



MetaCentrum

Aktuální stav

Martin Kuba

CESNET a ÚVT MU



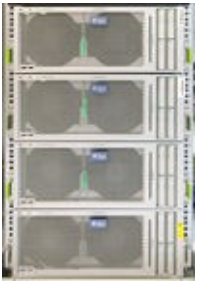
MetaCentrum

- sdružuje výpočetní prostředky
 - Superpočítačové Centrum Brno, MU, Brno
 - Univerzita Karlova, Praha
 - Západočeská Univerzita, Plzeň
 - NCBR + Loschmitovy laboratoře, PŘF MU Brno
 - Biofyzikální ústav AV ČR, Brno
 - Ústav nauky o dřevě, společně MZLU a VUT, Brno
 - PŘF Jihočeská Univerzita, České Budějovice
 - FEEC Vysoké Učení Technické, Brno
 - CESNET

- MetaCentrum je aktivita CESNETu
- celkem sdružuje přes 1000 (jader) CPU
- česká národní gridová digitální e-infrastruktura
- zdroje jednotlivých organizací jsou sdíleny v různé míře
 - MU, UK, CESNET sdíleny všem
 - ZČU si rezervuje část kapacity, zbytek všem
 - MZLU, NCBR atd. vyhrazeno jen pro ně

- uživatelé mohou být všichni zaměstnanci a studenti z akademické sféry a výzkumu
- vlastníci výpočetních zdrojů mají na svém vybavení prioritu nebo i výhradní přístup
- všichni mohou v případě potřeby využít cizí volné zdroje nebo „neutrální“ stroje CESNETu
- grid zlepšující využití zdrojů více organizací

Hardware

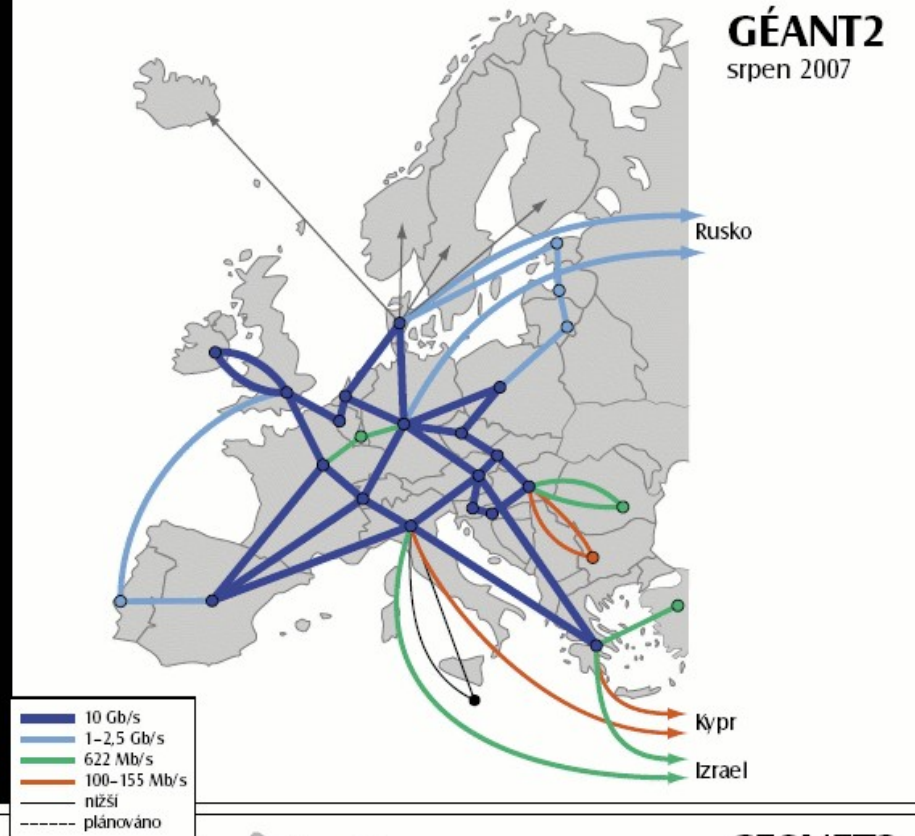


- linuxové clustery Intel/AMD
 - 1, 2, 4 CPU stroje
- 8 strojů 16CPU Opteron s 64GB RAM
- SMP (symmetric multiprocessing) stroje
 - 16 CPU Itanium, 48 CPU MIPS
- síťové karty 1Gb/s Ethernet, 2.5Gb/s Myrinet, 20Gb/s Infiniband

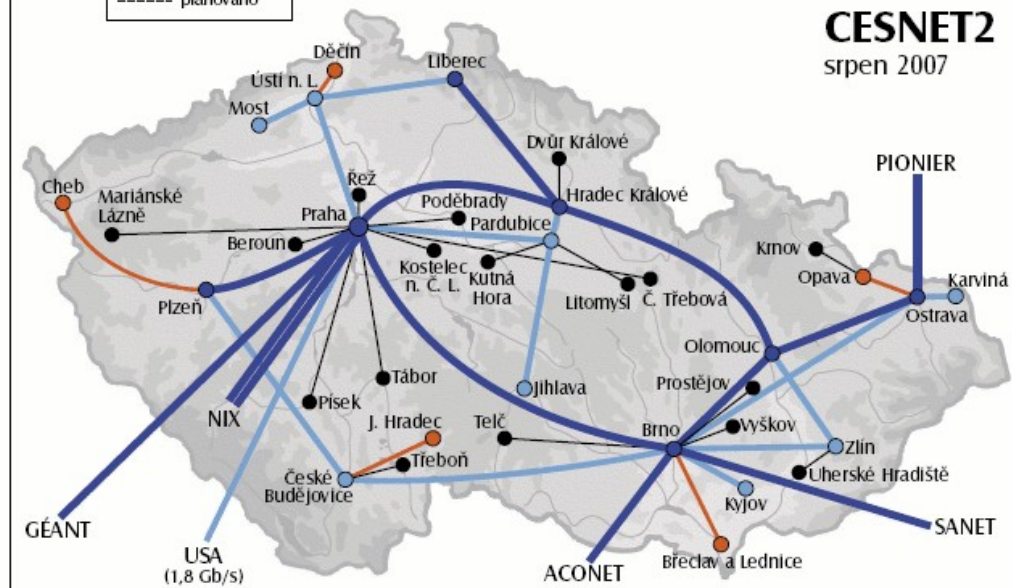


Topologie sítí GÉANT2 a CESNET2

GÉANT2
srpen 2007



CESNET2
srpen 2007



Úložné prostory

- lokální disky na výpočetních strojích
- sdílené disky pro jednotlivé clustery (NFSv3)
- 100TB diskové pole přístupné přes NFSv4
- dvě páskové knihovny s 1000 páskami po 400GB dat, tj. celkem 400TB
- zálohy i dlouhodobá archivace



Adresáře uživatelů

- /home/makub domovský sdílený přes cluster
- /scratch/makub velmi rychlý pro data právě běžící úlohy na každém výpočetním stroji
- /storage/home/makub sdílený přes NFSv4 přes (téměř) všechny stroje, neomezené místo
- /afs/ics.muni.cz/home/makub sdílený přes AFS přes všechny stroje
- možnost přístupu z osobního stroje přes AFS a NFSv4

Uživatelé



- výpočetní chemie
- materiálové a strukturní simulace
- simulace proudění
- rozpoznávání a generování řeči
- fyzikální geodézie
- ekologické modelování
- zpracování videa
- data mining
- analýza lékařských obrazů

Sjednocené prostředí

- jedna přihláška
- jedno přihlašovací jméno
- jedno heslo
- sdílené souborové systémy (AFS, NFSv4)
- systém plánování úloh (PBSPro)
- systém správy účtů (Perun)
- jeden portál **<http://meta.cesnet.cz/>**
- jedna komunikační adresa **meta@cesnet.cz**

System sdílení zátěže PBSPro

- přiděluje uživatelům **zdroje**
 - zdroje jsou: stroje, CPU, paměť, licence
- uživatelé specifikují množství požadovaných zdrojů a jejich vlastnosti
 - typ CPU, OS, síťový interface, geografické umístění
- úlohy se řadí do front
- přidělené zdroje lze použít **interaktivně i dávkově**

Portál METACentra

- jedna zapamatovatelná adresa
- <http://meta.cesnet.cz/>
- přehled zdrojů (HW, SW, síť), novinky, výpadky, dokumentace, návody, diskusní fórum
- správa účtu uživatele
- zobrazení aktuálního stavu
- historické záznamy o úlohách (accounting)



METACentrum

- O METACentru
- Vybavení
- Návody
- Přihláška
- Můj účet
- Aktuální stav
- Aktuality
- Dokumentace
- Interní část
- Mapa portálu
- Diskuzní fórum

RSS



Stav strojů

Stroj	Stav	Typ	Velikost	Uživatel
1	OK
2	OK
3	OK
4	OK
5	OK
6	OK
7	OK
8	OK
9	OK
10	OK



Přihláška

Číslo uživatele:

Heslo:

Osobní nastavení

Osobní údaje

Meno příjmení:

Titul:

Adresa:

Telefon:

E-mail:

- ### Výpadky
- 9. 9. 2008**
Přepojování clusterů na novou UPS
 - 3.9.2008**
nezálohované stroje
 - 03. 09. 2008**
Krátkodobý výpadek skirita

Diskuzní fórum

phobos

1. 1. 2008

2. 1. 2008

3. 1. 2008

4. 1. 2008

5. 1. 2008

6. 1. 2008

7. 1. 2008

8. 1. 2008

9. 1. 2008

10. 1. 2008

- ### Novinky
- 11.8.2008**
novinky a výpadky v preferovaném jazyce
 - 7. 8. 2008**
InfiniBand
 - 17.7.2008**
Nová verze programového balíku WIEN2k 08.2

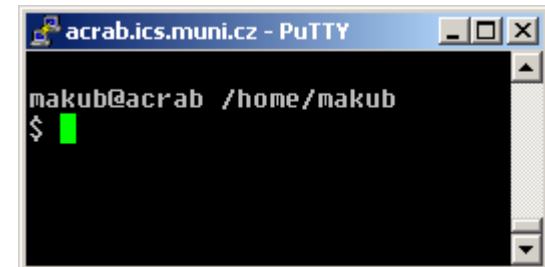
Vítejte na portálu METACentra!

Projekt METACentrum je aktivita výzkumného záměru **CESNET** zastřešující většinu aktivit souvisejících v České republice s Gridy, superpočítači, klastrovým nebo gridovým počítáním a/nebo výkonným počítáním obecně. Vlastním cílem projektu METACentrum je provoz a správa současných výpočetních zdrojů a v budoucnosti další rozšiřování výpočetní kapacity největších akademických center České republiky. Systémy projektem spravované vytváří virtuální distribuovaný počítač, v poslední době označovaný jako **Grid**.

- ### Projekty
- EGEE

Potřebné nástroje uživatele

- WWW prohlížeč kvůli portálu
- ssh klient (PuTTY) pro vzdálený příkazový řádek
- FTP nebo scp (WinScp) klient pro přenos souborů
- případně vzdálené sdílení disků (Samba, NFSv4, AFS)



Autentizace (prokázání totožnosti)

- single-sign on
- jednotné jméno a heslo, Kerberos
- volitelně i digitální X509 certifikát, použitelný i v evropském gridu EGEE
- uživatelé obvykle používají běžný ssh/PuTTY pro spojení na frontend stroj (skirit.ics.muni.cz) čímž tam získají Kerberos lístek platný 10 hodin
- v rámci METACentra pak všechno kerberizované (telnet, ssh, scp, AFS, NFSv4, Apache, Oracle)

Správa účtů

- prvotní přihláška
- souhlas s pravidly (podpis na papíru nebo autentizace proti MU, UK, ZČU, federaci)
- členství v METACentru je věčné
- účty na strojích mají platnost do konce roku
- v prosinci prodloužení účtů na strojích za podání zprávy o činnosti v uplynulém roce
- žádosti o účty na dalších strojích

Perun



- systém pro správu účtů Perun
- eviduje osoby, stroje, účty
- databáze, master část, slave části na všech strojích
- např. uživatel požádá o účet na novém clusteru, administrátor žádost schválí, Perun do pár minut přegeneruje seznamy uživatelů na všech strojích v clusteru

pbsmon, accounting

- webové rozhraní k plánovacímu systému PBS
- zobrazení strojů, úloh, front
- personalizovaný pohled – jen stroje a fronty na které má daný uživatel přístup
- PBS drží informace o úlohách do 24 hodin po jejich dokončení
- accounting zobrazuje data o starších úlohách (data z PBS a logu jader OS)



Tato stránka zobrazuje osobní pohled na PBS pro uživatele **makub**, tj. fronty a stroje uživateli přístupné.

Úlohy uživatele makub

uživatel	Počet úloh		Počet CPU	
	běžících	ostatních	používaných	ostatních
makub	0	0	0	0

Fronty dostupné uživateli makub

fronta	priorita	časové limity	nutná vlastnost	úloh					
				ve frontě	běžících/max	hotových	celkem	max. na uživatele	
gridlab	80	0 - 720:00:00	globus	0	0 / 4	0	0	4	
short	60	0 - 02:00:00	linux	0	0 / 12	0	0	8	
normal	50	02:00:01 - 24:00:00	linux	14	24 / 112	26	64	20	
quark	45	0 - 720:00:00	quark	0	0 /	0	0		
ncbr	45	0 - 720:00:00	per	4	22 / 120	1	29	16	
cpmd	45	0 - 0	onlycpmd	0	5 / 120	4	9	32	
orca	45	0 - 720:00:00	orca	0	0 / 120			pracující	80
orca16g	45	0 - 720:00:00	orca16g	0	0 / 120			sdíleně pracující	80
long	40	24:00:01 - 720:00:00	long	24	37 /			vyhrazeně pracující	40

Stroje dostupné z fronty "short"

ajax	hydra	hydra1	hydra2	hydra3	hydra4	hydra5	hydra6	nepřipojené
hydra7	hydra8	hydra9	hydra10	konos1	konos2	konos3	konos4	vypnuté
konos5	konos6	konos7	konos8	konos9	konos10	konos19-1	konos20-1	neznámý stav
konos21-1	konos22-1	konos24-1	konos26-1	konos27-1	konos28-1	konos29-1	konos30-1	údržba
konos31-1	konos32-1	konos33-1	konos34-1	konos35-1	konos36-1	konos37-1	manwe1	test

O METACentru

Vybavení

Návody

Příhláška

Můj účet

Aktuální stav

Produkční PBS

Osobní pohled

Stroje

Úlohy

Fronty úloh

Vlastnosti strojů

Uživatelé

Testovací PBS

Ganglia

Aktuality

Dokumentace

Interní část



O METACentru

Vybavení

Návody

Přihláška

Můj účet

Aktuální stav

Produkční PBS

Osobní pohled

Stroje

Úlohy

Fronty úloh

Vlastnosti strojů

Uživatelé

Testovací PBS

Ganglia

Aktuality

Dokumentace

Interní část

Úloha

- Základní informace
- Úplná tabulka informací

Základní informace

úloha	CPU	vyhraz. paměť	použitá paměť	jméno	uživatel	CPU čas	čas běhu	stav	fronta
527476.skirit-f.ics.muni.cz	10	10x 400mb = 4000mb	51mb	dca.sh	sopib1am	32:10:35	05:31:39	R - běží	normal
zdroje	10:ppn=1#brno#linux#p3								
vytvořena	Středa, 10. září 2008 9:53:52								
spuštěna	Středa, 10. září 2008 9:53:52								
poslední změna stavu	Středa, 10. září 2008 16:25:24								
pracovní adresář	/home/sopib1am								
použitý stroj/cpu	skirit1/0 skirit2/0 skirit3/0 skirit4/0 skirit5/0 skirit6/0 skirit7/0 skirit8/0 skirit9/0 skirit10/0								
seznam proměnných	PBS_O_HOME=/home/sopib1am PBS_O_LOGNAME=sopib1am PBS_O_PATH=/software/pbs-7.0.0/bin:/usr/heimdal/bin:/usr/heimdal/sbin:/usr/ PBS_O_MAIL=/var/mail/sopib1am PBS_O_SHELL=/bin/bash PBS_O_HOST=skirit.ics.muni.cz PBS_O_WORKDIR=/home/sopib1am PBS_O_SYSTEM=Linux PBS_O_QUEUE=normal								



Úlohy uživatele makub

Přehled posledních 10 záznamů

 Poslední(ch) záznamů

 Nalezeno 10 položek, Zobrazuji všech položek. 1 Počet na stránku:

Úloha	Procesory	Název	Fronta	Čas vytvoření	Čas spuštění	Čas ukončení	Walltime z PBSPPro	Celkový čas uživatele	Celkový čas systému	Pbs ser
9793.arien.ics.muni.cz	1	STDIN	short	08.08.2008, 19:36:04	08.08.2008, 19:36:26	08.08.2008, 19:46:24	0:09:59,00	0:02:44,08	0:00:00,14	arien.ics.m
9652.arien.ics.muni.cz	1	STDIN	short	05.08.2008, 10:55:24	05.08.2008, 10:55:27	05.08.2008, 11:53:45	0:58:18,00	0:00:00,00	0:00:00,00	arien.ics.m
482626.skirit-f.ics.muni.cz	2	STDIN	short	26.06.2008, 10:54:19	26.06.2008, 10:54:46	26.06.2008, 11:05:52	0:11:06,00	0:00:00,00	0:00:00,00	skirit-f.ics.mur
482625.skirit-f.ics.muni.cz	2	STDIN	short	26.06.2008, 10:52:30	26.06.2008, 10:53:05	26.06.2008, 10:54:13	0:01:27,00	0:00:00,00	0:00:00,00	skirit-f.ics.mur
482620.skirit-f.ics.muni.cz	1	STDIN	short	26.06.2008, 10:21:39	26.06.2008, 10:23:16	26.06.2008, 10:30:10	0:07:49,00	0:00:00,00	0:00:00,00	skirit-f.ics.mur
2218.arien.ics.muni.cz	1	STDIN	short	28.05.2008, 13:04:47	28.05.2008, 13:04:51	28.05.2008, 13:05:06	0:00:18,00	0:00:00,00	0:00:00,00	arien.ics.m
474516.skirit-f.ics.muni.cz	1	STDIN	short	28.05.2008, 13:04:21	28.05.2008, 13:04:27	28.05.2008, 13:04:42	0:00:16,00	0:00:00,00	0:00:00,00	skirit-f.ics.mur
2213.arien.ics.muni.cz	1	STDIN	short	28.05.2008, 09:14:26	28.05.2008, 09:07:52	28.05.2008, 09:00:01	0:14:26,00	0:00:07,52	0:00:01,14	arien.ics.m

Provozní a testovací prostředí

- potřebujeme zajistit zároveň stabilitu i rozvoj
- proto jsme zavedli dvě prostředí
 - produkční prostředí – maximální stabilita
 - testovací prostředí
 - rozšířeno o novinky v plánovači, jádru OS, Infiniband
 - obsahuje nové silné stroje jako odměnu dobrovolným testerům z řad uživatelů
- z pohledu uživatele se testovací prostředí liší jen nastavením proměnné `PBS_SERVER`

Request Tracking systém

- jednotná komunikační adresa meta@cesnet.cz
- příchozí emaily evidovány v RT jako tzv. lístky
- lístek má **číslo jednací, žadatele, vlastníka a stav**
- vlastník je odpovědný za vyřízení lístku
- každá změna lístku je e-mailem oznámena žadateli a všem lidem v uživatelské podpoře a provozu METACentra
- smyslem RT je evidence žádostí, evidence historie řešení a zastupitelnost pracovníků METACentra (kvůli dovoleným, nemocem, služebním cestám atp.)

Uživatelská podpora

- dohlíží na řešení lístků v RT
- odpovídá na dotazy uživatelů
- zajišťuje provoz portálu
- zajišťuje dokumentaci na portálu
- vyřizuje přihlášky a žádosti o další účty

Změny v poslední době

- přednostní přístup za publikace
- nasazení kerberizovaného NFSv4
- diskové pole 60TB a povýšeno na 100TB
- vysokorychlostní síť Infiniband
- manwe6 a manwe7 (64 a 128GB RAM)
- parametr „mem“ pro rezervaci paměti
- testovací prostředí
- diskuzní fórum
- přehled adresářů uživatele na portálu
- změna hesla přes federovanou autentizaci
- accounting od ledna 2008
- virtualizace, IPv6

Plánovaný rozvoj

- nový HW
 - na přelomu roku výměna clusterů nympha a minos na ZČU Plzeň za nové stroje (20x 8CPU, 16GB RAM)
 - nový cluster hermes z PřF JČU (10x 4CPU, 16GB RAM, 1TB disk)
 - nový cluster z FEEC VUT (4x 16CPU)
- decentralizace plánovacího systému a přechod z PBSPro na Torque

Plánovaný rozvoj (2)

- zavedení specializovaných čelních uzlů místo frontend uzlů clusterů (skirit)
- využití vznikající federace identit eduld.cz pro autentizaci na portálu
- využití virtualizace strojů pro virtuální clustery a větší kontrolu uživatele nad prostředím
- virtualizace sítí, IPv6

Konec

- děkuji za pozornost