

Výpočetní zdroje v MetaCentru a jejich využití

Miroslav Ruda

Cesnet
a
Masarykova Univerzita

Praha, 2008



Hardware I.

- jádro tvoří clustery na MU, Cesnetu, ZČU
 - 32 bit Xeon, 64bit AMD Opteron a 64bit Intel Xeon
 - lokální `scratch` filesystem
 - většinou jeden sdílený `home` přes město (NFS)
 - sdílený filesystem `AFS` přes celé MetaCentrum
- větší SMP stroje, vlastněné členy MetaCentra
 - SGI Altix (UK 16x Itanium II, ZČU 8x Itanium II)
 - Sun x4600 (8x MU 8xdual core AMD Opteron)
 - SGI Origin 2000 (UK 48x MIPS R10000)
- permanentní diskové prostory zálohované na páskové knihovny
 - MU, ZČU - 2x200 TB
- rychlejší diskové pole, nabízené přes NFSv4 (MU, 60 TB)
 - prototyp i pro další instalace
 - testovací instalace PVFS2 dostupná pro testy

Hardware II.

- další clustery
 - MetaCentra, pro vyhrazené použití (EGEE cluster na Cesnetu)
 - integrace s FZU na úrovni podpory
 - organizací, kterým poskytujeme správu jejich clusteru
 - Institute for Theoretical Computer Science (ITI) na ZČU
 - National Centre for Biomolecular Research (NCBR) na MU
 - organizací, které chtějí integrovat jejich cluster do MetaCentra
 - Ústav nauky o dřevě, MZLU a VUT
 - právě instalovaný cluster v ČB
 - plánovaný cluster na FEEC, VUT
- síť
 - Myrinet a Infiniband na clusterech na MU
 - clustery blízko Cesnetí páteře - rozumná latence i pře ČR
 - přístup ke speciálnějším sítím typu
 - lambda mezi MU a Cesnetem
 - virtuální L2 síť přes cele MetaCentrum
 - postupný přechod na IPv6

Software

- Perun - jednotná správa účtů
- Kerberos - centrální databáze hesel, single sign-on
- AFS - distribuovaný filesystem
- **PBSPro** - jednotný dávkový systém
- **virtualizace** - Xen, Magrathea

Použití I.

- zatím až na výjimky jednotná správa úloh přes celé MetaCentrum
- jediný systém PBSPro, jednotné účtování, uživatelská podpora, správa účtů
- "fair-share" pro uživatele
- prioritní přístup na nejžádanější zdroje, pro paralelní úlohy, podle "úrovně spolupráce s MetaCentrem"
- u clusterů vlastněných jinou organizací lepší priorita vlastníka
 - vyhrazena část jen pro jeho použití
 - lepší priorita pro úlohy vlastníka
 - omezení cizích úloh (např. příliš dlouhých)
- jednotná přihláška do MetaCentra
 - uživatelé jednoduše ovlivňují kde mají účty
 - prvotní ověření totožnosti přes federativní nástroje (Shibboleth) nebo proti autentizačním systémům domovské organizace
 - různé clusterly mohou vyžadovat podepsaní speciální licence/prohlášení

Použití II.

- licencovaný software
 - různé typy licencí (pro skupinu, pro organizaci, pro skupinu strojů, pro celé MetaCentrum) - přístup řízený podle přání vlastníka a pomocí přístupových práv v AFS
- typy úloh
 - klasický mix dlouhých dávkových úloh (měsíc) a kratších úloh (experimenty, high-throughput úlohy)
 - postupně přibývá požadavků na interaktivní úlohy a přístup
 - vyvíjíme rozšíření podle požadavků
- linuxová instalace
 - preferujeme instalaci "totožnou" s naší
 - náš linuxový obraz, debian 4.0
 - virtualizace přes Xen
 - integrováno do našeho monitoringu, dávkového systému, správy účtů
 - podpora odlišných instalací linuxu
 - požadavky EU projektů, speciální software
 - separátní cluster dostupný jen přes vyšší vrstvy gridového middleware
 - přes virtualizaci a náš systém pro přepínání aktivních domén

Použití (vývoj)

- postupný přechod na model s více dávkovými systémy
 - dedikovaný systém pro EGEE cluster
 - experimentální instance PBSPro, poskytující nově vyvíjené služby
 - v plánu jsou i speciální dávkové systémy pro největší skupiny
 - se standardním gridovým rozhraním (gLite, Globus) umožňujícím
 - integraci s evropskými projekty
 - dynamické sdílení fyzických strojů mezi virtuálními clustery
- využití virtuálních strojů na výpočetních uzlech
 - pro sdílení stroje různými instalacemi Linuxu (debian pro MetaCentrum, SLC pro EGEE, ...)
 - pro poskytnutí lepší priority pro vlastníka clusteru
 - ve vývoji
 - další využití pro interaktivitu
 - pro virtuální clustery, umožňující dynamické přesouvání výpočtů, změnu velikosti clusteru apod.

Připojení nového clusteru

- nová organizace
 - vyhovuje přihláška MetaCentra
 - vyhovuje naše generování účtů
- instalace a správa clusteru
 - vlastní nebo náš obraz linuxu
 - speciální požadavky (HW,SW)
 - diskové prostory
 - virtualizace
 - lokální podpora
- typy úloh
 - dávkové, interaktivní, prioritní
- dávkový systém
 - oddělený vs. centrální, nové fronty, priority
- přístup pro MetaCentrové uživatele
- přístup k síti
 - veřejné IP adresy vs. NAT, firewall
 - uživatelé z jiných organizací ve vnitřní síti