB RAM) and a server with two AMD-64 processors (together with the SuSE Linux operating system, as it is the only Linux system really supporting this kind of processors). The IBM computer will be installed in Brno and the AMD server in Pilsen. The experience with these computers will help us decide about further investments in 2004 as well. Furthermore, in 2003 we purchased additional 16 Myrinet cards and the corresponding switch extension card in Brno so all 32 nodes (64 Pentium 4 Xeon processors) will be interconnected with this high-speed network.

Acquisition of new software was limited, apart from the SuSE Linux operation system mentioned above, to the purchase of a supercomputing license for the *Gaussian'03* program. Computations with the Gaussian program account for almost 50 % of the total MetaCentre computing capacity, and besides the new version contains many extensions and upgrades requested by end users.

9.2.2 Operational statistics

More detailed statistical data about utilization of all MetaCentre capacities during the last year are available in the MetaCentre Annual Report, what follows are just selected data about utilization of the clusters.

While the annual average utilization of clusters is only 41 %, during the last six months the interest in cluster computing has been evidently increasing:

- Utilization in the last 6 months is 51 %
- Utilization in the last two months is 65 %
- Utilization in the last month (November) is 80 %

Most demanded are the powerful systems with Pentium 4 Xeon processors.

In total 24,524 jobs were processed representing almost 39.5 thousand days of computation (between January 1st and December 12th, 2003). Per-job average processor utilisation was 1.79, indicating that one- and two-processor jobs prevail. On the other hand several users were running their jobs simultaneously on 11–12 processors on average, so there definitely is certain interest in using the clusters for medium-parallel jobs as well (see also below).

MetaCentre has approximately 200 users, about 70 of them being highly active. The 7 most active users have consumed 45 % of all processing time for their computations (these statistics involve clusters, not big computers). Accounts are assigned for one year and for their extension we require the users to deliver a report about their utilization of MetaCentre resources during the previous calendar year.

9.3 Information services

The portal at *meta.cesnet.cz* is the basic information gateway of the MetaCentre. The portal provides both general information for casual visitors and specific information for registered users – the latter requires authentication.

The current trend in the domain of information services is to use directory services as "shopwindows" for important data. One of the basic objectives is to make the search and acquisition of information as simple as possible. This information may be distributed in different parts of the computing infrastructure, e.g., in Perun or the batch system). Therefore, in 2003 we have been work on an update of MetaCentre directory services, both in terms of technology and contents – in order to make it a viewport to the Perun system.

The update of information services involved the following changes:

Schema modification: Update and extension of the LDAP schema for the presentation of user data. The schema used is based on a standard schema extended by several specific attributes. For compatibility reasons, our strategy has been to reuse standard schemas as much as possible. In particular, we use the `eduPerson` schema and recommendations from the Internet2 project (concerning personal data) and `rfc2307` (concerning account and group data). Description of the new schema is included in the technical report about the Perun system [Kře04].

Promotion of the Czech language: The support for Czech diacritics in the clients of directory services is still far from satisfactory. Our solution tries to find a compromise between preserving compatibility and providing information in proper Czech, i.e., with all accents.

Mechanism of data distribution from the MetaDatabase: Independently of the data structure, this update allows to propagate database modifications to LDAP in real time. Some clients and applications can greatly benefit from this improvement.

Infrastructure of directory servers: The new mechanism of data distribution from the MetaDatabase will enable the replication of LDAP servers via fully standard mechanisms. The OpenLDAP supports optional Kerberos authentication during the replication.

9.3.1 Perun

The principal new feature of the Perun system for user account management is the implementation of the concept of homogeneous clusters, i.e., clusters where user accounts are identical on all nodes. The original system was proposed when these clusters were not widespread yet. Even though the system was able to administer them, the increase in the number of nodes beyond 50 started to cause considerable performance problems. The current solution which has eliminated repeated time-consuming generation of identical data files can accommodate clusters with thousands of nodes.

Another noteworthy achievement related to the administration of big clusters is the newly developed part of the administrative portal. It enables the administrator to observe the state of data propagation on individual nodes. Another extension of the administrative portal is a systematic interface for browsing and updating data about users, administered tools, accounts, etc.

In relation to our cooperation in international projects, the application data scheme has been extended with PKI infrastructure support. The system is able to keep track of the bindings of users to their X509 certificates and subsequently map them to particular accounts.

Data export from the Perun system to the MetaCentre information service (LDAP) was updated so that it is now compatible with standard schemas `inetOrgPerson`, `eduPerson`, and `rfc2307`. Recent features implemented in the OpenLDAP software enabled us to introduce an incremental propagation of modifications into the LDAP tree during full server operation. Modifications thus show up right after they take place, and not only after the regular overnight shutdown of the LDAP server as before.

Furthermore, we tested the possibility of using the SQL back-end of the LDAP server (i.e., translation of LDAP queries into on-line database queries), but it was found unsatisfactory (results are summarised in the bachelor thesis by Miloš Malík at FI MU).

During the second half of 2003 we migrated the central server of the Perun system from the SGI station, which did not meet the performance requirements anymore and started to suffer from hardware errors, to a new IA32-based server running Linux. At the same time we have replaced the Oracle 8.0 database server with the up-to-date 9.2 version. As a by-product of this update, full Kerberos authentication became available.

The current architecture of the entire Perun system architecture is described in an extensive technical report. At the same time we have revised and completely documented all the used database structures.

Needless to mention, throughout the year we continued the routine tasks – system and data maintenance, handling of errors and user support.

9.4 Applications

In 2003 we also aimed at supporting the development and deployment of really large parallel applications utilising synchronously a large number of nodes in individual clusters.

9.4.1 Browsing the state space

In 2003 we finished the development of the application that models conformational behaviour of molecules using the force feedback mechanism (installed in the HCI laboratory at FI MU). The basic part of the application is the computation of state space for the haptic (i.e., force feedback) system. Because of a high computational complexity, it is implemented as a distributed application over the MPI communication model. The computation is distributed to individual nodes using a method called "Transposition Table Driven Scheduling" distribution of computation on particular nodes, thus representing a distributed application with many small asynchronous messages. Such applications are generally appropriate for the cluster environment.

In addition to tuning the application itself [Kře03], we performed a series of experiments focusing on parallel computation properties. Results can be summarised as follows:

- On one homogeneous PC cluster the application scales in accord with the theoretical limit (given by a slight non-uniformity in the distribution of the computation) up to the available 64 CPUs.
- The results for a distributed computation run simultaneously on clusters in Brno and Pilsen do not differ from those obtained on only one of these clusters. It proves the original hypothesis that this type of application is not dependent on the latency of network interconnection.
- Experimental computations on more than 100 CPUs have proved that for a non-trivial input a linear scaling can be achieved even for this number of processors. However, exact verification of this statement will require to find a better distribution of computations taking into account the different performance of non-homogeneous clusters.

Furthermore, we have identified anomalies in MPI properties during the mentioned experiments:

- When dual-processor cluster nodes are fully utilised (i.e., two computation processes allocated on such nodes), a considerable part of their performance is consumed by MPI overhead, resulting in a noticeable degradation of total performance.

- After sending from several hundred to thousand messages we observe MPI congestion causing long execution times of MPI functions calls that are non-blocking by their defined semantics. It results in an incomplete utilization of the CPU, degrading further the total performance. This phenomenon has been observed only with the *ch_p4* communicator (i.e., the TCP connection), but not during a direct communication over Myrinet.

All these results were presented on the EuroPVM/MPI conference and are summarised in the article [KPM03].

9.4.2 Fluid dynamics

In 2003 we also tested the scalability of FLUENT (parallel computation software for numeric flow simulation) on PC clusters of the MetaCentre. The test objectives were to complete data from the end 2002 and provide information about the properties of this software using new hardware and/or a new version of the software.

To be able to compare the test results with those from the end of 2002, we used the same series of jobs representing a portfolio of typical jobs with different computing demands. We have also used several computing models and methods. For testing purposes on PC clusters only a few of computation iterations have been carried out. Apart from the duration of one iteration, which is the basic indicator of the computation procedure, we also simultaneously observed further parameters, such as for the start time of FLUENT on chosen parts of the cluster. For these tests we used the clusters in Brno (*skirit*) and in Pilsen (*nympha* and *minos*).

The benchmarks are run on a varying number of cluster processors using different network communicators and network devices. Either one or both processors of the cluster nodes are utilised, depending on the total number of cluster processors used – one for 4–15 processors and both for a higher number.

For 32 or more processors we use 4 parts of the available clusters (*nympha*, *minos*, *skirit* and *skurut* in Prague) and the number of allocated computing nodes is distributed in a uniform way.

Results for FLUENT 6.0.12

The dependence of computation times on the number of processors and type of communicator used for one of the more demanding jobs is shown in Figure 9.1 and Figure 9.2. An almost linear acceleration is reached for the Myrinet network interface, other interfaces (using 100 Mbps) give solid results as well. Contrary to our expectations the use of Network MPI (NetMPI), which is intended exactly for clusters, does not increase performance in any way, not even if we used Gigabit Ethernet on *minos* cluster.

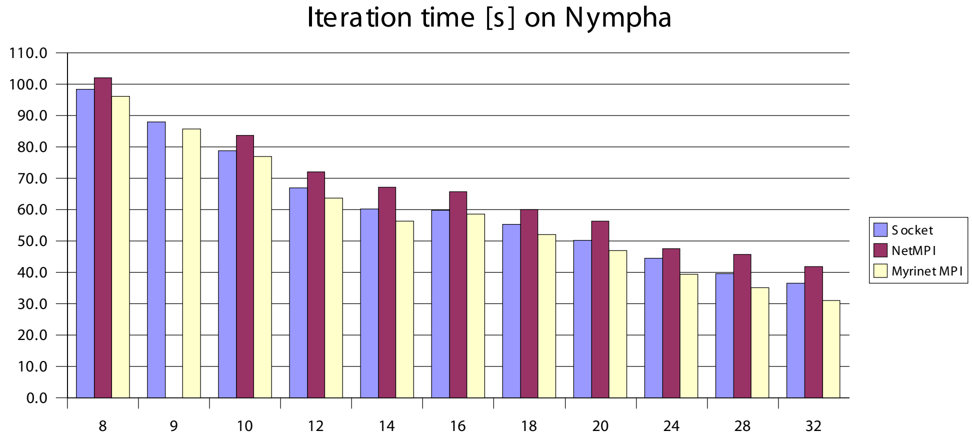


Figure 9.1: Iteration time in relation to the number of processors on Nympha

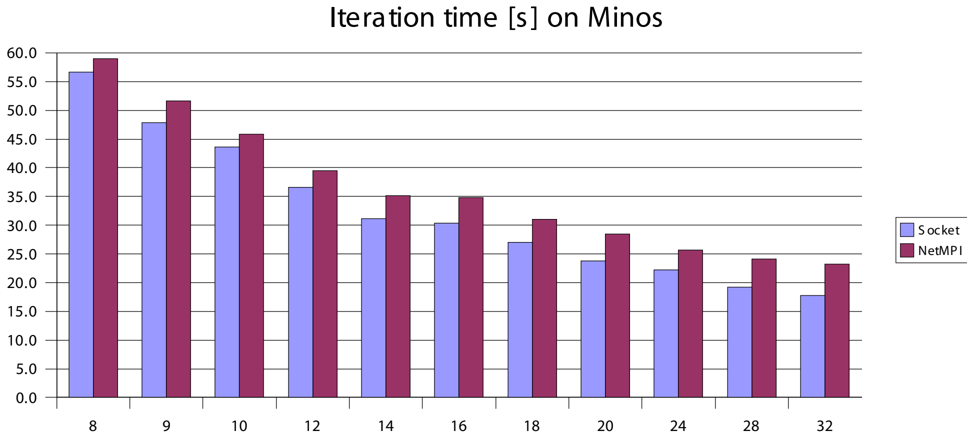


Figure 9.2: Iteration time in relation to the number of processors on Minos

The Figure 9.3 shows that for sockets a slight acceleration of the computation still takes place for 56 processors, then the performance starts to decrease. For Network MPI the situation is even worse.

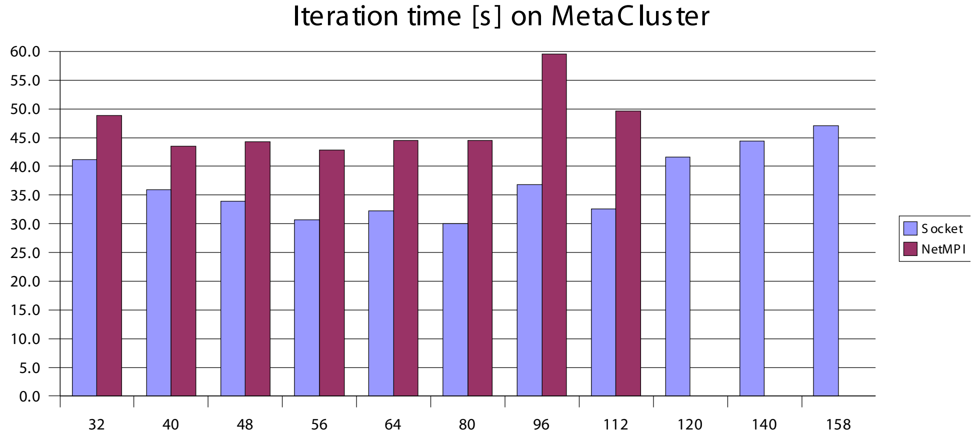


Figure 9.3: Iteration time in relation to the number of processors on MetaCentre

Results for FLUENT 6.1.22

In the second half of 2003 we installed a newer version of *FLUENT* on AFS in Pilsen. Ordinary users can access it by means of an appropriate module. Being limited by the number of available licenses, we performed additional mid-sized tests on new machines which form a part of the *skirit* cluster. We intended to compare *FLUENT* properties on Myrinet (2 Gbps) and on Gigabit Ethernet. Jobs have been started on cluster via the *PBSPro* batch system in the ordinary way without applying special privileges.

We found that the new *FLUENT* version does not pose so strict requirements on installed drivers for Myrinet. The preparation of configuration files for a *FLUENT* run on Myrinet has also been simplified. On the other hand, we had to carry out some modifications in the *FLUENT* installation in order to meet the specific requirements of the MetaCentre PC clusters, not envisioned by *FLUENT* developers.

In Figure 9.4 we can see iteration times for a job of a medium complexity for various communicators. Regrettably, this *FLUENT* version does not provide usual detailed reports about the parallel computation procedure when Myrinet is used. The reported values can thus be measured only indirectly and with potentially large errors.

Users can be pleased that Network MPI functions well, from the point of view of iteration process the communicators are equivalent (for all jobs), even though Network MPI and Myrinet have a slightly better performance. The next Figure 9.5 shows the acceleration of the computation with respect to the least number of processors used for the given benchmark (job). We can see that for jobs running on machines with the the second processor being free (up to 14 CPUs), the scalability is good, at the level of SMP supercomputers.

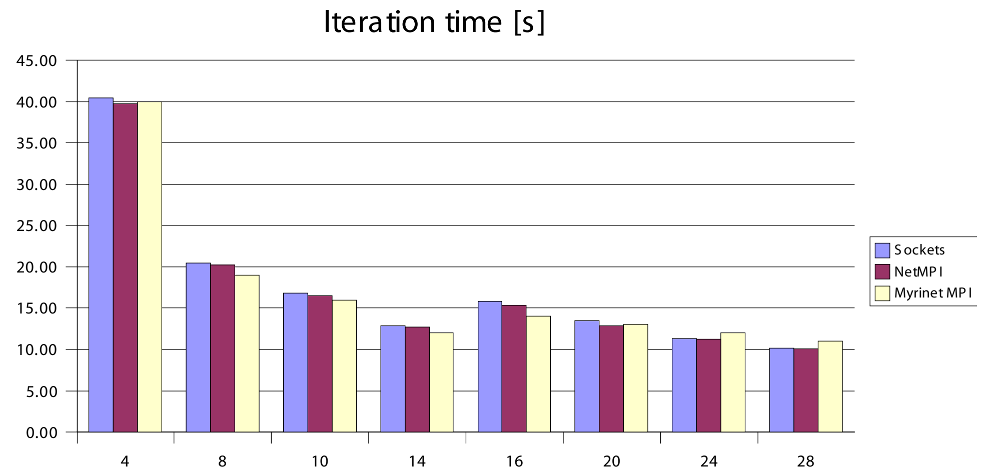


Figure 9.4: Iteration time in relation to the number of processors and to the type of used network interconnection on Nympha

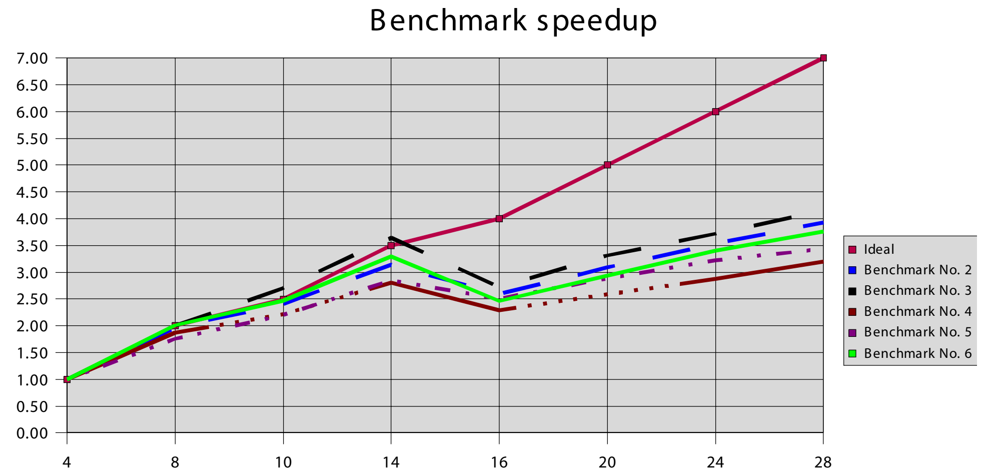


Figure 9.5: Speedup of the testing job

On the next Figure 9.6 we can see the startup time of all processes on all participating computing nodes. Here the Network MPI communicator shows an improvement in contrast to the previous version. The figure also gives the startup time of a job with medium demands and its distribution over individual computing nodes. Once again the use of Sockets is not appropriate because of a high number of processors. For comparison, Figure 9.8 shows the data volume that is transferred per iteration for particular jobs between computing nodes during the computation.

The *FLUENT* version 6.1.22 has reached a status of a wider usability and higher stability even for larger-scale PC clusters of the MetaCentre type.

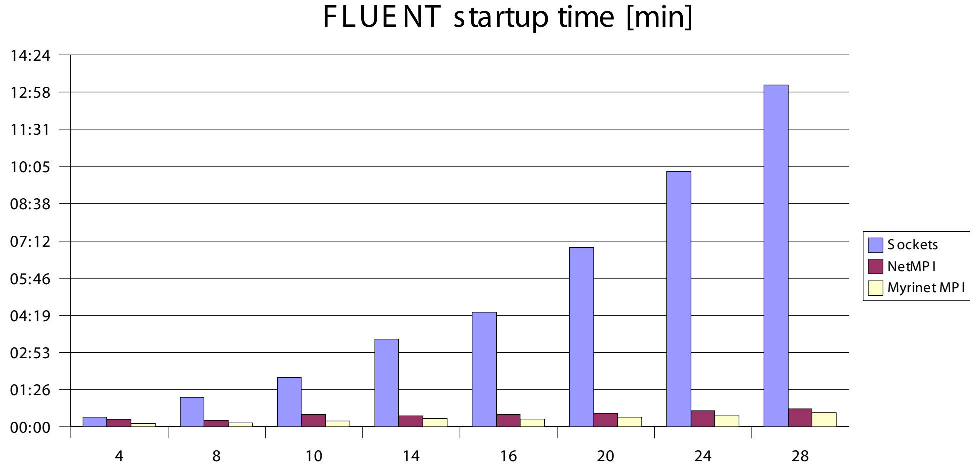


Figure 9.6: Startup time of data distribution in relation to the number of processors and to the type of used network interconnection on Nympha

All in all, we are able to answer the question whether an investment in a high-speed system of Myrinet type pays off for *FLUENT* computations or not: Gigabit Ethernet together with the Network MPI communicator provides the same performance, and for lower number of processors (up to 10) even the Sockets use is acceptable.

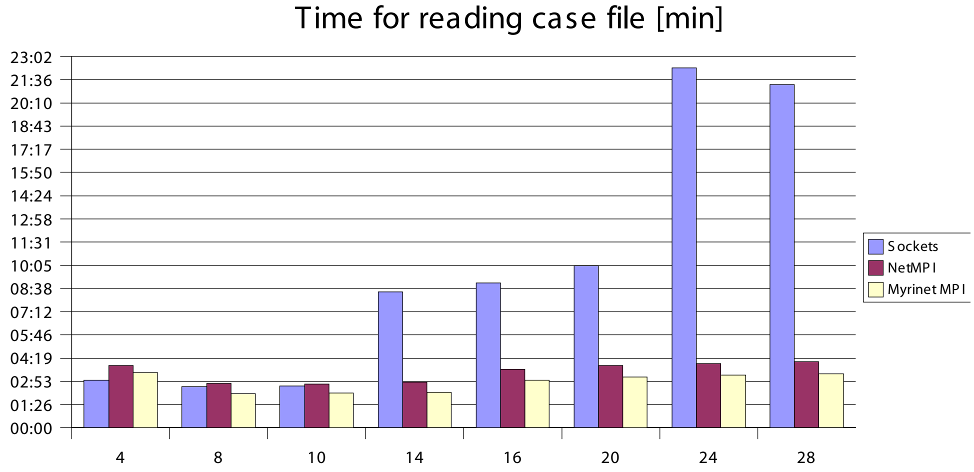


Figure 9.7: Time for reading case file in relation to the number of processors and to the type of used network interconnection on Nympha

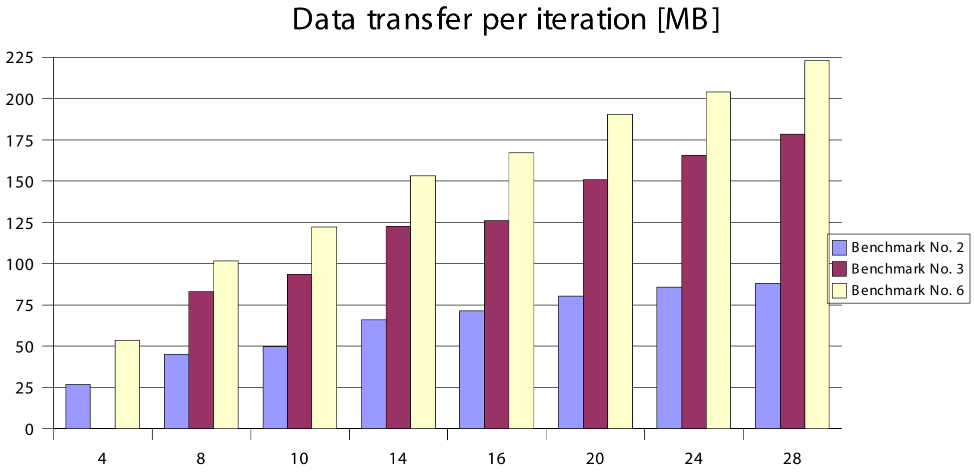


Figure 9.8: Data transfer per iteration

10 Voice Services in CESNET2 Network

10.1 Introduction

The project started in the second half of 1999. Its goal was testing usability of technologies for the voice and data network convergence. Since the beginning, the project directed its attention at the area of Voice over IP (VoIP). In retrospective, this direction has proved correct: alternative ways such as Voice over ATM and Voice over Frame Relay were less successful. After first two years, the value of this project increased significantly. Results of the experiments were successfully tested in operation and the project was moved among the strategic projects.

In 2001, the VoIP network was interconnected with the public switched telephone network (PSTN) which caused further expansion of the IP telephony network: by the end of 2003, most CESNET members participated in the project. Interconnection with the PSTN network allows significant price reduction of telephone charges thanks to low interconnection fees at a quality comparable with classic calls using the PSTN.

This project is building an advanced experimental platform usable also by other research projects; it also allows the CESNET2 network to support this progressive voice communication method. A general project goal is providing the Association members with support for connecting and using the CESNET IP telephony infrastructure. The research goal consists of IP telephony component and application certification and development.

Project outputs are tested in operations, regularly published and presented at conferences. In the area of international cooperation, VoIP interconnection with foreign institutions is being tested; we also play an active role in a TERENA project – the IP Telephony Cookbook publication.

10.2 Changes in the VoIP network infrastructure

Other Association members joined this project during 2003. New Voice Gateways (VoGW) were installed in the following network locations:

- Institute of Economy in Jindřichův Hradec – PBX connected to VoGW through 2×ISDN/BRI
- University of Ostrava – PBX through 4×ISDN/BRI

- University of West Bohemia in Pilsen – PBX through 1×ISDN/PRI
- Charles University Prague, Faculty of Paedagogy – PBX through 1×ISDN/PRI
- Academy of Sciences Prague – PBX through 1×ISDN/PRI
- Purkyně University Ústí nad Labem – PBX through 1×ISDN/PRI
- CESNET Prague, 2×ISDN/PRI – peering with the GTS Czech
- University of Veterinary and Pharmaceutical Sciences Brno – PBX through 1×ISDN/PRI (to be activated shortly)
- Tomáš Baťa University Zlín – PBX through 1×ISDN/PRI (to be activated in January 2004)

The following connectivity changes took place in 2003:

- Silesian University Opava – PBX connection changed from E&M to ISDN/BRI
- Institute of Physics Prague – FXS connection cancelled (connected to the Czech Academy of Sciences through ISDN)
- Masaryk University Brno – dialing prefix changed to 549 49x xxx.

In 2003, PBXs of all institutions cooperating in this project use only the ISDN technology. In the early stages of this project, other interconnection modules (E&M, FXS, FXO) were also used. The Silesian University in Opava and the Physical Institute in Prague had taken an advantage of an alternative connection. In 2001 we set up requirements for connecting PBXs into the infrastructure whereby only the digital connection using ISDN was preferred. This allows sending the caller identification which is important for charging calls, collecting statistics and tracing malicious calls. Currently, by the end of 2003, all PBXs of Association members satisfy the rules given in the project cooperation contract.

A significant change achieved this year is an upgrade of the PSTN conectivity through the GTS Czech operator. This connection uses 2×ISDN/PRI connected through a new CISCO 3745 voice gateway. Calls can be directed to PSTN via the Aliatel or GTS operators. This allows better flexibility and increases the service accessibility. The interconnection with Aliatel via NIX.CZ uses the IP protocol.

10.3 Current state of the IP telephony network

Over ten thousand participants from those Association members whose PBXs are connected via voice gateways (VoGW) use the IP telephony service in 2003. The VoGWs and PBXs are interconnected using ISDN; this allows keeping the call detailed records (CDR) in an SQL database at a RADIUS server.

Since its establishment, the CESNET IP telephony network has been oriented towards using the H.323 protocol (minimum version 2); internal network components (Gateways and Gatekeepers) are based on the Cisco platform. Its advantage has been proved during network upgrades and management. Connectivity of the CESNET Association non-members is provided by an external Gatekeeper *gk-ext.cesnet.cz* located in Prague and running on a Linux platform. From the hierarchy point of view, this Gatekeeper is linked above the internal Gatekeepers in Ostrava and Prague.

The internal Gatekeepers (GK) back up each other; all CESNET network VoGWs can connect to both GKs using different priorities dependent on their geographical positions. Ordinarily, the higher priority connection is active. Requests to connect outside the CESNET2 VoIP network are directed to the operator providing public telephone services; these calls are charged according to the price list which is a part of the contract between the CESNET Association and the connected organization. The network logical diagram is shown on Figure 10.1.

Calls are charged using 1-second increments; no minimum call duration is charged. Calls within the CESNET2 network between the VoIP project members are free of charge. The call fees are calculated using the *TAS-IP* application; the *IPTA* program developed by CESNET employees is used experimentally. Call details are recorded in databases of both applications.

Institutions listed in table 10.1 are registered on internal GKs in Prague and Ostrava. They can access the PSTN through the Aliatel GK remote zone or through the GTS VoGW.

Institutions listed in the following Table 10.2 are registered on the CESNET external GK; they cannot access the PSTN. However, communication between institutions registered on the internal GKs and the external GK is not limited. During 2003, SANET institutions were included in the CESNET H.323 external peering. As a result, one can call, e.g., the Slovak Technical University and Žilina University free of charge:

Figure 10.2 shows the IP telephony network by the end of 2003

Fees for calling the PSTN are advantageous because all external voice traffic is terminated in a single point in Prague and the telecom operator's interconnection costs are low. Some example costs in 2003: Prague 0.82 CZK/min in peak hours, 0.51 CZK/min off peak, Germany 1.97 CZK/min, USA 2.01 CZK/min. Instructions on IP telephony use are available at *http://www.cesnet.cz/iptelefonie/voip-cesnet.html*

Institution	*Prefix*	*Phone number*
Czech Technical University, Prague Institute of Chemical Technology, CESNET	94	22435xxxx
Technical University Ostrava	#0	59699xxxx, 59732xxxx
Masaryk University Brno	7	54949xxxx
University of South Bohemia Č. Budějovice	858	38777xxxx, 38903xxxx
Charles University Prague, Rectorate	98	224491xxx
Charles University Prague, Faculty of Paedagogy	94	221900xxx
University Pardubice	22	466036xxx, 466037xxx, 466038xxx
University Pardubice in Česká Třebová	83	465533006, 465534008
Technical University Liberec	40, 47	48535xxxx
Faculty of Pharmacy Hradec Králové	55	495067xxx
University Hradec Králové	56	495061xxx
Institute of Economy Prague	#0	224095xxx, 224094xxx
Institute of Economy in Jindřichův Hradec	#0	384417xxx
Silesian University Opava	*0	553684xxx
Silesian Univ. Karviná, School of Business Admin.	9	596398xxx
Academy of Sciences Prague	0**	26605xxxx
Technical University Brno	0*8	54114xxxx
Palacky University Olomouc	0*8	58563xxxx, 58732xxxx, 58744xxxx
Purkyně University Ústí nad Labem	–	47528xxxx
University of West Bohemia Pilsen	#0	37763xxxx
University of Ostrava	79	597460xxx, 596160xxx

Table 10.1: Member institutions in the IP telephony project

Institution	*Phone number*
CERN (*www.cern.ch*)	00412276xxxxx
FERMILAB (*www.fnal.gov*)	001630840xxxx
SLAC (*www.slac.stanford.edu*)	001650926xxxx
STU Bratislava, Rectorate	00421257294xxx
STU Bratislava, Fac. of Mechanical Engineering	00421257296xxx
STU Bratislava, Fac. of Civil Engineering	00421259274xxx
STU Bratislava, Fac. of Material Science and Technology	00421335511xxx
Žilina University	0042141513xxxx

Table 10.2: Foreign institutions registered on *gk-ext.cesnet.cz*

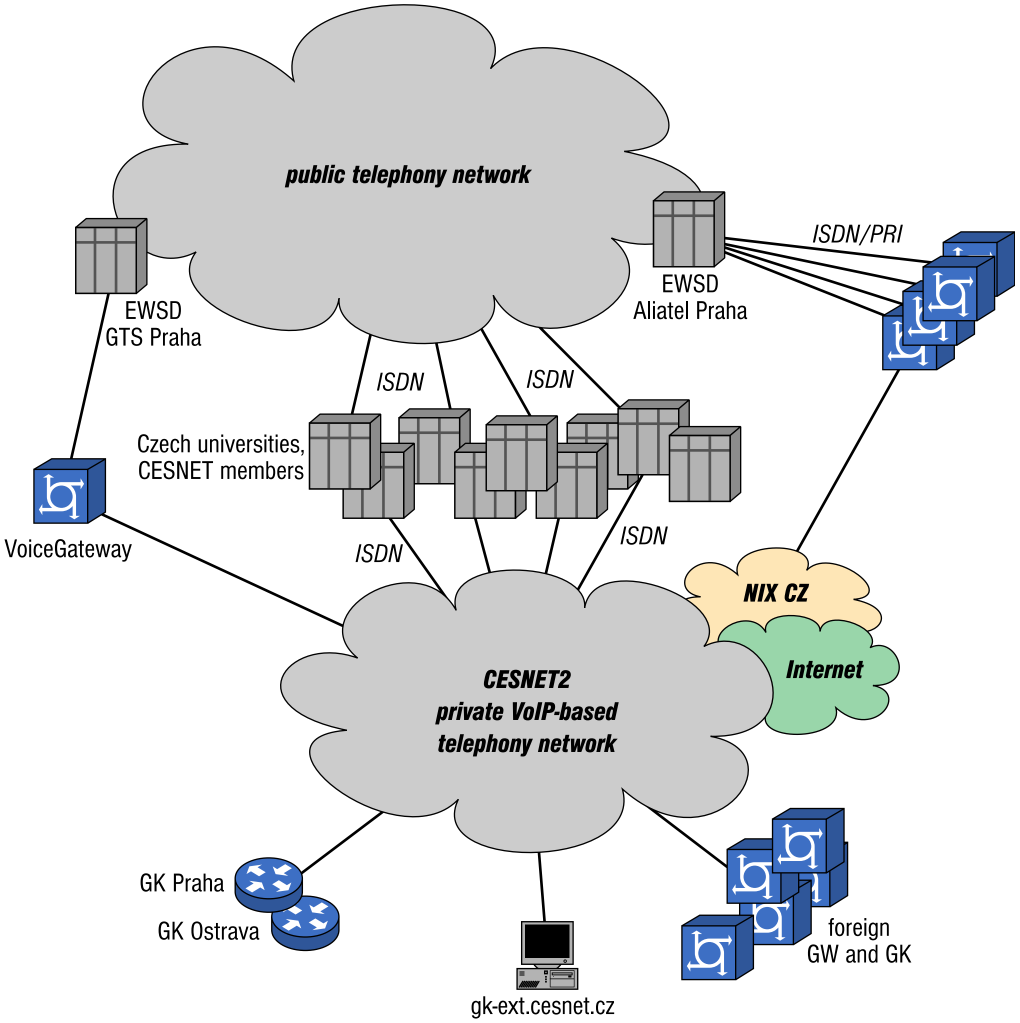


Figure 10.1: Logical diagram of the IP telephony network

10.4 Rules for cooperation in the IP telephony project

Organisations connected to the CESNET NREN and conforming to the technical requirements may cooperate in the IP Telephony project.

10.4.1 Technical requirements

- The PBX must be connected using a digital interface (ISDN BRI, ISDN PRI) only.
- PBX must provide the caller identification.

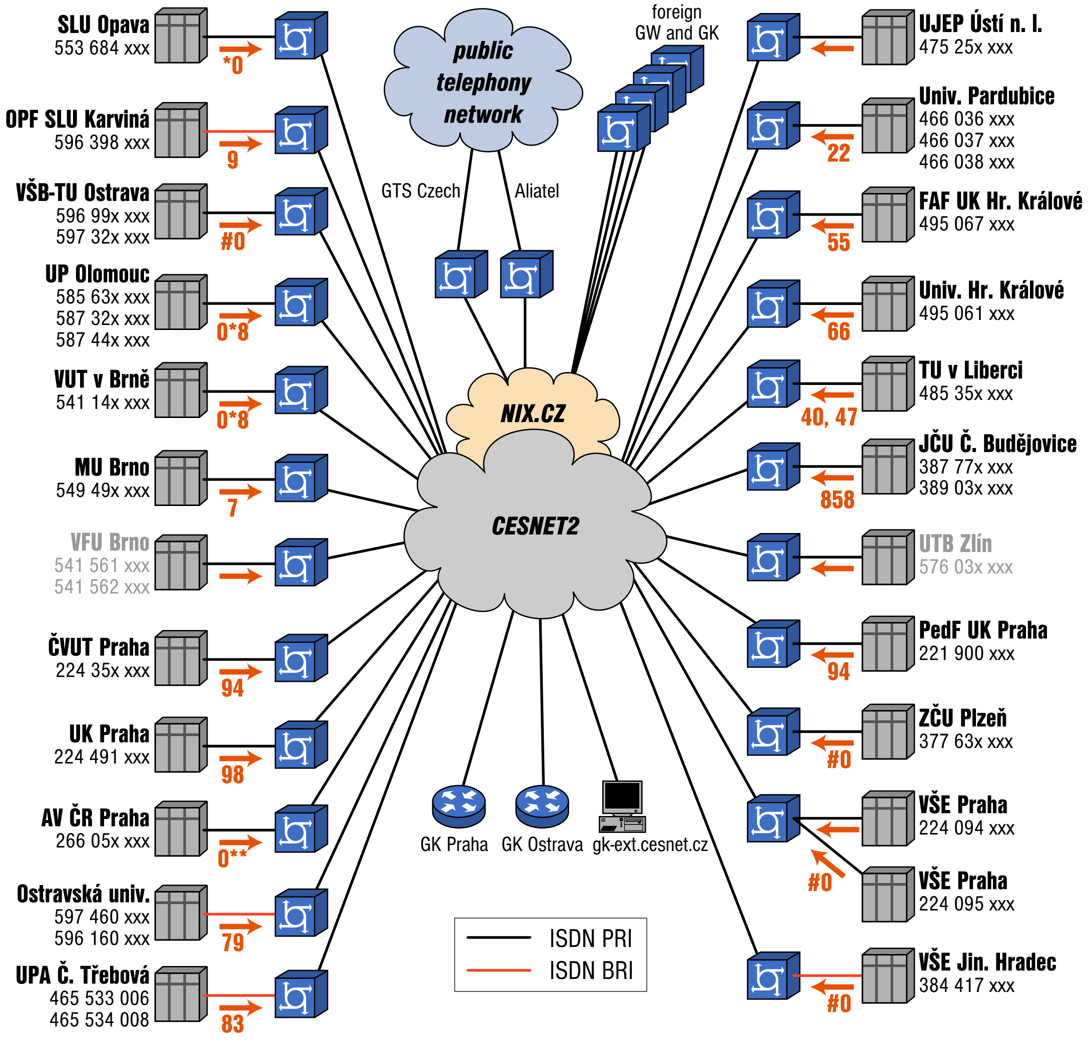


Figure 10.2: Internal diagram of the CESNET IP telephony network

- The PBX operator is responsible for adjusting the accounting within the PBX and charging application, as well as for allowing branches access the IP telephony.

- The end point of the VoIP network is a port of voice gateway for connecting to PBX. The PBX operator will connect the PBX to the VoIP network interface.

- Before the interconnection, the type and interface setting (type of interface ISDN/BRI or ISDN/PRI, setting Network-side or User-side, with/without CRC4 for PRI) must be specified.

- The voice gateway must be accessible remotely by the IP telephony supporting staff.

- The voice gateway must deliver the information on calls using the RADIUS protocol.

- Voice gateway must be compatible with the current VoIP technical solution – Voice Gateway must be Cisco-based (C17xx, C26xx, C36xx, MC3810, AS5xxx platforms). Voice gateway must be fully compatible with H.323 version 2 or higher on the Cisco platform. The reasons are compatibility of other applications (e.g. IPTA – IP Telephony Accounting), common format of CDR records and their sending to the RADIUS server. Other advantages of our solution are access authorisation through TACACS, helpdesk access, supervision and technical support provided by CESNET.

10.5 Applications for charging calls

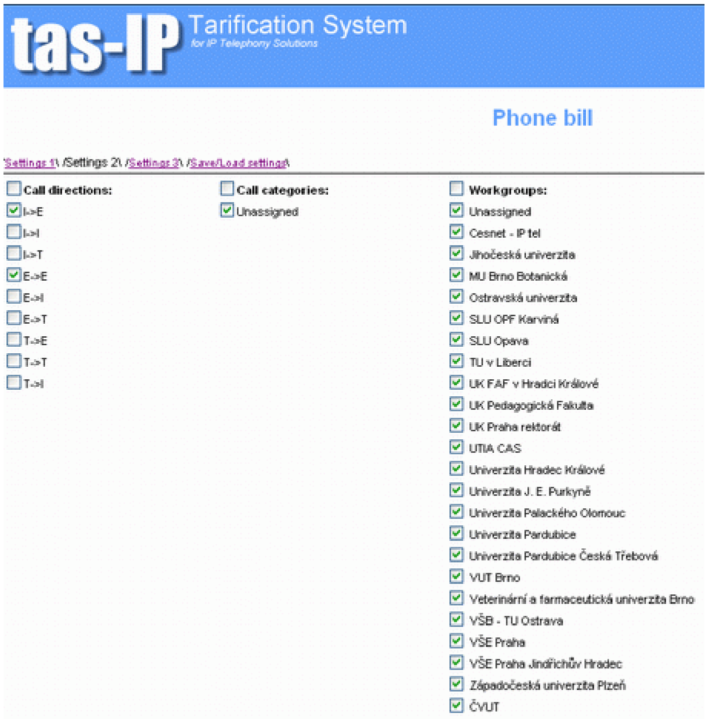
Call records are sent to RADIUS server to be used by the *IPTA* and *TAS-IP* applications. The first planned modification undertaken concerns the *IPTA* application: support for the SIP protocol calls (which requires processing postfix identificators using e-mail format, in addition to prefix identificators formatted as telephone numbers). Some 50 % of this task have been completed.

The second accounting application *TAS-IP* which was bought by the end of 2003 is also running routinely.

The Figure 10.3 shows a mask used for selecting an institution and an appropriate type of call. The application allows generating standard summaries as well as detailed printouts for selected time periods. Figure 10.4 shows monthly summary printouts of several selected institutions. The topmost part includes a monthly summary of calls within the CESNET2 network (these calls are free of charge). The bottom part shows a detailed printout of all calls from voice gateways including the total connection time.

Volumes of VoIP traffic within the CESNET network are depicted on the Figure 10.5 which shows the development during the last two years. New university PBXs are being connected and traffic will continue to grow. The figure shows a substantial growth during 2003. Comparison of latest data, November 2003, with those dated November 2002, shows a sixfold traffic growth.

During our project development, two protocol directions arose which are being solved in the project research section. Specifically, these are the H.323 protocol (based on the ITU-T standard for multimedia communication using data packet networks), and the IETF SIP/SDP protocol. Both standards are important in the VoIP field and the project team must pay appropriate attention to both of them.



Phone bill

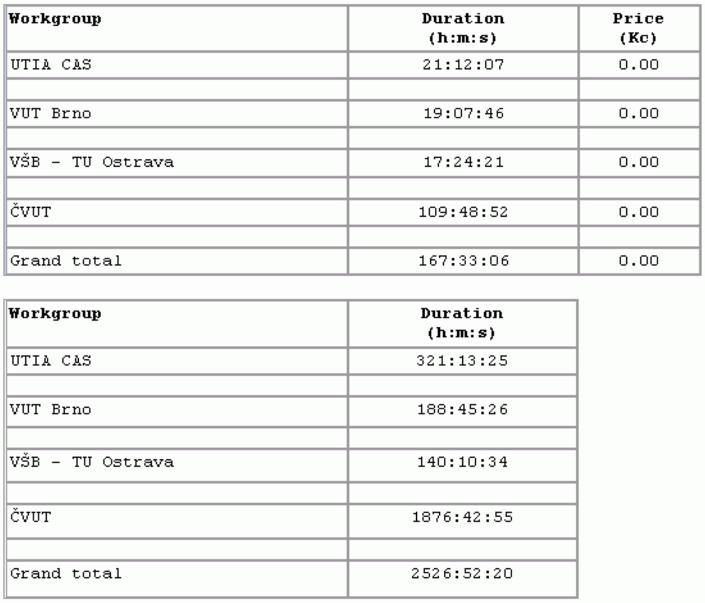
Settings 1\ /Settings 2\ /Settings 3\ /Save/Load settings\

□ **Call directions:**
☑ I->E
□ I->I
□ I->T
☑ E->E
□ E->I
□ E->T
□ T->E
□ T->T
□ T->I

□ **Call categories:**
☑ Unassigned

□ **Workgroups:**
☑ Unassigned
☑ Cesnet - IP tel
☑ Jihočeská univerzita
☑ MU Brno Botanická
☑ Ostravská univerzita
☑ SLU OPF Karviná
☑ SLU Opava
☑ TU v Liberci
☑ UK FAF v Hradci Králové
☑ UK Pedagogická Fakulta
☑ UK Praha rektorát
☑ UTIA CAS
☑ Univerzita Hradec Králové
☑ Univerzita J. E. Purkyně
☑ Univerzita Palackého Olomouc
☑ Univerzita Pardubice
☑ Univerzita Pardubice Česká Třebová
☑ VUT Brno
☑ Veterinární a farmaceutická univerzita Brno
☑ VŠB - TU Ostrava
☑ VŠE Praha
☑ VŠE Praha Jindřichův Hradec
☑ Západočeská univerzita Plzeň
☑ ČVUT

Figure 10.3: Selection mask of institution and call type in the TAS-IP application



Workgroup	Duration (h:m:s)	Price (Kc)
UTIA CAS	21:12:07	0.00
VUT Brno	19:07:46	0.00
VŠB - TU Ostrava	17:24:21	0.00
ČVUT	109:48:52	0.00
Grand total	167:33:06	0.00

Workgroup	Duration (h:m:s)
UTIA CAS	321:13:25
VUT Brno	188:45:26
VŠB - TU Ostrava	140:10:34
ČVUT	1876:42:55
Grand total	2526:52:20

Figure 10.4: Example of monthly summaries for selected institutions

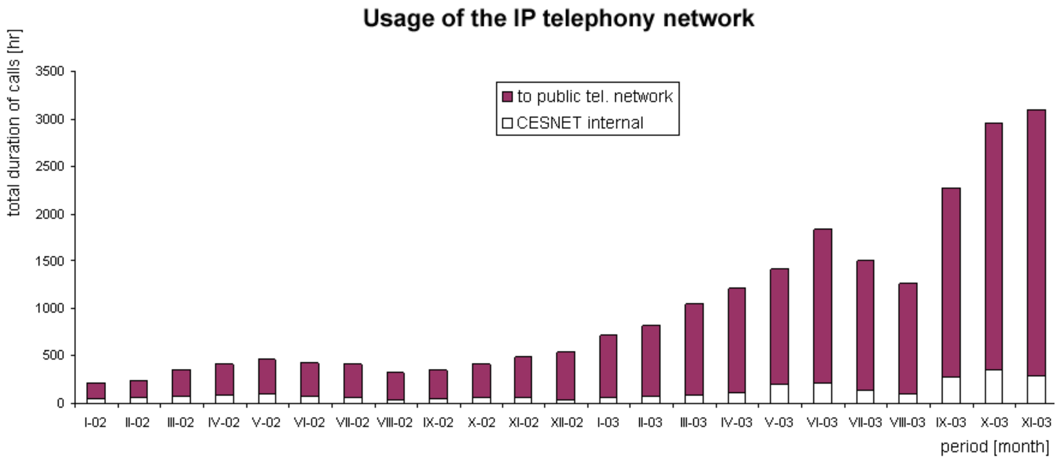


Figure 10.5: Mothly traffic volumes since January 2002

10.6 The H.323 area tasks

The H.323 signalisation network has been deployed successfully and it operates routinely now. Connecting the voice gateways is being offered to the CESNET members as a part of project cooperation. A part of the project included further tests of IP phones; this is still an experimental field.

10.6.1 Testing the H.450 services

In the area of H.323, tests of H.450 services have been completed: we attempted to test as many H.323 services as possible from all twelve defined in the current ITU-T H.450 recommendation. Tests were conducted using the *Siemens OptiPoint400 Standard* and *Welltech LanPhone 101* phones. We found out that the *Lan Phone 101* supports services H.450.1 to H.450.4 while the *OptiPoint 400 Standard* supports services from H.450.1 to H.450.4 plus H.450.7. A list of tested H.450 services follows:

- H.450.1 Generic protocol for H.323 supplementary services
- H.450.2 Call transfer supplementary service for H.323
- H.450.3 Call diversion supplementary service for H.323
- H.450.4 Call hold supplementary service for H.323
- H.450.7 Message waiting indication supplementary service for H.323

10.6.2 H.323 IP Phone and NAT

Another task included testing H.323 telephones on private addresses behind a Network Address Translator. The *OptiPoint400 Standard* and *Welltech LanPhone 101* IP phones were used again.

Figure 10.6: The OptiPoint400 Standard IP phone

Figure 10.7: The LAN Phone101 IP phone

The *Welltech LanPhone 101* gave better results: it allows configuring a shared public IP address in the “IP sharing” item. The IP telephone has a private address from the range of the network behind the NAT, of course, but after the shared public address is configured, the phone inserts this public address into the signalling messages. As a result, no H.323 support is required from the NAT.

Configuration command

```
    ifaddr -ipsharing 1 195.113.113.151
```

sets the shared address which the phone should use. On the other hand, the *OptiPoint400 Standard* requires the H.323 NAT support. For this purpose, routing the TCP ports 1710–1720 and UDP ports 5010–5017 to internal addresses was sufficient.

10.6.3 The Kerio Technologies Products

This Project team has been involved in long-term cooperation with the Kerio Technologies company, a developer of VoIP program applications. Project members test the Kerio products and in return, they are allowed to use these application within the CESNET2 network.

In the first half of 2003 we have tested the *Interactive Voice Response (IVR)* system which could be used, e.g., as a base for information call centre or as an information system. We created necessary scripts using the VoiceXML language which reproduce a selected text to users using voice synthesis. Callers can use the DTMF tone dialing to traverse the information tree; automatic switching to a phone number included in the script has been tested successfully, too.

Kerio Technologies suspended development and sale of its VoIP products in September 2003. Our external H.323 peering is based on a *Kerio Gatekeeper* (a Linux application); therefore, another suitable platform should be found to replace it by the end of 2004.

10.6.4 New version of H.323

In July 2003, the ITU-T H.323 version 5 recommendation was released officially. Considering the number of new features, we are glad that the H.323 standardising process is still active. The fifth version brings new management functions and new services; in adition, it is backwards compatible with previous versions.

10.6.5 H.323 interconnection with the CESNET2 network

Two possible interconnection modes exist: direct routing to VoGW, or through the central Gatekeeper (GK). Fundamental difference lies in the H.225.0 signalling mode which has two parts – RAS signalling and Q.931 (Call Signalling). If direct routing to VoGW is used, RAS signalling is not used – see the Figure 10.8.

Our preferred solution uses the CESNET GK *gk-ext.cesnet.cz*. The connection methods and actual configuration files can be found in the CESNET Technical report 28/2003.

10.7 Tasks solved in the SIP area

In 2003 we continued testing the infrastructure components which use the SIP signalling protocol.

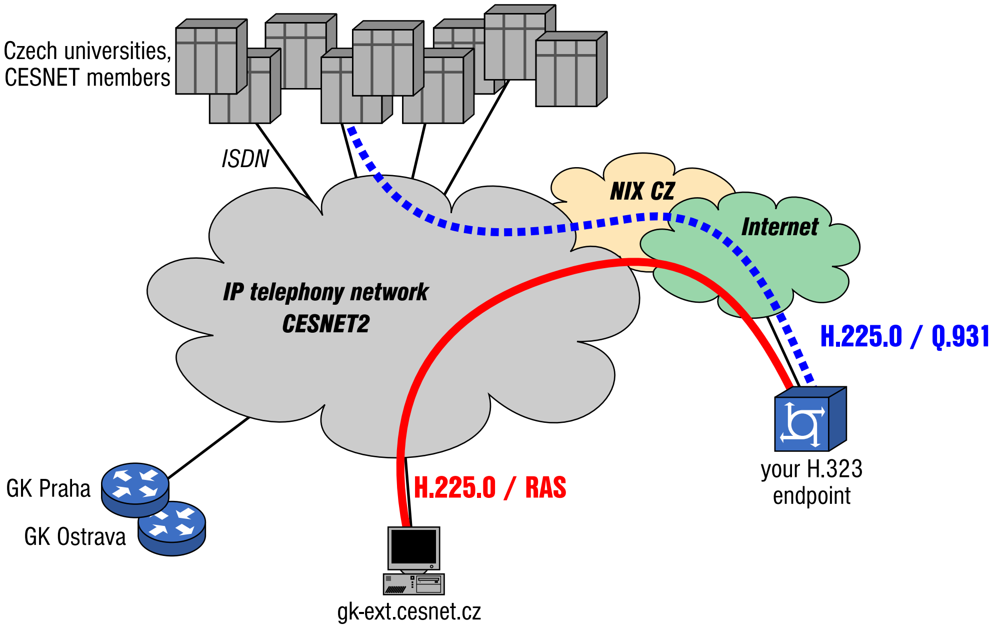


Figure 10.8: H.225.0 signalling routes

10.7.1 SIP Server and gateways

Core component of the SIP VoIP network is a proxy and registrar server *SIP Express Router (SER)* installed on a PC platform under a Linux RedHat operating system. At present, this server can process requests of directly connected test SIP clients and redirect calls to all VoGW of connected institutions. These VoGWs, used in the H.323 network, can process concurrently both the SIP and H.323 signalling protocols.

Rules for call forwarding to voice gateways according to telephone number prefixes are stored directly in the configuration file. Therefore, changing these rules without restarting the server is impossible. We are considering building an external routing data storage but this solution would require writing a new server module as well. This function is not necessary at present and so its implementation has been postponed.

Rather a more crucial problem is the inability to establish a SIP connection from PBXs of affiliated institutions. Users dial a special prefix, e.g., 94, to connect using the VoIP network. As a result, the call is connected using the H.323 signalling protocol. Creating another prefix for entering the SIP network would be absolutely inconvenient: it would only confuse the users and it would not improve the functioning and comfort of the VoIP service.

One possible solution would be using the ENUM services – translating phone numbers to URI using domain name server records: after receiving appropriate

ENUM records from the name server, the gateway should be able to decide which signalling protocol is to be used for the connection. Unfortunately, the ENUM services do not yet operate perfectly on the gateways; we endeavour to solve this issue by cooperating closely with the gateway manufacturer.

To demonstrate the SIP routing configuration, a part of the *ser.cfg* is presented. This determines the routing functions of the SIP Express Router (SER) application working as a proxy and registrar server. Called numbers can be routed using an international format starting with 420 or using the nine-digit national format.

```
# main routing logic
route {
    if (uri==myself || uri=~”[@:]cesnet\.cz([;:].*)*”) {
        # Save location contact
        if (method==”REGISTER”) {
            if (!www_authorize(”cesnet.cz”, ”subscriber”)) {
                www_challenge(”cesnet.cz”, ”0”);
                break;
            };
            save(”location”);
            break;
        };
        # Gateway forwarding section start
        if (uri=~”^sip:(420)?22435[0-9]{4}@”) {
           rewritehostport(”GWIPaddress:GWport”); # CVUT Praha
            route(1);
            break;
        };
        if (uri=~”^sip:(420)?59(699|732)[0-9]{4}@”) {
          rewritehostport(”GWIPaddress:GWport”); # VSB Ostrava
            route(1);
            break;
        };
        if (uri=~”^sip:(420)?54949[0-9]{4}@”) {
            rewritehostport(”GWIPaddress:GWport”); # MUNI Brno
            route(1);
            break;
        };
        #.
        #.
        #.
        # Other gateways
        # Gateway forwarding section end

       #  native SIP destinations are handled using our USRLOC DB
        if (!lookup(”location”)) {
            sl_send_reply(”404”, ”Not Found”);
            break;
        };
    };
    # forward to current uri now
    if (!t_relay()) {
        sl_reply_error();
    };
}
```

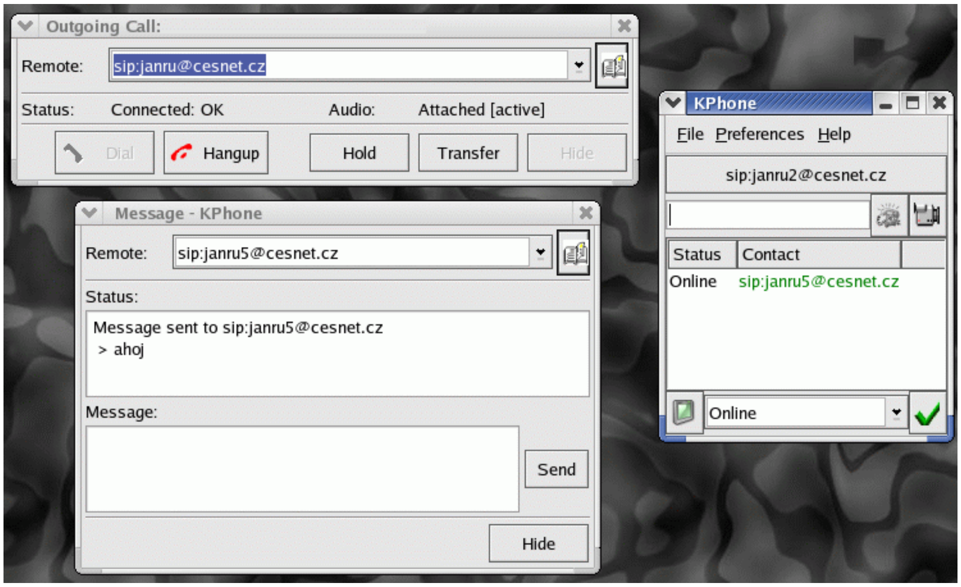
Further important function of the core SIP server is the authentication and the authorisation of IP phones and software clients. The server is equipped with modules which provide these functions by querying a MySQL database or a RADIUS server. However, most authentication and authorization functions within CESNET are resolved using the LDAP directory services. We managed to acquire the server LDAP modules from the development team and we are testing them. Unlike the core of the server, these modules are not available under the GPL license. As soon as the tests of AAA mechanisms are finished, accessing the PSTN via SIP clients will be possible.

Full integration of the SIP and H.323 networks requires an operating translation gateway of signalling protocols. The products we started testing by the end of last year proved unsuitable; therefore, tests of an *Asterisk* software PBX are being prepared.

10.7.2 SIP IP telephony clients

Here we concentrate particularly on testing the IP phone clients. In the field of hardware IP phones, we are concerned especially with the Cisco and Siemens products. An advantage of Cisco IP phones is the fact that an easy change of firmware allows connecting them to the Cisco CallManager. For the Siemens IP phones, three firmware versions are available: SIP, H.323, and HFA.

In addition to tests of hardware phones, we attempt to find a suitable software client for two favourite platforms – Linux and MS Windows. One of the agents for the Windows platform is the *Windows Messenger* (not to be confused with the *MSN Messenger*), currently in version 5. This version corrects major problems in signalling protocol implementation. This client allows voice calls, videoconferencing sessions as well as sending short text messages. A Linux client with similar features is the *Wirlab Kphone*, currently in version 3.14 – see Figure 10.9.



Remote: sip:janru@cesnet.cz

Status: Connected: OK

Audio: Attached [active]

Remote: sip:janru5@cesnet.cz

Status:

Message sent to sip:janru5@cesnet.cz
> ahoj

Message:

sip:janru2@cesnet.cz

Status	Contact
Online	sip:janru5@cesnet.cz

Online

Figure 10.9: The KPhone SIP client

We also test clients made by SJLabs, Xten and others. Our main requirements are stability, quality of user interface, sufficient choices and quality of codecs and a sufficient number of functions provided.

10.7.3 Interconnecting the SIP IP telephony networks

One of the criteria of a telephony network quality is the number of reachable workplaces. To localise the called participant, SIP uses the DNS service. Calls into and from outside the CESNET Association networks can be realised without additional settings or difficult management of special connecting rules. Using the SIP protocol, one can call the *iptel.org*, SANET, NASK, MIT and many other networks. The experimental service ENUM (RFC 2916) provides translation between the E.164 numbers and URIs. As a result, an E.164 number acts as a unified identifier for accessing several different services (H.323, SIP, e-mail, Web, etc.).

10.8 Further project goals

The team members prepare a set of rules of IP phone operation for members of the CESNET Association. These are intended especially for remote workplaces with an IP connectivity where few phone connections exist and where IP phones could replace the fixed phone lines. In present, IP phones are used mainly by the CESNET staff. A mass deployment of IP phones is expected in the second half of 2004.

To make the VoIP service extensible and manageable, the project team also prepares rules for using the IP phones to which applicants for an IP phone service with a public phone number must conform. The public phone numbers for IP phones have been allocated to CESNET by the GTS Czech operator. We are also negotiating with the Czech Telecommunications Office to be allocated an access prefix for the CESNET2 network.

The project team has the following priorities:

- supporting the cooperating project members
- replacing the Kerio Technologies external peering GK *gk-ext.cesnet.cz*
- testing new H.323 products and monitoring the Open H.323 solution development
- enhancing the SIP server by more services (Web), servicing more domains, backup
- completing the authentication and authorisation mechanisms for the SIP AAA, multi-domain authentication mechanism using the LDAP directory services
- solving the problem of SIP calls into the PSTN and VoGW
- testing the SIP clients
- implementing translation between the E.164 number and URI using ENUM
- interconnecting with foreign Universities
- writing an application which would determine the QoS of realized calls using the VoGW records
- publishing activities and presenting the project results regularly.

10.9 Project publishing activities

We participate in the TERENA international project *IP Telephony Cookbook*; our contributions become the Cookbook subchapters. Mr. Sven Ubik is the CESNET contributor. The project results are published regularly and presented at conferences.

11 End-to-end performance

This project investigates theoretical and practical aspects of end-to-end performance to provide high throughput and other qualitative communication characteristics required by applications communicating over wide-area high-speed networks.

Our results are presented on the project web pages[1], These include all published papers, technical reports, presented talks, experimental data and developed software. In this chapter, an overview of selected project results from 2003 as well as some interesting technical problems are presented.

11.1 Transferring large data volumes over large-scale high-speed networks

The Internet has been a large-scale network spanning long distances almost since its origin. However, two new characteristics have changed the Internet only recently. First, it has become a truly high-speed network with backbone links operating at 10 Gbps or even higher speeds. Second, researchers in fields such as physics or astronomy have started to transfer large volumes of data, from terabytes to petabytes.

As all these three characteristics (long distances, high speeds and large date volumes) have met, one has found that the communication protocols used so far, particularly the reliable TCP transport protocol carrying over 95 % of Internet traffic, as well as the data processing mechanisms on connected computers, no longer suffice to provide required communication qualities. Their considerable improvement is necessary in order to achieve high throughput and other qualitative characteristics, such as low delay fluctuation, required by current applications. This usually belongs to the *end-to-end performance* field.

In 2003, several papers on transfer of large data volumes in large-scale networks were presented on both domestic and international conferences. Some interesting technical details are presented here.

[1] *http://www.cesnet.cz/english/project/qosip/*

11.2 End station configuration

We have found that at speeds approaching some 100–300 Mbps, most performance problems result from suboptimal configuration of end stations. At higher speeds above some 300 Mbps, modifying the characteristics of communication protocols is usually necessary. In this section, we shall mention the most important end station configuration details which influence the throughput achievable.

11.2.1 Socket buffers

Socket buffers on both sides of a connection (sender and receiver) limit the TCP protocol window of outstanding data (which must fit in the smaller of these two buffers) and are therefore a critical factor which influences the achievable throughput. The window size limits the volume of data that can be transferred during one RTT (round trip time) interval. Default size of socket buffers in most operating systems ranges from 16 kB to 64 kB. Such a small window combined with RTT at the order of tens or hundreds of milliseconds, which is common in long-distance communication, limit the throughput to several or several tens of Mbps regardless of the bandwidth available in the network.

Socket buffer sizes can be adjusted either for all new connections opened in an operating system, or individually for each socket opened within an application. Some operating systems, such as Linux, provide a sort of autoconfiguration and window moderation which adjusts the buffer size according to current requirements and available memory. For example, in the Linux operating system, one can use the following command to set default sender and receiver socket buffer sizes to 2 MB for all new connections:

```
systctl -w net/ipv4/tcp_rmem=4096 2097152 16777216
systctl -w net/ipv4/tcp_wmem=4096 2097152 16777216
```

An example of socket buffer size influencing the throughput achievable between the CESNET2 (CZ) and UNINETT (NO) networks is shown in Figure 11.1. However, there are more details involved. Linux further modifies the requested buffer sizes according to values of some kernel variables, resulting in a TCP window limit different from the specified socket buffer size. Linux also includes several other TCP implementation specifics influencing performance. We described some of them in [UbC03] and [UbC03a]. A more detailed technical report describing more Linux internals, whose understanding is useful for the end-host performance tuning, is being prepared.

Unfortunately, we can also get into trouble by setting the socket buffers too large and allowing the window to grow too much. Big windows can fill up router

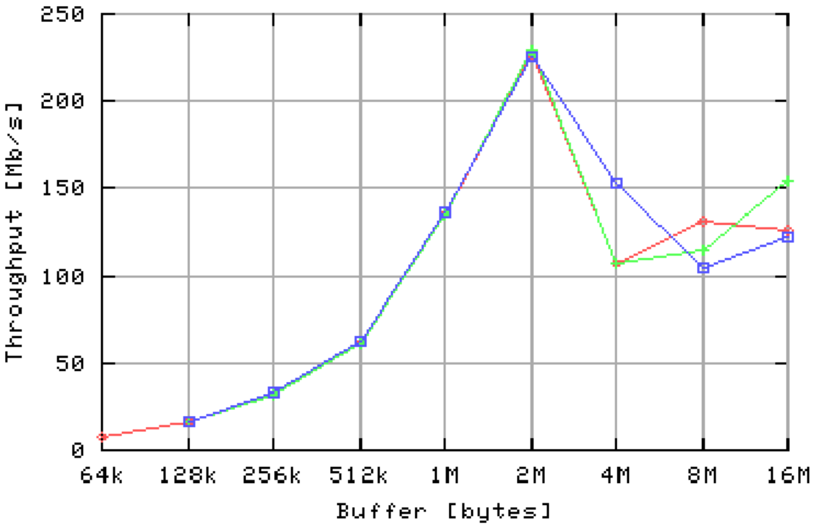


Figure 11.1: Relation between the achievable throughput and socket buffer size

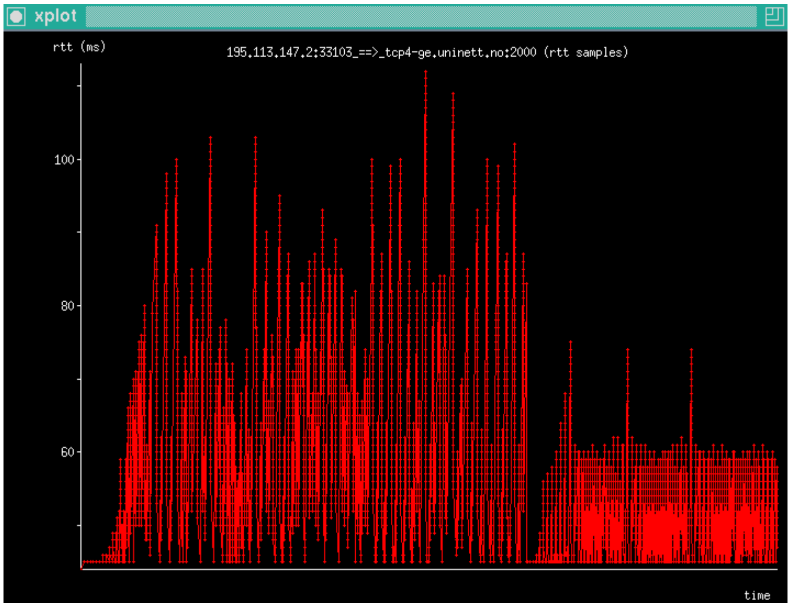

xplot
rtt (ms)
195.113.147.2:33103_==>_tcp4-ge.uninett.no:2000 (rtt samples)
100
80
60
time

Figure 11.2: RTT fluctuations during one TCP connection

queues, which together with traffic fluctuations increases probability of a queue overflow and packet loss. As a result, congestion control will react by reducing the data sending rate. We can try to predict this phenomenon by observing the relation between current throughput and window size, or by monitoring the RTT fluctuations. For example, we can see in Figure 11.2 that the RTT of a monitored connection reached up to several multiples of the basic RTT measured on an unloaded network which was about 40 ms. At some time, a packet was lost and throughput was reduced. Consequently, the RTT stabilised again. However, it turns out that in highly-multiplexed backbone circuits with complex traffic dynamics, determining conclusively that the RTT growth and fluctuations have been really caused by filling up the router queues is very difficult.

An important Linux networking component that requires proper configuration is a network adapter transmission queue (txqueue). Each network adapter has its own txqueue. Packets from all connections transmitting through a network adapter come to its txqueue before they are moved to the adapter and sent to the network. We discussed the txqueue behaviour in more detail in [UbC03a]. As a rule of thumb, the *ifconfig* command setting the txqueue to 1000 packets for a Gigabit Ethernet adapter can be used.

11.2.2 Application tuning

Throughput can also be limited by the application. For example, we noticed low throughput while copying files over a network using the well-known *scp* utility. Socket buffers, txqueue and other networking components in the operating system were configured properly. Processor load on both end stations was low.

We found that the problem was caused by the way the ssh protocol (used by the scp utility) handles its data: it is split into 32 kB blocks which are acknowledged at the application level; the default maximum number of outstanding blocks is four. Thus, the ssh protocol creates its own application window with a default maximum size of 128 kB above the TCP window. The size of this window can be set in the source code of the ssh distribution. The influence of increased ssh window on the throughput is shown in Figure 11.3. Of course, increasing the data rate to be ciphered and deciphered increases the processor load as well.

11.3 PERT

PERT (Performance Enhancement and Response Team) is an emerging international initiative which attempts to create technical and organisational framework to help users resolve their networking application performance problems. To some extent, PERT should enhance performance just like CERT improves security.

CESNET takes an active role in PERT preparation, presently within the TF-NGN Geant activity. In the second half of 2004, PERT should become a part of the proposed GN2 project. Our experience with PERT preparation has become a part of the D8.1 deliverable “Multi-domain monitoring and PERT” of the GN2 project.

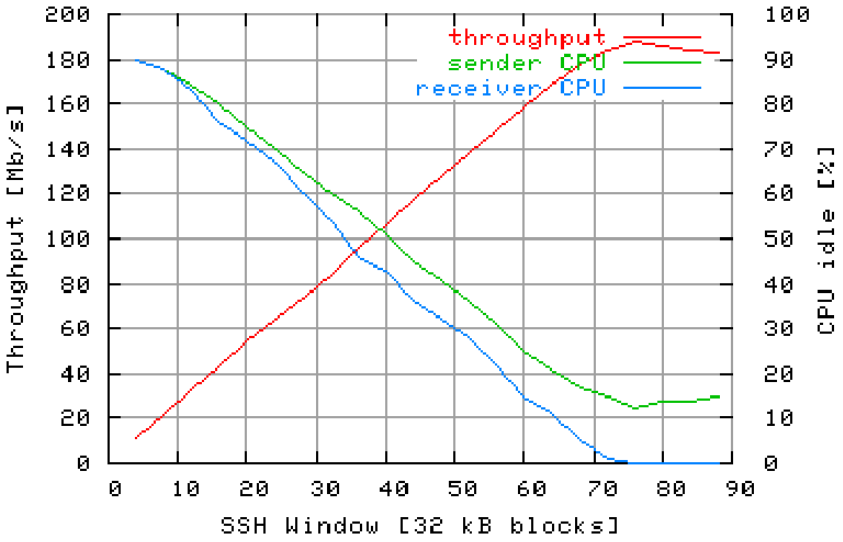


Figure 11.3: Relation between the throughput and internal ssh protocol window

We identified two groups of people interested in the PERT activities and willing to become pilot PERT users. The first group are the GRID researchers (particularly people from the Masaryk University); the other group are people taking care of the streaming video data transfer over the Internet.

We started to build the PERT web pages which should include three parts:

- PERT mission and operation
- frequently asked questions (FAQ), optionally updated by users
- database of known performance problems with user interface for case submission.

We proposed a structure of performance problem description and we created a trial database using MySQL and PHP4 scripts. After getting more experience and considering the requirements, we concluded that a new database version based on the *Request Tracker (RT)* will be needed. We identified the following requirements and motivations leadings to our decision to use the RT:

- RT including the LDAP authentication is already being used in CESNET.
- The database must include problem solution tracking (included in RT).
- The database must be accepted by the network operation people (who are accustomed to using the RT, we regard this as an important factor).
- User interface must be tailored for PERT purposes (in contrast to the generic RT user interface which includes several elements unnecessary or useless for PERT) and must be properly localised (in contrast to current mixture of English and Czech).
- The system must be distributed. The proposed structure includes one database for the backbone network (GN2) and one database for each NREN with optional case escalation. Detailed solution will require further

in-depth analysis. One problem is already known: RT currently does not support any distributed structure, but we anticipate that this can be solved after some research and development.
- People developing the CESNET RT should be involved in this project. As the GRID research team has some more requests on adding functionality, we plan to include RT development in the CESNET activities for 2004 as well as in the GN2 project.

We propose that during escalation, each case will be investigated first to determine the likely problem area and subsequently forwarded to the person responsible for resolving problems in that area. The following potential problem areas have been identified:

- Unix (TCP window tuning, etc.)
- Windows (driver problems, etc.)
- PC hardware (interrupts, component selection, etc.)
- local or remote LAN (switches, DHCP, DNS, etc.)
- local or remote metropolitan network
- local or remote NREN
- GN2 or another global network

Perhaps the most difficult task will be finding and training the right "front-line" people accepting cases and identifying problem areas, as well as people responsible for individual problem areas.

Another task critically important for the PERT success is availability of a good performance monitoring system. Requirements on such system are currently being specified based on experience from many individual performance measurements. The system should also be developed in the GN2 project framework.

11.4 Data link bandwidth estimation

Available bandwidth along a certain network path, i. e. the part of the installed bandwidth not currently used by existing traffic, is a very important dynamic network characteristics. It suggests what throughput can be expected for additional applications, whether any network segment is overloaded or failing, or whether network upgrade may be necessary.

Available bandwidth measurement tools, such as *iperf*, try to completely fill all remaining bandwidth by sending data as fast as possible and measuring the achieved throughput. Obviously, this method affects the existing traffic significantly and may be used only for a short time.

In contrast, the free capacity estimation tools send only several carefully scheduled packets and try to estimate the bandwidth available by analysing the sending and receiving times of testing packets.

11.4.1 Classification of bandwidth estimation tools

As the prospect of estimating the available bandwidth without stressing the existing traffic appears attractive, we have decided to investigate if these tools can also be used in large high-speed networks. Previous studies were mostly limited to lower speeds or simple network topologies. The bandwidth estimation tools can be classified according to the following criteria:

- whether it can determine the bandwidth of the bottleneck or of all links along a path
- whether the installed or free bandwidth is reported
- whether it is based on observing the RTT changes of single testing packets or on observing delay dispersion of a set of testing packets
- whether installation on the sender, receiver or on both sides is required.

We classified several known tools representing different approaches in Table 11.1.

	Every link	*Installed vs.*		
Tool	*vs. bottleneck*	*free bw*	*Method*	*Location*
Clink	bottleneck	installed bw	RTT	sender
Sprobe	bottleneck	installed bw	dispersion	sender + receiver
Pchar	every link	installed bw	RTT	sender
Pathchar	every link	installed bw	RTT	sender
Pathrate	bottleneck	installed bw	dispersion	sender + receiver
Pathload	bottleneck	free bw	dispersion*	sender + receiver
ABwE	bottleneck	free bw.	dispersion	sender + receiver

* relative one-way delay

Table 11.1: Classification of bandwidth estimation tools

The *pathload* tool reports IP-level available bandwidth, whereas the *ABwE* tool reports free bandwidth normalized for the TCP protocol.

11.4.2 Observation summary

We summarized our observations on behaviour of bandwidth estimation tools in the CESNET technical report 25/2003. We shall mention several interesting findings here.

The *pathload* tool, as distributed, can estimate bandwidth up to some 120 Mbps. After tuning some of its internal constants, we managed to make it work at some 800 Mbps. However, tests on our testbed with traffic bandwidth generated by a packet stream showed that *pathload* could provide only very coarse estimates in this range. When accuracy of 100 Mbps was requested, all results fitted in; however, when accuracy of 10 Mbps was requested, most results were out of range.

Within another experiment, a set of several bandwidth measurement and estimation tools was deployed for a period of one month on two paths over the Géant network, consisting of more than 10 routers and OC-48 or Gigabit Ethernet links. Every hour, one set of traffic measurements and estimations by each tool took place:

- *TCP iperf* with various socket buffer sizes
- parallel *TCP iperf* with five data streams
- *UDP iperf*
- *Pathload*
- *ABwE*
- *TCP iperf* with socket buffer size adjusted according to the ABwE results

A sample of measured results in one four-day period is shown in Figure 11.4.

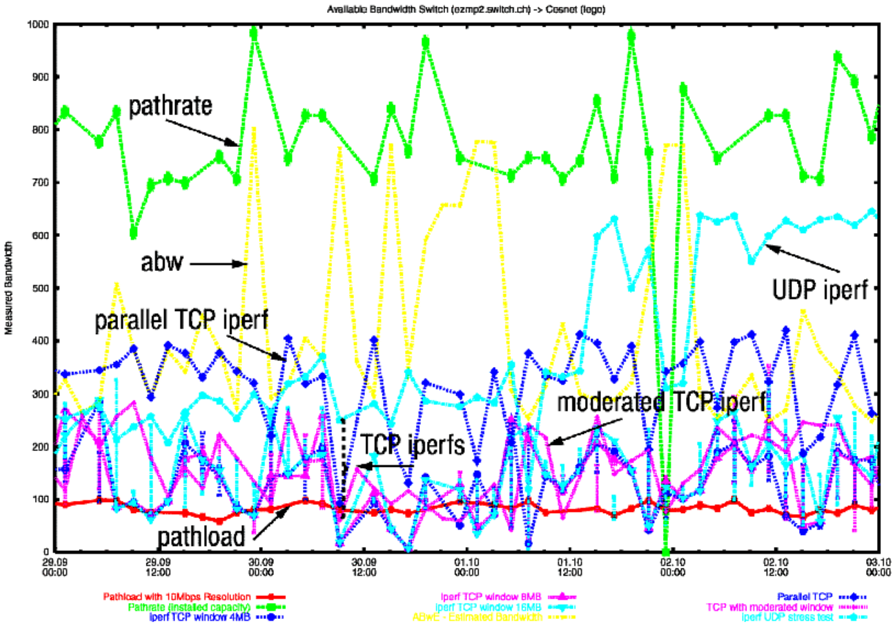


Figure 11.4: Bandwidth measurement and estimation

One can see that values produced by different tools vary significantly and concluding which value is close to the real available bandwidth is difficult. We can assume that parallel *TCP iperf* or *UDP iperf* are more likely to fill the available bandwidth, but they also more stress the existing traffic and so they can report results higher than bandwidth really available. The *pathload* command is very unreliable and often systematically underestimates the available bandwidth. A more detailed discussion of our observations can be found in an internal project report [UKr03].

11.5 Computer network simulations for congestion control research

Computer network simulation and emulation allows researchers to conduct experiments on models of computer networks in order to evaluate protocol behaviour and compare alternatives under defined and repeatable conditions, which would not be possible on real networks with unpredictable traffic dynamics. The most widely known network simulator is the *ns2*. Our experience with using *ns2* for congestion control research as well as our additions and enhancements to this simulator have been published in the CESNET technical report 26/2003. In this section, a summary of some of our findings and recommendations for use of *ns2* follows.

The *ns2* is a freely available discrete-event object-oriented network simulator which provides a framework for building a network model, input data specification, output data analysis and result presentation. Source code is also available which allows users to add new features to the simulator, such as support for new communication protocols, monitoring tools, etc.

In real networks, four components make up the end-to-end packet delay. The *ns2* tool simulates all these delay components except for the processing delay:

- Serialisation delay – time needed to put the packet on network link
- Propagation delay – time needed for the energy representing a single bit to propagate along network links, bounded with the speed of light
- Queueing delay – time that the packet waits in network node queues to be served
- Processing delay – time needed to process a packet in network nodes

11.5.1 Installation and simulation scripts

The *ns2* tool is implemented in C++ and Tcl and should run on any Posix-like operating system (tested on FreeBSD, Linux, SunOS and Solaris) and on Microsoft Windows. The *ns2* uses several other software packages (Tcl/Tk, xgraph, etc.) which can be installed either separately or together with *ns2* from the "ns-allinone" package. Some of these packages are mandatory, while others are optional, such as the *nam-l* for animation of a simulation run.

Once the *ns2* is installed, a simulation task is specified by a simulation script written in Tcl. This script describes the network topology (nodes and their interconnection), communications protocols (e.g., TCP) and events (scheduling of data streams to be sent). Lengths of packet queues attached to links and maximum size of TCP window can also be specified. Creating the simulation scripts is a complex task which requires understanding of the *ns2* object classes and Tcl programming.

11.5.2 TCP in ns2

There are two flavours of TCP in *ns2*. The first is a one-way TCP which uses objects of different classes on the sender and receiver sides. For the sender side, several classes are available for TCP: Tahoe, Reno, Newreno, Vegas and Sack or Fack, supporting selective acknowledgements. For the receiver side, three classes are available for TCP receiver: without delayed acknowledgements, with delayed acknowledgements and with selective acknowledgements. Subclasses can be derived from these supplied classes to implement modifications to the standard TCP congestion control. The second flavour is a two-way TCP which uses objects of the same class on both the sender and the receiver sides. One-way TCP is used more frequently than the two-way TCP which implements only the Reno congestion control and is considered under development.

TCP in the *ns2* differs from real TCP implementations in several aspects that need to be considered during simulations, such as absence of flow control or sender blocking calls. It also does not include any throughput indication needed for almost any simulation. Our observations of TCP in *ns2* have been published in the project report [UbK03].

11.5.3 Example of simulation using ns2

One of the network topologies frequently used in simulations is shown in Figure 11.5. Hosts connected to router R1 send data to hosts connected to router R2. The sum of data rates produced by source hosts is usually bigger than throughput of the link between router R1 and router R2, making it a bottleneck link. This link has also a specified non-zero packet loss rate and one-way delay while the links between hosts and routers usually are lossless and have fixed one-way delay and throughput.

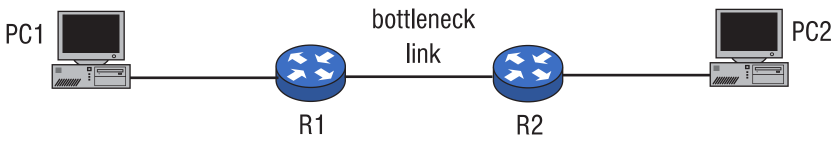


Figure 11.5: A simple simulation topology

The following steps must be taken:

1. Create an object for the *ns2* simulator.
2. Create objects for network nodes, links and queues attached to links and specify their parameters, thus creating the network topology.
3. Create objects for the TCP sender and TCP receiver and specify their maximum window sizes.
4. Create objects for the sending and receiving applications and attach them to the TCP sender and TCP receiver objects, respectively.
5. Schedule events, such as start and end times of data streams and when the simulation should stop.
6. Start the simulation.

An example simulation script implementing the previous steps (refered to by corresponding numbers in comments) on the given network topology can look as follows:

```
# 1. Create an object of the ns2 simulator
set ns [new Simulator]

$ns color 0 Red
$ns color 1 Blue

proc finish {} {
        exit 0
}

# 2. Create objects for network nodes, links and queues attached to links
#   and specify their parameters, thus creating the network topology
set pc1 [$ns node]
set pc2 [$ns node]
set r1 [$ns node]
set r2 [$ns node]
set em [new ErrorModel]

# Set link characteristics
$ns duplex-link $pc1 $r1 90Mb 20ms DropTail
$ns duplex-link $r1 $r2 50M 100ms DropTail
$ns duplex-link $r2 $pc2 90Mb 20ms DropTail

$ns queue-limit $pc1 $r1 6000000
$ns queue-limit $r1 $r2 300000

$ns duplex-link-op $pc1 $r1 orient right
$ns duplex-link-op $r1 $r2 orient right
```

```
$ns duplex-link-op $r2 $pc2 orient right

$em unit pkt
$em ranvar [new RandomVariable/Uniform]
$em set rate_ 0.0001
set streams 5
set segsize 1500

for {set i 0} {$i < $streams} {incr i} {
# 3. Create objects for the TCP sender and receiver and specify maximum
#    window sizes
 set tcpz($i) [new Agent/TCP/Reno]
 set tcpc($i) [new Agent/TCPSink]
 $ns attach-agent $pc1 $tcpz($i)
 $ns attach-agent $pc2 $tcpc($i)
 $tcpz($i) set fid_ 0
 $tcpc($i) set fid_ 1
 $ns connect $tcpz($i) $tcpc($i)
 $tcpc($i) listen
 $tcpz($i) set window_ 500
 $tcpz($i) set segsize_ $segsize
# 4. Create objects for the sending and receiving application and
#   attach them to objects for the TCP sender and receiver, respectively
 set snd($i) [new Application/FTP]
 set rcv($i) [new Application/TCPCNT]
 $snd($i) attach-agent $tcpz($i)
 $rcv($i) attach-agent $tcpc($i)
}

set null [new Agent/Null]
$em drop-target $null
$ns lossmodel $em $r1 $r2

# 5. Schedule events, such as the start and end times of data streams
#    and when the simulation is to stop
for {set i 0} {$i < $streams} {incr i} {
 $ns at 0 ”$snd($i) start”
}
$ns at 0 ”$rcv(0) settimer 0.1”
$ns at 0 ”$tcpc(0) settimer 0.1”

for {set i 0} {$i < $streams} {incr i} {
 $ns at $TIME ”$snd($i) stop”
}

$ns at $TIME ”$rcv(0) stop”
$ns at $TIME ”finish”

# 6. Start simulation
$ns run
```

11.5.4 Memory requirements

The volume of memory required by the *ns2* for a simulation depends on the number of packets within the simulated network and on the number of packet headers maintained for each packet. In fast long-distance networks, which are often a subject of current research in congestion control, the number of packets within the network can be some tens or hundreds of thousands and the volume of memory required can grow to several gigabytes. The memory requirements can be lowered by first removing all packet headers and then adding only the required headers. For example, the following commands can be added at the beginning of a simulation script:

```
remove-all-packet-headers
add-packet-header TCP IP
```

11.5.5 Scripts for batch processing

To evaluate the congestion control mechanisms under various network conditions, a set of simulations of a selected network topology must be run where the network characteristics of the bottleneck line are varied. These include the link bandwidth, packet loss rate and one-way delay. We may also wish to experiment with different packet sizes, number of parallel streams, as well as changing the test duration and time granularity for computing the resulting characteristics, such as the achieved throughput.

We added logging of TCP connection characteristics and created a set of scripts to simplify the use of *ns2* for simulation of common experimental scenarios with various link characteristics and protocol parameters. The inter-relations of individual scripts are illustrated in Figure 11.6:

The sequence of script actions can be described as follows:

- Script *sim1.tcl* describes the network topology and simulation task
- Script *simrun* runs a simulation with some parameters:
  - Script *simrun* creates the scripts *simtemp.tcl* and *simtemp.gpl*
  - Script *simrun* calls the *ns2* for the *simtemp.tcl* script

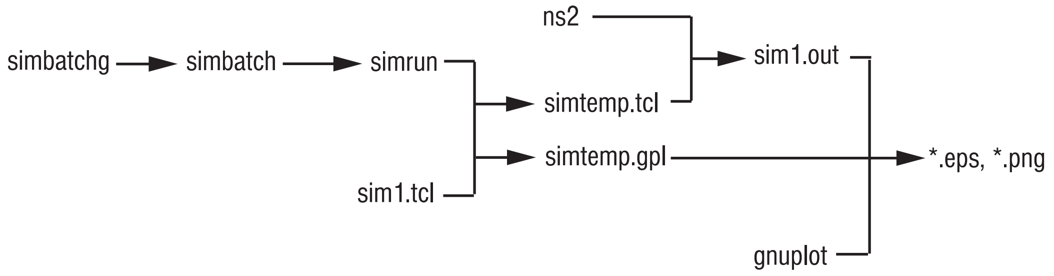


Figure 11.6: Scripts for batch simulation processing

  - *Ns2* creates the output file *sim1.out*
  - Script *simrun* calls the *gnuplot* for the *simtemp.gpl* script and *sim1.out* file
  - *Gnuplot* creates diagrams in PNG format

- Script *simbatch* calls repeatedly the *simrun* script with different parameters
- Script *simbatchg* can create the *simbatch* script according to specified criteria

11.5.6 Throughput measurement

To monitor the throughput at the application level, we created a new class *Application/TCPCNT*; to monitor throughput at the TCP level, we modified the class *Agent/TCPSink*. A description of these enhacements can be found in [UbK03].

11.5.7 Reaction to a change of available bandwidth

In order to study responses of a congestion control mechanism to increased or decreased available bandwidth, we created a sender-side application class *Application/TCPFTP* which generates periodic bursts of packets. To start the application, the following commands in the simulation script can be used:

```
set snd [Application/TCPFTP]
$snd set interval_ n
$snd set burstsize_ m
$snd start
```

where *n* is the period in seconds and *m* is the number of MSS-length packets to be sent in each period. The application must be attached to the TCP sender – see the example simulation scripts. To stop the application, the following command in the simulation script can be used:

```
$snd stop
```

11.5.8 Adjusting the AIMD parameters

In the original *ns2* TCP, the congestion control parameters within the slow start as well as congestion avoidance phases are fixed. The latter is based on AIMD(1, 0.5). To be able to experiment with recent proposals of Fast TCP, changing the AIMD parameters should be possible. We have modified certain *ns2* classes so as to be able to adjust both the slow start and congestion avoidance parameters. A detailed description of these enhancements can be found in [UbK03].

11.5.9 Asynchronous monitoring of TCP characteristics

The *Ns2* can synchronously monitor the TCP charakteristics (cwnd, ssthresh,. . .) after any of them is changed. In some cases, an asynchronous monitoring (recording the values of all characteristics in a given time interval) may bring clearer results. Therefore, we modified the *ns2* to allow asynchonous monitoring as well. A detailed description can also be found in [UbK03].

11.5.10 Difficulties we ran into

We encountered several symptoms of unexpected behaviour and ran into some problems when using the *ns2*:

- TCP stops sending data for a few seconds sometimes
- slow start occurs in TCP Reno after the router queue fills up
- throughput diagrams show fluctuations in fine-timescale
- non-numeric artefacts appear in the simulation log.

These phenomena are presented together with explanations for some of them in [UbK03].

At the present time, the *ns2* simulator is used for research in congestion control for long-distance high-speed networks. A paper on this topic is being prepared.

11.6 Developed software

In 2003 we created the following software packages:

11.6.1 Evaluation of bandwidth measurement and estimation tools

A set of scripts used for evaluation of bandwidth measurement and estimation tools. The obtained results were presented in the CESNET technical report 25/2003.

11.6.2 Analysis of time and geographical characteristics of network traffic

A set of tools for analysing time and geographical characteristics of network traffic from netflow records. These tools were used to analyse the CESNET international traffic. A technical report on this topic is being prepared.

11.6.3 Linux kernel monitoring

A patch for configuring and monitoring certain events in Linux kernel that influence performance of TCP bulk transfers. Particularly, it allows setting up the AIMD speed as well as enabling, disabling and monitoring the CWV and CWR mechanisms. This patch is being used for congestion control research; a paper on this topic is being prepared.

11.6.4 NIST Net deterministic patch

A patch that provides deterministic packet loss and queue length for a popular emulation package NIST Net. Our enhancements will be described in a technical report on our experience with network emulation.

11.7 Other project activities

Together with the *Optical networks and their development* project we conducted an experiment using the Intel 10 Gigabit Ethernet PC adapters in order to evaluate the feasibility of providing a 10-Gigabit Ethernet connectivity up to the end stations. The results were presented in CESNET technical report 10/2003.

Building productive relationships with international partners leading to motivating proposals of further research activities within several planned 6th Framework Programme projects is also regarded as an important project result.

11.8 Planned activities

In 2004 we plan to concentrate on three research areas. The first area is congestion control in long-distance high-speed networks. We managed to gain a lot of experience in this field and we are working on several papers and technical reports on this topic. The second area is performance monitoring. Our intention is to implement the results of the SCAMPI project which is developing a programmable monitoring platform for the high-speed Internet. The third area is the Performance Enhancement and Response Team.

Part III

International Projects

12 GÉANT and GN2

12.1 The GÉANT Project

GÉANT is the most significant international project of the European Union 5th Framework Programme in which the CESNET Association takes part. The project was launched on November 1, 2000; a Consortium of 27 European National Research and Education Networks (NRENs) operators coordinated by the DANTE Ltd. set its goal to design, build and make operational a pan-European infrastructure connecting the NRENs of European countries by October 31, 2004 when the project is to end. This infrastructure – the Géant network – must allow the European researchers the following:

- transferring large amounts of data in a short time
- make use of advanced network applications such as grid computing
- cooperate on joint projects in real-time.

This goal should be reached especially as follows:

- Backbone lines of the Géant network will have a minimum bandwidth of 2.5 Gbps; the network core will be formed by circuits with a transfer speed of 10 Gbps.

- The Géant network will ensure the interconnection of more NRENs than its predecessor, the TEN-155 network. It means especially the connection of new countries accessing the EU, e.g., Bulgaria, Estonia, Latvia, Lithuania, Romania, and Slovakia. Malta will also be connected to the Géant network.

- High-quality access to Research and Education Networks outside Europe will be provided for the Géant network users.

- In addition to standard IP services, the Géant network will provide also the Premium IP, guaranteed bandwidth service, virtual private networks, multicast and other services based on the developments in the field of communication technologies.

Since the beginning of the project, the CESNET Association has been participating in design of network topology currently interconnecting twenty-eight national and regional Research and Education Networks in thirty-three countries and serving more than thirty thousand research institutions. Due to an active participation in this task, the Czech Republic could connect directly to the network core whose lines have a transfer capacity of 10 Gbps. Furthermore, one of the network core nodes is located directly in the premises of the CESNET Association in Prague. It is connected to other GÉANT network nodes as follows:

- Frankfurt a. M., Germany – 10 Gbps
- Bratislava, Slovakia – 2.5 Gbps
- Poznań, Poland – 2.5 Gbps

Locating the network node directly in the premises of the CESNET Association significantly increases the reliability of our connection to the Géant network and it minimizes the expenses for this connection as well.

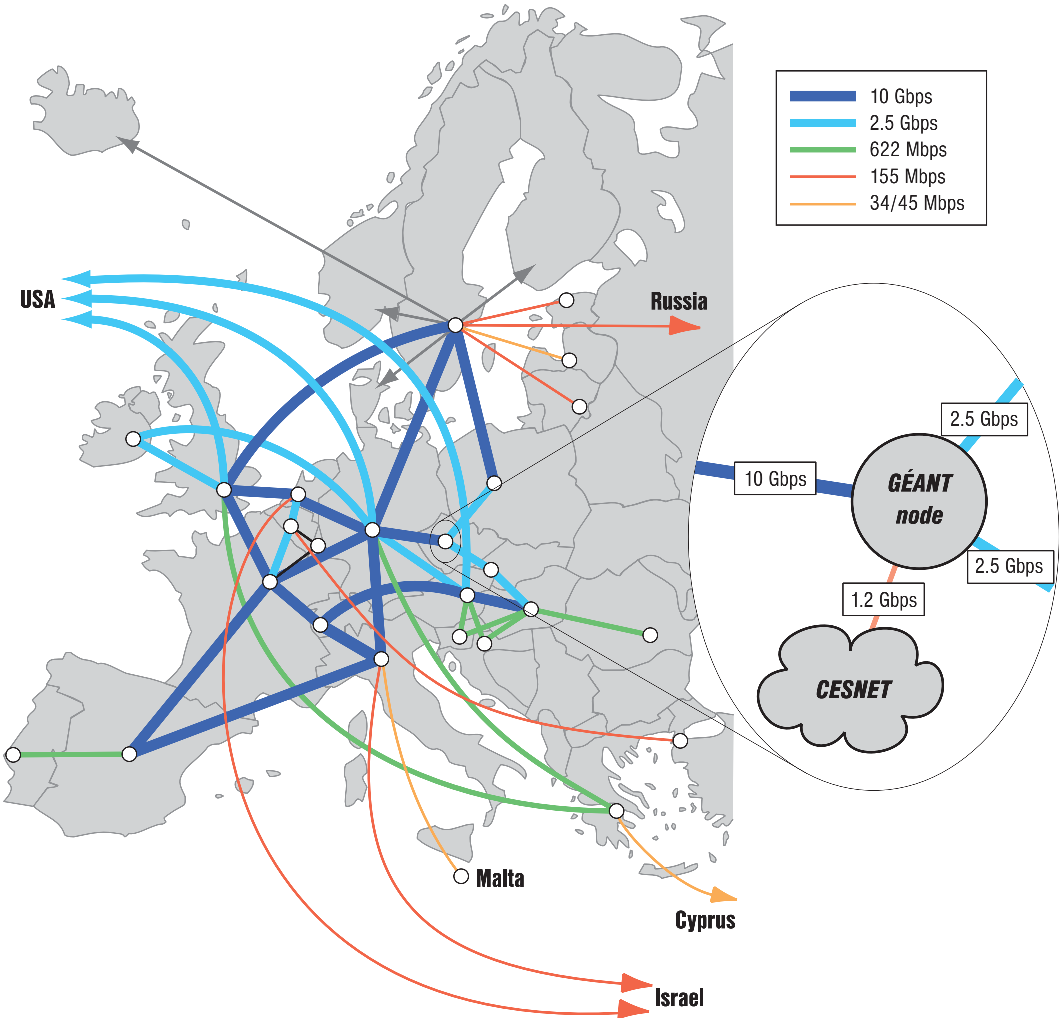


Figure 12.1: Topology of the Géant network in November 2003

A general rule in the European National Research Networks, including the Czech network CESNET2, says that the infrastructure for the research institutions and operated services themselves are research objects due to the non-standard requirements regarding their parameters. Therefore, an integral part of the GÉANT project is the research in the field of information and communication technologies conducted by teams composed of experts from individual NRENs, called the *TF-NGN (Task Force – Next Generation Network)*. The experts from the CESNET Association applied their experience above all in the fields of tool development for monitoring large-scale networks and assessing the network status, implementation of the IPv6 protocol, multicast implementation and in the field of Quality of Service.

Further information can be found at *http://www.dante.net/geant/index.html*

12.2 Preparation of the GN2 Project

As the Géant project ends in October 2004, a proposal of a new project called *Multi-Gigabit European Academic Network (GN2)* was elaborated and submitted within the call "Integrated Infrastructure Initiative" of the 6th Framework Programme in order to keep the continuity of the European Academic Network.

The aim of the project, which will start around November 2004, is to build during the next four years a modern and highly efficient infrastructure which will allow users accessing their working environment (in the sense of information resources, computing powers, etc.) in real-time anywhere within the ERA (European Research Area). The service support ensuring end-to-end performance and mobility requirements will be emphasized. Thirty-one organisations involved in the issue of high-speed research networks will take part in the project. The total planned budget is approximately 180 million EUR; the EU subsidy makes 93 million EUR.

Project activities are divided into three groups:

a) **Projecting, construction and operation of the backbone network:** These activities include the user support and the support of the National Research and Education Networks (NRENs) development. The goal is to remove the differences in the technical level of NRENs in individual European countries and make ready for implementation of services ensuring direct connection of the network end devices upon request. A part of this group activities will include those focused on technical, organizational and economical aspects of implementing the up-to-date information and communication technologies for planning the research network development and coordination of research activities of individual NRENs in the field of high-speed networks and their services.

b) **Specific services:** This group includes both activities related to the network operation and supervision and activities ensuring quality of services and end-to-end performance. Support of grid technologies is expected as well. The GN2 will also provide connectivity to research networks in other world regions.

c) **Research activities:** The Consortium will follow the tradition of research activities related to the high-speed networks, which started already during work on the TEN-34 project in cooperation with the TERENA Association. The research in the GN2 project will be focused especially on:

- development of new monitoring tools for supervising large-scale high-speed networks,
- development of tools ensuring the network security,
- development of new services, especially in the field of authentication and authorisation,
- realisation of testing environment for new technologies as well as for other 6th Framework Program projects,
- research in the field of mobility and interoperability.

The CESNET Association is involved in preparation of the project and it plans to participate significantly in the network part as well as in the research activities, especially in the field of monitoring, network security, authentication and authorisation, mobility and end-to-end performance.

12.3 Conclusion

Participation of the CESNET Association in the GÉANT project allows the Czech Republic to be involved actively in building the international infrastructure connecting research workplaces in Europe and thus ensuring high-quality access to information sources abroad for the CESNET2 users under very advantageous financial conditions. The CESNET Association has also gained a very good reputation thanks to its contribution to research activities of the GÉANT project which allowed CESNET to get involved with the preparation of the GN2 and other projects. The submitters of the GN2 project successfully attended the European Commission hearings and the specification of individual activities is currently proceeding. The final version of the project will be submitted to the European Commission in the beginning of 2004.

13 DataGrid

13.1 General information

2003 was the third year of the DataGrid project, which is a part of the 5th EU Framework Programme. The project has more than 20 participants coordinated by CERN. This year, the aim was especially to create a more stable grid environment with the possibility of testing even more large-scale application computations.

The CESNET Association team is involved in the activities of work package 1, which is responsible for resource management and for the development of a complex Workload Management Service (WMS). We are specifically responsible for the logging and bookkeeping service and security mechanisms in use. CESNET also operates the Certification Authority whose certificates are admitted by all partners of this project and other European Grid projects as well. An unfunded activity is CESNET participation in operating the DataGrid testbed together with the Institute of Physics of the Academy of Sciences of the Czech Republic.

13.2 Logging service

In 2003, the team was involved in following basic activities:

1. Gradual implementation and support of the operating version 2.0.
2. Continued development of the next version 2.1 (with the prospect of version 3).
3. Integration of the logging service with R-GMA (Grid monitoring architecture).

13.2.1 Operating version 2.0

In February 2003, the DataGrid project passed a successful second review and the project management decided on using the new version 2 in the project's production testbed. This decision enabled the implementation of a conceptually new WMS architecture before the end of the first semester, including also advanced logging service properties. Beside the standard asynchronous mode of event delivery it supports priority-based and synchronous transfers (immediate event delivery to the stipulated bookkeeping server) and application events as well.

The new WMS architecture consists of components that transfer the job control to each other through a network connection, disk queue of requests or through a direct call of the corresponding procedure. All these control transfers are registered in the form of events by the logging service. Events are logged by both the transferring and accepting component, which, in addition to increasing robustness, enables very detailed post-mortem analysis of unexpected states (loss of task information, race conditions etc.). The logging state automaton that processes events and restores task state was also accordingly modified.

13.2.2 New extensions

The new production version had a much higher stability and enabled already large-scale tests of whole grid environment. Files with several tens of thousands jobs got through with a very high success rate – more than 95–97 % of jobs finished as expected. This stress tests exposed the limits of the existing architecture but also generated new user requirements on the functionality of the whole WMS, the logging service in particular. The tests have contributed to a quality improvement of the developed software.

As opposed to the previous version, the state automaton does not have a buffer memory function but immediately calculates the new state of each job, as soon as the server accepts an event about the job. The result of the computation is stored in database, thus avoiding a recalculation of the event state even in case of its crash. This version of the state automaton supports multiple jobs at the same time, with the dependency between particular sub-jobs being described by a directed acyclic graph (DAG).

The increasing stability of WMS and further *DataGrid* middleware components has caused an increased interest in statistical data collection about the whole DataGrid and its efficiency. The best source of such data (e.g., the ratio of successful and all input tasks, waiting time in the queue, ratio of waiting in queues to the duration of the computation) is the LB service. However, a direct access to this information through the user interface causes an unacceptable load on the database itself. Therefore, we had to propose and implement a pair of commands *dump*/*load*, which allow to create a copy of all events in the database in a controlled way. The database dump can be loaded into a separate database where independent very complex searches can be performed without any impact on the logging and bookkeeping service. A sequence of successive *dump* commands generates an exact copy of the original database, capturing its development in time. It also preserves finished jobs which are otherwise purged form the active database (after some grace period). Operation *purge* can be used for deleting all completed jobs from the database, whose data were already uploaded by users. Physical data removal of items marked by the *purge* commands will be performed only after the next *dump* command, ensuring the completeness of data provided in this way.

13.2.3 R-GMA and the logging service

During the year 2003 we finished the integration of the connection of L&B service with *R-GMA*, or *Relational Grid Monitoring Architecture*. The main problems of the long development cycle were permanent modifications and instability of the R-GMA code. At present, the extension of bookkeeping server is available and the server is now able to continually send information about new states of jobs to the basic R-GMA architecture. Data arrive to the *StreamProducer*, which continues to send them to other layers of the R-GMA infrastructure. Each higher layer registers at the previous one and defines a selection function (an SQL expression) indicating the kind of received data. At the end this chain there is either a user (e.g., receiving all states of his task) or a simple notification service sending an e-mail or SMS message to the user if such a state occurs.

Unfortunately, the current R-GMA implementation does not provide all building blocks necessary for a full utilization of this infrastructure. First, no security is available, data are transferred in an open form among particular nodes and even these nodes are not authenticated in any way (let alone authorized). Such an infrastructure is too vulnerable to attacks in order to be used on a production Grid.

Further implementation deficiency is the absence of some persistent components. LB sends data to R-GMA permanently and relies on the assumption that R-GMA does not lose any information – which is not true. Furthermore, there is no possibility to query R-GMA about the latest state value of a particular task.

Consequently, towards the end of the year 2003 we started to work on a proper implementation of the R-GMA infrastructure, which would utilize LB services components and provide persistence and full security.

13.3 Security

The security in the context of *DataGrid* (like the majority of grid projects) is based on the Public Key Infrastructure and its certificates. Certificates are always issued for a limited period, which complicates the situation for tasks waiting in the queue or running on computing nodes for too long. The certificate can expire prematurely and the task can be rejected from further processing. On the other hand, certificates with a too-long validity are more prone to theft and abuse. The solution is to extend the validity of certificates before they expire – we have already worked on this solution last year.

At present we extend proxy certificates for tasks that are known to the WMS (running or waiting in a queue). Following our proposals and corrections, the *Myproxy* server has been modified to support certificates renewal. The *Globus* job manager has been also modified in order to enable certificate extensions even for running jobs. The modifications were tested by *Condor* system developers and accepted to the stable *Globus* version together with their changes. We take care of the certificate renewal for WMS and *Condor* handles the transfer of new certificates to machines with running jobs.

In the framework of the *DataGrid* project, an authorization service was implemented in 2003 through the so-called *Virtual Organization Management Service (VOMS)*. It keeps basic authorization information and provides it to entities in the form of attributed (and de facto authorization) certificates. We have extended the *Myproxy* service so that it queries also the VOMS server during the certificate renewal and ensures an update of the attributed (authorisation) certificate.

Being so-called VOMS "early adopters", we use authorization information for authorising access to the LB data. We support common manipulations with the ACL (Access Control List), where we accept user DN or VOMS groups etc. Regrettably, so far we have been the only users of this information within work package 1 and otherwise just a small group of work package 4 uses it. Consequently, for the present the WMS is not able to offer and support advanced authorisation operations like for example the possibility of cancelling other than the proper task. Provisionally, we don't distribute VOMS information to the R-GMA.

The logging service infrastructure is ready to consistently utilize authentication and authorization information, e.g., inter-logger certificate control etc. However, security requirements in the framework of the DataGrid project so far have not been that high.

13.4 The EGEE project

During the last year we were already involved in the preparation of a pan European EGEE project (Enabling Grids for E-science and industry in Europe) within the 6th EU Framework Programme, together with almost hundred participants from all European countries, Russia and USA. The aim of this project, coordinated by CERN as well, is to create a genuine production and stable pan-European grid infrastructure. The project passed successfully the initial reviews and is expected to start on the 1st of April, 2004. With a budget of almost 32 million Euros for two years, the project intends to interconnect all European national, regional and thematically oriented grids into a uniform European Grid infrastructure, which should then be available to all academic users asking for computing or data capacities. At the same time it should further intensify cooperation on the European and global level as well.

The project coordinator is CERN, centre of European research in the domain of high energy physics. In total around 70 institutions should be involved in this project, Czech Republic being represented by CESNET. Like CR, many countries are involved through their national research and education network or national Grid agency. Other partners are research institutions and universities. Russia is explicitly involved as well and EU is still investigating the most appropriate form of including USA and Japan (the condition is their financial participation). We expect each partner will bring national, regional or thematic grid infrastructures to the project (computers, storage capacities, Internet connectivity) and *EGEE* will provide the money specifically for whole Grid system management and operation and to a certain extent for necessary development and re-engineering of indispensable program facilities. Most project partners will assume the role of a regional support centre with duties in the domains of training and disseminating information about the Grid technology and *EGEE*.

In 2004, the Grid infrastructure should be built on 25 nodes with a total capacity of approx. 5,000 processors and 50 TB of disk space. At the end of the biennial project around 100 nodes with 50 thousand processors and one petabyte of disk capacity should be involved in the *EGEE* Grid. *EGEE* plans to provide a simple and controllable way for accessing all these capacities to European research community in the broadest sense of the word.

From a certain point of view, we can regard the *EGEE* project as a natural continuation of the soon-to-be-finished *DataGrid* project of the 5th EU Framework Programme. We suppose that during the first phase the *EGEE* will utilize exactly the software that has been developed by the *DataGrid* project and is now being adapted for the needs of CERN and its users. Apart from the primary target group of high energy physicists, new applications are expected in the domain of bioinformatics, later Earth sciences (remote sensing), astrophysics, chemistry and others. First three domains were mentioned in the project proposal, the involvement of users from further domains is nonetheless one of explicit aims of the *EGEE* project.

The project itself comprises several interrelated activities. Beside the European Grid operation mentioned above and activities in the domain of training and information dissemination, the following 4 areas have been identified, each with its own development potential:

1. Grid middleware development and integration
2. resulting software quality assurance
3. security
4. specific network services

Especially due to the successful contribution to the *DataGrid* EU project, CESNET, as the only institution in the Central Europe, was accepted as a partner in one of the directly financed involvement development activities, namely in further middleware development. This can be seen as an explicit recognition of the outstanding abilities of the CESNET grid team.

Along with the accession to EU we can thus look forward to a new pan-European infrastructure in the domain of large-scale distributed systems. Explicit and extensive involvement of CESNET (the most extensive from all partners in Central Europe) will mediate a direct access to the European Grid for all interested users in the Czech Republic.

14 SCAMPI

14.1 Introduction

SCAMPI (Scaleable Monitoring Platform for the Internet) is an IST family project supported by the European Commission (IST-32404). CESNET is one of the principal contractors participating in the project since the proposal preparation in the beginning of 2001. The project has started on April 1, 2002; its total duration is 30 months. 2003 represented the most important year in the SCAMPI project.

The project goal is design and development of high-speed (i.e., up to 10 Gbps) network traffic monitoring architecture. The highest layer of this architecture are applications (e.g., QoS monitoring, SLS auditing, DoS detection, accounting). The middle layer is a universal MAPI (Monitoring API) which represents an interface to different hardware platforms. Supported hardware includes commodity NICs, data formats provided by routers, and especially a specialized hardware adapter being developed within the SCAMPI project.

14.2 Progress in year 2003

Four meetings of the project members and two project reviews took place in 2003. Reviews consist of recent result presentation, checking the spent budget and project progress. A Project Officer (a clerk selected by the European Union) and a team of oponents selected by him conduct the project review.

The first review in April 2003 stated that the project fulfilled majority of tasks except for the development of specialized monitoring adapter. This development was the task of the Greek company 4PLUS. Unfortunately, this company 4PLUS had resigned completely this task but refused to return the allocated resources. We presented an alternative solution based on the COMBO6 adapter which is being developed together with several other members of the CESNET Association. The Masaryk University in Brno plays a principal role.

The review resulted in a request to write four additional documents justifying the proposed change of monitoring adapter. These documents convinced our opponents to accept the COMBO6-based monitoring adapter. As a result, the 4PLUS company was excluded from the SCAMPI project and the Masaryk University joined this project. Moreover, the CESNET budget was increased by 3 man-months.

The second review took place in November 2003. Opponents consented to all submitted documents and expressed their satisfaction with the project progress.

14.2.1 Overview of the 2003 events

Members of the SCAMPI project wrote five planned documents (deliverables)

- D1.2 – SCAMPI Architecture and Component Design
- D1.3 – Final Architecture Design
- D2.1 – Preliminary Implementation Report
- D2.2 – SCAMPI Prototype Implementation Report
- D3.2 – Experimental Plans and Infrastructure Setup

as well as four additional deliverables:

- E1.1 – Comparison between the 4Plus and COMBO6 Adapters
- E1.2 – Reallocations of Tasks and Funds
- E1.3 – COMBO6 Adapter
- E1.4 – Description of the Applications and Demonstrators

In addition, two activities focused on SCAMPI presentation took place. The first one was the *1st SCAMPI Workshop* held in January in Amsterdam. The second one was the *SCAMPI Monitoring and Measurement BoF* where a summary of papers related to SCAMPI goals was presented. It was held as a part of the TNC-2003 conference in Zagreb in May 2003.

We are co-authors of most SCAMPI deliverables and main authors of the D3.2 and E1.3 deliverables.

14.3 The COMBO6 monitoring adapter

The COMBO6 card is based on FPGA and therefore it is a flexible device which can be adapted for various applications. High-speed monitoring, as considered in the SCAMPI project, requires an additional daughter card with Ethernet link modules (1 Gbps and 10 Gbps) and a time unit for precise timestamp generation.

The internal structure of the monitoring adapter consists of three levels: hardware, firmware (i.e., the VHDL program defining the structure implemented in FPGA) and the Linux driver.

Hardware: The monitoring adapter is based on a COMBO6 card containing a Virtex II FPGA, SRAM and CAM memories for packet header filtering, DDRAM memory to store selected packets, PCI bus interface and a daughter card connector. The clock unit contains a precise TCXO (Temperature compensated quartz oscillator) and a synchronising circuit using a PPS (Pulse per Second) signal from an external source, (e.g., from a GPS receiver). The clock provides timestamps with the resolution in order of 10 ns, which can generate a unique timestamp for each received packet at a 10 Gbps speed. The absolute accuracy of the clock depends on the synchronisation method; it is better then 5 μs if a PPS signal is available and about 1 ms when the NTP protocol is used.

The VHDL program: The VHDL program is the core of the adapter. It implements individual functional blocks in the FPGA structure. Several blocks have been designed as machines called nanoprocessors whose characteristic feature is a limited instruction set; their complexity lies between a finite state machine and a RISC processor. The nanoprogram is interpreted by a firmware block and is stored in SRAM.

Driver: Linux was chosen as the platform for the SCAMPI software. The driver must allow communication between the adapter and higher layers of the SCAMPI architecture. It includes download of the VHDL program, adapter configuration and especially an optimised data transfer from the adapter through the PCI bus. The adapter RAM module is mapped to the user address space to improve the transfer speed.

The development of the adapter was split into two phases:

Phase I: A current version of the COMBO6 and daughter cards with 1 Gbps modules is used in Phase I. An IPv4 packet header filtering unit providing a basic monitoring function is implemented in firmware. The timestamp generation is provided by a separate PCI card.

Phase II: An improved version of the COMBO6 card with higher throughput and a 64-bit PCI bus will be designed and developed. New daughter card with 10 Gbps link modules must be designed, too. Firmware will contain units for application support.

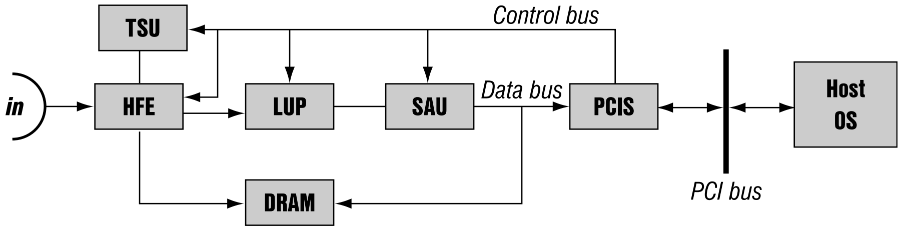


Figure 14.1: Adapter structure, Phase I

Structure of adapter functional blocks in both phases is shown in figures:

HFE (Header Field Extractor): This block preprocesses packet headers and transforms them to a unified form.

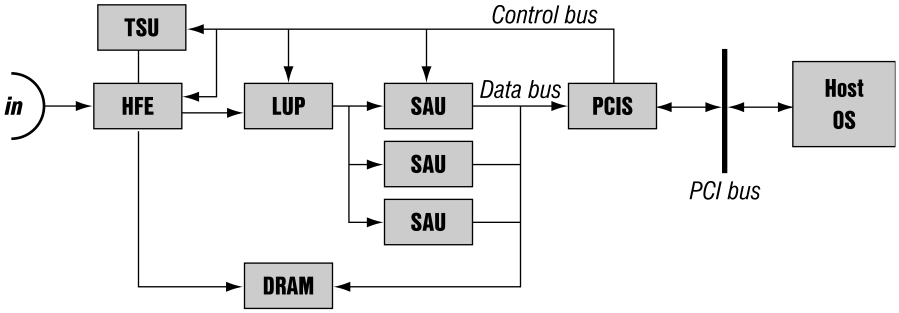


Figure 14.2: Adapter structure, Phase II

LUP (Look Up Processor): This block is designed as a nanoprocessor; its task is filtering according to the packet header. A 272-bit wide CAM is used for pattern matching and a unit for arithmetical comparison.

TSU (Timestamp Unit): This block generates timestamps with a 10 ns resolution. Optionally, clock may be synchronised using a PPS (Pulse per Second) signal, e.g., from a GPS receiver.

DRAM (Dynamic RAM): Selected packets are stored in dynamic RAM mapped to the user memory space of the operating system; this allows direct access to data from any monitoring application.

SAU (Sampler Unit): This block implements both deterministic and probabilistic sampling of input data.

STU (Statistical Unit): The statistical unit consists of 256 counter sets for evaluating up to 256 traffic classes. Each set contains a packet counter, accumulators for sum of values and for sum of squared values, as well as registers for maximum and minimum values. The value may represent a packet length or a time distance between two subsequent packets.

PCK (Payload Checker): This block implements filtering based on packet payload content. It can find up to 500 strings, each up to 16 bytes long.

14.4 Planned Activities for 2004

2004 is the last year of the SCAMPI project. Our task will consist mainly of participating in new COMBO6 card design (e.g., the structure of functional blocks, Linux driver, MAPI library). Other important task is organising and testing all layers of the SCAMPI architecture within the framework of WP3 (workpackage 3). CESNET is the leader of this workpackage, responsible for writing the D3.4 – Description of Experiment Results deliverable.

15 6NET

In 2002 CESNET became a partner in the *6NET* project (IST-2001-32603), which is a part of the EU 5th Framework Programme. Our contribution is mainly hardware and software development of an IPv6 PC-based router (see Chapter 7) within the work group 3 *(Basic Network Services)*.

Apart from testing and evaluation of technologies, this work group also has ambitions and already a few interesting results in the area of IPv6 multicast. It is well-known that IPv6 multicast can so far only work within a single administrative domain, where all routers can learn the address of the rendezvous points (RP). IPv4 uses MSDP (Multicast Source Discovery Protocol), which is not likely – and for good reasons! – to be standardised for IPv6 at all. A partial solution for inter-domain IPv6 multicast can be SSM (Source-specific Multicast). However, it is not suitable for all types of multicast sessions and also requires that routers and switches of last-hop networks support the protocols IGMPv3 (for IPv4) and MLDv2 (for IPv6). The WP3 group thus took the advantage of the redundancy in the IPv6 addresses and designed a mechanism that encodes the RP address directly in the IPv6 address of the respective multicast group. This mechanism is known as *embedded RP* and has already been submitted to IETF for standardisation.

Active development of protocols and mechanisms is often hampered by the inflexibility and closed character of commercial routers. CESNET in cooperation with the Norwegian NREN UNINETT (and with support of other members of the 6NET consortium) thus prepared a new activity for the year 2004, whose primary aim will be a high-performance open source platform for IPv6 multicast protocol development. This platform will be based on the COMBO6 card and a selected PIM-SM daemon.

CESNET also participates, although with considerable less capacity, in other 6NET work packages:

- WP 1 – Build and operate the IPv6 network
- WP 2 – IPv4-IPv6 coexistence, interworking and migration
- WP 7 – Dissemination and exploitation of results (here we contributed a CVS server)

A very distinguishing feature of the experimental 6NET network is the absence of IPv4. In February 2003 we connected our national IPv6 backbone *natively* to 6NET through an STM-1 circuit. The current 6NET topology is shown in Figure 15.1.

A more detailed information about the 6NET project is available from its home page[1].

[1] *http://www.6net.org/*

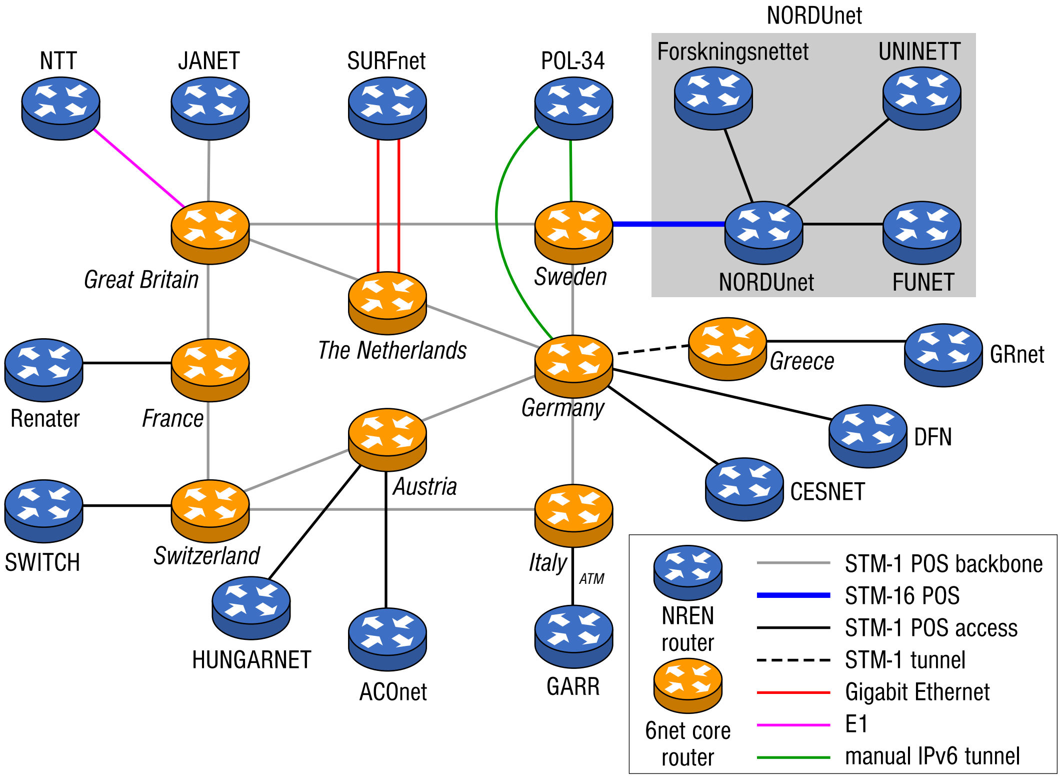


Figure 15.1: 6NET topology

Part IV

Other Projects

16 Infrastructure and Technologies for On-Line Education

The fundamental goal of this project is implementation of Education electronic support (eSupport) for full-time students, firstly at the Department of Telecommunication Engineering (Czech Technical University in Prague, Faculty of Electrical Engineering), with possible sharing with other departments within the CTU and at partner technical universities.

In 2003, our team resumed the activities and followed up with the results from 2002 when we had proposed and implemented a conception for support of modern education forms within a full-time university study; the pilot examples served as an inspiration for further steps and applications.

During the first half of 2003, with respect to the aims of the project, we had been finishing implementation of the environment (poly-functional WWW portal) for support of education, adding and optimising its functions. We have been continually and systematically supplementing educational materials of various types for substantial part of courses taught at our department. At present time, the system appears to be a comprehensive source of quality materials for education support.

The workplace for acquisition of video streams and recordings was supplemented with several necessary parts (namely, a mixer for editing analog tributary signals in real time, on-head and lapel wireless microphone system). This opened way for recording integral blocks of lectures (in fact, our ten recorded lectures almost completely cover one semester of *Transmission Systems* course) as well as for experimental live broadcast of the said lectures over the Internet – using the equipment (file server and streaming server) of the CESNET Association in both cases. Concurrently we had been testing the quality of video and audio outputs using various transmission speeds; we concluded that the three used alternatives should satisfy most of our users.

On the departmental level, intensive educational activities among teachers were performed in order to make them regularly use the system and contribute to its content, primarily with text documents and possibly, in the future, with recordings of their own lectures.

During the last semester week we carried out an extensive survey among the users of the system (both students and teachers); the results clearly showed that positive responses and constructive attitudes prevailed. We have made

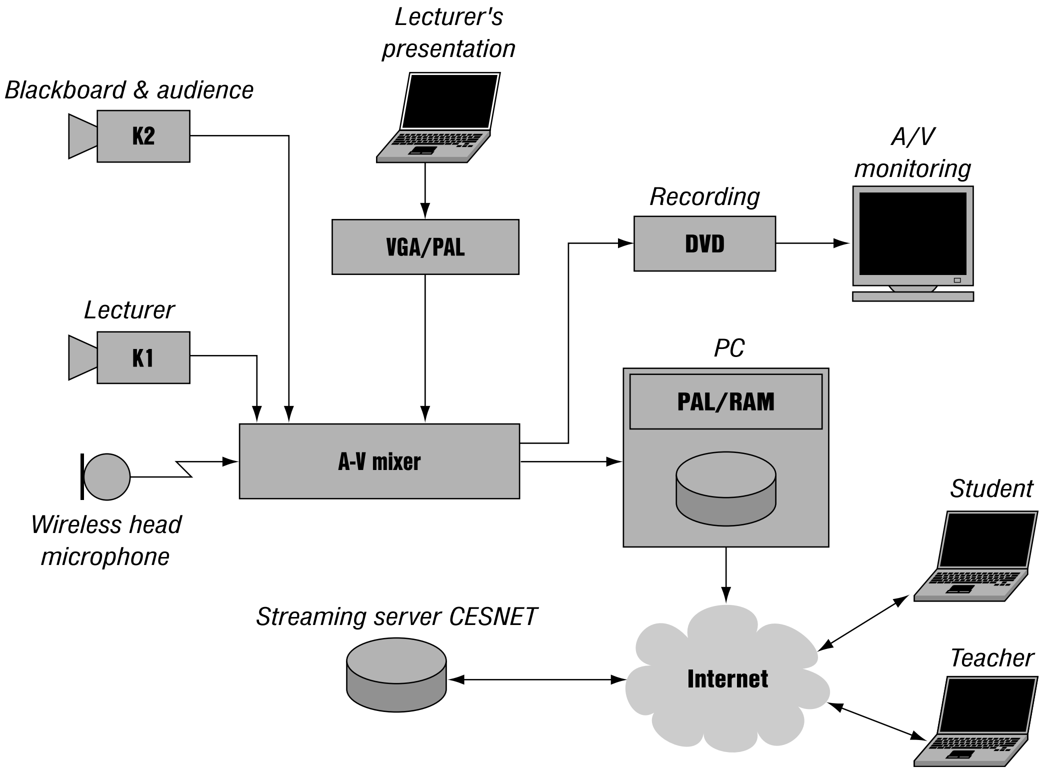


Figure 16.1: Arrangement of the Recording and Streaming Workplace

functional and logistic changes in the system implementation according to the users' remarks.

From the beginning of the winter semester 2003/2004 we have been regularly recording and broadcasting on-line a wider selection of lectures; to be specific, this concerns the courses *Transmission Systems II* and *Program Control of Switching Systems* as well as *Digital Filter Design* in cooperation with the Faculty of Transportation Sciences, and also regular lectures and seminars of the Research and Development Center (RDC). Thus, the project moved from a departmental level to a university level. We also abandoned the originally used RealVideo format and accepted the Windows Media format as a new standard, which the students welcomed. Our departmental WWW portal that provides access to the published materials for education support is now running steadily, being continuously optimised and upgraded.

The expansion of our activities and transition from experiments to the regular broadcast of lectures described above was possible after we had completed and optimised the mobile workplace. The quality of recordings substantially improved thanks to a DVD recorder use; photography and workplace servicing took an advantage of a new video camera and tripod with motor-driven

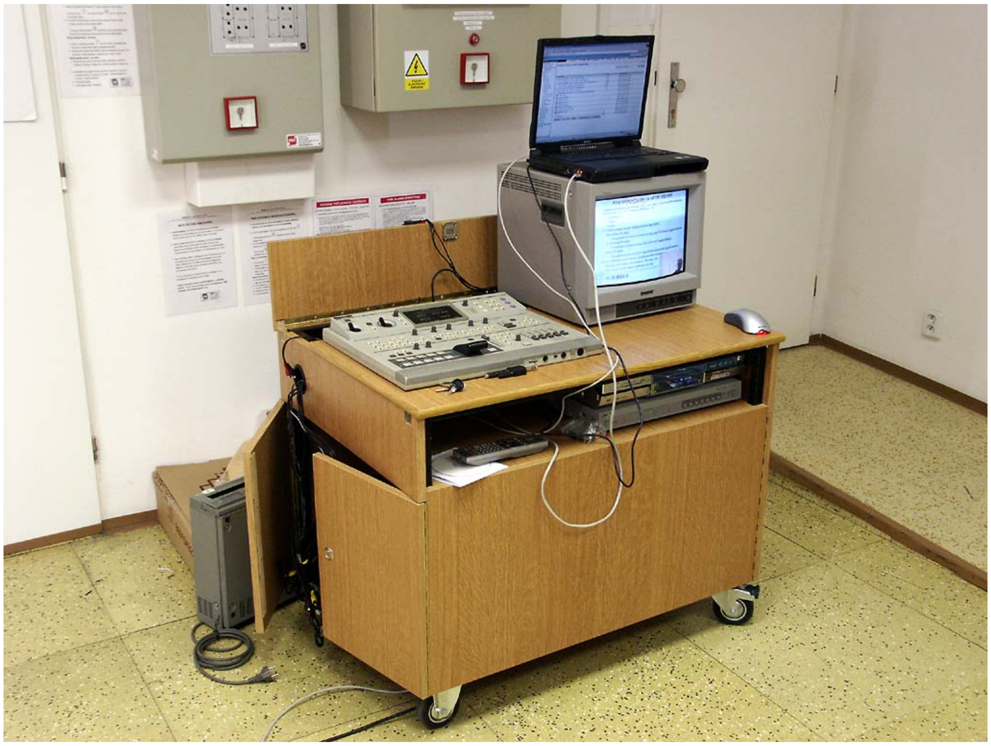


Figure 16.2: Mobile Subsystem Implementation

head; a notebook computer acting as an output streaming device improved the overall system mobility. We have also built a mobile table that integrates all necessary devices into one functional unit, thus eliminating the necessity of time-consuming repeated plugging and unplugging the devices for each lecture. The next experimental step should be a self-serviced operation of the entire system.

At the end of October we organized an experimental videoconferencing lecture (dedicated particularly to streaming technologies) in which our colleagues (teachers and students) from the partner department at the Slovak University of Technology in Bratislava took part remotely. Thus, the project entered an international dimension. The connection used three ISDN BRI lines (i.e. six B-channels). The lecture was also recorded on a DVD. Furthermore, we discussed videoconferencing issues, including IP-based connection and quality of on-line streamed video, within a videoconferencing session with a partner department at the Technical University of Ostrava.

Currently we prepare routine collaboration in education using shared lectures in selected courses together with technical universities in Bratislava and Ostrava.

Another successful experiment of this type was an evaluation of a UMTS Internet on-line streamed lecture quality and of its recording quality. This lecture was broadcast from the Alcatel SEL headquarters in Stuttgart, Germany. Feedback from participating experts was very positive in all cases. We also performed experiments with dial-up connection using standard modems for those students who do not have broadband Internet access. We confirmed that the transmission of sound from streamed lectures was good enough – keeping in mind that students usually have access to static electronic presentations (e.g., PDF files) published on the WWW portal, this is a satisfactory result.

Emphasis was put on publishing and dissemination of our experience. Partial results of the project were introduced at the following conferences: *COFAX–Telekomunikácie 2003* (Bratislava, Slovakia), *eLearning ve vysokoškolském vzdělávání 2003* (Zlín, Czech Rep.), *EAEEIE* (Gdańsk, Poland), *ICEE* (Valencia, Spain), *RTT* (Bratislava and Častá-Píla, Slovakia), *ICETA* (Košice, Slovakia) and *VIEWDET* (Vienna, Austria). At all these forums the project was presented and new contacts and impulses for our work were found.

Sharing the experience at least within the Czech Technical University is regarded as the most important; other priorities include an appointment of a common strategy and distinct proclamation of support for modern eLearning technologies (especially eSupport), expanding the capabilities of the Centre for Educational Support, as well as determining the roles of the Computing Centre and CESNET Association in this process. Coordinating the project activities with other universities would be advisable; this could result in higher effectiveness by sharing the experience and preparing eSupport materials and in common strategy in applying for European projects with respect to the relevant EU priorities.

17 Distributed Contact Centre

Progressive evolution of voice services in the VoIP area led in the previous years to the idea of implementing IP telephony in the CESNET2 network and organising the necessary support for its users. The particular aim of this project was to pilot a new voice service – a contact centre. As a technology for this pilot test we chose the solution offered by Cisco Systems, which is a part of the modular product named Cisco Architecture for Voice, Video and Integrated Data (AVVID). We decided to deploy the contact centre in a fully redundant configuration consisting of at least two nodes. The project prepared the scene for a broad utilisation of IP-only voice services in the CESNET2 network. We are now able to support each CESNET member who wants to implement an infrastructure for such voice service.

17.1 Characteristics of the System

IP Contact Centre (IPCC) is a modular set of products, which are used either stand-alone or as an integrated system providing all necessary functions of a contact centre based on specific user requirements. It is also possible to extend the system by including new functions like output modification, monitoring and archiving call details, charging and CRM. The administrator of contact centre can effectively manage the entire system and modify promptly its configuration according to the actual needs.

The basis for the IPCC is the *Intelligent Contact Management Server (ICM)*, the control and monitoring component of the contact centre. This product has the following functions:

- monitoring and control of peripherals (ACD, IVR and other communications channels)
- human resources monitoring, i.e., management of operators and supervisors, which ensures an efficient distribution of incoming calls among them.
- definition of conditions that express relationships between the consumer and service provider.
- consolidation of data about the operation of whole contact centre.

ICM system allows to build large-scale solutions with the possibility of using different peripherals. The peripherals are controlled through the Peripheral Gateways (PG), which communicate with the ICM core through a protocol based on the CSTA standard and with each peripheral through a proprietary protocol. If necessary, interfaces for new peripherals or communication channels can be easily implemented. Another integral parts of the IPCC are a software phone

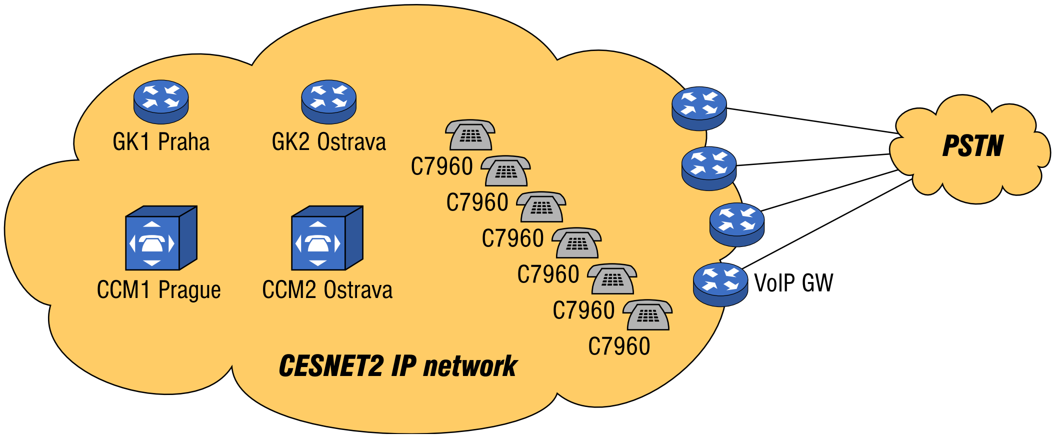


Figure 17.1: Logical scheme of call centre structure

exchange *Cisco CallManager (CCM)* and a system for automatic communication with the customer – *Interactive Voice Response server (IP IVR)*.

The key benefit of contact centres based on IP and VoIP lies in their flexibility – their set-up is relatively simple and they can also easily be integrated in the operational environment of the application at hand. The specific advantages are:

- easy implementation of new customised operator workplaces.
- possibility of distributing one contact centre over more locations with centralised control, using just the existing network infrastructure.
- effective possibility of using home-based operators.
- new options of voice and data integration in operators' applications.
- the administration and development of a complete solution can be entrusted to the consumer.

ICM guarantees routing of incoming and outgoing contacts to operators who have the most relevant information with respect to customer and contact characteristics, such as:

- way of contact – by phone, e-mail etc.
- origin of contact – e.g., phone number of the caller, address of the e-mail sender, web page
- contact place – called phone numbers, URL web pages visited by the client, e-mail address of the receiver
- duration of contact
- information about the client (area of interests etc.) obtained from e.g., DTMF choice at the IVR system, text of an e-mail or data entered to web forms.
- information extracted from the existing customer databases that are available to the contact centre.

- number of calls being currently served by the operators.

Call routing rules are defined in a graphical user interface by means of block diagrams and then fine-tuned off-line by internal ICM components. Only afterwards they are activated for the operations in real time.

IP contact centre enables the concept of Computer Telephony Integration (CTI), i.e., the possibility of integrating telephony and contact centre management with existing external applications. To this end, the applications must be equipped with the CTI application interface. IPCC offers several solutions starting from ready-to-run applications through development environment to third-party solutions integrated by leading manufacturers of CRM systems.

Apart from the phone contacts, IPCC supports other communication channels like e-mail or interactive communication via web pages. In the latter case, the Peripheral Gateway (PG) assigns the contact an appropriate application.

Cisco E-mail Manager (CeM) is a module capable of handling large volumes of e-mails. It is implemented as an e-mail client communicating with mail servers through SMTP, POP3 or IMAP4 protocols. CeM analyses incoming e-mails and classifies them according to set of predefined rules. The operators than process the categorised e-mails in a web application. For example, they can answer an e-mail using one of standard templates, access the archive of previous e-mails, forward the e-mail to other operators for further processing (including internal comments) and so on.

Cisco Collaboration Server (CCS) supports the interactive communication between contacts and operators by providing a guided web browsing. The CCS can also use common web technologies for complementing the phone communication by visual presentation means. The communicating parties can thus exchange or share information in many different forms:

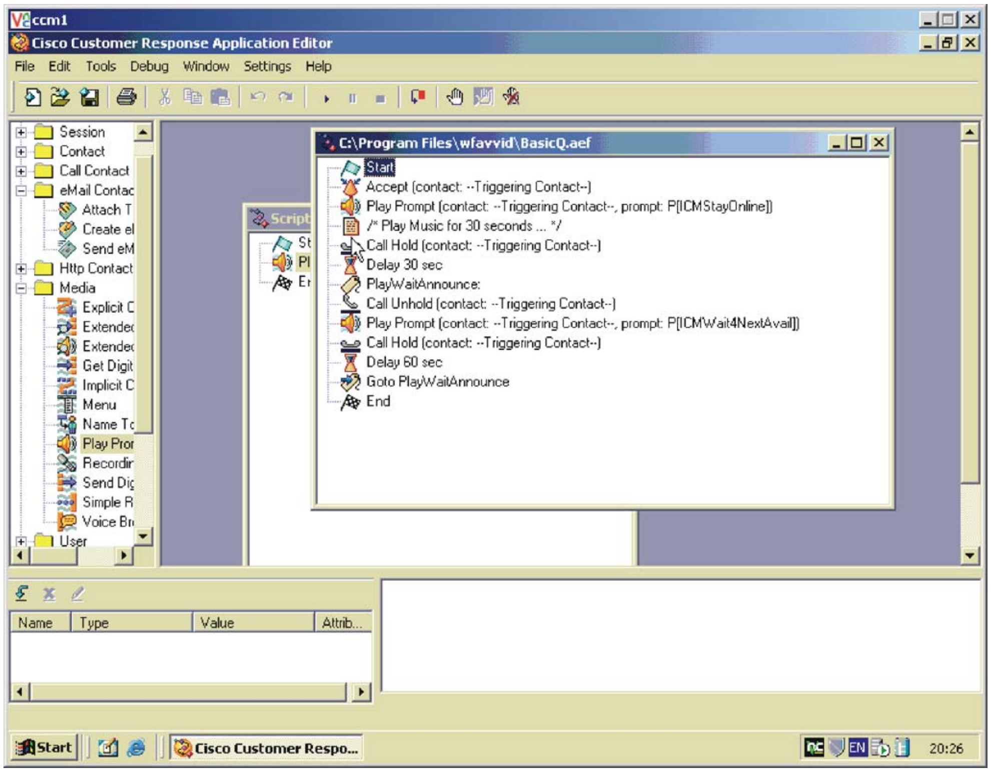
- cross-sharing of web pages – customers can send selected pages or automatically follow the navigation through web pages by the partner.
- cross-sharing of forms – e.g., the operator can help to the customer with filling in difficult parts of an order and the customer can just supply his or her personal data.
- text-based conversation (chat).
- virtual table for drawing diagrams and writing text.
- sharing the window of an arbitrary application with other parties, including the possibility of keyboard and mouse sharing.

The CCS can also be coupled with the traditional voice operation of the contact centre. For example, the client can use a simple web form to request a voice call-back from the contact centre. The IP contact centre then mediates a phone call between an operator and the customer. Both participants can then communicate by voice as well as by web page sharing.

17.2 Changes in system configuration

The contact centre is configured as a full redundant system of two branch points. One branch point is in Prague (CESNET) and the other in Ostrava (VSB-TUO). The Prague CCM is set up as a publisher and the one in Ostrava as a subscriber (back-up server). During 2003 we upgraded the the whole contact centre to new versions – CCM 3.3 and ICM 5.0. While doing the upgrade, we again have to cope with the complexity of system configuration, where software components from Cisco Systems must cooperate properly with the underlying Microsoft Windows 2000 operating system. We had to lead extensive consultations with the vendor of the IPCC in order to resolve these difficulties.

In the second half of 2003, the second PG within the ICM component was put into operation. This PG supports the CEM and CCS applications. Under this setup we tested the co-operation with voice recording service of *Kerio Voice Mail*. Everything worked without any complications and the voicemail system was able to access all mailboxes. However, during this test we found out that the Gatekeeper control component (GK) must be configured in the GRC mode. Currently only *gk-ext.cesnet.cz* can be configured this way.



```
Start
Accept (contact: --Triggering Contact--)
Play Prompt (contact: --Triggering Contact--, prompt: P[ICMStayOnline])
/* Play Music for 30 seconds ... */
Call Hold (contact: --Triggering Contact--)
Delay 30 sec
PlayWaitAnnounce:
Call Unhold (contact: --Triggering Contact--)
Play Prompt (contact: --Triggering Contact--, prompt: P[ICMWait4NextAvail])
Call Hold (contact: --Triggering Contact--)
Delay 60 sec
Goto PlayWaitAnnounce
End
```

Figure 17.2: Simple script editing in IPCC

During 2003 we also tested different types of phone sets and and their compatibility with CCM. These tests resulted in the purchase of IP phones *Cisco 7940*. These phones support both the Skinny Station protocol and the SIP protocol. The current operating version of the CCM does not support the SIP protocol. Consequently, we intend to launch the new version 4.0. during next year. The new version will support all VoIP protocols we need for both operational and testing purposes.

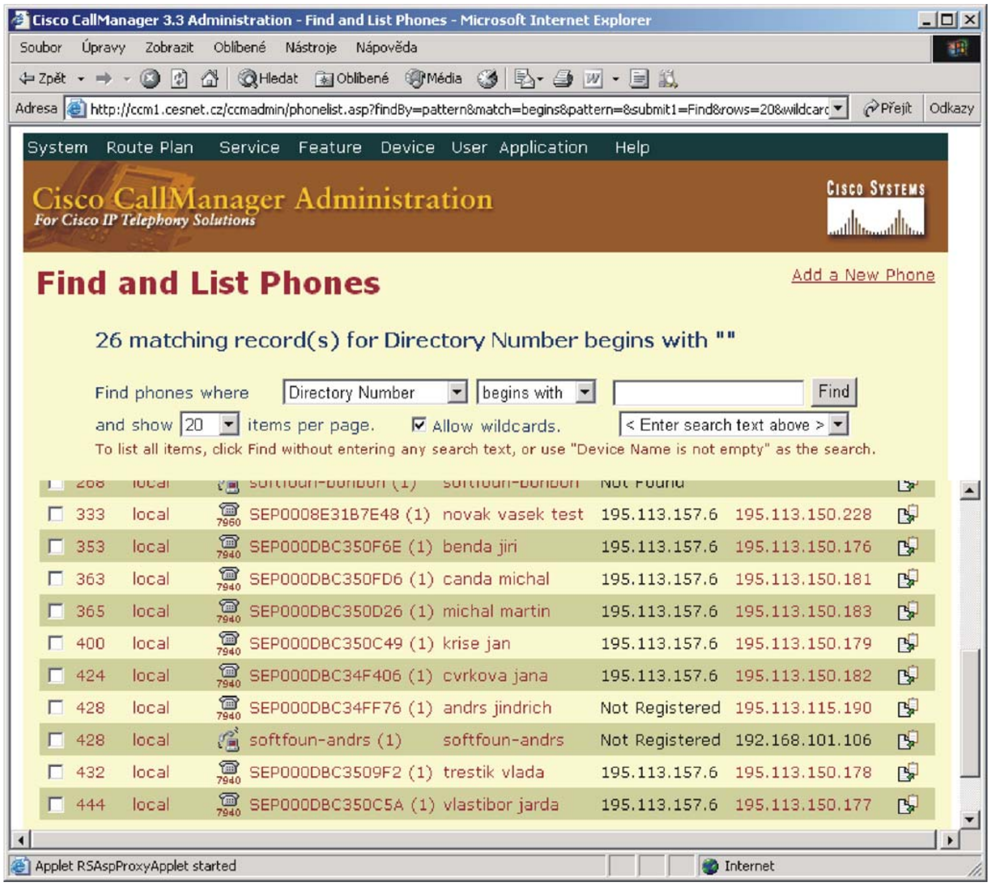
We also bought *Welltech LanPhone 101* phones for H.323 testing. These phones can communicate from a private network through a NAT device. Another useful feature offered by some IP phones, for example *Cisco 7920*, is the WiFi support. We tested their compatibility with the whole CCM system. In all tested cases the IP phones worked flawlessly within the whole system and it is thus possible to use them in a production environment. As a consequence of further developments in IP telephony we expect to see a number of new manufacturers of phone's components in the near future. As a result, more compatibility tests may be necessary for ensuring that the new phone set can be integrated into our system and our IP network. To enable better mobility of our customers, we also test the access to the CCM using software applications like *SoftPhone*.

All phones we purchased were distributed to the operating staff in PoPs of the CESNET2 network in order to help reduce the cost of network administration.

For calling through CCM, we acquired a public prefix from the GTS operator. The format of the prefix is 234 680 xxx. This step enables us to create conditions for opening the CESNET contact centre to the public telephone network. This step also rectified the previous interim solution, which utilised internal phone numbers of VSB-TUO for the contact centre. The assigned number space is divided in accord with the rest of voice service in the CESNET2 network as follows:

- 234 680 1xx – for contact centre needs, call distribution and information services
- 234 680 [2–4]xx – for connecting Call Manager phone sets
- 234 680 5xx – available for CESNET videoconferences
- 234 680 [8–9]xx – SIP testing

This allocation is already in use and all phones registered to each CCM were assigned a branch line following the rules. The same holds for lines that are being tested for the needs of the contact centre and individual information services. The whole system of redundant CCM, which has undergone a detailed testing, is now ready to accommodate CCM systems of CESNET members. In the first half of 2004 we intend to start a pilot project investigating a setup in which the central CCM system will be the head of subordinated CCM systems operated by selected CESNET members. In the context of this project we intend to submit an application to the Czech Telecommunication Office in order to receive an access code into the CESNET2 network.


Cisco CallManager 3.3 Administration - Find and List Phones - Microsoft Internet Explorer
Soubor Úpravy Zobrazit Oblíbené Nástroje Nápověda
Zpět Hledat Oblíbené Média
Adresa
Přejít
Odkazy
System Route Plan Service Feature Device User Application Help
Cisco CallManager Administration
For Cisco IP Telephony Solutions
Cisco Systems
Add a New Phone
Find
Applet RSAspProxyApplet started
Internet

Adresa http://ccm1.cesnet.cz/ccmadmin/phonelist.asp?findBy=pattern&match=begins&pattern=&submit1=Find&rows=20&wildcarc

Find and List Phones

26 matching record(s) for Directory Number begins with ""

Find phones where Directory Number begins with

and show 20 items per page. ☑ Allow wildcards. < Enter search text above >

To list all items, click Find without entering any search text, or use "Device Name is not empty" as the search.

□	268	local		softfoun-bonbon (1)	softfoun-bonbon	Not Found		
□	333	local	☎ 7960	SEP0008E31B7E48 (1)	novak vasek test	195.113.157.6	195.113.150.228	
□	353	local	☎ 7940	SEP000DBC350F6E (1)	benda jiri	195.113.157.6	195.113.150.176	
□	363	local	☎ 7940	SEP000DBC350FD6 (1)	canda michal	195.113.157.6	195.113.150.181	
□	365	local	☎ 7940	SEP000DBC350D26 (1)	michal martin	195.113.157.6	195.113.150.183	
□	400	local	☎ 7940	SEP000DBC350C49 (1)	krise jan	195.113.157.6	195.113.150.179	
□	424	local	☎ 7940	SEP000DBC34F406 (1)	cvrkova jana	195.113.157.6	195.113.150.182	
□	428	local	☎ 7940	SEP000DBC34FF76 (1)	andrs jindrich	Not Registered	195.113.115.190	
□	428	local		softfoun-andrs (1)	softfoun-andrs	Not Registered	192.168.101.106	
□	432	local	☎ 7940	SEP000DBC3509F2 (1)	trestik vlada	195.113.157.6	195.113.150.178	
□	444	local	☎ 7940	SEP000DBC350C5A (1)	vlastibor jarda	195.113.157.6	195.113.150.177	

Figure 17.3: List of phones registered in CCM

During 2003 we managed to configure and launch new services for the users of IP phones, e.g., caller identification that provides the receiving side with auxiliary information about the caller or the phone book of Czech Telecom with flexible search capabilities.

17.3 Future aims of the project

Several tests of the system are now underway, where we adjust the routing scripts according to the suggestions of our users. The outcome should be a set of optimised scripts tailored to real operational conditions of the CESNET2 network.

We also prepare rules for service provisioning with respect to connecting terminal components (either separately or through external CCMs). This will require extensive tests of the CCM configuration under conditions that may potentially occur in real operation but cannot be simulated in advance.

18 Intelligent NetFlow analyzer

The outcome of this project is the *NetFlow Monitor* program. It is a modular distributed system for analysis of network traffic based on NetFlow statistics. NetFlow Monitor is able to analyze the traffic almost in real time, including intelligent filtering, aggregation and statistical evaluation of the data. It also provides a number of options for selections based on multiple criteria (e.g., source/destination IP address, protocol, port, etc.) with the granularity down to individual flows.

The entire system consists of three blocks:

- executive part – NetFlow Collector,
- user interface – NetFlow Monitor,
- warning subsystem – NetFlow Event.

18.1 NetFlow Collector

The first part is written in the C programming language and performs the processing of received data. The system currently supports the export formats of NetFlow versions 1, 5, 6 and 7. Other functionality of the NetFlow Collector is available in the form of simple modules.

One module is the NetFlow Forwarder, which is designed for redirection of the data flow to other destination(s) specified by IP addresses or ports. Another module then allows filtering of incoming data according to access lists (ACL) – we can thus define the sources from which NetFlow exports are accepted.

An important module takes care of storing the received and evaluated NetFlow exports into a database. This export module not only stores data from internal buffers into a *MySQL* database, but also aggregates certain information about data flows. All information about data flows is stored in such a manner that non-aggregated information is stored in hourly tables and from them daily, weekly, monthly and yearly tables with aggregated information are generated.

The number of tables is theoretically limited only by the available disk space, but an explicit limit can also be set in the configuration if necessary. For example, we can instruct the system to keep the hourly tables for the last 3 days, the daily tables with aggregated data for the last 14 days, the weekly tables for the last 6 weeks and the monthly tables for the last two years.

The combination of tables with aggregated and non-aggregated information gives the user a considerable flexibility since the same data can be viewed from many different perspectives. Typically, for a network administrator it is often important to have access to detailed information about network activities of the last hour at the level of particular data flows.

18.2 The NetFlow Monitor

The second part of the system for NetFlow data processing is written in the PHP programming language. Its purpose is to present measured values in a user-friendly form, produce various graphs and statistics and, last but not least, allow an easy configuration of the *NetFlow Collector* through a web interface.

In the first half of the year 2003 we rewrote the system interface starting from the original core that was created in the previous year. The advantage of the *NetFlow Monitor* is the simplicity of the web interface, which allows a comfortable manipulation with measured statistics. The new web interface contains six menu items: Main, Profiles, Archive, Statistics, Options and Help.

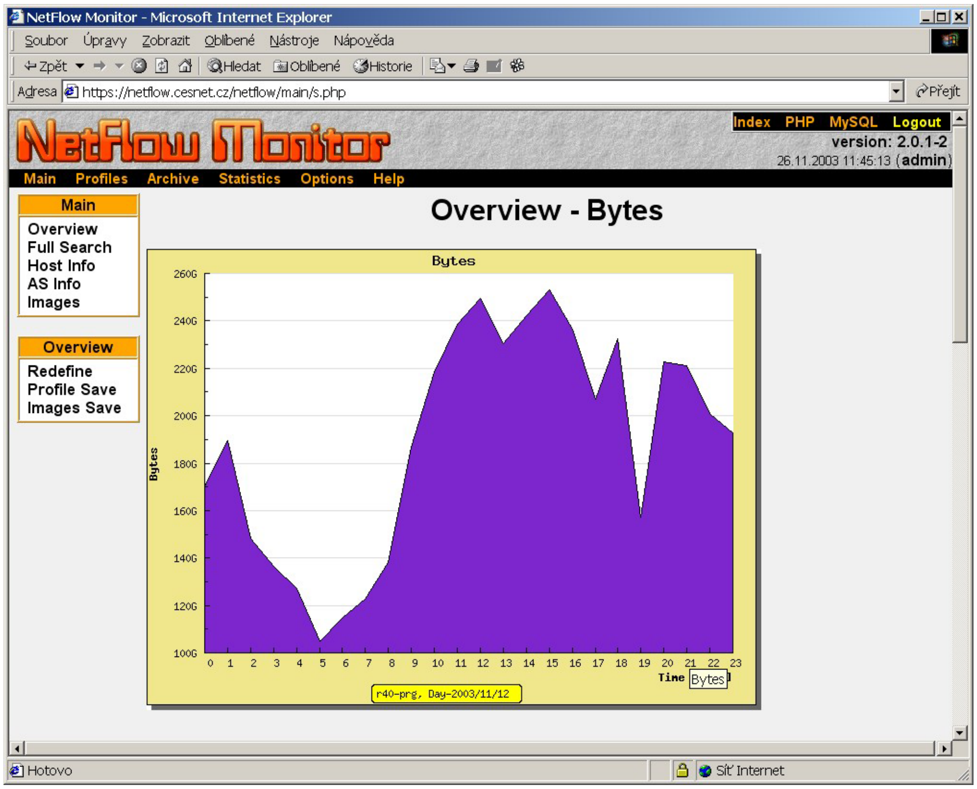


Figure 18.1: Statistics of transferred data volume example

The *Main* menu contains two basic search functions (Overview and Full Search), the next two menu items provide information about both communicating sides and their autonomous systems, respectively, and the last item shows a list of generated graphs (which can be exported or printed at user's discretion). The *Profiles* menu contains user-defined queries. The *Archive* menu enables the backup of individual tables or the system configuration. The *Statistics* menu provides information about the status of particular processes, size and state of individual tables, state of the database etc. The *Help* menu contains brief information about system installation, the distribution version, license conditions etc.

Since the NetFlow system has been designed for a distributed architecture, it is necessary to set up a *NetFlow Unit* item for each participating *NetFlow Collector*. The *NetFlow Unit* means just a single server – more *NetFlow Collectors* can run under one *NetFlow Unit*. When setting up a *NetFlow Collector*, the user has to specify a port on which the NetFlow exports from routers are awaited and the number of tables that are to be stored.

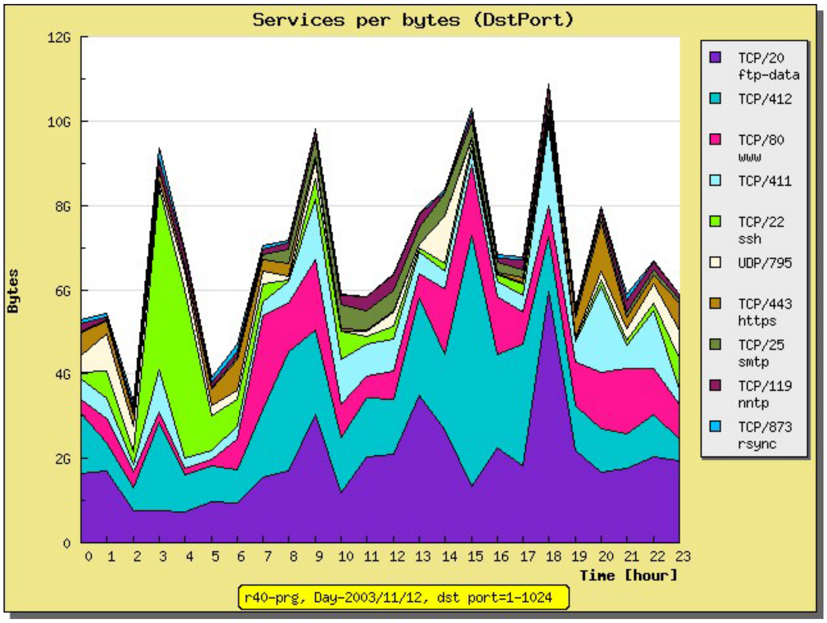
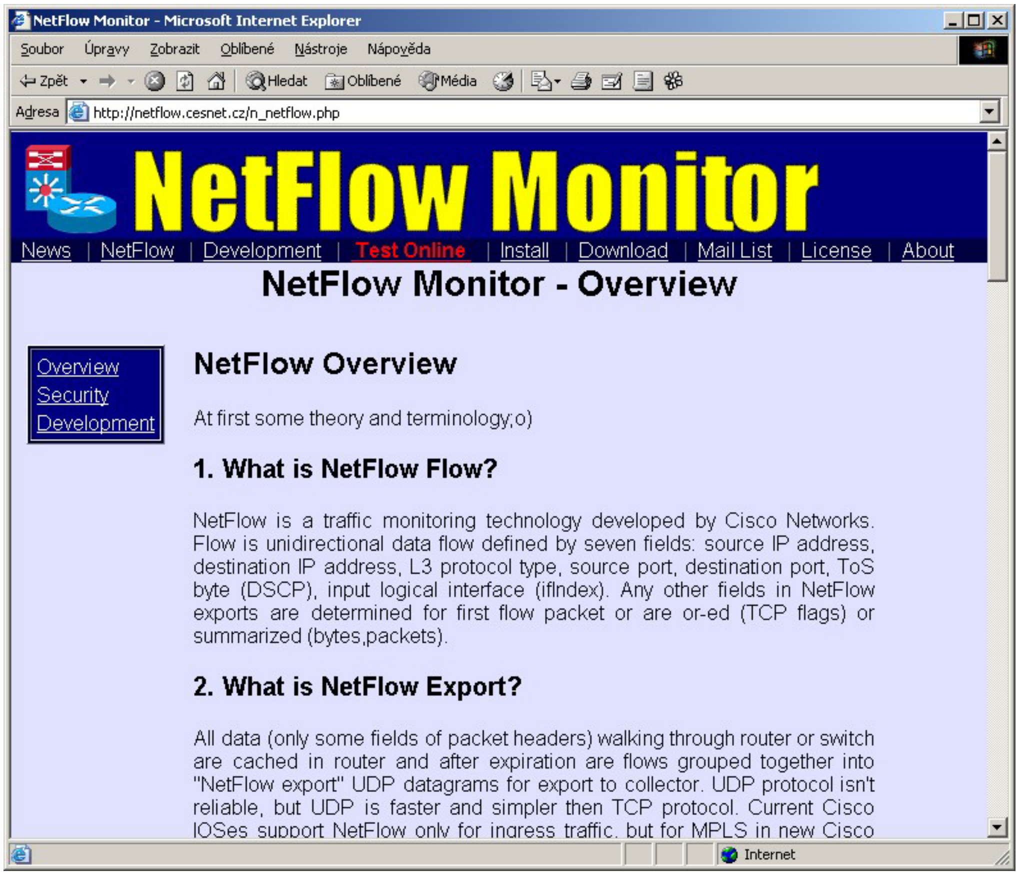


Figure 18.2: Statistics of the distribution of the used L4 ports (applications) example

The *NetFlow Monitor* supports 18 different statistics such as the summaries of transferred bytes, 10 most active hosts, distribution of protocols and L4 ports. All statistics can be defined for the whole network, a subnetwork, a selected host etc.

18.3 The NetFlow portal

During the year 2003 we finished the new portal, which provides news about *NetFlow Monitor* development. At present the portal offers the visitors a live test of the software (currently two OSR7600 routers with international connectivity and one GSR12000 backbone router are monitored). After completing the registration form, the entire distribution can be also downloaded directly from the portal. The URL of the portal is *netflow.cesnet.cz*. In order to establish a wider user base, two system versions are offered, one for the recommended operating system GNU/Debian Linux and another for other Linux distributions.


NetFlow Monitor - Microsoft Internet Explorer
Soubor Úpravy Zobrazit Oblíbené Nástroje Nápověda
Zpět Hledat Oblíbené Média
Adresa http://netflow.cesnet.cz/n_netflow.php
NetFlow Monitor
News | NetFlow | Development | Test Online | Install | Download | Mail List | License | About
Overview
Security
Development
Internet

NetFlow Monitor - Overview

NetFlow Overview

At first some theory and terminology;o)

1. What is NetFlow Flow?

NetFlow is a traffic monitoring technology developed by Cisco Networks. Flow is unidirectional data flow defined by seven fields: source IP address, destination IP address, L3 protocol type, source port, destination port, ToS byte (DSCP), input logical interface (ifIndex). Any other fields in NetFlow exports are determined for first flow packet or are or-ed (TCP flags) or summarized (bytes,packets).

2. What is NetFlow Export?

All data (only some fields of packet headers) walking through router or switch are cached in router and after expiration are flows grouped together into "NetFlow export" UDP datagrams for export to collector. UDP protocol isn't reliable, but UDP is faster and simpler then TCP protocol. Current Cisco IOSes support NetFlow only for ingress traffic, but for MPLS in new Cisco

Figure 18.3: The new portal *netflow.cesnet.cz*

18.4 Conclusion

During the year 2003 we designed and implemented a usable production version of a network monitoring system – the *NetFlow Monitor*. It is now being used by more than 1000 organisations from more than 60 countries all over the world. The most frequent users are Internet service providers (ISP), universities and telecommunication companies. The system is distributed under the GNU General Public License (GPL).

19 Storage over IP (HyperSCSI)

19.1 Introduction

The HyperSCSI protocol (see [HSC]) is a network protocol whose goal is transporting SCSI commands and data over a network. It was developed in year 2000 at the Data Storage Institute (DSI) as an alternative to the iSCSI protocol (see [SSC03]), which again as an alternative to the Fiber Channel did not work well on Ethernet networks. It evolved from a research of possible SCSI encapsulation into Ethernet frames.

Similarly to iSCSI, the HyperSCSI (hereinafter ”HSCSI”) mainly tries to offer a cheaper variant for building Storage Area Networks (SANs) compared with the Fiber Channel (hereinafter ”FC”). Compared with the iSCSI, HSCSI tries to deliver higher performance, which in both cases is worse than that using the FC technology.

Goals of the project were

- verifying the HSCSI protocol usability in the Linux operating system (both as a client and as a server) and Microsoft Windows OS (client only)
- testing the data throughput using the HSCSI protocol
- testing the HSCSI protocol usability with transport security mechanisms
- attempt to directly compare the HSCSI and iSCSI protocols.

Experience gained during testing within the project was described in the Technical report 23/2003.

19.2 HSCSI features

The protocol itself is designed for two variants of network transport. The first one actually implemented and described here is called HS/Eth and is based on data transport directly over the Ethernet. The authors claim that this variant works also on 802.11b networks. The second variant is meant to transport data over the IP (called HS/IP). Unfortunately the HS/IP variant is not yet available and it is not known when it will be.

Only a software implementation for operating systems Linux and MS Windows 2000 is currently available. The client–server approach is used. The project is Open Source but only for the Linux platform. Both source codes and binary packages of server and client are available for Linux. For the MS Windows platform, only a binary trial version of the client is available, without any source codes.

The HSCSI protocol also deals with data transport security, using the following mechanisms:

Unauthorized access to data protection: The client has to authenticate itself to the server using the correct login/password combination. Unique combination can be assigned for each exported volume.

Securing the data integrity: Only the SHA1 HMAC algorithm is currently implemented but many others can be used in future (when implemented). The architecture is modular, so other mechanisms can be easily added.

Encrypting transported data: Only the AES (Rijndael) algorithm is currently implemented. The 128-bit variant is used. Just as with data integrity, plans for using other algorithms exist, benefitting from the modular architecture.

19.3 Comparison of the HSCSI and iSCSI protocols

As the only available implementation of the HSCSI protocol is the Ethernet variant, the main disadvantage is the impossibility of routing HSCSI and the resulting impossibility of wider deployment. Another important disadvantage of HSCSI is the fact that this protocol is no official standard (unlike, e.g., iSCSI) and so it is unsupported in any way by manufacturers developing hardware data storage solutions. No manufacturer even plans to support and/or implement this protocol currently.

The main advantage of HSCSI compared to iSCSI is especially a lower network load as well as an end system (server and client) load. It results from both the protocol design and implementation and from the fact that for data transport, iSCSI uses the TCP/IP protocol with a bigger overhead. Another advantage when compared to iSCSI is a fully functional software implementation of both client and server (unlike iSCSI with server implementation problems); as a result, HSCSI can be used for a solution built on commonly available hardware, no expensive and specialized hardware is needed. Therefore, HSCSI can be used for building small and cheap SANs.

19.4 Practical experience

Testing general functionality of client and server: We have tested both single- and multi-client (server) variants. Tested versions using the Intel/Linux platform were functional but only with kernel versions 2.4.18

and higher. Older versions of HSCSI support even older kernel versions. We tested only the Ethernet variant because of the unavailability of the IP variant implementation. The client was connected directly to the server (and over a gigabit Ethernet for the performance test) or through a switch (for multi-client tests). Both the multi-client/single server approach as well as a setup where a single client connected to devices on multiple servers were functional.

Functionality tests over the 802.11b wireless networks: Though the tested versions allowed the client to access the remote device over the wireless network, all attempts to transport a contiguous data block failed, producing errors that terminated not only the data transport prematurely but also disallowed the client to connect to the server any more.

Testing the security mechanisms (integrity and encryption): We examined the functionality of mechanisms for data integrity control as well as for data encryption during the transport using HSCSI. The implementation includes only one algorithm for integrity control (SHA1 MAC) as well as for encryption (128-bit AES). Though the authors warned that encrypting transported data may cause data corruption, we could not reproduce this. Unfortunately, data encryption increased the system load up to three times. Versions prior to 20030903 could not provide the data integrity control, probably because of incorrect implementation, although the authors claimed this feature to be fully functional. The data integrity control worked flawlessly in the last tested version (20030930) although no change was mentioned in the change log.

Functionality of access to exported SCSI hard drives: The HSCSI protocol worked flawlessly with SCSI hard drives.

Functionality of access to exported IDE drives (both hard and optical): IDE disk drives worked natively. The IDE optical drives had to be used as SCSI drives using the IDE-SCSI emulation layer on both client and server sides. This applied also to the read-only CD-ROM drives. Writing on CD-RW drives worked as well. This implementation included a bug that caused the server to send the contents of previous medium to the client after a media change.

Testing the MS Windows 2000 Professional client implementation: We tested the usability of the first client available for a non-Linux operating system. Compared to the Linux client, features of the MS client are reduced considerably. No data integrity control and no data encryption is implemented, sessions close incorrectly and only a single-server variant is available. The problems with incorrect session closing affected only the hard disk drives, not the CD-ROM drives. Generally it is clear that this was the first version which is not as advanced as the Linux one.

Testing the client functionality on a handheld device: Unfortunately we could not compile a kernel for a handheld device with HSCSI support and so we could not test it. Another problem was that the connection should have been using a wireless network which had turned out as inappropriate.

19.5 Performance tests

For measuring the disk operation performance, a benchmarking program *IOzone*[1] was used. We used the function *close()* and the results were written to a binary *MS Excel* file.

IOzone parameters:

```
iozone -Rb hscsi.wks -n 4g -g 4g -z -c -a -i 0 -i 1
```

The graphs show that while the data transport integrity control has no impact on performance (graph 19.1 and graph 19.2), the data transport encryption has a considerable negative impact (graph 19.1 and graph 19.3).

The fact that the RedHat optimised kernels deliver better performance than standard (vanilla) kernels was confirmed (graph 19.1 and graph 19.4).

19.6 Conclusion

Unfortunately, the iSCSI and HSCSI protocols can not be compared directly: while the iSCSI server is satisfactorily implemented only in a specialised hardware solution, the HSCSI server is available only as a software implementation. Nevertheless, the following conclusions can be made:

1. The HSCSI over Ethernet protocol in the Linux implementation seems to be a stable and usable solution for building local SANs. Its main disadvantage still remains the restriction to Ethernet and so limiting the overall reach of a concrete solution. In contrast to solutions based on the iSCSI or Fibre Channel, the HSCSI price can be more attractive. Compared to iSCSI, the HSCSI offers a better performance with smaller system load. It seems that it could be a useful complement to smaller Linux-based computer clusters.

[1] *http://www.iozone.org/*

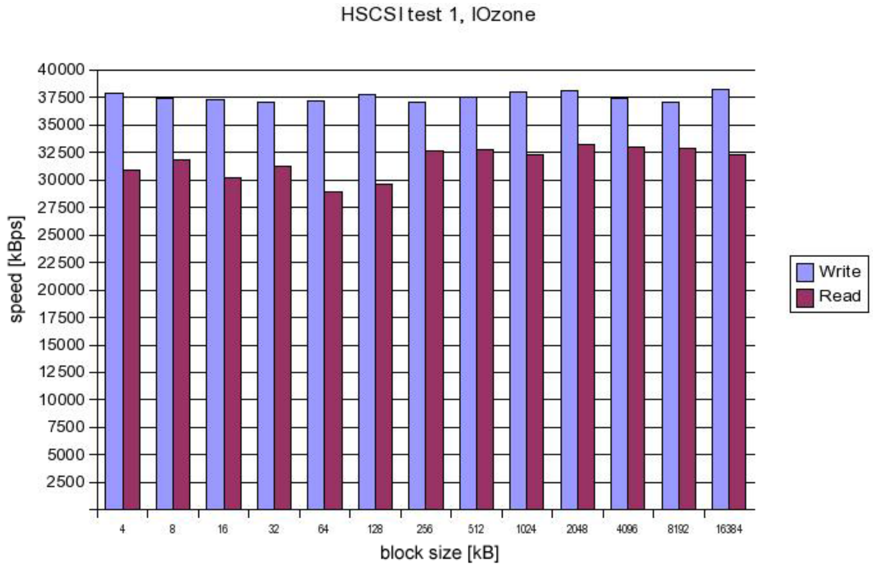


Figure 19.1: Results of the HSCSI benchmark

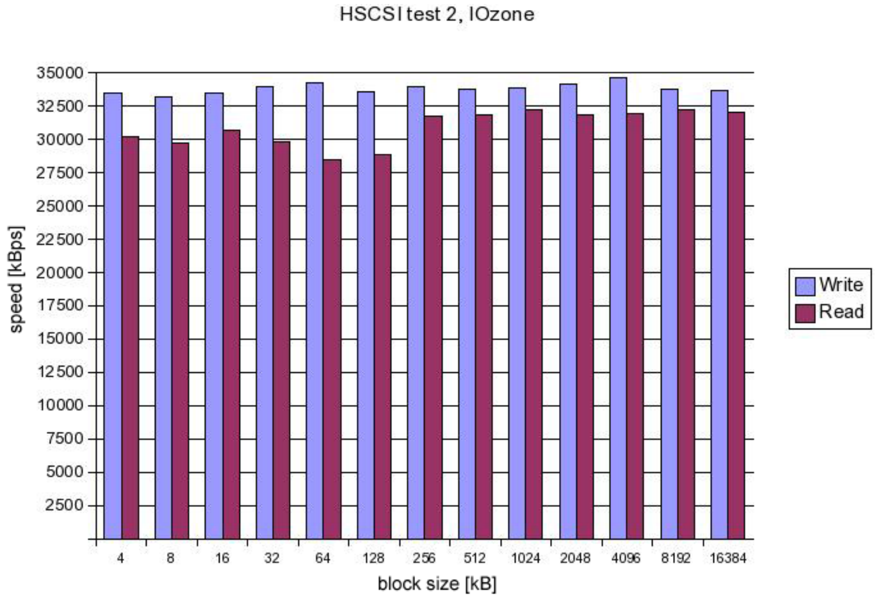


Figure 19.2: Results of the HSCSI benchmark using data integrity control

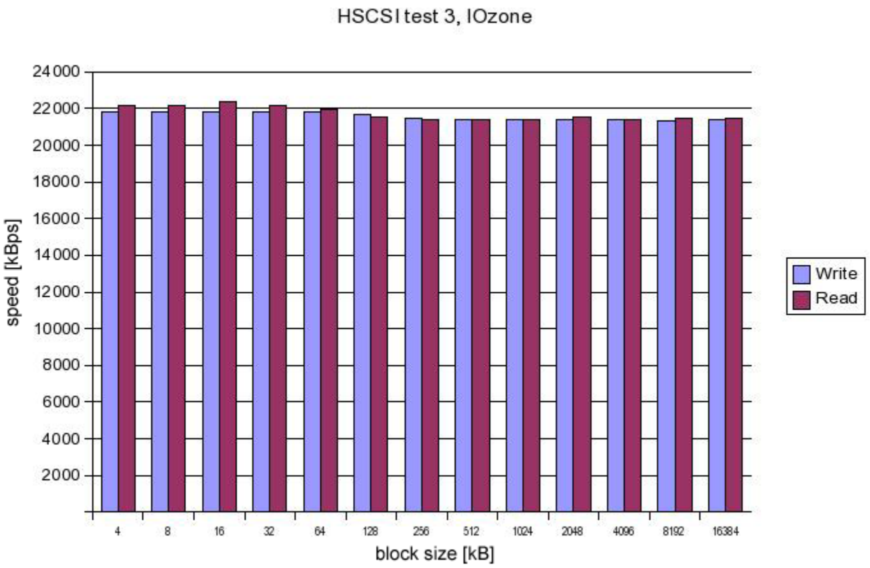


Figure 19.3: Results of the HSCSI benchmark using data encryption (128-bit AES)

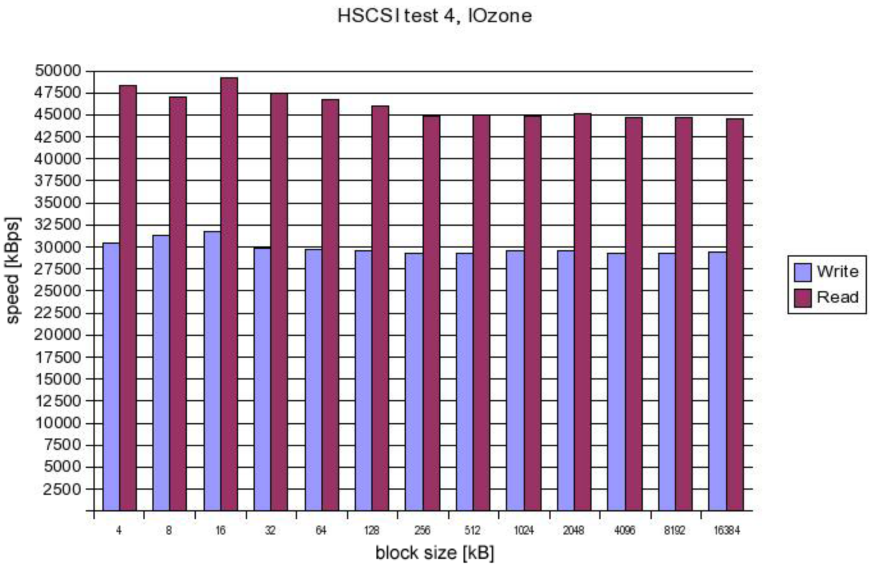


Figure 19.4: Results of the HSCSI benchmark using a RedHat optimised kernel

2. The data transport security is very well designed and implemented but its system performance requirements are considerably higher (at least for the currently implemented AES algorithm variant). The data transport encryption will become more important with the IP variant.

The project ended with these conclusions and no further work is planned.

20 Presentation

The task of the *Presentation* project is to publish information about the activities related to the research plan and results achieved. Beside our own publications we also support presentation activities of other researchers. Because of the character of the research plan we focus primarily on electronic presentation forms.

20.1 Web Server

The main platform for the electronic presentation of the research plan are the servers *www.cesnet.cz* (Czech version) and *www.ces.net* (in English). Majority of our results are available there – either directly, or as links to other servers operated by CESNET.

The server has three different interfaces:

Czech public: The most extensive one. Here we provide public information on the CESNET2 network, research and other activities of the association.

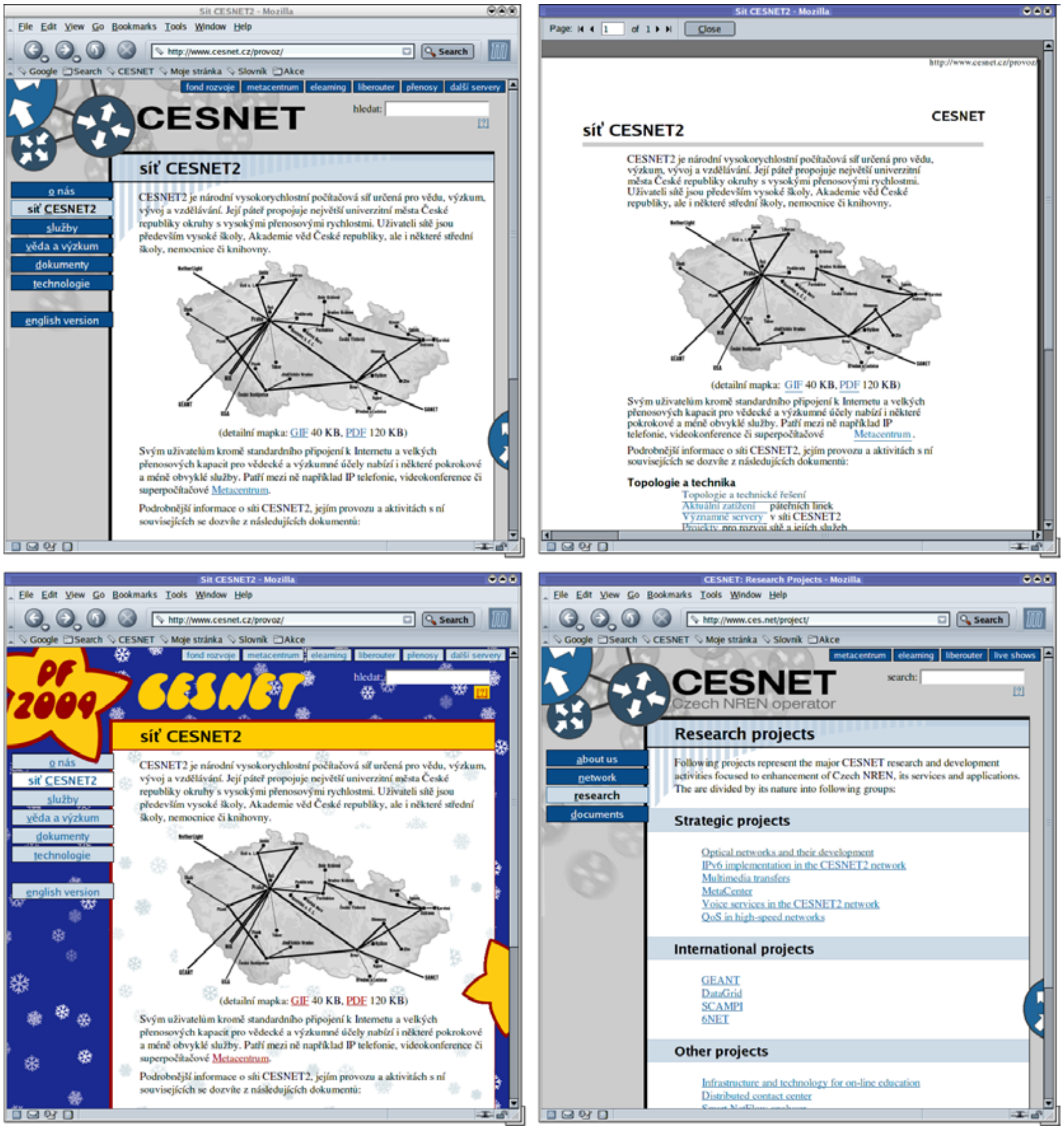
English public: Intended for our international partners and interested persons. The contents are similar to the previous one, although its extent is smaller. During 2003 we moved this interface to a separate virtual server named *www.ces.net*.

Private: It is used for the internal communication of participating researchers. Access is limited to users authenticated by the CAAS system.

We re-worked the whole server in 2003. For the most part, this change covered the translation of all pages from HTML into XHTML 1.0, actually to the strict version of XHTML 1.0. Due to the lack of formatting tools in this language version, all formatting is implemented using Cascading Style Sheets (CSS).

The combination of XHTML and CSS allows to easily modify the design of the whole server. It offers the opportunity to create a media-specific designs (for example a version for print) and deploy temporary looks created for specific occasions. For example, around the turn of the year we deployed a “PF 2004” design (see Figure 20.1).

The change of the code also motivated an adaptation of the structure of our WWW presentation. We tried to simplify it and help the visitors with orientation in the contents. To this end, we decreased the number of items in the main menu, reorganised and enriched some sections, introduced a bar at the top


Sit CESNET2 - Mozilla
File Edit View Go Bookmarks Tools Window Help
http://www.cesnet.cz/provoz/
Search
Google Search CESNET Moje stránka Slovník Akce
fond rozvoje metacentrum elearning liberouter přenosy další servery
CESNET
hledat:
o nás
síť CESNET2
služby
věda a výzkum
dokumenty
technologie
english version
Page: 1 of 1
Close
http://www.cesnet.cz/provoz/
PF 2004
CESNET: Research Projects - Mozilla
http://www.ces.net/project/
metacentrum elearning liberouter live shows
Czech NREN operator
search:
about us
network
research
documents

síť CESNET2

CESNET2 je národní vysokorychlostní počítačová síť určená pro vědu, výzkum, vývoj a vzdělávání. Její páteř propojuje největší univerzitní města České republiky okruhy s vysokými přenosovými rychlostmi. Uživateli sítě jsou především vysoké školy, Akademie věd České republiky, ale i některé střední školy, nemocnice či knihovny.

(detailní mapka: GIF 40 KB, PDF 120 KB)

Svým uživatelům kromě standardního připojení k Internetu a velkých přenosových kapacit pro vědecké a výzkumné účely nabízí i některé pokrokové a méně obvyklé služby. Patří mezi ně například IP telefonie, videokonference či superpočítačové Metacentrum.

Podrobnější informace o síti CESNET2, jejím provozu a aktivitách s ní souvisejících se dozvíte z následujících dokumentů:

Topologie a technika

Topologie a technické řešení
Aktuální zatížení páteřních linek
Významné servery v síti CESNET2
Projekty pro rozvoj sítě a jejích služeb

Research projects

Following projects represent the major CESNET research and development activities focused to enhancement of Czech NREN, its services and applications. The are divided by its nature into following groups:

Strategic projects

Optical networks and their development
IPv6 implementation in the CESNET2 network
Multimedia transfers
MetaCenter
Voice services in the CESNET2 network
QoS in high-speed networks

International projects

GEANT
DataGrid
SCAMPI
6NET

Other projects

Infrastructure and technology for on-line education
Distributed contact center
Smart NetFlow analyser

Figure 20.1: Server *www.cesnet.cz*/*www.ces.net* – standard page, its version for print (upper right), special seasonal design (lower left) and sample of English web (lower right)

containing links to other important CESNET servers and added alternate paths to important information. The steps described above led to a total change in the server design.

Naturally, the contents of the server have been maintained. We added the following significant materials in 2003:

- Annual report on research activities in 2002 in the Czech and English languages.
- 2002 annual report of CESNET Association (the report is bilingual).

- Materials form public workshops we organized, typically containing the program, presentations and video feeds of individual contributions.
- Technical reports written by CESNET researchers.

The server is available over both IPv4 and IPv6 protocols.

20.2 Publishing Activities

The most extensive publication issued in 2003 is the annual report on research activities in 2002 containing 235 pages. This year we decided for the first time to create two language mutations of this report – the Czech one (published January 15th) and the English one (published February 7th).

The preparation of the English version turned out to be rather difficult. Given the limited time [1], we decided to use a professional translation agency. The resulting translation had unfortunately varying quality and required a lot of additional work.

Nonetheless, the existence of the English report has been appreciated by our foreign partners. We thus plan to translate this report too[2]. However, we decided to change the method of its preparation, namely to ask the researchers to translate their contributions themselves and consequently proof-read these materials.

Our collaboration with the *Lupa* magazine continued. *Lupa* is an on-line magazine focusing on the Internet in the Czech Republic. We published a few dozen articles there during 2003 covering advanced networking technology, services, and interesting projects in this area.

Specific results on individual projects are documented in technical reports. In 2003 we published 30 reports – more than two thirds of them being in English and one bilingual. The number and language structure of technical reports improved significantly compared to previous year.

We continued publishing of the *Datagram* bulletin in which we inform CESNET members and other institutions about new features of CESNET2 network and activities related to its development. Three regular issues and one special[3] have been published in 2003.

All the documents mentioned above are available in electronic form from the server *www.cesnet.cz*, see especially the “Documents” section.

[1] We needed the report as a basis for consultation with foreign specialists during the preparation of new research plan.

[2] As you can see, the plan has been realized ;-)

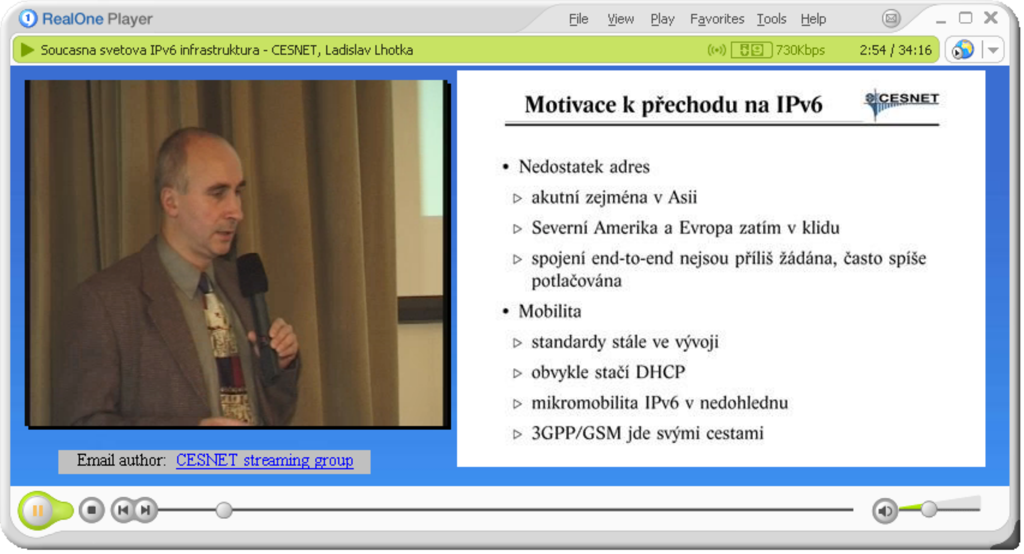
[3] dedicated to the call for projects for the CESNET development fund

20.3 Workshops and conferences

In 2003 we organised two workshops with the goal of bringing our research results to a broader audience.

On February 20th we held the workshop *New trends in the development of high-speed networks and their applications*. Speakers included representatives of European commission, Czech ministries of informatics and education, and foreign NRENs. The workshop was focused on current trends and perspectives of future development of leading edge networks and applications using them. Support of such activities – various projects and programs – was also discussed.

The second workshop, on October 22nd, had the title *IPv6 – development and deployment*. We invited primarily relevant specialists working for Czech Internet providers and larger networks operators and tried to share with them our experience in the IPv6 area in the hope that it will aid the deployment of IPv6 in real networks. Speakers were researchers who work on our project *Implementation of IPv6 in CESNET2 network*.


RealOne Player
File View Play Favorites Tools Help
Soucasna svetova IPv6 infrastruktura - CESNET, Ladislav Lhotka
730Kbps
2:54 / 34:16
Email author: CESNET streaming group
CESNET

Motivace k přechodu na IPv6

- Nedostatek adres
  - ▷ akutní zejména v Asii
  - ▷ Severní Amerika a Evropa zatím v klidu
  - ▷ spojení end-to-end nejsou příliš žádána, často spíše potlačována
- Mobilita
  - ▷ standardy stále ve vývoji
  - ▷ obvykle stačí DHCP
  - ▷ mikromobilita IPv6 v nedohlednu
  - ▷ 3GPP/GSM jde svými cestami

Figure 20.2: Video from the IPv6 workshop

This workshop has been appreciated by the participants who expressed their interest in this kind of workshops. Therefore, we plan further workshops focusing on modern networking technology – like IP telephony, network measurement and evaluation, quality of service and so on.

Both workshops were broadcasted to the Internet and presentations and recordings of all contributions are available on *www.cesnet.cz*.

As in previous years, in collaboration with Palacký University in Olomouc we organised the conference *Broadband networks and their applications*. Our colleagues contributed six papers to the conference programme:

- Liberouter, IPv6 router (Antoš, Novotný)
- Security in broadband networks (Denemark, Hladká)
- Usage of a high-speed networks for the need of medical applications (Dostál, Slavíček, Petrenko)
- Communication using the AccesGrid Point technology (Hladká, Holub, Rebok)
- Lectures from the recording (Hladká, Liška)
- Usage of IP telephony in CESNET2 network (Vozňák)

21 Security of local CESNET2 networks

21.1 Summary

The CESNET National Research and Education Network consists of a number of local networks containing computers with various operating systems. Ensuring the security of large heterogeneous networks brings about great demands for the work capacity of their administrators, and it is known that especially large university networks often include insufficiently secured machines. The objective of this project was to make the unenviable job of administrators easier by providing them with access to a very useful software – partly available under the GPL licence, partly created within this project framework.

21.2 Security audit using the NESSUS program

The *NESSUS* program, its auxiliary programs as well as their use within the CESNET Association networks have been described in the last year's annual report already. The *NESSUS* program has been upgraded significantly; its authors have corrected many of its errors and its installation has become much easier – running a single installation program is sufficient. *NESSUS* security tests have become plentiful; they numbered around 1900 in mid-December 2003.

The *NESSUS* auxiliary programs developed within this project framework did not have to undergo any significant changes.

21.3 Intrusion Detection System using the SNORT program

The *SNORT* program is generally regarded as the most widely used one in its field (network traffic monitoring at zero costs). It has been described sufficiently in the last year's annual report, too. No *SNORT* auxiliary programs have been developed within this project.

21.4 Intrusion Detection System using LaBrea

The previous annual report has already stated that *LaBrea* slows down and limits the propagation of computer viruses and network worms – moreover, almost without any need for human intervention. One should add that *LaBrea* also checks for any received SYN-ACK packets which signify that someone is trying to set up a TCP connection using our forged source IP address – most likely, to attack the target using a Denial of Service attack. However, *LaBrea* responds to SYN-ACK by sending a RST packet and as a result, TCP connection is terminated and this attack fails. This proves again the usefulness of the *LaBrea* program for the overall Internet community.

The usefulness of the *LaBrea* program can be further extended significantly using several auxiliary programs developed within this project's framework. This year, our efforts have concentrated on them to make the life easier not only for the *LaBrea* administrators but also for those network or domain administrators who would be receiving our reports on their potentially compromised machines. These programs process the data recorded in the *LaBrea* output files; this way, *LaBrea* turns into a simple and very effective system for detecting attacks which target our network. In comparison with standard Intrusion Detection Systems, *LaBrea* with our supplementary programs is characterised by these important differences:

- No complex and difficult configuration is necessary.
- No physical servers have ever existed on the IP addresses monitored by the *LaBrea* system. As a result, noone but network viruses, worms and hackers can have any reason to attempt connecting there – every connection attempt can be rightfully classified as an attack.
- *LaBrea* records only data from received packet headers; no data from received packets are recorded. This means that we know that an attack has occured as well as its source and destination addresses; however, we cannot tell the kind of virus or the sort of attack.
- *LaBrea* records only the source and destination TCP addresses (including the port numbers) as well as the ICMP ECHO packets. All other protocols including UDP are ignored. However, we do not have to regret this; the UDP packets are forged too easily and too often.

21.5 LaBrea auxiliary programs

21.5.1 LaBreaBackEnd

This program can be terminated prematurely, as well as restarted later, without losing any results.

The LaBreaBackEnd operator can limit the number of resulting output files created; in addition, a limit count can be set on the command line. This limit count determines the number of connection attempts at which further processing terminates. The default limit count value is the minimum number of connection attempts detected; usually, this value equals ”1”.

To make the program run faster and to minimise the communication with WHOIS servers, results of previous WHOIS queries are cached. Therefore, no further WHOIS queries for the same IP address space or for the same domain are necessary. (Still, administrators of the APNIC WHOIS server have asked the CESNET director to confirm that this was an official project of the CESNET Association.)

Program ignores all connection attempts originating from those address ranges which IANA had not assigned to any Regional Internet Registry yet. (Therefore, administrators should monitor official reports about new assignments of IP space to the RIRs.)

If a reverse record of the originating (source) IP address exists, program checks if it matches the address record. Should they differ, no domain data are sought.

This program can be given a list of IP ranges which are of some special interest to its operator – e.g., all networks belonging to our autonomous system. Records about connection attempts originating within theses IP ranges can be processed separately – e.g., more often or in a different way than the records about the rest of attacks.

The *LaBreaBackEnd* can also be tailored according to wishes of domain or network administrators in the following ways:

- to ignore attacks coming from given single IP addresses or IP address ranges
- not to send the reports to given e-mail addresses
- to replace given administrator address(es) by another – this is useful especially for those administrators who cannot update their contact records in their WHOIS databases
- to send administrators of certain networks a single report for each recorded attacking IP address instead of the usual bulk report, if their RTT system requires this.

21.5.2 LaBreaReport

We can send the *LaBrea* IDS reports to a better set of CESNET network and domain administrators than to those found in the public WHOIS databases. Because of that, the *LaBreaReport* program has been modified to run in two passes. The first pass puts aside the reports addressed to the CESNET administrators to be checked manually; reports intended for the rest of Internet are sent in the usual way. (Usually, only several reports addressed to the CESNET administrators exist.) In the second pass, reports for the CESNET administrators are sent; unnecessary generic addresses, e.g., *abuse@cesnet.cz*, are removed automatically.

The program checks the administrator addresses to determine if the report is to contain only the Czech language message, English language message, or both.

Two different test modes can be selected on the command line: in the first one, no mail is sent; in the second one, all mail is sent only to a single preselected test address.

21.5.3 LaBreaReportDetailed

will be appreciated especially by those *LaBrea* IDS system managers whom some network or domain administrators will have asked for a detailed information about attacks. This program allows looking up all records about attacks coming from a single IP address (in this case, the IP address will be given directly on the command line), or about attacks coming from several IP addresses (the name of a file containing their list will be given on the command line). Also, attacks coming from whole networks size /24 or /16 can be looked up. The resulting file gives human readable attack dates and times (as opposed to the *LaBreaReport* which gives the “epoch time”), as well as a complete log of attacks (unlike the *LaBreaReport* which reports only the first and last ones). Just like the *LaBreaReport*, the *LaBreaReportDetailed* program masks out a part of the target IP addresses to improve the system security.

21.6 SYSLOG-NG

The Technical University of Ostrava network includes a new *SYSLOG-NG* server, so far operating in a pilot phase. Selected Unix servers as well as network elements send it their data to be logged; these are saved in separate files according to their fully-qualified domain names and “syslog facility”. Purpose-made programs analyse these files constantly; selection of the most suitable program for on-line analysis of logged data is still going on.

Some kinds of attacks threatening the system or network security may be logged and sent to the *SYSLOG* server. The on-line analysis program running on the *SYSLOG* server should notice these incidents and react accordingly. In addition to security incidents, other incidents of varying importance are recorded together with other facts which may influence the functionality of network systems and their services. To be able to put the *SYSLOG* server into regular operation, the following tasks must be solved first:

- securing the communication between the network servers and *SYSLOG* server
- interface for reliable and timely sending of warning messages using E-mail or SMS (via the attached SMS gateway)
- adding proper filters for detection of important events
- proper selection of one or several programs for on-line analysis of log files.

When selecting appropriate hardware for the *SYSLOG* server, the most important parameters are the CPU speed, RAM size, hard disk capacity and speed. We selected an Optiplex GX150 made by DELL. The *Debian Woody* distribution with standard kernel 2.4 was chosen for the operating system. A GSM modem *Siemens MC35i* connected to the server operates as a SMS gateway; this can send appropriate warning messages to the administrators who can react fast even if not present at their terminals. This solution is not complete yet but even now it is clear that it improves the functionality and security of the whole network.

21.7 Results and Experience

21.7.1 NESSUS

Some of our colleagues we had relied upon were too busy to fulfill the requirements of the CESNET Network Operations Department: to display the security audit results (created by *NESSUS* and *WebBackEnd*) using the XML format, making it more user-friendly (clearer arrangement, optionally hiding some results, etc.). However, security audit of machines connected to the Prague-Dejvice CESNET network was performed every two weeks for the whole year 2003; its results were made available on a CESNET HTTPS server for each authorised user who also had received an e-mail summary of results beforehand.

Security audit of the Technical University of Ostrava machines has been using the *NESSUS* program in combination with the *PTS* tool described in last year's report. Results of security audits are distributed directly to the administrators of appropriate systems. Security audits are performed once a week.

Results of *NESSUS* security audits within the Academy of Sciences network confirm a well-known fact: very few machines running the Microsoft Windows operating system can be regarded as safe. One cannot depend on ordinary users to properly look after the security of their machines; as a result, these machines must be managed centrally. Therefore, *NESSUS* runs only occasionally in the Academy of Sciences network, mainly to discover those machines whose security updates were overdue. The security audit results are made available only to the appropriate network administrators.

21.7.2 SNORT

No auxiliary programs are necessary to run the *SNORT* program. It just needs a proper set-up to report as few false alarms as possible; achieving this is not easy – it depends on the administrator's experience and network traffic. An important benefit is the availablility of complete data which can be used for a detailed attack investigation.

SNORT ran as an Intrusion Detection System on all three workplaces of this project, especially at the Academy of Sciences. It proved its usefulness when detecting external network penetration attacks, recording attacks targeting single TCP or UDP ports, as well as detecting any "suspicious" internal network traffic.

21.7.3 LaBrea

The *LaBrea* server in the Prague-Dejvice network has been running over a year. It has been distributing some 500 to 1500 e-mails per week to those network or domain administrators where the connection attempts had originated. Absolute majority of those administrators who responsed personally was grateful for these notices; their reactions and suggestions helped us improve the *LaBrea* auxiliary programs.

Summer 2003 has witnessed a surge of extremely active network viruses; their activity is still going on and confirms the importance of the *LaBrea* server – see Figure 21.1. Note: Smaller bandwidth recorded between October 13–October 20 resulted from changed configuration (the "–p" parameter), not from a smaller number of attacks.

Comparison of the latest data with those published in the previous annual report shows an alarming fact: The data flow from the attackers to the *LaBrea* server reached saturation (2 KBps) within some 14 days in Autumn 2002. One year later, in Autumn 2003, this data flow reached saturation (now 8 KBps) within mere 7 hours. . .

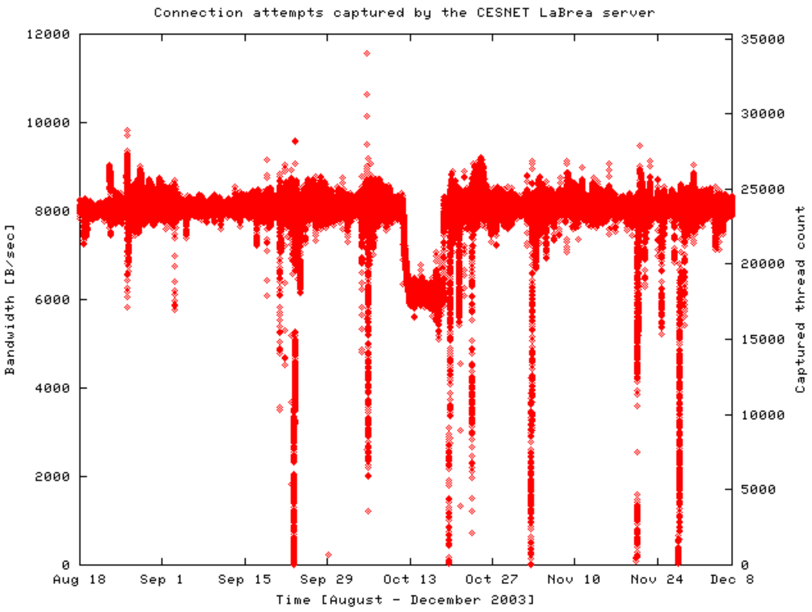


Figure 21.1: Attacks recorded by the LaBrea IDS system (18. 10.–8. 12. 2003)

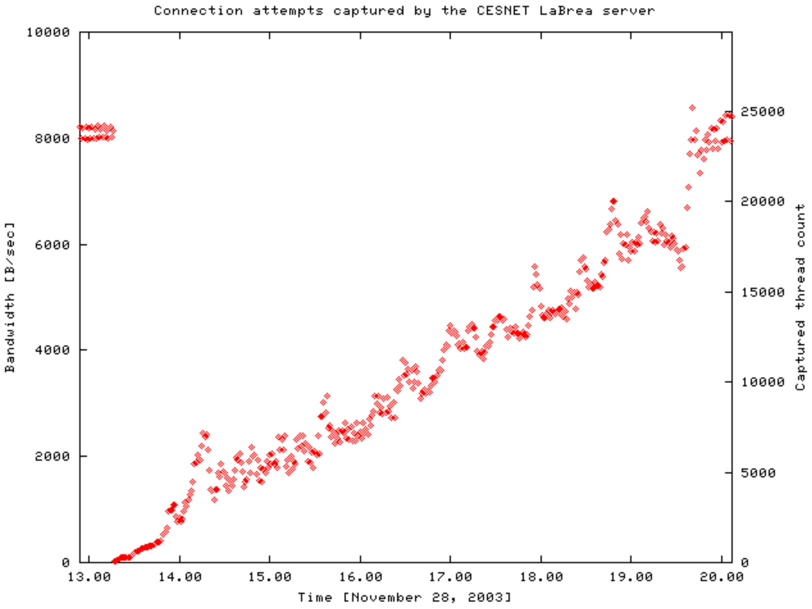


Figure 21.2: Detailed recording of attacks in November 28, 2003

21.8 Conclusion and plans

All software developed within this project framework has been published regularly on our FTP server *ftp://ftp.cesnet.cz/local/audit/*. Up-to-date information on this project progress as well as on new software versions are made available to every subscriber of the *AUDIT-L@cesnet.cz* list.

21.8.1 Future plans and further progress

Auxiliary programs for*LaBrea* work quite well. We plan to continue maintaining them but most of work seems to be finished.

We do not inted to develop any auxiliary software for *SNORT*; this program is useful but we plan to run it only should all other problem-solving methods fail.

The *NESSUS* program has been running routinely for a long time – together with the *PTS* or *WebBackEnd* program according to the administrator preferences. Both alternatives provide the operators of audited machines with all necessary information using the text mode. An improved user comfort which the CESNET Network Operations Department prefers would require using the XML format and a radically reworked user interface. If sufficient work capacity is available, no obstacles should impede the progress of this project.

22 NTP server controlled by the national time etalon

22.1 Introduction

Main goal of the project is the development and operation of a time server (called TimeCZ) controlled by the Czech national time etalon. The project has started with cooperation of the Institute of Radio Engineering and Electronics of the Academy of Science of the Czech Republic (IREE), which operates the time etalon.

The main feature of this time server, compared with other time servers, is its independence from any third party time signal (e.g., navigation systems or public time services). The national time etalon is a trustable source with a metrologically defined relation to the UTC time.

22.2 Structure of the time server

The server consists of three main functional blocks:

- computer running *ntpd* daemon,
- control system KPC,
- microprocessor system FK.

Detailed description of components, processes and relations between them was presented in the 2002 report. After one year of experimental operation we see no need to change the design and we are thus confident the server can be used in a production environment.

22.3 Progress in 2003

In 2003 we focused on experimental verification and improvement of server features.

22.3.1 New version of the FK system

We designed and manufactured a new microprocessor system FK in order to replace the former development version. The new version has the following features:

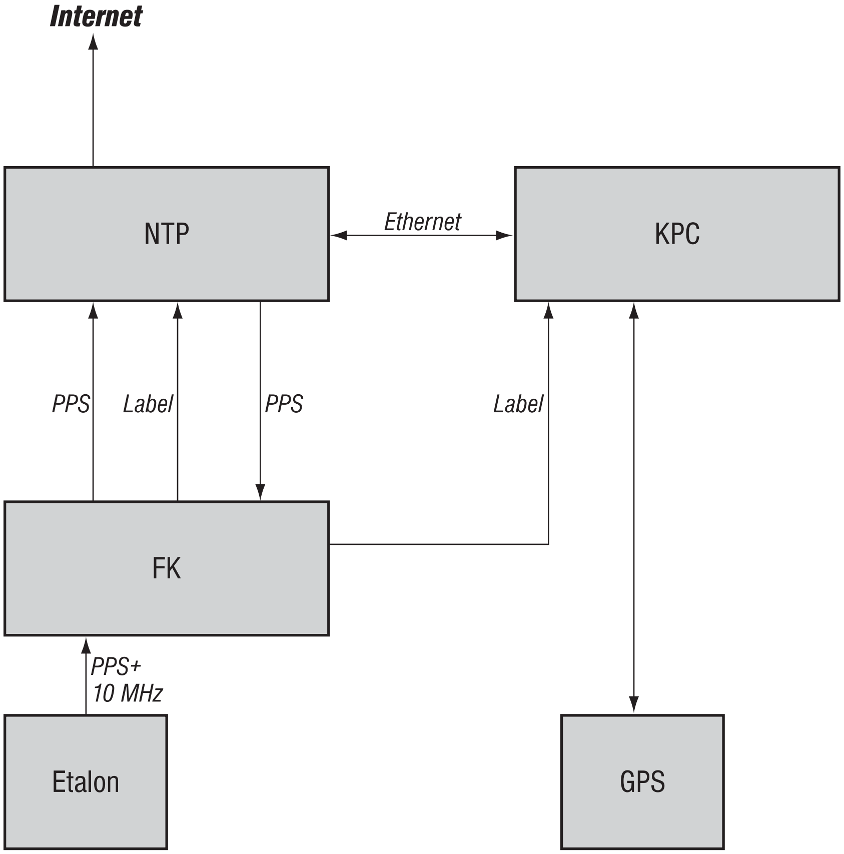


Figure 22.1: Functional blocks of the server

- more stable and reliable main board,
- generation of leap second flag,
- more accurate measurement of the difference between server time and internal time.

22.3.2 New version of software

We wrote new software for the KPC control system. Our operational experiences showed the system was not sufficiently tolerant to packet loss and so we decided to improve the communication algorithm between KPC and other components.

22.3.3 WWW pages

As the time server is intended for public usage, it is also important to prepare appropriate documentation for both casual and expert users. We started to collect all necessary information on WWW pages, which now contain an explanation of server principles, user instructions, server characteristics and links to recommended NTP clients. Unfortunately, we did not succeed to finish the WWW pages yet and so the task continues.

22.4 Characteristics of server

The server checks continually the difference between its internal time and the national etalon. In addition to that, the internal time is compared to other independent time sources – a GPS receiver and an external NTP server. The control systems can block server output and switch it effectively off if the difference indicates a system malfunction.

Apart from checking against the second label, the internal time is also checked with a resolution of 100 ns. It is necessary to do both checks, as the etalon output also consists of two signals: the second label and the PPS (Pulse per Second) signal. We have to eliminate the theoretical situation when the time provided by the server differs from the exact time by an integer multiple of a second.

The complete list of checks is as follows:

- matching server time against the second label of the FK system,
- matching server time against the GPS time,
- matching server time against an independent external NTP server,
- measuring the difference (with 100 ns accuracy) of PPS signals between the server and the etalon.

The output of the time server is blocked no later than 3 seconds after a conflict is discovered between second labels of the server and the system FK or as soon as the measured difference of PPS signals exceeds 30 μs. When a conflict is found between the server time and GPS or external NTP time, a warning is generated without blocking the server output. In this case the error may actually be in the external system.

22.4.1 Accuracy of the time server

The Figure 22.2 shows the measured absolute difference between the server and etalon times. The 90-minute interval shown in the figure represents a typical behaviour of the measured value, with the absolute error lying in the range of 500 ns around the mean value of 1 microsecond. As the mean value was stable throughout the observed period (i.e., several months), it is possible to compensate for it and thus reach the absolute accuracy about 500 ns.

The accuracy is better than expected. It is mainly due to the usage of a temperature compensated oscillator (TCXO) and a special card for processing the

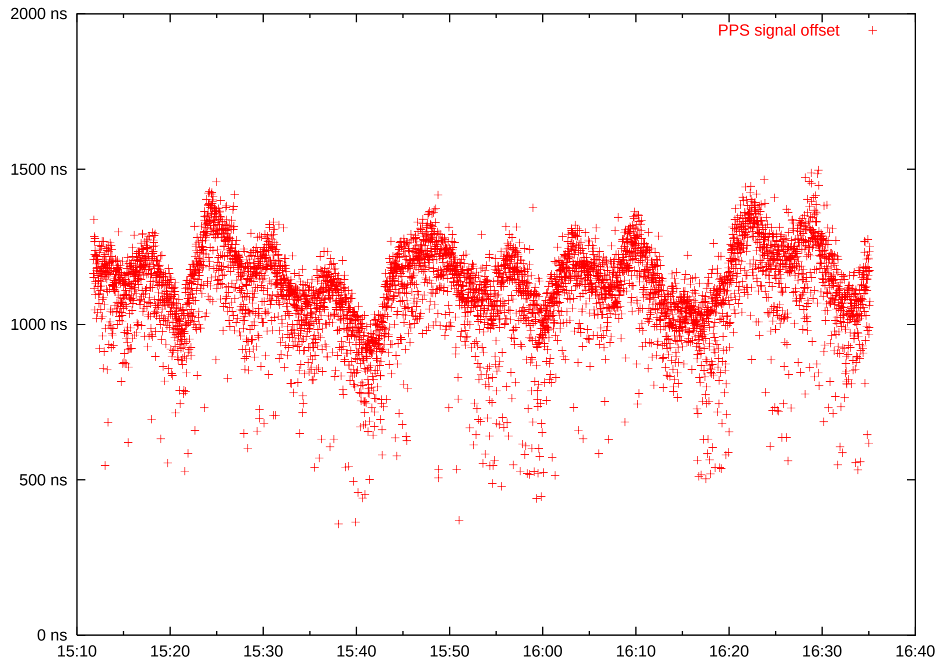


Figure 22.2: Absolute error of the time server

PPS input without any latency. Placing the server in a well air-conditioned room probably helped as well.

22.5 Plans for future

We have to finish the WWW pages of the server in a near future. We intend to continue the observation and measurement of the server characteristics. We consider to develop a real-time hardware module having the synchronization algorithm implemented independently of the main processor.

23 Platforms for video transmission and production

Streaming platform built by CESNET in previous years reached a production status at the end of 2002. Consequently, in the year 2003 we focused on making its use more effective and widespread.

23.1 International cooperation

In 2003 we joined the *Netcast Taskforce (TF-Netcast)* activity within the TERENA Technical Programme framework. TF-Netcast, being the successor of TF-STREAM, focuses on gathering information about the status of media-streaming applications in the national research networks, coordinating activities in this area and developing tools for an effective use of streaming media in the high-speed networking environment.

23.1.1 Development

We concentrated our development effort primarily on extending the functionality of the Announcement portal at *http://live.academic.tv/* by implementing support for multiple languages and automated access.

The multilingual support enables access for the pan-European academic community. While the portal is now configured to announce events in all supported languages, a multilingual documentation and an active interface still to be added. To this end, we modified portal to allow for adding a new language just by sending a text file with translated sentences. At the same time we asked representatives of the NRENs participating in TF-Netcast to provide these translations. Thanks to their active cooperation we now have our portal in nine different languages.

An automated access was mandated by the requirement of interoperability with announcements issued by other networks, irrespective of whether they are public or not. As we want our portal to play an integrating role functionality, we must be able to accept data from others. Internal data structures of announcement portal are based on the XML language and so it was quite natural to use XML for automated access as well. The access module is realised as an application gateway based on the SOAP protocol – data format is identical to the data format for user access. Each remote system is authenticated in local user database and authorised for operating only upon own data (i.e., those submitted earlier by itself).

23.1.2 Live transmissions

One of the goals of TF-Netcast is to provide an infrastructure for live transmissions from significant conferences with a pan-European impact. During the year 2003, TF-Netcast set up two ad-hoc Content Delivery Networks. These networks distributed live transmissions from RIPE 45 and RIPE 46 meetings. We participated in both networks and have been the only node distributing that contents over IPv6.

23.2 Metadata Indexing

With continuously growing volumes and quality of digital multimedia (not only in the research and development area) the need for an effective organisation of the material and flexible search capabilities becomes more pressing. Instead of a direct similarity-based search in the stored material (sound sample or picture), one can search the metadata that describe the stored material. We were not able to find any public Internet search engine capable of metadata searching.

Because metadata are typically represented as text, we adopted a common full-text search engine (*Jyxo*[1]) for this purpose. While aiming primarily at searching the Czech part of the Internet, we tried to make the system architecture scalable and portable. Input data (URLs pointing to multimedia data) are obtained from the crawler, which is a part of the full-text search engine. Our *distiller* component then extracts metadata from files pointed to by those URLs and sends them back to the search engine (the format of metadata is XML). Finally, the search engine processes these XML files and stores them into its full-text database. The presentation part relies on standard WWW interfaces.

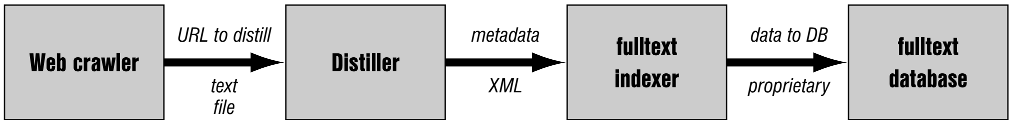


Figure 23.1: Structure of the multimedia indexing

All communication between components is asynchronous and independent of particular user requests. The system is ready for parallelisation, which can be useful for indexing large amount of data. Because of the open data format and open interfaces, it is possible to integrate our component into any other search engine.

The system we developed is able to search the majority of multimedia files that are publicly available in Czech Internet (about 23 thousand files by now). As

[1] *http://www.jyxo.cz/*

a side-effect of this activity, we also solved the issue of full-text search in the CESNET video archive, volume of which is also steadily growing.

23.3 Streaming Platform

The base for all our activities is the streaming platform we built in the previous years. The system supports the following formats: Real Video, Windows Media, QuickTime (up to version 4) and MPEG-4 (3GPP profiles only).

An important part of our activities is the support we provide to the scientific and research community in the area of media streaming.

23.3.1 IPv6

The only new component is the streaming system for IPv6. This server shares data files with the other components so that we are able to send the same data using various transports and, if necessary, add new storage and networking capacities without endangering data integrity.

On the basis of our previous evaluations we selected Windows Media 9 (based on Windows 2003 Server) as the only usable platform for IPv6 streaming. This server has been connected over both IPv4 and IPv6 to the CESNET2 network. It is thus possible to redistribute the streams between both environments (we used this feature during live streaming from the RIPE 46 meeting and Megaconference V).

23.3.2 Live transmissions

During 2003 we continued the methodological and technical support of live transmissions from professional conferences. We supported more than 20 events, the most important being:

- Winter School of Computer Graphics 2003
- RIPE 45
- Objects 2003
- Cryptofest 2003
- RIPE 46
- ATLAS Overview Week 2003
- Ostrava Linux seminars
- IPv6 – development and implementation (CESNET seminar)
- New trends in the development of high-speed networks and their applications (CESNET seminar).

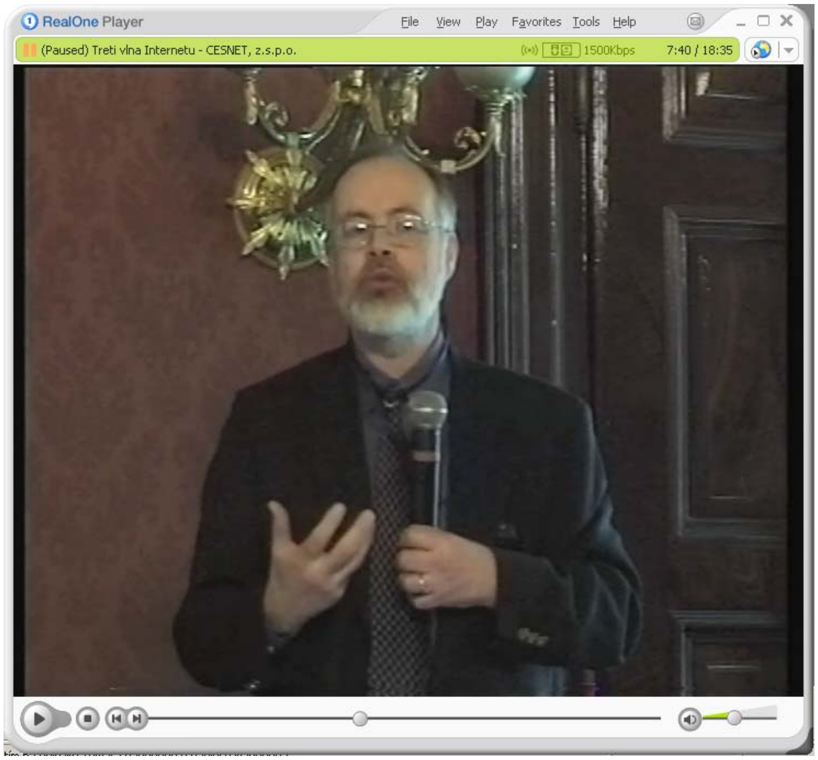

RealOne Player
File View Play Favorites Tools Help
(Paused) Treti vlna Internetu - CESNET, z.s.p.o.
1500Kbps
7:40 / 18:35

Figure 23.2: Recording from the workshop *New Ways in Development of High-speed Networks and their Applications*

From the viewpoint of networking technologies, the most important event was the live transmission of Megaconference V, a worldwide videoconference organised by Ohio State University with support from Internet2. Our streaming server has been the only server outside Internet2 infrastructure and the only one streaming video over the IPv6 protocol.

We continued our support for lecture recording at the Masaryk University and Czech Technical University. In addition, we started a new cooperation with the audiovisual centre of the Student union of Czech Technical University, which produces a lot of high-quality content (mainly in the area of natural sciences and IT).

23.3.3 Video archive

The CESNET video archive contains around 200 hours of video material and 800 individual contributions.

So far, a considerable manual effort has been necessary for maintaining the presentation side of the video archive. We thus migrated the archive system to a database platform with the aim of eliminating most of the manual work. The migration happened successfully at the end of 2003.

23.4 Announcement portal

Along with the development of the announcement portal we have been taking care about its routine operation throughout the year. In 2003 we registered more than 200 announcements through this portal.

In addition to announcing events organised by CESNET, the portal was also used for announcing live transmissions from a number of lectures on diverse topics or academic events.

Approximately half of the active users are from the international research community (mostly from the TF-Netcast group). These users have been announcing 20–30 % of live transmissions.

23.5 Portal streaming.cesnet.cz

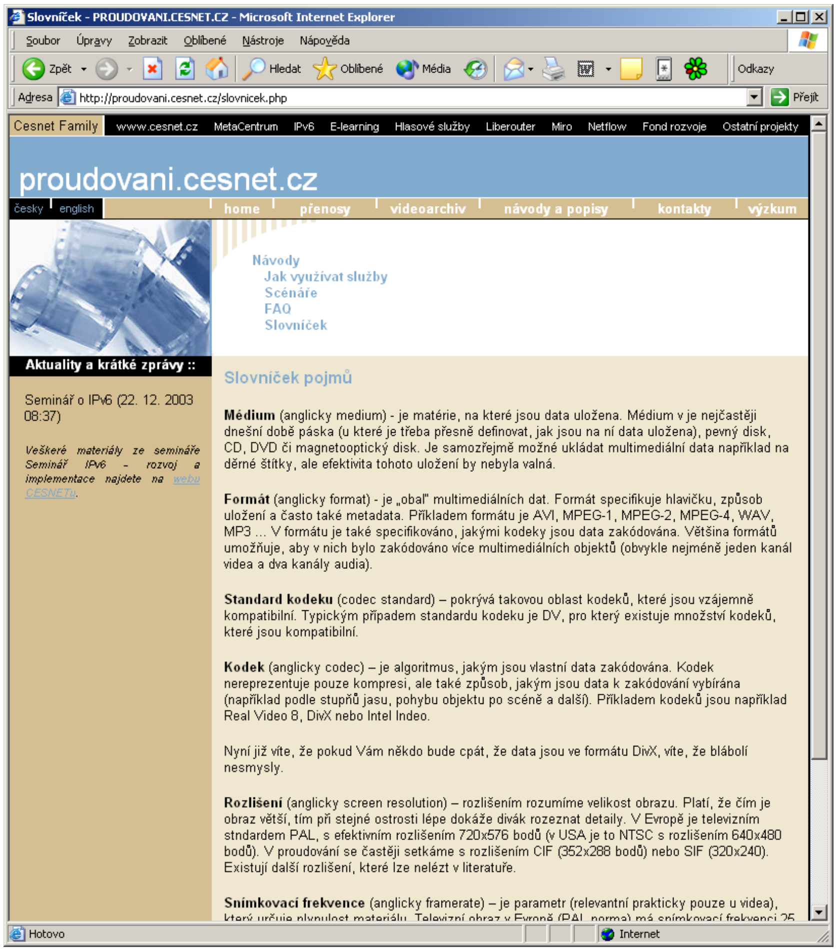
Due to the active dissemination of project results, we had to cope with an increasing demand for methodical support of multimedia services. In order to make such a cooperation more effective and also to present our know-how in a summarised form, we decided to start a new portal *streaming.cesnet.cz*.

The portal integrates our experience and recommendations in the area of multimedia streaming with our publications and a video archive. The portal contents are supposed to evolve in the future in order to reflect the state of the art.

23.6 Conclusions

During 2003 we successfully continued the development of the CESNET media streaming platform and actually delivered more than was planned. An important decision was to concentrate our capacities on the cooperation within TF-Netcast, instead of developing a distributed transcoding system.

Our contacts with other national research networks in Europe and in USA also became more intensive in 2003.



proudovani.cesnet.cz

Aktuality a krátké zprávy ::

Seminář o IPv6 (22. 12. 2003 08:37)

Veškeré materiály ze semináře Seminář IPv6 – rozvoj a implementace najdete na webu CESNETu.

Slovníček pojmů

Médium (anglicky medium) - je matérie, na které jsou data uložena. Médium v je nejčastěji dnešní době páska (u které je třeba přesně definovat, jak jsou na ní data uložena), pevný disk, CD, DVD či magnetooptický disk. Je samozřejmě možné ukládat multimediální data například na děrné štítky, ale efektivita tohoto uložení by nebyla valná.

Formát (anglicky format) - je „obal" multimediálních dat. Formát specifikuje hlavičku, způsob uložení a často také metadata. Příkladem formátu je AVI, MPEG-1, MPEG-2, MPEG-4, WAV, MP3 ... V formátu je také specifikováno, jakými kodeky jsou data zakódována. Většina formátů umožňuje, aby v nich bylo zakódováno více multimediálních objektů (obvykle nejméně jeden kanál videa a dva kanály audia).

Standard kodeku (codec standard) – pokrývá takovou oblast kodeků, které jsou vzájemně kompatibilní. Typickým případem standardu kodeku je DV, pro který existuje množství kodeků, které jsou kompatibilní.

Kodek (anglicky codec) – je algoritmus, jakým jsou vlastní data zakódována. Kodek nereprezentuje pouze kompresi, ale také způsob, jakým jsou data k zakódování vybírána (například podle stupňů jasu, pohybu objektu po scéně a další). Příkladem kodeků jsou například Real Video 8, DivX nebo Intel Indeo.

Nyní již víte, že pokud Vám někdo bude cpát, že data jsou ve formátu DivX, víte, že blábolí nesmysly.

Rozlišení (anglicky screen resolution) – rozlišením rozumíme velikost obrazu. Platí, že čím je obraz větší, tím při stejné ostrosti lépe dokáže divák rozeznat detaily. V Evropě je televizním stndardem PAL, s efektivním rozlišením 720x576 bodů (v USA je to NTSC s rozlišením 640x480 bodů). V proudování se častěji setkáme s rozlišením CIF (352x288 bodů) nebo SIF (320x240). Existují další rozlišení, které lze nelézt v literatuře.

Snímkovací frekvence (anglicky framerate) – je parametr (relevantní prakticky pouze u videa), který určuje plynulost materiálu. Televizní obraz v Evropě (PAL norma) má snímkovací frekvenci 25

Figure 23.3: Server *streaming.cesnet.cz*

24 MeDiMed

The following terms, not too common in the ICT field, should be introduced here:

- PACS = Picture Archiving and Communication Systems (a special kind of an information system)
- DICOM = Digital Imaging Communications in Medicine (A standard for data interchange between PACS systems, supported by NEMA – the National Electrical Manufacturers' Association)
- Modality = Equipment producing an imaging information for PACS
- RIS/HIS = Radiology Information System/ Hospital Information System

24.1 Introduction

The Institute of Computer Science at the Masaryk University has been cooperating closely with Brno hospitals in implementing information and communication technology in the field of taking, transporting, archiving and presenting digital picture medical data since 1999. This cooperation consists of activities and projects whose goal is a development of a *metropolitan archive of medical imaging information* obtained from hospital modalities, diagnostic equipment like ultrasound (US), digital mammograph (DMG), computer tomography (CT), magnetic resonance (MR), etc., as well as providing access to this archive via computer network. The goal is to improve the quality of medical operations and common medical care and to enhance the environment for medicinal research and student education by using an up-to-date information technology and medical informatics.

The project includes support of imaging data transfer between individual workplaces or hospitals which a patient visits during his/her treatment, including optional consultations by remote specialists. As a result, correct diagnosis is easier and faster to achieve, repeated checkups are eliminated, time is saved for both the patient and doctor which saves finances as well. This project can also be regarded as a pilot project for other regions in our country.

24.1.1 Recapitulation

At the end of 1999, a PACS system was purchased for real-time processing, transfer and archiving of imaging data (both static and dynamic) as a basis for the archive. In the same year, first ultrasound modalities were connected to this system. These were the workplaces within the Brno Faculty Hospital (Obilní trh – Maternity Hospital and Černopolní – Paediatric Hospital) to allow consulting the ultrasound checkups by physicians specialised in prenatal (intrauterine) diagnostics of foetuses as well as diagnostics of new-born children's organs, especially cardiologic. Data between these locations are transferred using dedicated fibres of the Brno Academic Computer Network.

Originally, the ultrasound systems used in these locations provided analogue output only; therefore, converting the analogue output to the DICOM format was necessary. During the year 2000 we learned about the properties and limitations of the installed PACS system. The system underwent various modifications according to practical usage requirements and an imaging database was being built gradually.

One of the first key results of this project was an agreement reached among all collaborating subjects on the necessity of using the DICOM standard. Afterwards, all diagnostical equipment purchased later was DICOM-conformant. In the second half of 2000, a magnetic resonance equipment located in the St. Anne University Hospital was connected to the system. This MR equipment has already contained the DICOM output. Several other DICOM-conformant modalities belonging to the Masaryk Memorial Cancer Institute were connected in 2001. These were: a digital mammograph, computer tomograph and three ultrasound devices.

24.1.2 Research fields

The complex solution of building the metropolitan archive of medical imaging information and its usage covers three basic fields: legislative, technological, and financial.

In *the legislative field* the key problem is securing the medical information from any abuse. This is a very sensitive problem which is scrutinised very closely by the hospitals, and the rules for securing the data both in its origin and during its transport and archival are very strict.

Because the solution of this problem could not be solved just within the scope of our project but on the other hand it was a precondition for its further development, a meeting of directors of all Brno hospitals, deputy mayor of Brno, rector of the Masaryk University, a representative from the Ministry of Health care as well as a representative of the project team from the Masaryk University Computing Centre which coordinates all activities concerning the archive project. On this meeting, basic principles were agreed upon and this allowed further coordinated development of this system.

The field of technology covers research and implementation of practical solutions in two subareas regarding the hardware and usage of the whole archiving system.

The first subarea deals with proper data archiving: archive server and its safe operation, as well as selection of archiving media with respect to their cost, functional parameters and capacity. Problems of data rate, security and reliability of network transmission, access rights etc., also belong to this subarea.

The second subarea covers video display devices. One understands that different equipment parameters may be required depending on modalities used; in general, they may differ significantly from those common in other fields. E.g., sharpness of displays necessary for the field of magnetic resonance is so high that it touches (or surpasses) the technical limits of contemporary display units. This applies, e.g., to the flat panel displays commonly available.

The Financial subarea includes search for financial resources for the system (in budgets of participating hospitals as well as in domestic and international grant programmes), but also search for a balance between requirements, technical facilities and solution costs in particular areas of deployment.

Very close collaboration with medical specialists is necessary for success because only they are able to tell which display units are appropriate and sufficient for a given purpose and modality. For example, ultrasound – unlike magnetic resonance – does not require high-quality (and costly) monitors. A similar problem is selecting an optimum archiving method which allows immediate access to stored data: depending on the modality, one checkup may generate several, tens or hundreds of images.

24.1.3 The central archive

A remarkable goal reached in 2001 was the *central archival server site* built in a secure area of the new ICS MU computer room at Botanická 68a street (the computer room itself with this site was officially opened on December 6, 2001). To reach a maximum physical security of sensitive data on our servers and data storage devices, this site was installed in a standalone, separate and locked section of the computer room. This site allows interconnection of diagnostic modalities located in Brno hospitals; at the time of writing this article it contained graphic documentation consisting of 980,000 images. All checkups produced 1.3 TB of data stored on 300 DVD discs of the long-term archive.

The central archive site consists of the following equipment:

- *A central operational server for metropolitan PACS*: HP Netserver LH 3000 with two 1 GHz processors, a 764 GB RAID, 4 GB RAM and a Linux RedHat operating system. Both a short- and long-term archive are connected to this server.

- *The short-term archive of production data*, with its 400 GB capacity is logically split to two 200 GB parts for improved redundancy. The first one, so-called “acquisition” part, functions as an operational storage of latest generated pictures. The second one, so-called “pre-fetch” part, serves as a cache of selected images from the long-term archive to speed up the access time.

24.1.4 Display devices

In parallel with development and operation of the PACS archive, evaluating the quality of available display devices is under way. For objective evaluation, a sufficient data quantity must be obtained first; this will be available only after more modalities of different types are connected during the current and next year.

Our experience confirms that an universal type of display devices or display software suitable for all modality types and areas of their use does not exist currently. In some special cases (e.g., tomography pictures of brain structures), properties of contemporary digital displays fall behind the classical snapshots.

However, the problems to be solved are not only technical: another obstacle is an inertial and conservative attitude, distrust and inflexibility of some medical specialists.

24.1.5 Current fields of collaborations

Based on practical experience, these seven partial medical applications from various medical fields were included in the metropolitan PACS archive activities:

Magnetic resonance in the Masaryk Memorial Cancer Institute (MMCI): Transfer of MR checkup results from the St. Anne University Hospital to the MMCI radiology clinic and archival of images in context of CT diagnoses.

Mammodiagnostics: Transfer of breast DMG and US images from the MMCI to the St. Anne University Hospital, Department of Oncology.

Breast diagnostics at the MMCI: Electronic archival and transfer of selected patients’ graphic documentation (studies and presentations) of the MMCI Radiology Department which includes the DMG, CT and US investigation.

Brain diagnostics at the St. Anne University Hospital: Electronic archival and transfer of graphic documentation of the central nerve system pathology (studies and presentations which include the CT, MR and AG checkup results) from the St. Anne University Hospital, Department of viewing methods. These images are shared electronically with the Departments of neurosurgery, radiation oncology and neurology of this hospital.

Paediatric oncology: Electronic archival of all paediatric oncology patients' checkup results (especially X-ray, CT, MR, AG and histology) including those found in the hospital information system, which can be shared electronically or transferred to other departments of paediatric oncology in the Czech Republic.

Neonatal cardiology: Electronic archiving of US images of newborns' heart defects which can be exported and consulted with the cardiosurgery department experts.

Pathology: Development of consultations network of pathology specialists – transport of images for consultation and second opinion between pathology specialists of member hospitals – including transfer of complex graphic documentation of selected patients.

24.1.6 Project technical solution

First medical images were transferred using a completely standalone network interconnecting only the PACS server and appropriate modalities. This solution provided a very easy solution of both security and IP addressing network plan. After these first steps we moved forward to a production environment; as a result, a need for accessing the PACS from ordinary hospital workstations in the internal hospital network followed. A problem with IP addressing followed because the private address spaces of participating hospitals overlapped. Another problem was the security of the central PACS archive. Both problems were solved by deploying NAT firewalls. An interconnection diagram of Brno hospitals is shown on the Figure 24.1.

Central PACS servers are separated from the carrier infrastructure by a firewall. In addition to traffic filtering, this firewall also translates the IP addresses of central PACS servers. Therefore, each hospital can see these servers having an IP address belonging to its own address space.

Another NAT firewall is connected to the hospital network; its output is connected to the hospital border router. This ensures both the PACS system security (this firewall is under control of the MU ICS) and the hospital internal network protection from potential mistakes on our side (all traffic passes through a router under control of the hospital network specialists). This solution allows accessing both the PACS archive system and the Internet from the same user station at the same time. The hospital firewall translates the hospital IP addresses of workstations and modalities to another address space which does not overlap with that of other participants or servers.

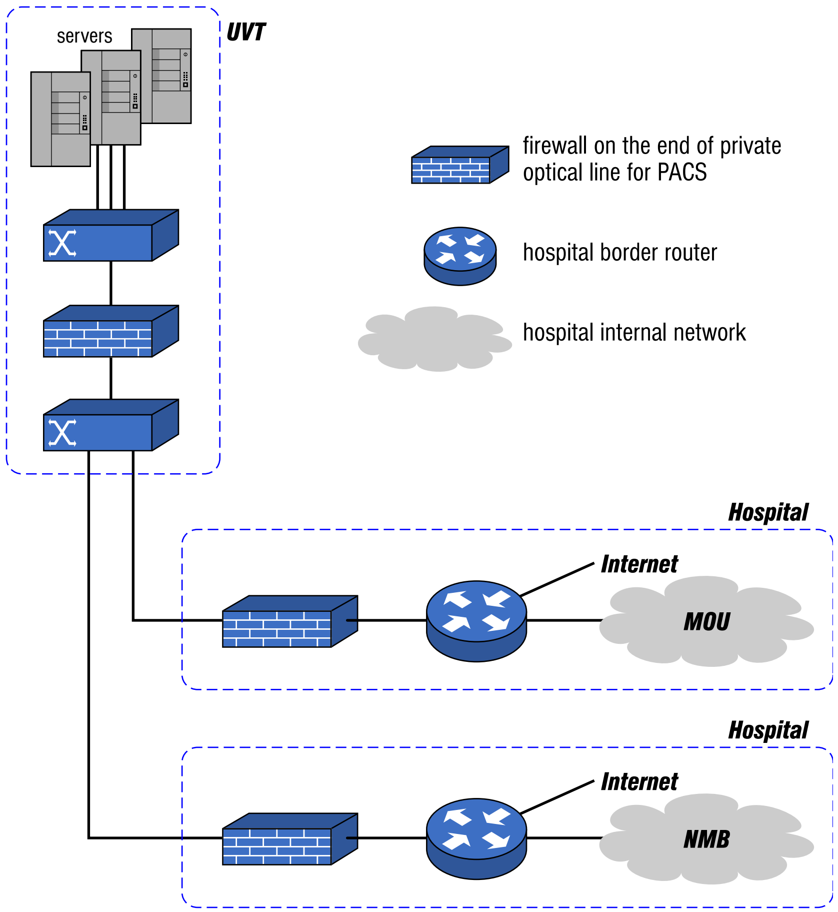

servers
UVT
firewall on the end of private optical line for PACS
hospital border router
hospital internal network
Hospital
Internet
MOU
Hospital
Internet
NMB

Figure 24.1: Interconnection diagram of Brno hospitals

24.2 Project progress in 2003

24.2.1 Related projects

In 2003, research and implementation of the PACS system development continued. The first part of system implementation was supported by a Ministry of Education grant project. Its main goal was to obtain authentic information on necessary bandwidth, volumes of stored data, experience with image processing, image information system features, possibilities and limitations of connecting various information resources (US, CT, MR, etc.).

The second part, funded by the Ministry of Health care grant agency, was oriented towards obtaining information about separate modalities' requirements for bandwidth and response time during routine operation, radiologists' experience with display monitors including the selection of a suitable viewer for education purposes, etc. Gradually, all of the Brno hospitals joined this project:

they are connected to the central PACS archive via dedicated fibre optic lines in a star-shaped topology for security and speed. This topology diagram is shown in Figure 24.2.

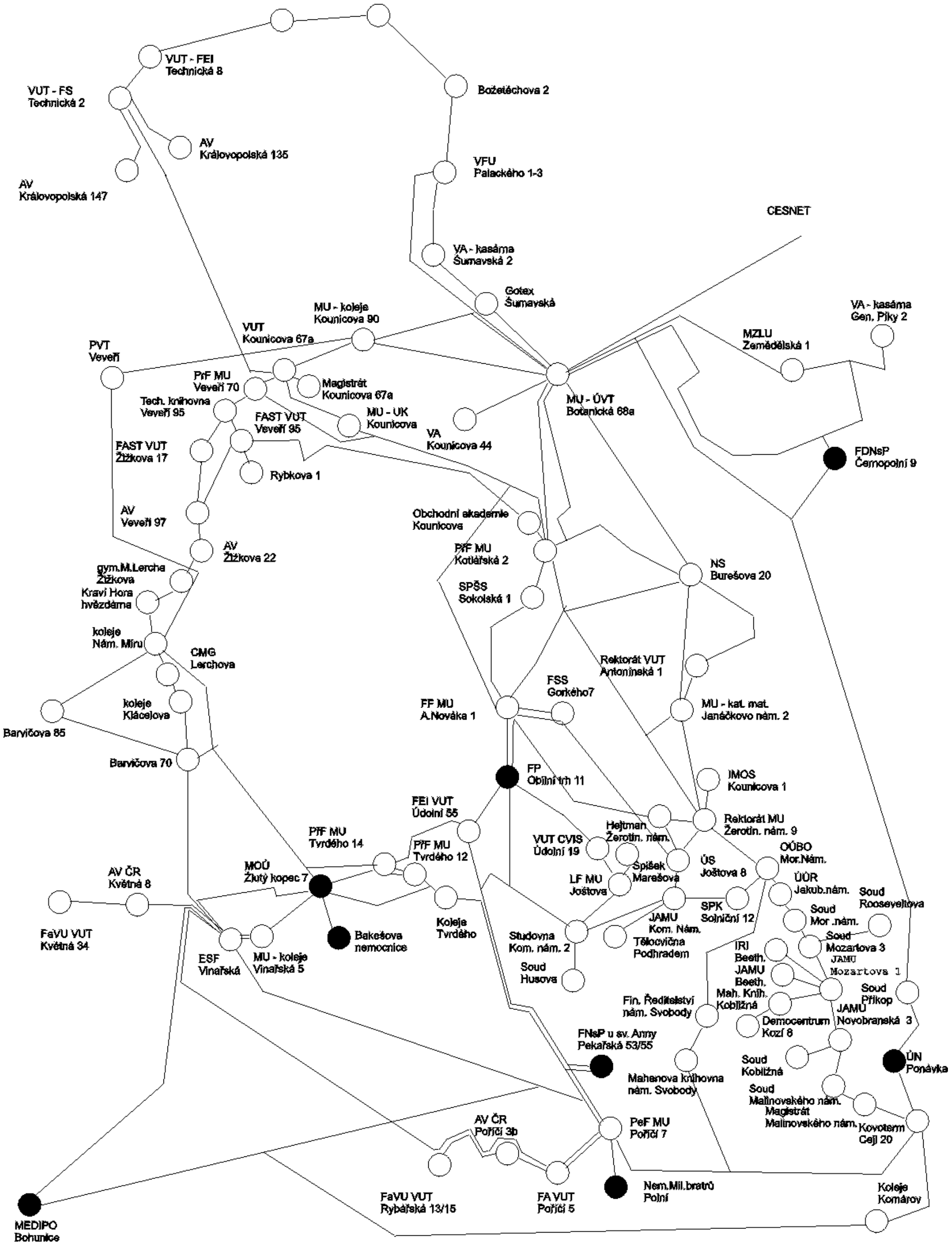


Figure 24.2: The Brno Optical Academic Computer Network

Currently, CESNET, z. s. p. o. is significantly participating in these activities. The main reason is that other organisations outside of the Brno region connect to the metropolitan PACS archive. This brings rather high requirements on the quality of data links as well as several other problems which our project must solve. The goal is getting experience with transferring large data volumes among various modalities and remote locations over public data network where appropriate security measures must be taken using encryption and other technologies.

24.2.2 Extending beyond the Brno region

We started connecting the extra-Brno participants in 2003. The principal problem for remote hospitals is getting sufficiently fat data pipes for communication with the central PACS servers located in Brno.

The optimum solution is using the CESNET2 National Research and Education Network. This network provides sufficient bandwidth for transferring medical images. The only problem remaining is the security of transferred data. To solve this, we decided to use IPSEC tunnelling which is implemented using the Cisco PIX firewalls. On the MU ICS side, a PIX 525 is used to terminate the IPSEC tunnels from remote hospitals; it is connected to the firewall separating the PACS servers from the rest of the network. On the remote hospital side, a PIX 515E is used. Its functions are the IPSEC tunnel termination as well as network address translation and filtering the traffic between the hospital and PACS. Currently, the 3DES encryption algorithm is used; in near future, the AES algorithm should be used. A general diagram of extra-Brno hospital connection is shown in Figure 24.3.

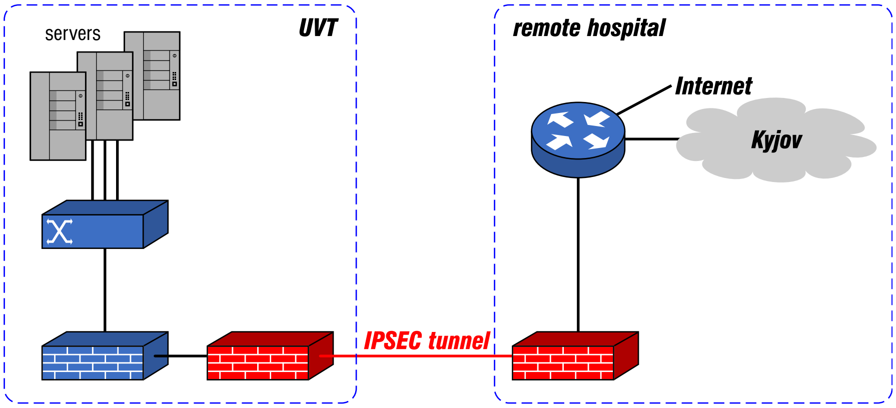


Figure 24.3: Interconnection diagram of extra-Brno hospitals

In the beginning of 2003, a hospital in Kyjov, a South Moravian town, was connected to the PACS via the CESNET2 network at a 10 Mbps speed. At the same time, two display monitors were installed in this hospital allowing viewing images from this hospital's own CT as well as communication with collaborating

Brno hospitals. The Kyjov district hospital is not using our PACS system for long-term archival currently but gradually, it is getting to participate in the interchange system of medical image information exchange.

The experience we had gained while connecting the Kyjov hospital became useful for connecting the hospital in Jihlava, another Moravian town. This hospital is connected via an 8 Mbps leased line. This bandwidth is not quite sufficient because this hospital stores all of its images in the central PACS archive in Brno. Therefore, a faster link using the CESNET2 network is being discussed.

All routine operation of this hospital is dependent on the central PACS system. The Jihlava hospital was equipped gradually with six diagnostic stations which can process the X-ray images as well as interchange this information with hospitals in Brno. Long-term archival is provided for the Jihlava hospital by the PACS system and we are ready to help it with eventual transition to a fully digital image processing.

24.3 Expected development

The work planned for 2004 can be divided to the following areas: connecting more participants, upgrading the network bandwidth and improving the system reliability.

24.3.1 Connecting more participants

Several new hospitals are interested in participation in our PACS system. As soon as an appropriate data link is available, the IKEM health care institute in Prague should be connected using the CESNET2 network. From the security point of view, this connection will be equivalent to using a public data network; therefore, data encryption will be necessary.

The data link used for connecting the Jihlava hospital is of insufficient bandwidth. As the bandwidth upgrade of long-distance lines is expensive, we expect that the Jihlava hospital will also use the CESNET2 network for its connection to the PACS system. The proposed topology is illustrated in Figure 24.4.

24.3.2 Reliability improvement

With respect to growing utilisation of the central PACS system and expected participation of more hospitals and health care organisations in Brno and elsewhere, improved reliability of both the PACS system itself and networking infrastructure is necessary. Reliability of the PACS system itself will be improved by building a backup PACS centre: this backup site will be situated in the Comenius Square

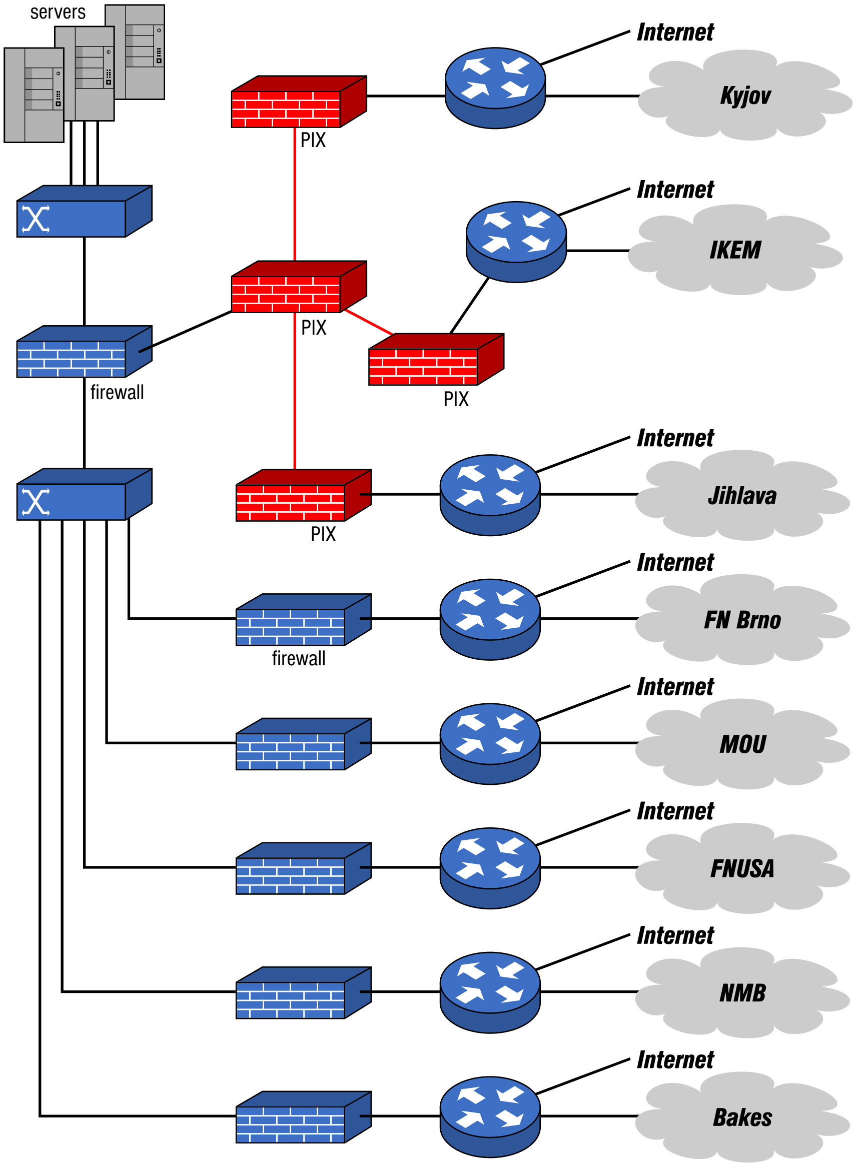


Figure 24.4: Connection of individual project participants

where the university central computer hall open 24 hours is located. To minimise system downtime caused by software failures, the backup PACS system will run a different software. Data interchange between the primary and backup centre will be based on the DICOM standard.

To improve reliability of networking infrastructure, we plan to build independent fibre optics links from all Brno hospitals to both the primary and backup centres. This solution is possible owing to our own large fibre optics network. The planned topology is shown in Figure 24.5.

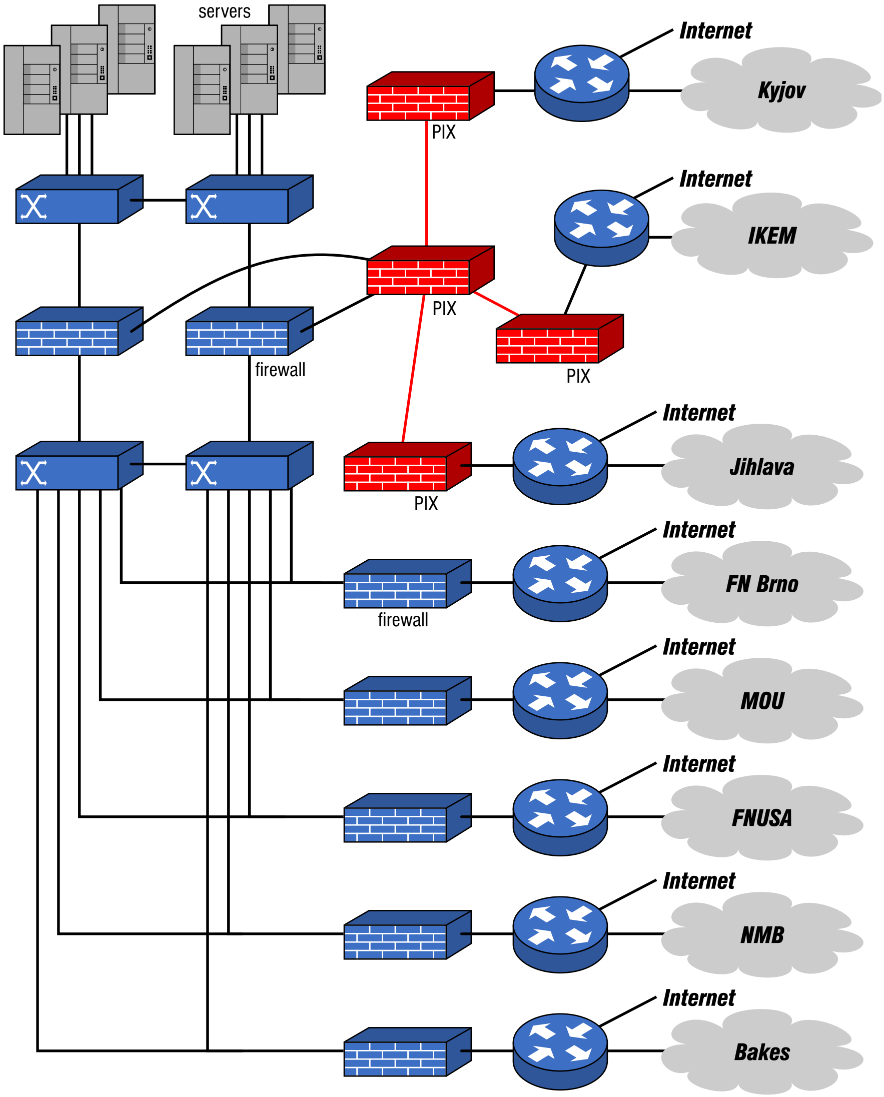


Figure 24.5: Logical network topology after the backup centre is built

The situation of health care institutions outside Brno is much more difficult. Deployment of two independent leased data links of sufficient bandwidth is too expensive. This is why we plan testing the backup capabilities of a number of dialup lines where traffic will be split in parallel. This solution will require also some kind of control provisions: as the backup connection will provide a smaller bandwidth than that of the primary data link, a suitable method for image transfer must be found so that this data will be available on a local caching system before the medicians need it.

24.3.3 New applications

In cooperation with the *Multimedia transmissions* project we plan to study the feasibility of on-line voice and video communications of the PACS users. We assume that we will be able to support remote consultations concerning the images stored in PACS.

24.4 Results

All of the above mentioned activities allow or will allow increasing the number of health care organisations participating in the PACS system. This will increase the amount of stored information for scientific and research tasks and of course for improvement of medical student education.

Lessons drawn from selected applications and general solutions applied in the *MeDiMed* project have been published and presented at domestic and foreign conferences. Publications are listed in the Appendix of this report.

Part V

Conclusion and Annexes

25 Conclusion

2003 is the last year of solving the current research plan; therefore, already by the end of 2002, the CESNET Association management dealt with a question of how to ensure financing of its main activity, i.e. research and development in the field of information and communication technologies, in the coming years. The Ministry of Education, Youth and Sports issued a long-awaited call for submitting proposals of research plans for the period 2004–2008 (or 2010). Therefore, on 27 February 2003, the CESNET Association submitted its proposal identification code MSM6383917201 for a seven-year research plan *Optical National Research Network and Its New Applications*.

The proposal was based on an evaluation of current status of world research networks and on a prognosis of development in this field, although the detailed plans were elaborated only for years 2004–2007 with regard to a fast development in this field.

It has been found recently that using optical fibres and lambdas on a national and global scale is strategically important for development of certain important science and development fields. It allows carrying out some projects which could not be realised in time and for adequate expenses using regular telecommunication carrier services. Some telecommunication companies (e.g., LEVEL3) react to this situation by leasing fibers or by enabling NRENs to participate in laying down the optical cables.

In 2003, research and educational networks characterised by either experimental or production features have been evolving rather differently. Production networks (e.g., Géant, Abilene, ACONET and CESNET2) provide services for research and education communities. Provision of network services for research and education has specific features – the network is not built and operated to make profit. Therefore, it can provide those services necessary for development of research and education in various fields which common ISPs do not provide because these services would either bring no profit to the providers or they would be unacceptably expensive. Currently, this is the case, e.g., for gigabit local links.

At the same time, success of these networks brings a new problem: users in many countries create pressure to make the NREN services, originally of a research and experimental character, as stable and reliable as those provided by the best ISPs – but still for a low price and on a higher technical level. It means that the original opportunity for research and experiments in the field of information and communication technologies (and especially of large-scale computer networks) would be significantly reduced or totally eliminated. This brings a possible risk: in the future, production networks would differ from the ISP networks especially by their non-profitability; this would bring the only advantage to their users rather than any differences in their technical level (the same routers and switches could be deployed only after the development of hardware and software terminates and these are supplied to production networks only if their profitable deployment among ISPs is expected).

Certain improvement of the situation can be reached by selecting some parts of the network or services to have a production character and some others to have a research and experimental character. However, this attitude requires a rather demanding coordination work in the project phase, in implementation and operation of the network; it is also very demanding on the qualification of the network operators. Another alternative is using one network as a production network (e.g., SURFNET5) and building another one (e.g., SURFNET6).

An independent construction of experimental networks brings even better possibilities for network research and development. This method has been used for a longer time (see, e.g., the production and experimental CENIC networks in California), but it has significantly expanded during 2003. Similar federal NationalLightRail and NationalLambdaRail projects as well as subsequent projects financed by individual states (e.g., FloridaLambdaRail) originated in the United States. A global experimental network TransLight is emerging; CESNET is one of its participants.

The goal of the research plan *Optical National Research Network and Its New Applications* is a design of an integrated network environment suitable for specific requirements of the academic community and verification of its characteristics in actual operation. Furthermore, experience gained by operating the academic networks shows that having sufficient free bandwidth is only one of the demands put on the academic network; implementing other advanced services to operate a high-quality academic network is also necessary. As a result, the research team will focus, apart from the research in the field of infrastructure and network protocols, also on the fields of applications and network services (so-called middleware) linking the application and network layers.

The goals and basic strategies of the research plan being prepared were consulted with important foreign experts who accepted the invitation by CESNET Association and participated in the meeting within the CESNET Association premises on February 20, 2003. The goals of the new research plan resulted from this discussion so as to comply with the world trends in the field of information and communication technologies.

For proposing and constructing the new generation infrastructure, we will focus on using optical technologies while emphasizing the use of leased optical fibers fully controlled by us, equipped with our own devices and capable of providing more channels in each fiber. The research subject will also include studies of building long-distance intercity optical routes without optical regenerators on the line. In the field of routing, development and deployment of PC-based gigabit routers is expected; network migration to the IPv6 protocol is also planned, including applications and services. Naturally, both protocols (the current IPv4 and the coming IPv6) will be operated simultaneously for some time.

The research team also wants to make the most of the opportunities resulting from our participation in the TransLight network and a very affordable cost of leasing the optical fibre in the Czech Republic. The key issue will be solving the optical transmission system for TransLight in the Czech Republic and enlarging the experimental application of the TransLight network.

In the field of applications, we intend to focus especially on the grid development, i.e., an environment for cooperation of distributed entities, whether they are people, groups of people or machines. The following belong among grids:

- computing grid for demanding scientific calculations consisting of a large number of geographically distributed computers
- storage grid – environment for data saving and accessing remotely
- access grid – standardized environment for a cooperation using videoconference and multimedia applications and tools for cooperation on shared documents.

Another area of interest will be the development of IP telephony, videoconferencing tools and tools for streaming multimedia contents. We intend to pay a lot of attention to the distance learning matters.

Research in the area of network services as links between infrastructure and applications will include:

- development of instruments for monitoring and evaluation of network operation,
- development of instruments for monitoring network performance characteristics and tools for their optimisation,
- research in the field of authentication and authorization mechanisms to access the network resources.

The involvement of our experts in international activities, especially in the 6th Framework Program projects, will be an integral part of our research plan. The CESNET Association's research team currently takes part in several international projects of the 5th Framework Program – it has gained its prestige by contributing to their solution. Therefore, the CESNET Association does not have to beg to be accepted into the consortiums of newly submitted projects – on the contrary, it is asked to take part in them.

In order to fulfill all these goals, the planned capacity of the research team is significantly extended in comparison to the previous research plan. We expect to use the project management method which proved useful during the solution of the current research plan.

A List of Connected Institutions

A.1 CESNET Members

institution	*connection [Mbps]*
Academy of Performing Arts in Prague	100
Academy of Sciences of the Czech Republic	1000
Academy of Fine Arts in Prague	10
Czech University of Agriculture in Prague	1000
Czech Technical University in Prague	1000
Janáček Academy of Musical and Dramatic Arts in Brno	1000
University of South Bohemia in České Budějovice	1000
Masaryk University in Brno	1000
Mendel University of Agriculture and Forestry in Brno	1000
University of Ostrava	1000
Silesian University in Opava	100
Technical University of Ostrava	1000
Technical University in Liberec	1000
University of Hradec Králové	1000
University of Jan Evangelista Purkyně in Ústí nad Labem	1000
Charles University in Prague	1000
Palacký University in Olomouc	1000
University of Pardubice	1000
Tomáš Baťa University in Zlín	1000
University of Veterinary and Pharmaceutical Sciences in Brno	1000
Military Academy in Brno	100
Purkyně Military Medical Academy in Hradec Králové	1000
Institute of Chemical Technology in Prague	1000
University of Economics in Prague	1000
Academy of Arts, Architecture and Design in Prague	100
Military College of Ground Forces in Vyškov	34
Brno University of Technology	1000
University of West Bohemia in Plzeň	1000

A.2 The Most Important Connected Institutions of Research and Education

institution	*connection [Mbps]*
Technical and Test Institute for Constructions Praha	0.128
State Technical Library in Prague	10
National Library of The Czech Republic	155
Moravian Library in Brno	10
Research Library in Hradec Králové	10
Research Library of South Bohemia in České Budějovice	100
Research Library of North Bohemia in Ústí nad Labem	2
Tábor Public Library	10
Masaryk Hospital in Ústí nad Labem	34
Research Institute of Geodesy, Topography and Cartography	2
Nuclear Research Institute Řež	2
Observatory and Planetarium of Prague, Štefánik Observatory Centre	0.064
General University Hospital	10
University Hospital Olomouc	100
Research Institute of Agricultural Economics	10
Na Homolce Hospital	2
Moravian-Silesian Research Library in Ostrava	2
Research Library in Liberec	10
New York University in Prague	0.512
Food Research Institute Prague	2
University Hospital Brno	2
Research Library in Olomouc	10
University Hospital Bulovka	10
Institute for Information on Education	10
University Hospital in Hradec Králové	155
Higher Professional School of Information Services in Prague	10
University Hospital Královské Vinohrady	10
Hospital Liberec	34
Anglo-American College	2
University Hospital Plzeň	155
Jiří Mahen Library in Brno	10
University Hospital u Svaté Anny in Brno	10
Traumatological Hospital in Brno	10
Private Higher Professional School of fiscal counselling STING	0.256

Research Institute for Organic Syntheses	2
Research Institute for Veterinary Medicine	0.512
University of New York in Prague	2
Pilsen City Library	10
Observatory and Planetarium in Plzeň	10
Central Military Hospital Praha	34
Masaryk Institute for Oncology	4
CARITAS – Higher Professional Social School in Olomouc	10
District Library in the town of Olomouc	1
Municipal Library in Prague	10
University Hospital with Policlinic Ostrava	10
Moravian Museum	2
Kerio Technologies	100
Hospital and Ambulance in Kutná Hora	1
District Hospital Kyjov	1
Science and Technology Park Ostrava	1
District Library in the town of Hradec Králové	1
Eastern Bohemia Museum in Hradec Králové	1
City Museum in Dvůr Králové nad Labem	1
Milosrdní Bratři Hospital	1
Bakeš Surgery Hospital	0.256
Research Institute for Vegetal Industry in Olomouc	1
Josef Podsedník Higher Medical School in Brno	0.512
Centre of Traffic Research	1
BIC Plzeň – Science and Technology Park	10
J. A. Komenský University	2
Museum of Arts, Architecture and Design in Prague, Library	0.064
Centre of Cardiovascular and Transplant Surgery Brno	4
Centre of University Studies	10
University Hospital in Motol	10
Karel Engliš University in Brno	0.256
National Institute for Professional Education	4
Hospital Děčín	2
TESTCOM – Technical Centre of Telecommunications and Post Prague	10
LOM Prague – Research Technical Institute for Aviation	10
City Library in Cheb	1
Higher Professional School of Economy in Liberec	100

B List of Researchers

Adamec Petr	Technical University in Liberec
Altmannová Lada, Ing.	CESNET
Andres Pavel, MUDr.	Masaryk Memorial Cancer Institute
Andrš Jindřich, Ing.	FPh Charles University in Hradec Králové
Aster Jaroslav	Masaryk University
Bartoníček Tomáš, Ing.	University Pardubice
Bartoňková Helena, MUDr.	Masaryk Memorial Cancer Institute
Boháč Leoš, Ing.	Czech Technical University
Buchta Martin	University of Economics Prague
Burian Jiří	DELTAx Systems
Cápík Jaromír	Brno University of Technology
Cimbal Pavel, Ing.	Czech Technical University
Crha Luděk	Brno University of Technology
Čejka Rudolf, Ing.	Brno University of Technology
Černý Martin, Ing.	CESNET
Černý Pavel, Ing.	Czech Academy of Sciences
Denemark Jiří	Masaryk University
Diviš Zdeněk, Prof. Ing., CSc.	Technical University of Ostrava
Doležal Ivan, Ing. BcA.	Technical University of Ostrava
Dostál Otto, Ing., CSc.	Masaryk University
Dušek Václav, Ing.	University Pardubice
Fabuš Pavol	Masaryk University
Faltýnek Pavel	Brno University of Technology
Fanta Václav, Ing., CSc.	Institute of Chemical Technology
Friedl Štěpán	Brno University of Technology
Frýda Michal, Ing.	Czech Technical University
Fučík Otto, Dr. Ing.	Brno University of Technology
Furman Jan, Ing.	CESNET
Gruntorád Jan, Ing., CSc.	CESNET
Haluza Jan, Ing.	Technical University of Ostrava
Hažmuk Ivo, Ing.	Brno University of Technology
Hladká Eva, RNDr.	Masaryk University
Hofer Filip	Masaryk University
Holeček Jáchym	Masaryk University
Holeček Jan	Masaryk University
Holub Petr, Mgr.	Masaryk University
Holý Radek, Ing.	Charles University
Hrad Jaromír, Ing.	Czech Technical University
Hrb Jan	–
Hulínský Ivo	CESNET
Indra Miroslav, Ing., CSc.	Czech Academy of Sciences

Jasiok Lumír, Bc.	Technical University of Ostrava
Javorník Michal, RNDr.	Masaryk University
Kácha Pavel	CESNET
Kalix Igor, Ing.	Regional Hospital Kyjov
Karásek Miroslav, Ing., DrSc.	Czech Academy of Sciences
Kmoch David, Mgr.	Technical University in Liberec
Kňourek Jindřich, Ing.	University of West Bohemia
Komárková Jitka, Ing., Ph.D.	University Pardubice
Kopecký Dušan	University Pardubice
Košňar Tomáš, Ing.	CESNET
Kouřil Daniel, Mgr.	Masaryk University
Král Antonín	PRAGONET
Kratochvíla Tomáš	Masaryk University
Kretschmer Petr, Ing.	Technical University in Liberec
Kropáčová Andrea	CESNET
Krsek Michal, Bc.	–
Krupa Petr, Doc. MUDr., CSc.	University Hospital Brno
Krutý Petr	Masaryk University
Křenek Aleš, Mgr.	Masaryk University
Kubec František, Ing., CSc.	Czech Academy of Sciences
Ledvinka Jaroslav, Ing.	Masaryk University
Lhotka Ladislav, Ing., CSc.	CESNET
Lipovčan Marek	Masaryk University
Liška Miloš	Masaryk University
Lokajíček Miloš, RNDr., CSc.	Czech Academy of Sciences
Marek Jan, Ing	University of South Bohemia
Martínek Tomáš	Brno University of Technology
Mašek Josef, Ing.	University of West Bohemia
Matuška Miroslav	University of Economics Prague
Matyska Luděk, Doc. RNDr., CSc.	Masaryk University
Míchal Martin, Ing.	CESNET
Mikušek Petr	Brno University of Technology
Minaříková Kateřina	Masaryk University
Mulač Miloš	Masaryk University
Nejman Jan, Ing.	CESNET
Neuman Michal, Ing.	Czech Technical University
Novák Petr	Masaryk University
Novák Václav, Ing.	CESNET
Novakov Ivan	Czech Technical University
Novotný Jiří, Ing.	Masaryk University
Pazdera Jan	Brno University of Technology
Pospíšil Jan, Ing.	University of West Bohemia
Pospíšil Marek	Masaryk University

Potužník František, Mgr.	Charles University
Pustka Martin, Ing.	Technical University of Ostrava
Radil Jan, Ing.	CESNET
Rebok Tomáš	Masaryk University
Rohleder David, Mgr.	Masaryk University
Roškot Stanislav, Ing.	Czech Technical University
Ruda Miroslav, Mgr.	Masaryk University
Růžička Jan, Ing.	CESNET
Rybka Tomáš	Masaryk University
Řehák Vojtěch	Masaryk University
Salvet Zdeněk, Mgr.	Masaryk University
Satrapa Pavel, RNDr., Ph.D.	Technical University in Liberec
Sitera Jiří, Ing.	University of West Bohemia
Slavíček Karel, Mgr.	Masaryk University
Slíva Roman, Ing.	Technical University of Ostrava
Smítka Jiří, Ing.	Czech Technical University
Smotlacha Vladimír, Ing. RNDr.	CESNET
Sova Milan, Ing.	CESNET
Staněk Filip	Technical University of Ostrava
Studený Daniel, Ing.	CESNET
Sverenyák Helmut, Ing.	CESNET
Šafránek David	Masaryk University
Šíma Stanislav, Ing., CSc.	CESNET
Šimák Boris, Doc. Ing., CSc.	Czech Technical University
Šimeček Pavel	Masaryk University
Šimek Pavel, Ing.	University of West Bohemia
Škrabal Jiří, Ing.	Masaryk University
Šmejkal Ivo, Ing.	University of Economics Prague
Šmrha Pavel, Dr. Ing.	University of West Bohemia
Švábenský Mojmír, MUDr.	Regional Hospital Kyjov
Tobola Jiří	Brno University of Technology
Tomášek Jan, Ing.	CESNET
Třeštík Vladimír, Ing.	CESNET
Ubik Sven, Dr. Ing.	CESNET
Vachek Pavel, Ing.	CESNET
Vaněk Milan	Škoda Auto
Verich Josef, Ing.	Technical University of Ostrava
Veselá Bohumila, Ing.	University of Economics Prague
Veselá Soňa, Ing.	CESNET
Víšek Jan, Mgr.	Charles University
Vlášek Jakub	University of West Bohemia
Voců Michal, Mgr.	Charles University
Vodrážka Jiří, Ing., Ph.D.	Czech Technical University

Vojtěch Josef, Ing.	Czech Technical University
Vozňák Miroslav, Ing.	Technical University of Ostrava
Wimmer Miloš, Ing.	University of West Bohemia
Zeman Tomáš, Ing., Ph.D.	Czech Technical University
Zemčík Pavel, Doc. Dr. Ing.	Brno University of Technology
Zloský Ondřej	Masaryk University
Žádník Martin	Brno University of Technology

C Own Publishing Activities

C.1 Standalone Publications

Team of authors: *High-speed National Research Network and its New Applications*. CESNET, 2003, ISBN 80-238-0166-4

C.2 Contributions in Proceedings and other Publications

Antoš D., Kořenek J., Řehák V.: *Vyhledávání v IPv6 směrovači implementovaném v hradlovém poli (Lookup in an IPv6 router implemented in a gate array)*.
in proceedings *EurOpen, Sbornik prispevku XXIII. konference*, EurOpen, 2003, page 91–102, ISBN 80-86583-04-X

Antoš D., Novotný J.: *Liberouter, IPv6 routeri (Liberouter, IPv6 Router)*.
in proceedings *Širokopásmové sítě a jejich aplikace*, Univerzita Palackého v Olomouci, 2003, page 75–81, ISBN 80-244-0642-X

Denemark J., Hladká E.: *Bezpečnost v aktivních sítích (Security in Active Networks)*.
in proceedings *Širokopásmové sítě a jejich aplikace*, Univerzita Palackého v Olomouci, 2003, page 138–142, ISBN 80-244-0642-X

Dostál O., Javorník M., Slavíček K., Petrenko M.: *MEDIMED – Regional Centre for Archiving and Interhospital Exchange of Medicine Multimedia Data*.
in proceedings *Proceedings of the Second IASTED International Conference on Communications, Internet, and Information Technology*, International Association of Science and Technology for Development – IASTED, 2003, page 609–614, ISBN 0-88986-398-9

Dostál O., Slavíček K., Petrenko M.: *Využití vysokorychlostní sítě pro potřebu medicínských aplikací (Usage of a High-speed Network for Medical Applications)*.
in proceedings *Širokopásmové sítě a jejich aplikace*, Univerzita Palackého v Olomouci, 2003, page 70–74, ISBN 80-244-0642-X

Hladká E., Denemark J.: *On-Demand Secure Collaborative Support*.
in proceedings *Italian Association in Telemedicine and Medical Informatics*, ITIM 2004, 2003, page 50–50

Hladká E., Holub P., Rebok T.: *Komunikace s technologií AccesGrid Point (Communication using the AccessGrid Technology).*
in proceedings *Širokopásmové sítě a jejich aplikace*, Univerzita Palackého v Olomouci, 2003, page 65–69, ISBN 80-244-0642-X

Hladká E., Liška M.: *Přednášky ze záznamu (Lectures from the Recordings).*
in proceedings *Širokopásmové sítě a jejich aplikace*, Univerzita Palackého v Olomouci, 2003, page 130–133, ISBN 80-244-0642-X

Höfer F.: *Vývoj nanoprogramů pro procesory implementované v hradlovém poli (Development of nanoprograms for processors implemented in a gate array).*
in proceedings *Sborník příspěvků z XXIII. konference EurOpen.CZ*, EurOpen.CZ, 2003, page 103–110, ISBN 80-86583-04-X

Hrad J., Zeman T.: *Infrastructure for Support of Education.*
in proceedings *ICETA 2003 – Conference Proceedings*, IEEE, 2003, page 251–254, ISBN 80-89066-67-4

Hrad J., Zeman T.: *New Trends in Designing of WWW Portals for Support of Education.*
in proceedings *COFAX – Telekomunikácie 2003*, D&D STUDIO, s. r. o., 2003, page 139–142, ISBN 80-967019-4-0

Hrad, J., Zeman, T.: *Portál pro podporu on–line výuky v prezenční formě studia (Portal for Support of On-Line Education for Full-Time Study).*
in *eLearning ve vysokoškolském vzdělávání 2003 – Sborník příspěvků*, Univerzita Tomáše Bati, 2003, page 64–69, ISBN 80-7318-138-X

Hrad J., Zeman T.: *Streaming of Lectures? Advance or Loss?.*
in proceedings *VIEWDET 2003*, Vienna University of Technology, 2003, page L13–L13, ISBN 3-85465-013-2

Karásek M., Radil J., Boháč L.: *Optimalizace přenosu NRZ dat s rychlostí 10 Gbit/s po vláknech G.652 bez linkových zesilovačů: simulace a experiment (Optimisation of NRZ Data Transport on 10 Gbps over G.652 Fibres without In Line Aplifiers: Simulation and Experiment).*
in proceedings *Optické komunikace 2003*, 2003, page 77–82, ISBN 80-86742-03-2

Karásek M., Radil J., Boháč L.: *Transmission limits of 10Gbit/s NRZ data over.*
in proceedings *Proceedings of the 7th Conference on Telecommunications*, 2003, page 765–768, ISBN 953-184-052-0

Křenek A., Peterlík I., Matyska L: *Building 3D State Spaces of Virtual Environments with a TDS-based Algorithm.*
in proceedings *Recent Advances in Parallel Virtual Machine and Message Passing Interface*, Springer–Verlag, 2003, page 529–536, ISBN 3-540-20149-1

Lhotka L.: *Netopeer – konfigurační systém pro směrovače a sítě IP (Netopeer – a configuration system for IP routers and networks).*
in proceedings *Sborník příspěvků XXIII. konference EurOpen*, EurOpen.CZ, 2003, page 117–127, ISBN 80-86583-04-X

Novotný J., Fučík O., Antoš D.: *Liberouter – New Way in IPv6 Routers.*
in proceedings *ICETA 2003 2nd International Conference Proceedings*, elfa, Košice, 2003, page 153–158, ISBN 80-89066-67-4

Novotný J., Fučík O., Antoš D.: *Project of IPv6 Router with FPGA Hardware Accelerator.*
in proceedings *Field-Programmable Logic and Applications, 13th International Conference FPL 2003, Proceedings*, Springer Verlag, 2003, page 964–967, ISBN 3-540-40822-3

Radil J., Boháč L., Karásek M.: *Optically Amplified Multigigabit Links in CESNET2 network.*
in proceedings *TERENA Networking Conference*, TERENA Organization, 2003

Růžička J., Ubik S.: *VoIP service in CESNET network.*
in proceedings *PIONIER 2003 Polski internet opticzny: technologie,uslugi i aplikacje*, Poznanske Centrum Superkomputerowo–Sieciowe, 2003, page 43–47, ISBN 83-913639-4-5

Vozňák M.: *Comparison of H.323 and SIP Protocol Specification.*
in proceedings *Research in Telecommunication Technology 2003*, FEI STU Bratislava, 2003, page 45, ISBN 80-227-1934-X

Vozňák M.: *Hlasové služby v síti CESNET2 (Voice Services in CESNET2 Network).*
in proceedings *EurOpen XXIII. konference*, EurOpen, Prague, 2003, page 85, ISBN 80-86583-04-X

Vozňák M.: *Využití IP telefonie v síti CESNET2 (Usage of IP Telephony in CESNET2 network).*
in proceedings *Širokopásmové sítě a jejich aplikace*, Univerzita Palackého v Olomouci, 2003, page 44, ISBN 80-244-0642-X

Zeman T., Hrad J.: *Smart WWW Portal for Support of Education.*
in proceedings *International Conference on Engineering Education Proceedings*, Universidad Politécnica de Valencia, 2003, page 4428–4428, ISBN 84-600-9918-0

Zeman, T., Hrad, J.: *Streaming přednášek (Streaming of Lectures).*
in *eLearning ve vysokoškolském vzdělávání 2003 – Sborník příspěvků*, Univerzita Tomáše Bati, 2003, page 261–266, ISBN 80-7318-138-X

Zeman, T., Hrad, J.: *Universities and Streaming of Lectures.*
in proceedings *COFAX – Telekomunikácie 2003*, D&D STUDIO, s. r. o., 2003, page 305–306, ISBN 80-967019-4-0

Zeman T., Hrad J.: *WWW and Streaming – New Ways in Education*.
in proceedings *RTT 2003 – Proceedings*, FEI, Slovak University of Technology, 2003, page 259–262, ISBN 80-227-1934-X

Zeman, T., Hrad, J., Šimák, B.: *Support of Online Education*.
in proceedings *14th EAEEIE International Conference on Innovations in Education for Electrical and Information Engineering*, PTETiS, 2003, page A57–A57, ISBN 83-918622-0-8

Zeman, T., Vodrážka, J., Hrad, J.: *What Do Students Expect from Universities?*.
in proceedings *COFAX – Telekomunikácie 2003*, D&D STUDIO, s. r. o., 2003, page 303–304, ISBN 80-967019-4-0

C.3 Articles in Specialized Magazines

Hladká, E., Holub, P.: *Jak se dělá divadlo na FI MU (How Theatre is Made on FI MU)*.
in journal *Zpravodaj Ústavu výpočetní techniky Masarykovy univerzity v Brně*, number 5, 13, page 4, ISSN 1212–0901

Hladká E., Liška M.: *Přednášky ze záznamu na FI MU (Lectures from the Recordings on FI MU)*.
in journal *Zpravodaj Ústavu výpočetní techniky Masarykovy univerzity v Brně*, number 4, 13, page 3, ISSN 1212–0901

Hladká E., Matyska L.: *Aplikace nad vysokorychlostními sítěmi (Applications over High-speed Networks)*.
in journal *IT-Net*, number 3, 3, page 3, ISSN 1212–6780

Krsek M.: *Konferujte po drátech (Confer over Wires)*.
in journal *Ekonom*, number 22, 2003, page 1, ISSN 1210–0714

Matyska, L., Holub, P., Hladká, E.: *Jak přednášet na dálku aneb virtuální přednáška v Koreji (How to Lecture Remotely or Virutal Lesson in Korea)*.
in journal *Zpravodaj Ústavu výpočetní techniky Masarykovy univerzity v Brně*, number 1, 14, page 3, ISSN 1212–0901

Radil J., Boháč L., Karásek M.: *Optical networking in CESNET2 gigabit network*.
in journal *Annales of telecommunications*, number 11–12, 58, page 1

Ubik S.: *Přenosy velkých objemů dat v rozlehlých Gigabitových sítích (Transmissions of Large Data Volumes in Wide Gigabit Networks)*.
in journal *Sdělovací technika*, number 5, 2003, page 1, ISSN 0036–9942

C.4 Technical Reports

Antoš D., Kořenek J., Minaříková K., Řehák V.: *Packet header matching in Combo6 IPv6 router*.
technical report number 1/2003, CESNET, 2003

Bartoň J.: *Performance Testing Tools*.
technical report number 18/2003, CESNET, 2003

Denemark J., Holub P., Hladká E.: *RAP - Reflector Administration Protocol*.
technical report number 9/2003, CESNET, 2003

Doležal I., Illich M., Krsek M.: *Vyhledávání v multimediálních datech na Internetu / Internet search in multimedia data*.
technical report number 19/2003, CESNET, 2003

Frýda M.: *Architektura portálu eLearning (Architecture of the eLearning Portal)*.
technical report number 6/2003, CESNET, 2003

Hejtmánek L., Holub P.: *IBP Deployment Tests and Integration with DiDaS Project*.
technical report number 20/2003, CESNET, 2003

Höfer F.: *Packet Analysis for IPv6 Router Implemented by a PCI Acceleration card*.
technical report number 5/2003, CESNET, 2003

Jasiok L., Krsek M.: *Stav podpory protokolu IPv6 platformách pro proudování multimédií (Status of IPv6 Support in Multimedia Streaming Platforms)*.
technical report number 4/2003, CESNET, 2003

Kratochvíla T., Řehák V., Šimeček P.: *Verification of COMBO6 VHDL Design*.
technical report number 17/2003, CESNET, 2003

Krsek M.: *IPv6 proudování Megaconference V (Megaconference V IPv6 Streaming)*.
technical report number 27/2003, CESNET, 2003

Kutina T., Svoboda J., Soudek L., Lešner M., Kněz P., Čížek P., Sobotka A., Stánek P., Krsek M., Lederbuch P.: *Testování kodérů MPEG4 – výsledky projektu (MPEG4 Encoders Testing – the Results)*.
technical report number 8/2003, CESNET, 2003

Lhotka L.: *XML schema for router configuration data: An annotated DTD*.
technical report number 2/2003, CESNET, 2003

Matuška M.: *IOSConvert: IOS to XML router configuration file converter – command reference*.
technical report number 30/2003, CESNET, 2003

Michalík P., Šmejkal I., Veselá B.: *Použití digitální nelineární střižny pro tvorbu multimediálních výukových lekcí (Use of Digital Non-Linear Editing Machinery for Creating Multimedia Education Lessons).*
technical report number 11/2003, CESNET, 2003

Novák P.: *XML parser API library for Netopeer.*
technical report number 22/2003, CESNET, 2003

Novotný J., Fučík O., Bardas R.: *Schematic of COMBO-4SFP Card.*
technical report number 12/2003, CESNET, 2003

Novotný J., Fučík O., Bardas R.: *Schematic of COMBO-4MTX Card.*
technical report number 13/2003, CESNET, 2003

Novotný J., Smotlacha V., Bardas R.: *Schematic of COMBO-PTM.*
technical report number 15/2003, CESNET, 2003

Novotný J., Šíma S., Bardas R.: *Schematic of COMBO-BOOT card.*
technical report number 14/2003, CESNET, 2003

Rebok T., Holub P.: *Synchronizing RTP Packet Reflector.*
technical report number 7/2003, CESNET, 2003

Sitera J.: *Aktuální stav LDAPu MetaCentra (Current Status of MetaCentrum LDAP).*
technical report number 29/2003, CESNET, 2003

Staněk F., Haluza J.: *Protokol HyperSCSI (HyperSCSI Protocol).*
technical report number 23/2003, CESNET, 2003

Šmejkal I., Veselá B.: *Malé digitální studio (Small Digital Studio).*
technical report number 16/2003, CESNET, 2003

Ubik S.: *Field trial with Intel 10 Gigabit Ethernet adapters for PC.*
technical report number 10/2003, CESNET, 2003

Ubik S., Klaban J.: *Experience with using simulations for congestion control research.*
technical report number 26/2003, CESNET, 2003

Ubik S., Král A.: *End-to-end Bandwidth Estimation Tools.*
technical report number 25/2003, CESNET, 2003

Ubik S., Vojtěch J.: *QoS in Layer 2 Networks with Cisco Catalyst 3550.*
technical report number 3/2003, CESNET, 2003

Vozňák M., Neuman M.: *The Peering with CESNET VoIP Network.*
technical report number 28/2003, CESNET, 2003

Wimmer M.: *Vysílání a přenos audio signálu ve velmi vysoké kvalitě do sítě Internet (Broadcasting and Transport of Very High Quality Audio Signal over Internet).*
technical report number 24/2003, CESNET, 2003

Zloský O.: *API description for Netopeer repository library.*
technical report number 21/2003, CESNET, 2003

C.5 On-line Publications

Krsek M.: *Digitální TV – mnoho zakopaných psů? (Digital TV – too many problems?).*
on server *Lupa*, 18. 11 2003, ISSN 1213–0702

Krsek M.: *Maminko, co je to ten eNum? (Mom, what's the eNum?).*
on server *Lupa*, 29. 7. 2003, ISSN 1213–0702

Krsek M.: *Má distribuce obsahu budoucnost? (Has the Contents Distribution some Future?).*
on server *Lupa*, 4. 7. 2003, ISSN 1213–0702

Krsek M.: *Pořádání přímého internetového přenosu I (Making of Live Internet Broadcast I).*
on server *Lupa*, 22. 4. 2003, ISSN 1213–0702

Krsek M.: *Pořádání přímého internetového přenosu II (Making of Live Internet Broadcast II).*
on server *Lupa*, 29. 4. 2003, ISSN 1213–0702

Krsek M.: *Pořádání přímého internetového přenosu III (Making of Live Internet Broadcast III).*
on server *Lupa*, 7.5. 2003, ISSN 1213–0702

Krsek M.: *Potřebujeme DRM? (Do we need DRM?).*
on server *Lupa*, 21. 2. 2003, ISSN 1213–0702

Krsek M.: *Přenos Megaconference V (Megaconference V Broadcast).*
on server *Lupa*, 18. 12. 2003, ISSN 1213–0702

Krsek M.: *Proudování videa – katedrála, nebo bazar? (Video Streaming – Cathedral or Bazaar?).*
on server *Lupa*, 20. 10. 2003, ISSN 1213–0702

Krsek M.: *Třetí vlna Internetu (Third Internet Wave).*
on server *Lupa*, 24. 6. 2003, ISSN 1213–0702

Krsek M.: *Vyhledávání v multimédiích na Internetu (Multimedia Search in Internet).*
on server *Lupa*, 31. 10. 2003, ISSN 1213–0702

Krsek M.: *Zlikvidují ISP výměnné systémy? (Will ISPs Kill the Exchange Systems?).*
on server *Lupa*, 28. 1. 2003, ISSN 1213–0702

Satrapa P.: *Bezpečné DNS (Secure DNS).*
on server *Lupa*, 22. 5. 2003, ISSN 1213–0702

Satrapa P.: *Bezpečné DNS 2 (Secure DNS 2).*
on server *Lupa*, 5. 6. 2003, ISSN 1213–0702

Satrapa P.: *DNS pro IPv6 – konec schizmatu (DNS for IPv6 – the End of the Schizm).*
on server *Lupa*, 23. 10. 2003, ISSN 1213–0702

Satrapa P.: *FiberCo (FiberCo).*
on server *Lupa*, 4. 12. 2003, ISSN 1213–0702

Satrapa P.: *IBP aneb síťové disky úplně jinak (IBP or Totaly Different Networks Disks).*
on server *Lupa*, 23. 1., 2001, ISSN 1213–0702

Satrapa P.: *IPv6 adresy budou jednodušší (IPv6 Addresses will be Simpler).*
on server *Lupa*, 6. 11. 2003, ISSN 1213–0702

Satrapa P.: *IPv6 v evropských akademických sítích (IPv6 in European Academic Networks).*
on server *Lupa*, 17. 4. 2003, ISSN 1213–0702

Satrapa P.: *M6Bone6 (MBone6).*
on server *Lupa*, 19. 6. 2003, ISSN 1213–0702

Satrapa P.: *Mezinárodní konektivita v Evropě (International Connectivity in Europe).*
on server *Lupa*, 20. 2. 2003, ISSN 1213–0702

Satrapa P.: *Nejrychlejší internetový přenos: 38 420,5 terabit-metrů/s (Fastest Internet Transmission: 38,420.5 terabit-meter/s).*
on server *Lupa*, 20. 11. 2003, ISSN 1213–0702

Satrapa P.: *NEPTUNE – Ethernet na dně mořském (NEPTUNE – Ethernet on the Seashore).*
on server *Lupa*, 6. 3. 2003, ISSN 1213–0702

Satrapa P.: *OBS – kříženec optického a elektronického přepínání (OBS – Hybrid of Optical and Electronic Switching).*
on server *Lupa*, 20. 3. 2003, ISSN 1213–0702

Satrapa P.: *OptIPuter (OptIPuter).*
on server *Lupa*, 4. 4. 2003, ISSN 1213–0702

Satrapa P.: *Prestižní univerzity zdarma (Prestigious Universities for Free).*
on server *Lupa*, 9. 5. 2003, ISSN 1213–0702

Satrapa P.: *Stav DHCPv6 na světových tocích (Current Status of DHCPv6 in the World).*
on server *Lupa*, 6. 2. 2003, ISSN 1213–0702

Satrapa P.: *Temná vlákna v CESNETu (Dark Fibres in CESNET).*
on server *Lupa*, 19. 12. 2003, ISSN 1213–0702

C.6 Presentations in the R&D Area

Altmanová L., Radil J., Šíma S.: *Fibres and advanced optical devices for a new networking strategy.*
Cambridge, 16. 9. 2003
http://www.terena.nl/tech/task-forces/tf-ngn/presentations/tf-ngn12/20030915_SS_Optical.pdf

Gruntorád J.: *How CESNET Got Access to Dark Fibres.*
SERENATE NREN Workshop, Amsterdam, únor 2003

Gruntorád J.: *Customer Empowered Fibre Network – CESNET Experience.* TERENA Workshop on NREN-controlled fibres, Copenhagen, říjen 2003

Gruntorád J., Šíma S.: *Customer Empowered Optical Networks.*
Internet2 Member Meeting, Arlington, USA, duben 2003

Lhotka L., Novotný J.: *Liberouter: a PC–based IPv6 Router.*
Zagreb, Croatia, TERENA, 21. 5. 2003

Radil J., Šíma S., Altmanová L.: *Optical networking developments in CESNET.*
TF-NGN, Rome
http://www.terena.nl/tech/task-forces/tf-ngn/presentations/tf-ngn10/20030206_JR_optical.ppt

Smotlacha V.: *Čas ve světě počítačů a sítí (Time in the World of Computers and Networks).*
Nové Hrady, 25.–28. 5. 2003

Smotlacha V.: *One–way Delay Measurement Using NTP Synchronization.*
Zagreb, Croatia, TERENA, 19.–22. 5. 2003

Ubik S.: *Přenosy velkých objemů dat v rozlehlých Gigabitových sítích (Transmissions of Large Data Volumes in Wide Gigabit Networks).*
Olomouc, CESNET 28.–29. 5. 2003

Ubik S., Cimbál P.: *Achieving Reliable High Performance in LFNs.*
Zagreb, Croatia, CESNET, 19.–22. 5. 2003

Ubik S., Cimbál P.: *Debugging end-to-end performance in commodity operating systems.*
Geneve, Switzerland, CERN, 3.–4. 2. 2003

Ubik S., Vojtěch J.: *QoS in Layer 2 Networks with Cisco Catalist 3550.*
Prague, CESNET, duben 2003

Vozňák M.: *IP telefonie v praxi (IP Telephony in Practice).*
Prague, WIRELESSCOM, s. r. o., 24.–26.11.2003

D Bibliography

[HSC] *Introduction to HyperSCSI.*
Modular Connected Storage Architecture Group, Network Storage Technology Division, 2002
http://nst.dsi.a-star.edu.sg/mcsa/hyperscsi/nspd2.pdf

[ICN04] Antoš D., Řehák V. a Kořenek J.: *Hardware router's lookup machine and its formal verification.*
accepted at *3rd International Conference on Networking*, 2004

[Hof03] Höfer F.: *Packet analysis for IPv6 router implemented by a PCI acceleration card.*
Bachelor thesis, Fakulta informatiky MU, Brno, 2003

[Gru03] Gruntorád J. a kol.: *Vysokorychlostní síť národního vyzkumu a její nové aplikace: průběžná zpráva o řešení výzkumného záměru 2002 (High-speed National research network and its New Applications: 2002 annual research report).*
CESNET, Prague, 2003, ISBN 80-238-9978-3

[Kře03] Křenek A.: *Haptic rendering of complex force fields.*
In Proc. *9th Eurographics Workshop on Virtual Environments*, ACM Press, 2003, str. 231–240, ISBN 3-905673-00-2

[KPM03] Křenek A., Peterlík I., Matyska L.: *Building 3D State Spaces of Virtual Environments with a TDS-based Algorithm.*
In *Recent Advances in Parallel Virtual Machine and Message Passing Interface*, Springer-Verlag, 2003, str. 529–536, ISBN 3-540-20149-1

[Kře04] Křenek A. et al.: *Systém Perun verze 2.2 (Perun System, version 2.2).*
technical report, CESNET, 2004, prepared

[MTV01] Martini L., Tappan D., Vogelsang S., Malis A. G., Vlachos D. S.: *Transport of Layer 2 Frames Over MPLS.*
draft IETF *draft-martini-l2circuit-trans-mpls-05*, August 2001

[NAF03] Novotný J., Antoš D. a Fučík O. *Project of IPv6 Router with FPGA Hardware Accelerator.*
in Cheung P.Y.K., Constantinides G.A. and and de Souza J.T. (Eds.), *Field-Programmable Logic and Applications*, 13th International Conference FPL 2003. Lecture Notes in Computer Science 2778, Springer, Berlin-Heidelberg, 2003, pp. 964-967. ISBN 3-540-40822-3.

[SSC03] Satran J., Sapuntzakis C., Chadalapaka M., Zeidner E.: *iSCSI*.
draft IETF *draft-ietf-ips-iscsi-20*, January 2003

[Sov02] Sova M.: *Middleware (Middleware)*.
in *Vysokorychlostní síť národního výzkumu a její nové aplikace 2001 (High-speed National Research Network and its New Applications 2001)*, Prague 2002, ISBN 80-238-8174-4

[UbK03] Ubik S., Klaban J.: *Experience with simulations for congestion control research*.
internal report of the End-to-end performance project, CESNET, 2003

[UbC03] Ubik S., Cimbál P.: *Debugging end-to-end performance in commodity operating systems*.
Pfldnet2003, CERN, Geneve, Switzerland, February 2003

[UbC03a] Ubik S., Cimbál P.: *Achieving Reliable High Performance in LFNs*.
TNC 2003, Zagreb, Croatia, May 2003

[UKr03] Ubik S., Král A.: *End-to-end Bandwidth Estimation – Progress Report*.
internal report of the End-to-end performance project, CESNET, 2003