

HP-SEE Klasteri

www.hp-see.eu

Mihajlo Savic

Elektrotehnicki fakultet Banja Luka



HP-SEE

High-Performance Computing Infrastructure
for South East Europe's Research Communities

Paralelno procesiranje



HP-SEE

High-Performance Computing Infrastructure
for South East Europe's Research Communities

- Multitasking OS
- Problem paralelizacije programa
- Vrste paralelnih mašina
 - Višeprosesorske mašine
 - Klasteri
- Dijeljena memorija
- Komunikacija porukama

Multiprocesorske mašine



HP-SEE

High-Performance Computing Infrastructure
for South East Europe's Research Communities

- SIMD, MIMD, SPMD
- Hyper-threading
- Multicore CPU
- Višestruki CPU
- Masovno-višeprocesorske mašine
- Problemi sa dijeljenjem memorije
- NUMA

Zašto klasteri?



HP-SEE

High-Performance Computing Infrastructure
for South East Europe's Research Communities

- Povećanje stabilnosti
 - Failover
- Povećanje performansi
 - Load balancing
 - High performance
- Gdje povući granicu?

Failover klasteri



HP-SEE

High-Performance Computing Infrastructure
for South East Europe's Research Communities

- Povećanje stabilnosti
- Umnogostručavanje resursa
- Dva ili više identičnih računara međusobno sinhronizovanih
- Način otkrivanja otkaza primarnog računara i preuzimanje njegovih zadataka

Load-balancing klasteri



HP-SEE

High-Performance Computing Infrastructure
for South East Europe's Research Communities

- Raspoređivanje opterećenja
- Mogu biti heterogeni
- Tipična upotreba za Internet servise
- Primitivno raspoređivanje
- DNS bazirano
- SPOF – Single Point Of Failure

High-performance klasteri



HP-SEE

High-Performance Computing Infrastructure
for South East Europe's Research Communities

- Klasteri visokih performansi
- Ko su krajnji korisnici?
- Vrste – hardversko-ekonomski
 - Custom-made
 - Beowulf
 - COW – Cluster Of Workstations

HPC – Custom-made



HP-SEE

High-Performance Computing Infrastructure
for South East Europe's Research Communities

- Klasteri pravljani po narudžbi
- Komponente klastera posebno razvijene za dati klaster
- Vrhunske performanse po vrhunskoj cijeni
- Obično koriste poseban OS za rad čvorova
- Front-end OS – tipično Unix

HPC – Beowulf i COW



HP-SEE

High-Performance Computing Infrastructure
for South East Europe's Research Communities

- COTS pristup – standardno dostupne komponente (Commercial Off-The-Shelf)
- Eventualno odstupanje u mrežnoj opremi
 - Ethernet, Myrinet, Infiniband, SCI, ...
- Tipično “farme” običnih računara (COW) ili računara u minimalnoj konfiguraciji (Beowulf)

HPC – Mrežno povezivanje



HP-SEE

High-Performance Computing Infrastructure
for South East Europe's Research Communities

- Specijalizovane mreže
 - Myrinet i Infiniband – najpolularniji
- Ethernet
 - Hubovi – loše performanse
 - Switched ethernet
 - Fast/Giga-Ethernet
 - 10GigEthernet – TCP/IP Offloading Engine

HPC – Ethernet topologije



HP-SEE

High-Performance Computing Infrastructure
for South East Europe's Research Communities

- Zvijezda
 - Za manje klastere
- Stablo
 - Tree ili Fat-tree
- Stack-ring
 - Software forwarding
- Stack-mesh
 - Ortogonalne topologije – n-dimenzionalne mreže

HPC – komunikacione biblioteke



HP-SEE

High-Performance Computing Infrastructure
for South East Europe's Research Communities

- Potpuno prilagođena rješenja
 - Specifični problemi – specifična rješenja
- PVM – Parallel Virtual Machine
- MPI – Message Passing Interface
 - MPICH
 - LAM/MPI
 - MPILite
 - MPI 2.0 – RMA, parallel file I/O, ...

Raspoređivanje procesa



HP-SEE

High-Performance Computing Infrastructure
for South East Europe's Research Communities

- Job Management System
 - Upravljanje resursima,
 - sistem lista čekanja,
 - raspoređivač procesa
- Condor – High Throughput Computing
- PBS – Portable Batch System (NASA)
- Torque – Open Source – *PBS kompatibilnost
- Maui – novije, naprednije rješenje

Klasteri i virtuelne mašine



HP-SEE

High-Performance Computing Infrastructure
for South East Europe's Research Communities

- Tipična upotreba u fail-over klasterima
- Umjesto rada na nivou aplikacija, radimo na nivou operativnog sistema – virtuelni diskovi, virtuelne mrežne kartice
- Dijeljeni storage
 - Eksterni – NAS/SAN
 - Replikacija lokalnih diskova u realnom vremenu - DRBD



- Pristup grid klasteru BA-01-ETFBL
- Pristup od 8 do 24 CPU iz klastera
- MPI programiranje
 - Tri cjeline
 - Neophodno predznanje: C, kompajliranje
 - Uvod u MPI programiranje
 - Realizacija praktičnog problema
 - Primjena gotovih rješenja i biblioteka