

## HARDWAROVÁ A SOFTWAREVÁ ŘEŠENÍ PRO VYSOCE VÝKONNÉ POČÍTÁNÍ

ING. PAVEL STRACHOTA, PH.D.

### Úvod

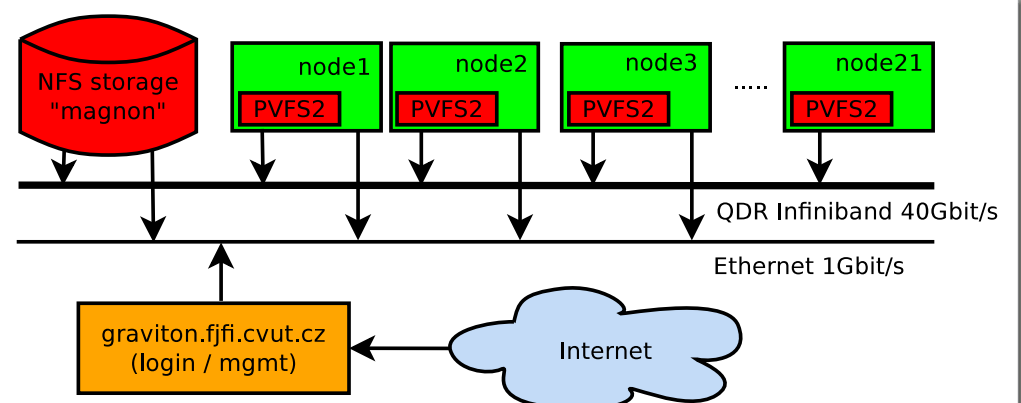
Současné systémy pro vysoce výkonné paralelní počítání (HPC - *High Performance Computing*) jsou zpravidla tzv. výpočetní svazky (clusters) složené z velkého počtu výpočetních uzlů (*compute nodes*) spojených rychlou sítí. Data jsou ukládána v úložišti tvořeném diskovými uzly (*storage nodes*) a optimalizovaném pro paralelní přenos dat. Systém je doplněn přihlašovacími uzly (*login nodes*) pro interaktivní práci, servery zajišťující správu a potřebnými síťovými prvky. Softwarové prostředí pak obsahuje celou řadu komponent, které efektivně a spravedlivě rozdělují výpočetní prostředky uživatelům a umožňují masivně paralelní výpočty. Na FJFI je v provozu cluster **Hyperion**, který kromě využití pro výpočty umožňuje studentům seznámit se s architekturou HPC systémů a podílet se na jejím rozvoji.

### Výpočetní cluster Hyperion na KIPL / KM v roce 2017



- ▶ architektura **AMD Opteron** (16-core Opteron 6272)
- ▶ **32 jader a 64 GB**, resp **128GB RAM** v 1 uzlu
- ▶ jeden uzol: GPGPU NVIDIA GeForce TITAN GK110
  - ▶ 2668 CUDA cores, 6 GB GDDR5 VRAM
- ▶ celkem **672 jader, 1600 GB RAM**
- ▶ interconnect: QDR **Infiniband 40Gbit/s** + gigabit Ethernet
- ▶ storage:
  - ▶ **10 TB** scratch - paralelní souborový systém *OrangeFS*
  - ▶ **80 TB** home - přes NFS, RAID6 + 800GB SSD cache
  - ▶ **500 GB** local scratch - pevný disk na každém uzlu
- ▶ operační systém **GNU/Linux** (*CentOS 6.8*)
- ▶ paralelní programování: *OpenMPI, OpenMP, CUDA*
- ▶ úlohový systém *OGS/GE* (Open Grid Scheduler / Grid Engine)
- ▶ konfigurace pomocí *cfengine*, SW konfigurace pomocí *modules*

### Schéma clusteru Hyperion



### Nabízená témata

- ▶ nasazení, testování a ladění **distribuovaných souborových systémů**
  - ▶ Lustre, GlusterFS, OrangeFS, ...
- ▶ instalace, nastavení a analýza funkčnosti **úlohového systému**
  - ▶ Torque, SGE, SLURM, ...
  - ▶ nastavení front, optimalizace nastavení scheduleru, simulace provozu
- ▶ instalace a konfigurace software pro **správu a monitoring** clusteru
  - ▶ cfengine, modules, Red Hat Cluster Suite, Ganglia, ...
- ▶ konfigurace **operačního systému**
  - ▶ hromadná instalace a globální správa OS, správa uživatelů, zabezpečení
- ▶ možnosti **ladění a profilování** paralelních algoritmů
  - ▶ LWM2, Scalasca, CUBE, Intel Parallel Studio ...
- ▶ **pokročilá paralelizace algoritmů**
  - ▶ topologie procesů, MPIIO, hybridní paralelizace, ...

### Postup a přínos práce (Bc. / Ing.)

- ▶ seznámení s hardwarovým řešením produkčních výpočetních systémů
- ▶ zvládnutí administrace systémů UNIX / Linux
- ▶ seznámení se softwarovými produkty pro provoz clusteru
- ▶ experimenty ve virtualizovaném prostředí
- ▶ návrh a funkčního řešení
- ▶ realizace na produkčním systému

Téma: **Hardware a software pro HPC systémy**

Autor: **Ing. Pavel Strachota, Ph.D.**

Kontakt: **pavel.strachota@fjfi.cvut.cz**

Homepage: **http://saint-paul.fjfi.cvut.cz**

MMG Homepage: **http://mmg.fjfi.cvut.cz**

Copyright © 2017 MMG FJFI ČVUT v Praze. Všechna práva vyhrazena.

