

Posrednici umreženih sustava

Prof. dr. sc Siniša Srbljić

Dr. sc. Ivan Benc

Fakultet elektrotehnike i računarstva,
Zavod za elektroniku, mikroelektroniku, računalne i inteligentne sisteme

Posrednički sustavi u grozdovima i spletovima računala

**Dipl. ing. Emir Imamagić
(emir.imamagic@srce.hr)**

Fakultet elektrotehnike i računarstva,
Zavod za elektroniku, mikroelektroniku, računalne i inteligentne sustave

Sadržaj predavanja

- **Računalni grozd**
- **Računalni splet**
- **Literatura**

Računalni grozd



Paralelno računarstvo

- **Dijeljenje problema u manje zadatke koji se mogu izvoditi paralelno**
- **Motivacija**
 - fizičke granice sklopolja (brzina svjetlosti, elektrona, ...)
 - ekonomski razlozi (viša složenost – viša cijena)
 - granice učinkovitosti (dvostruka brzina ne znači dvostruku učinkovitost)
 - veličina aplikacije (prevelika količina memorije i vremena za izvođenje)

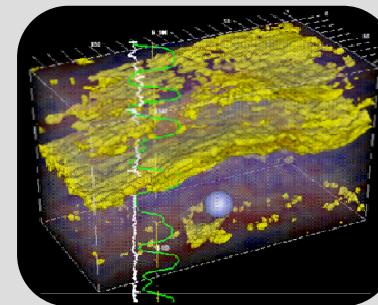
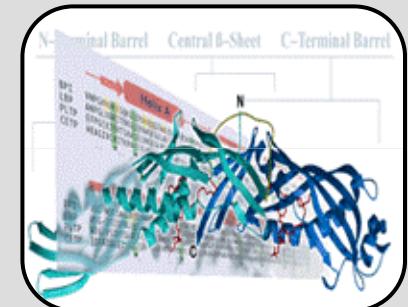
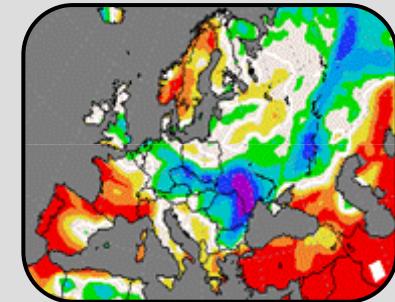
Paralelno računarstvo

- **Prednosti**
 - veća brzina izvođenja aplikacije
 - bolje iskorištenje dostupnih resursa
- **Mane**
 - složenost razvoja aplikacija
- **Tri velike prepreke**
 - sklopoljje
 - programska potpora
 - algoritmi

Paralelno računarstvo

- **Područja primjene**

- astronomija
- meteorologija i modeliranje klime
- životne znanosti
- znanosti o Zemljiji
- inženjerski problemi
- geologija
- vojne primjene
- ekonomsko modeliranje



Paralelno računarstvo

- **Superračunala**
 - izvođenje paralelnih programa
- **Vektorska superračunala**
 - Cray SV, NEC SX
 - Cray-1: 1976, 160.00 MFLOPS
8.8 M\$
- **Velika paralelna računala**
 - MPP (engl. *massively parallel processors*)
 - Earth Simulator, Blue Gene
 - Blue Gene/L: 2005, 360.00 TFLOPS, **~100 M\$**



Računalni grozd

- **Računalni grozd (engl. *Cluster*)**
- **“... skup umreženih samostalnih računala koji djeluju kao jedinstven računalni resurs”**
R. Buyya
- **Ekonomična alternativa za superračunala**
 - standardne sastavnice

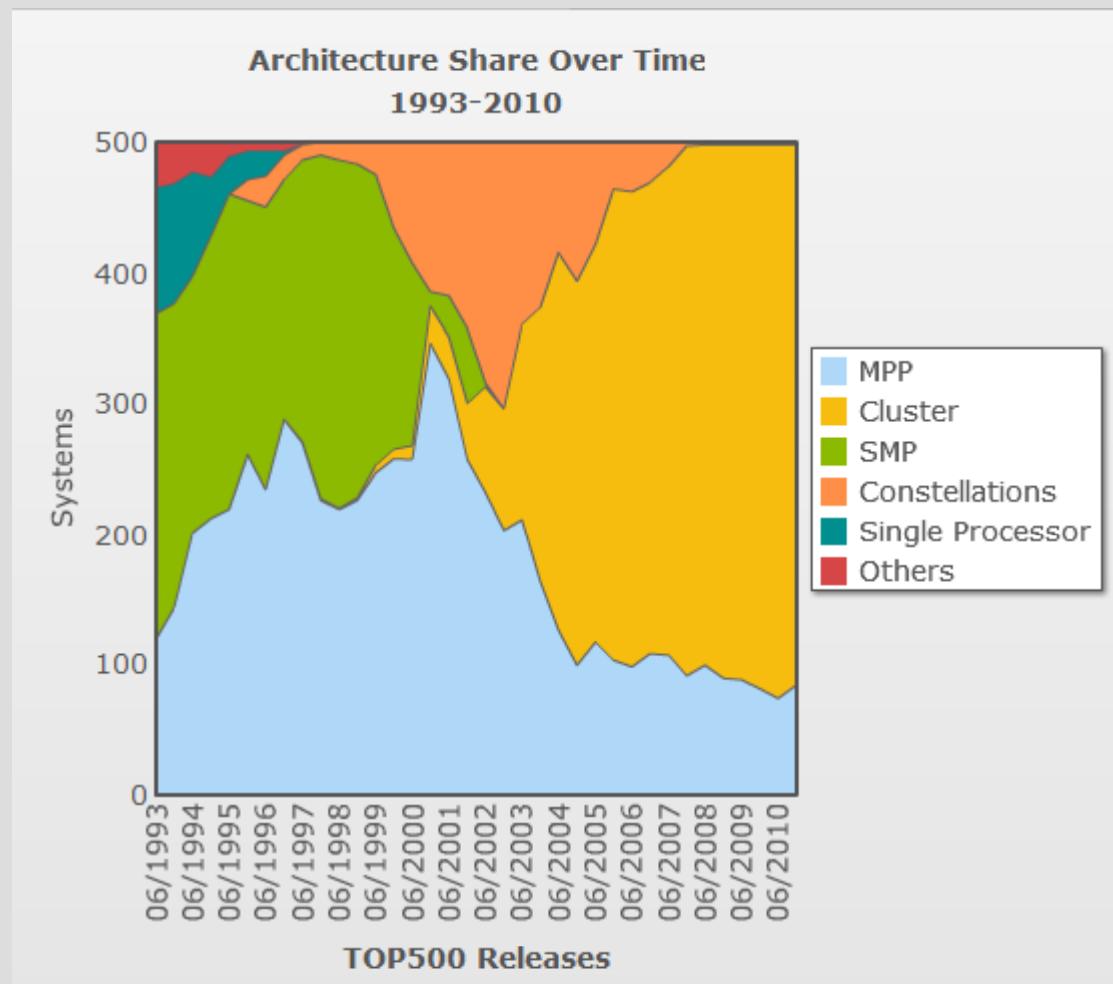


Računalni grozd

- **Ključni faktori (80-tih)**
 - mikroprocesori
 - brze mrežne tehnologije
 - Parallel Virtual Machine (PVM)
- **Beowulf arhitektura grozda (1994)**
 - T. Sterling & M. Baker
 - NASA Ames Centre

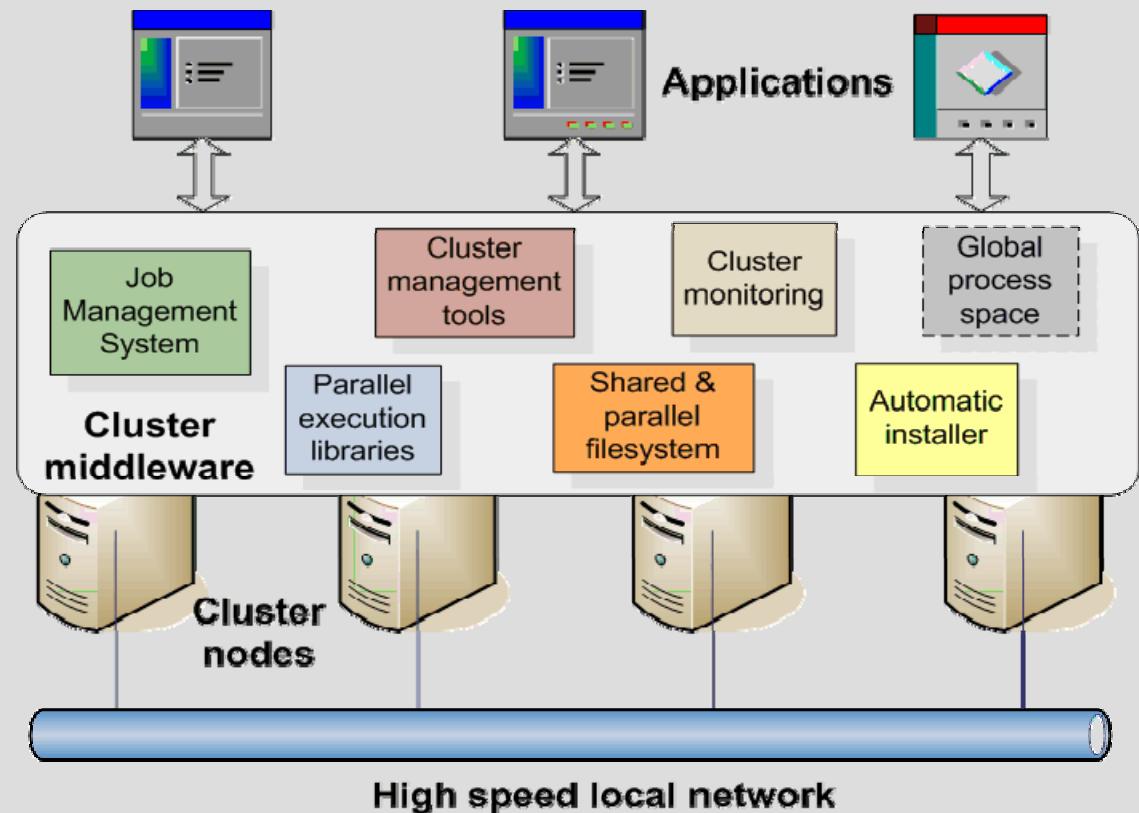
Računalni grozd

- **Top 500**
 - popis najmoćnijih računala
 - program za mjerjenje - Linpack
 - 82,8 % popisa čine grozdovi
 - 1. Tianhe-1A, China, 2.57 pflops (11.2010) **88 M\$**



Arhitektura grozda

- **Sastavnice**
 - računalni čvorovi
 - mreža
 - operacijski sustav (OS)
 - **posrednički sustavi**



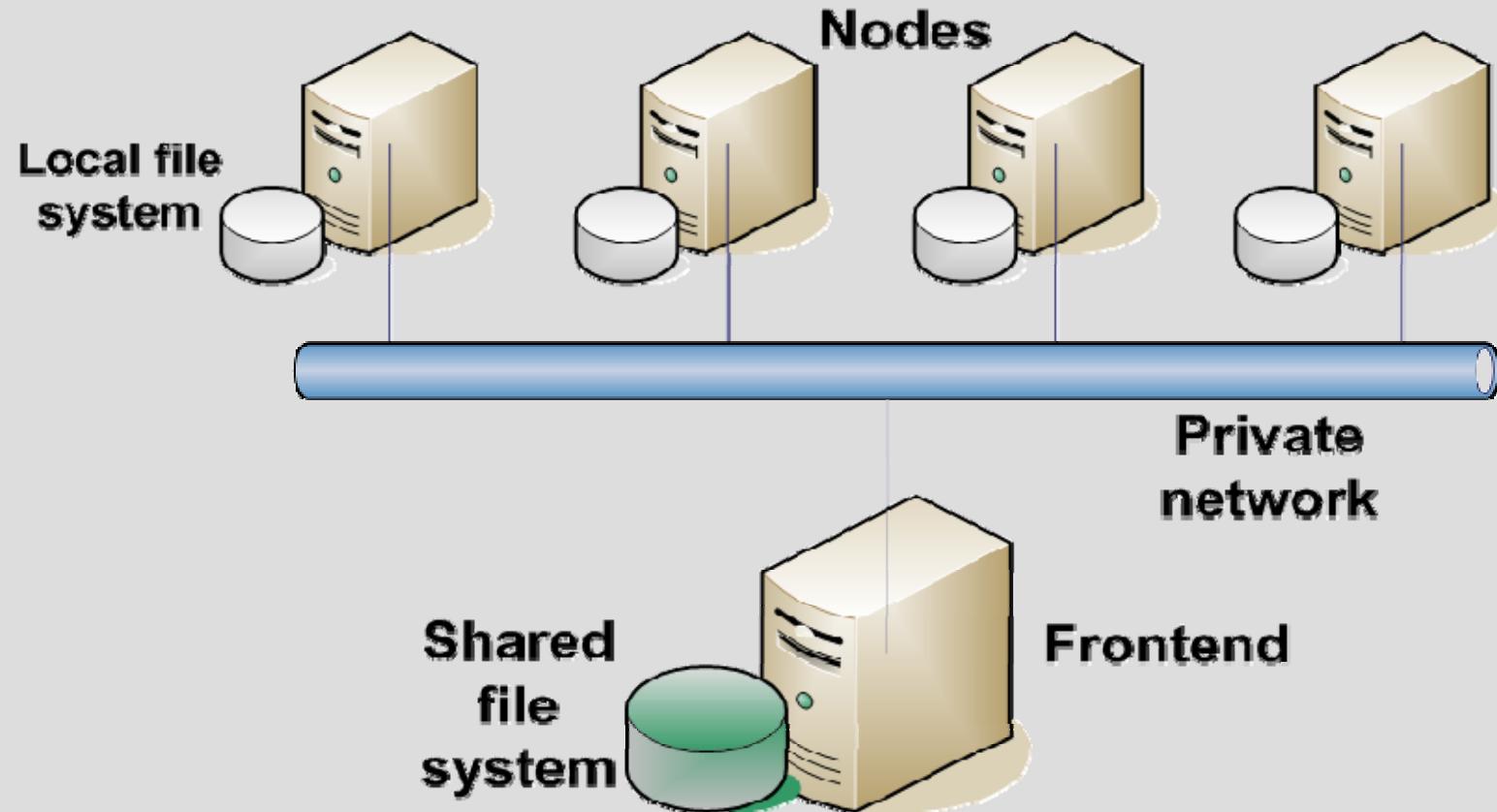
Arhitektura grozda

- **Računalni čvorovi**
 - poslužitelji, radne stanice, osobna računala
 - 32-bitna i 64-bitna arhitektura procesora
 - Intel Xeon, Itanium, AMD Opteron, Alpha
- **Mrežne tehnologije**
 - Gigabit Ethernet, Infiniband, Myrinet, Quadrics QsNet II
- **OS**
 - Linux, HPUX, Solaris, MS Windows

Arhitektura grozda

- **Posrednički sustavi**
 - **sustav za upravljanje poslovima**
 - nadzorni sustav
 - **dijeljeni i paralelni datotečni sustavi**
 - paralelne programske knjižnice i okoline
 - sustav za instalaciju i održavanje grozda

Beowulf arhitektura



Beowulf arhitektura

- **Pristupno računalo**
 - engl. *frontend*
 - pristup pomoću protokola SSH (engl. *Secure Shell*)
 - središnji servisi posredničkih sustava
 - dijeljeni diskovni prostor (npr. NFS, *Network File System*)

Beowulf arhitektura

- **Čvorovi**
 - dedicirana računala za izvođenje poslova
 - povezani privatnom mrežom ili mrežama
 - lokalni diskovni prostor
- **Privatne mreže**
 - visoka propusnost
 - malo kašnjenje
 - moguće korištenje više mreža

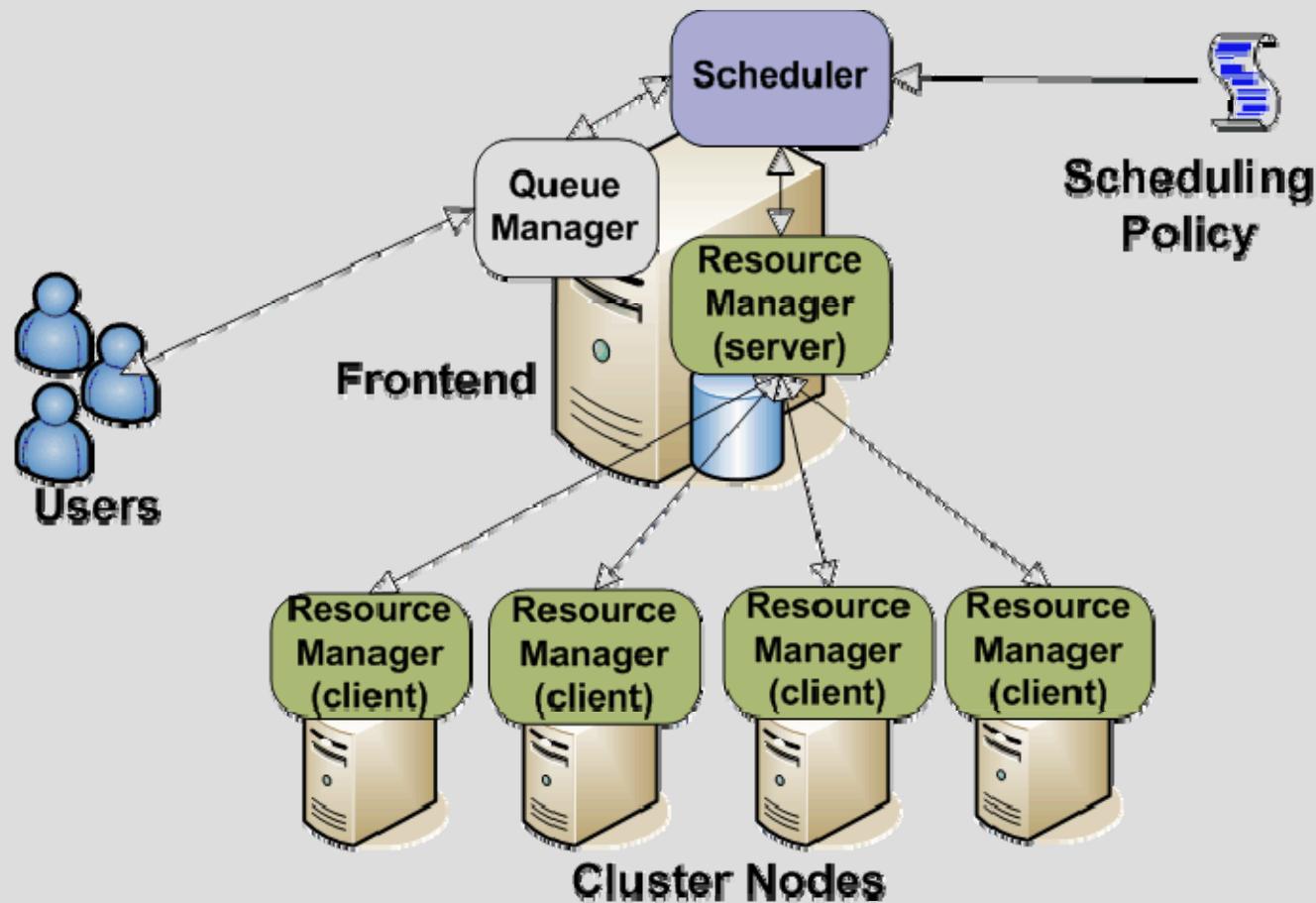
Sustav za upravljanje poslovima

- **Upravlja izvođenjem korisničkih programa na grozdu**
 - engl. *Job Management System (JMS)*, *batch system*
- **Korisničko sučelje za podnošenje i upravljanje poslovima**
 - podnošenje, upravljanje, dohvati detaljnih informacija o izvođenju
- **Omogućava definiranje politike korištenja klastera**
 - npr. *fair-share*, rezervacije, prioriteti, ...
- **Optimizira korištenje čvorova**
 - raspoređivanje poslova po čvorovima
 - praćenje opterećenja čvorovova

Sustav za upravljanje poslovima

- **Napredne funkcionalnosti**
 - spremanje stanja poslova (engl. *checkpointing*)
 - selidba poslova između čvorova (engl. *process migration*)
 - *preemption*
 - rezervacija čvorova (engl. *advance reservation*)
 - otpornost na greške u radu čvorova (engl. *fault tolerance*)
- **Rješenja**
 - Torque, Maiu, Sun Grid Engine

Sustav za upravljanje poslovima



Dijeljeni i paralelni datotečni sustavi

- **Dijeljeni datotečni sustavi**
 - omogućavaju transparentan pristup datotekama smještenim na udaljenom računalu
 - rješenja: **NFS, GFS, AFS, Coda**
- **Paralelni datotečni sustavi**
 - paralelno spremanje datoteka na više podatkovnih elemenata
 - učinkovito spremanje velikih datoteka
 - rješenja: **Lustre, Gluster, Fraunhofer FS**

Vrste grozdova

- **Grozdovi s visokom učinkovitosti**
 - engl. *High Performance Computing*
- **Grozdovi s visokom propusnošću**
 - engl. *High Throughput*
 - veliki broj neovisnih zadataka
 - nije bitna učinkovitost čvorova nego broj
- **Grozdovi s visokom dostupnošću**
 - engl. *High availability*
 - izbjegavanje SPOF (engl. *Single Point Of Failure*)
 - komercijalne primjene

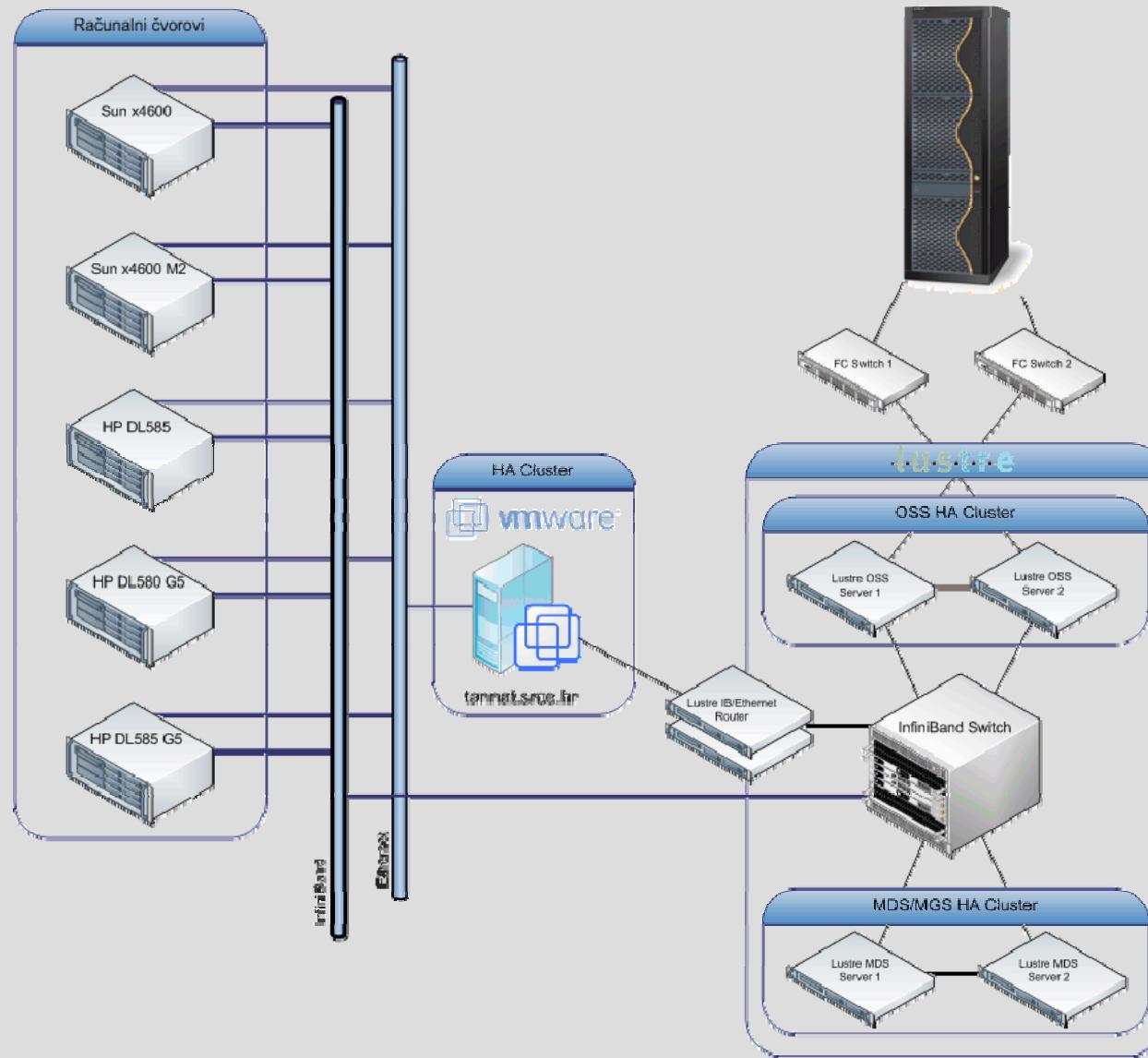
Vrste grozdova

- **Grozdovi za ravnomjerno opterećenje**
 - engl. *Load Balancing*
 - ostvaruju ravnomjerno opterećenje aplikacije na čvorovima
 - primjer: web poslužitelj
- **Hibridni grozdovi**
 - sadrže specifične uređaje
 - npr. matematički uređaji, *Graphics Processing Unit* (GPU)

Primjer grozda – Isabella

- **Računalni grozd na Srcu**
 - dostupan cijeloj akademskoj zajednici
 - <http://isabella.srce.hr>
- **Svojstva**
 - 44 čvora
 - 704 procesorske jezgre
 - 1,2 TB RAM
 - 2 TB dijeljenog diskovnog prostora SAN
- **Sun Grid Engine**
- **Dijeljeni datotečni sustav Lustre**

Primjer grozda – Isabella



Računalni splet

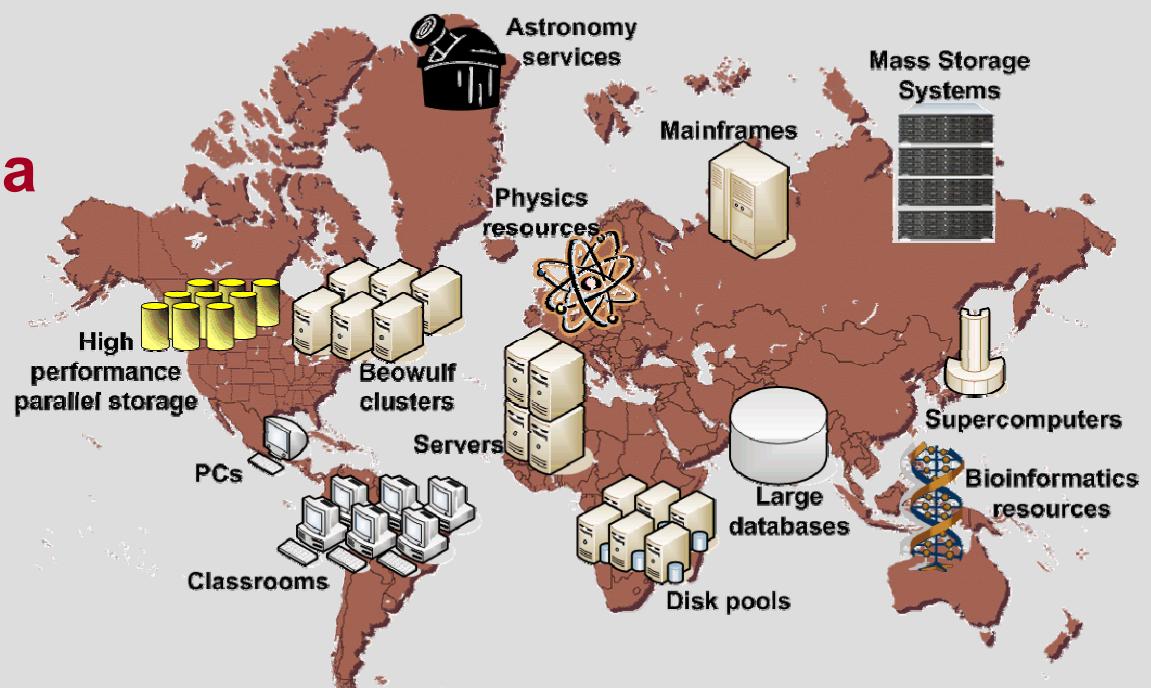


Splet

- **Splet (engl. *grid*)**
- **“složeni računalni sustav koji omogućava ujedinjavanje proizvoljnih računalnih resursa (tzv. grid sjedišta) povezanih globalnom mrežom Internet.”**
- **Analogija s električnom mrežom (engl. *grid*)**
 - računalna moć – usluga poput struje, plina, vode...

Splet

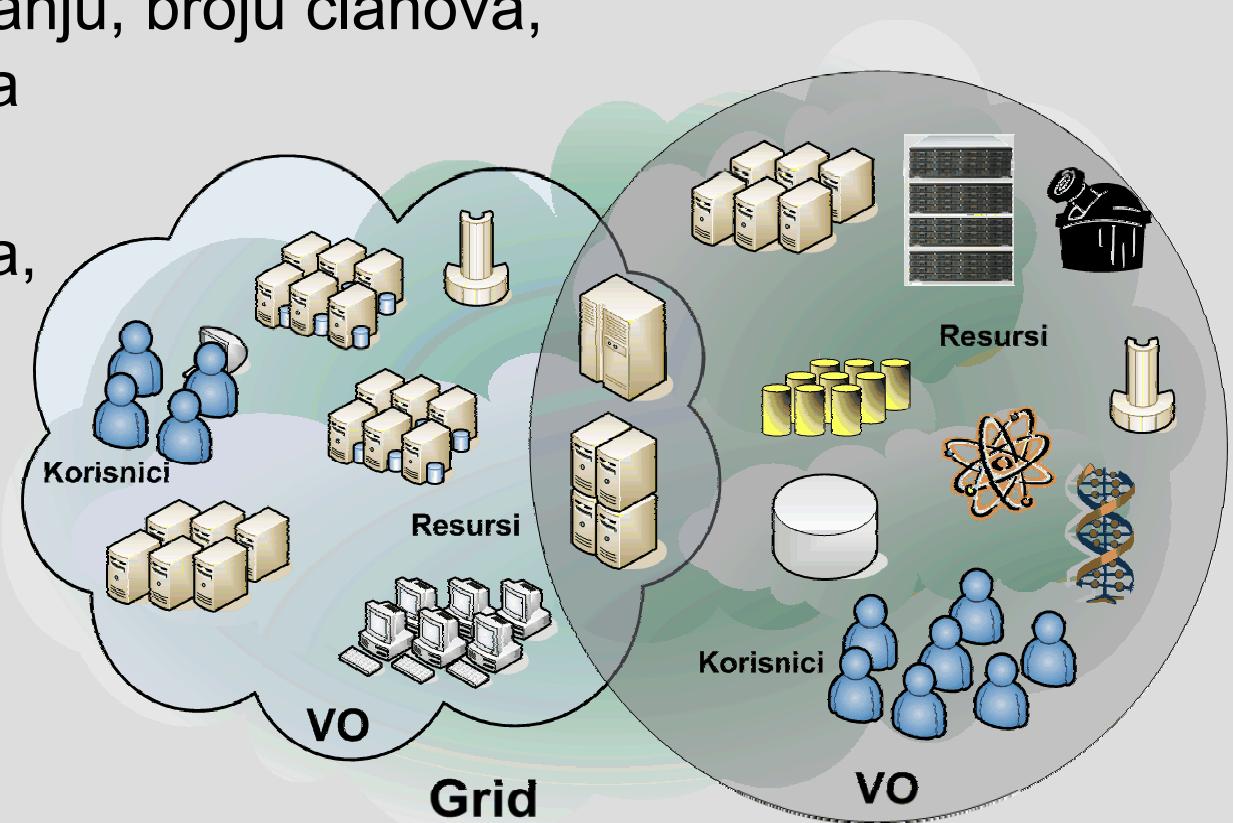
- **Resursi**
 - zemljopisno raspodijeljeni, heterogeni
 - računalni resursi (procesorska moć, spremište)
 - ostali resursi (instrumenti, senzori)
- **Ujedinjavanje resursa**
 - stvaranje privida superračunala
 - sigurno, koordinirano, kontrolirano



Splet

- **Virtualne organizacije**

- organizacije, osobe, resursi s istim interesima
- dinamične u trajanju, broju članova, ulogama članova
- npr. astrofizika, računalna kemija, biomedicina, itd.



Usporedba grozda i spletta

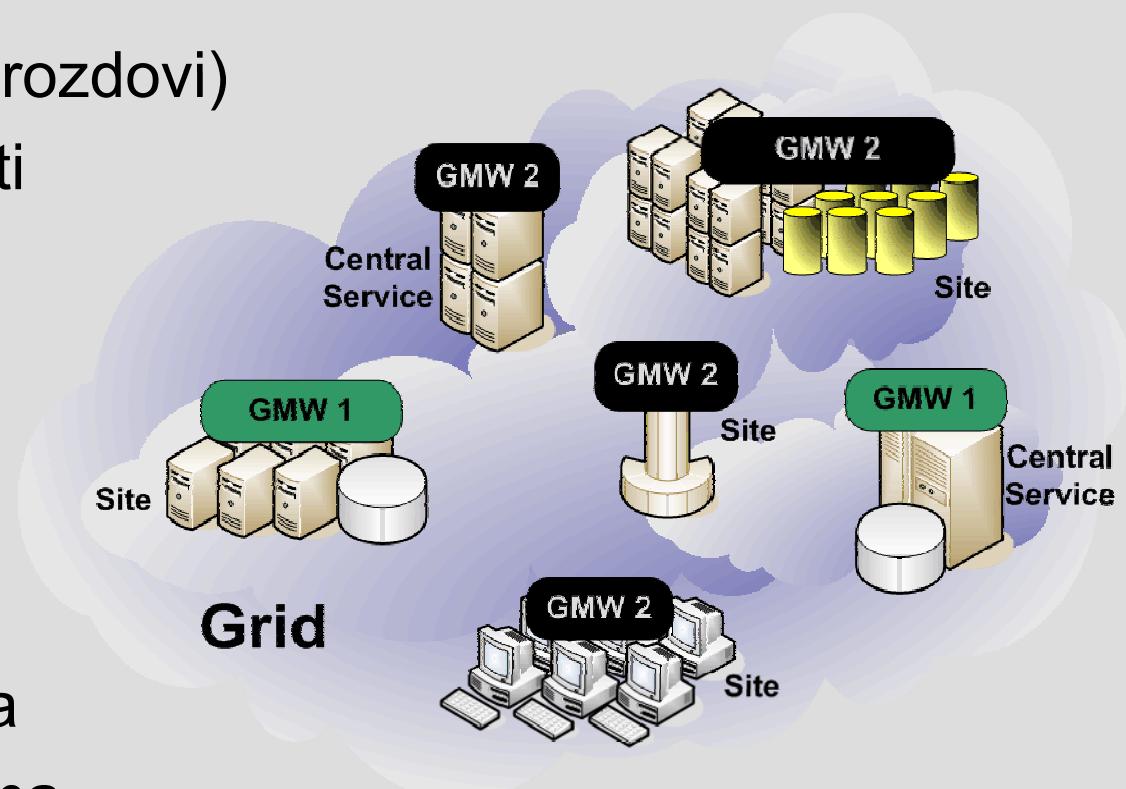
Kategorija	Grozd	Splet
Okolina	<ul style="list-style-type: none">- resursi usko povezani (lokalna mreža)- središnje upravljanje- resursi isključivo namijenjeni grozdu- jedinstvena politika korištenja	<ul style="list-style-type: none">- raspodijeljeni resursi- nema jedinstvene točke upravljanja- resursi nisu isključivo namijenjeni spletu- različite politike korištenja
Sigurnost	otvaranje računa na klasteru	digitalni i zastupnički certifikati, virtualne organizacije

Usporedba grozda i splet

Kategorija	Grozd	Splet
Upravljanje poslovima	jedinstven i centraliziran sustav za upravljanje poslovima (npr. Torque i SGE)	nema potpunog rješenja, primjeri: Condor-G, GridWay, gLite WMS, GRAM2, GRAM4
Upravljanje podacima	NFS na frontendu, lokalni disk na čvorovima	GridFTP, gLite alati
Pristup	SSH	pristupni čvorovi (UI), grid portali, osobna računala

Arhitektura spletta

- **Sjedišta**
 - računalni resursi (grodzovi)
 - podatkovni elementi
- **Središnji servisi**
 - sigurnosni servisi
 - nadzorni servisi
 - raspoređivanje i upravljanje izvođenjem poslova
 - upravljanje podacima
 - pristupni čvorovi



Standardi

- **Open Grid Forum (OGF)**
 - organizacija za standardizaciju (bivši GGF)
- **Open Grid Service Architecture (OGSA)**
 - općenita arhitektura spleteta zasnovana na web servisima
- **Web Service Resource Framework (WSRF)**
 - aktualni standard za ostvarenje spleteta
 - ne koristi se isključivo za spletove
 - još uvijek nije prihvaćen u praksi

Posrednički sustavi

- **Programska potpora koja omogućava rad spletom**
- **Funkcionalnosti**
 - sigurnost
 - nadzor i dohvati informacija
 - upravljanje podacima
 - upravljanje poslovima

Posrednički sustavi

- **Globus Toolkit 2**
 - de facto standard
 - sigurnost – GSI (engl. *Grid Security Infrastructure*)
 - nadzor – MDS (engl. *Monitoring and Discovery System*)
 - prijenos podataka – GridFTP
 - izvođenje poslova – GRAM (engl. *Grid Resource Allocation Manager*)

Posrednički sustavi

- **Globus Toolkit 4**
 - zasnovan na specifikaciji WSRF
 - sigurnost – servisi Authorization & Delegation
 - nadzor – WS-MDS
 - prijenos podataka – RFT (engl. Reliable File Transfer)
 - izvođenje poslova – WS-GRAM
 - prekinut razvoj u listopadu 2009.

Posrednički sustavi

- **UNICORE**
 - zasnovan na Javi
 - ostvarenje zasnovano na specifikaciji WSRF
- **gLite**
 - veliki broj komponenata
 - ostvaruje složene funkcionalnosti
 - upravljanje virtualnim organizacijama
 - globalno raspoređivanje poslova
 - globalno upravljanje podacima

Sigurnost

- **Veliki sigurnosni rizik**
 - direktni pristup vrijednim resursima
 - pristup različitim informacijskim izvorima
- **Integracija**
 - s postojećim sigurnosnim sustavima
 - s lokalnim sigurnosnim politikama
- **Dinamične virtualne organizacije**
- **Privatnost podataka**
- **Sociološko, politički aspekt spleta**

Sigurnost

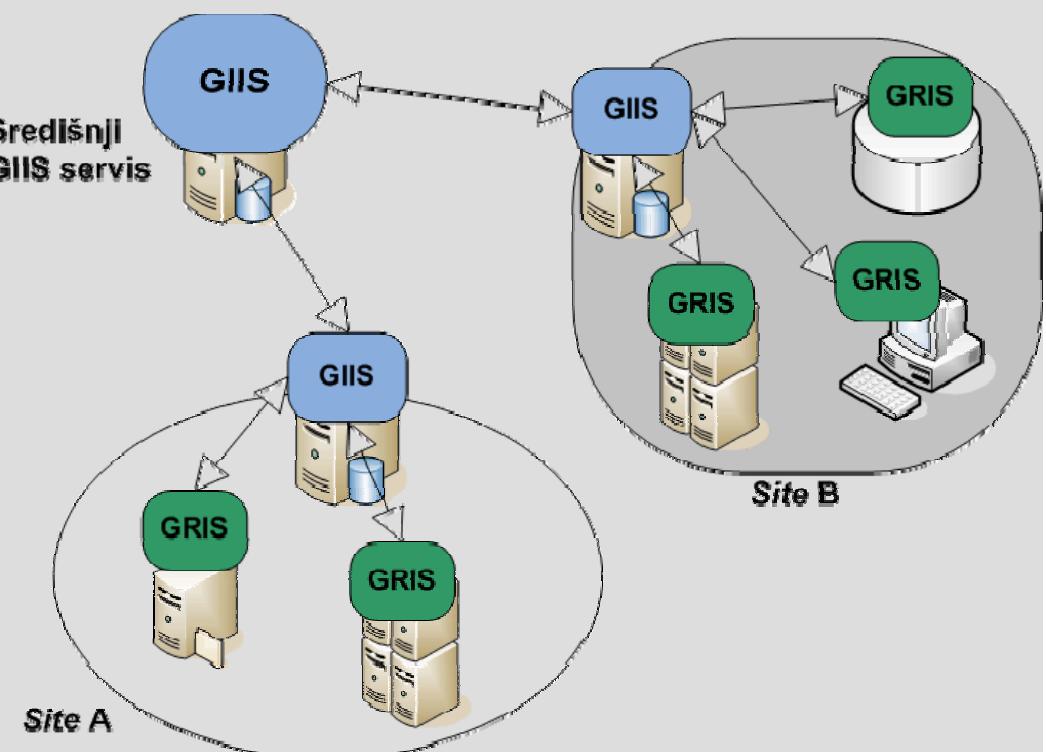
- **X509 digitalni certifikati**
 - poslužiteljski i korisnički certifikati
- **Svi posrednički sustavi zasnovani na certifikatima**
- **International Grid Trust Federation (IGTF)**
 - EUGridPMA, APGridPMA, TAGPMA
 - certifikati Certifikacijskih autoriteta (CA) članova prihvaćeni na većini spletova

Nadzor

- **Koliko čega je raspoloživo?**
 - računala, komponente, programska potpora (~statično)
- **Stanje sredstava**
 - koji poslužitelji/servisi su dostupni?
 - da li dostupni servisi izvršavaju svoje funkcije?
- **Iskorištenost i učinkovitost rada sustava**
- **Tko je i koliko koristio sustav?**
 - engl. *accounting*
- **Tko je pristupio ili pokušao pristupiti računalnim sredstvima?**
 - engl. *auditing*

Nadzor – WS-MDS

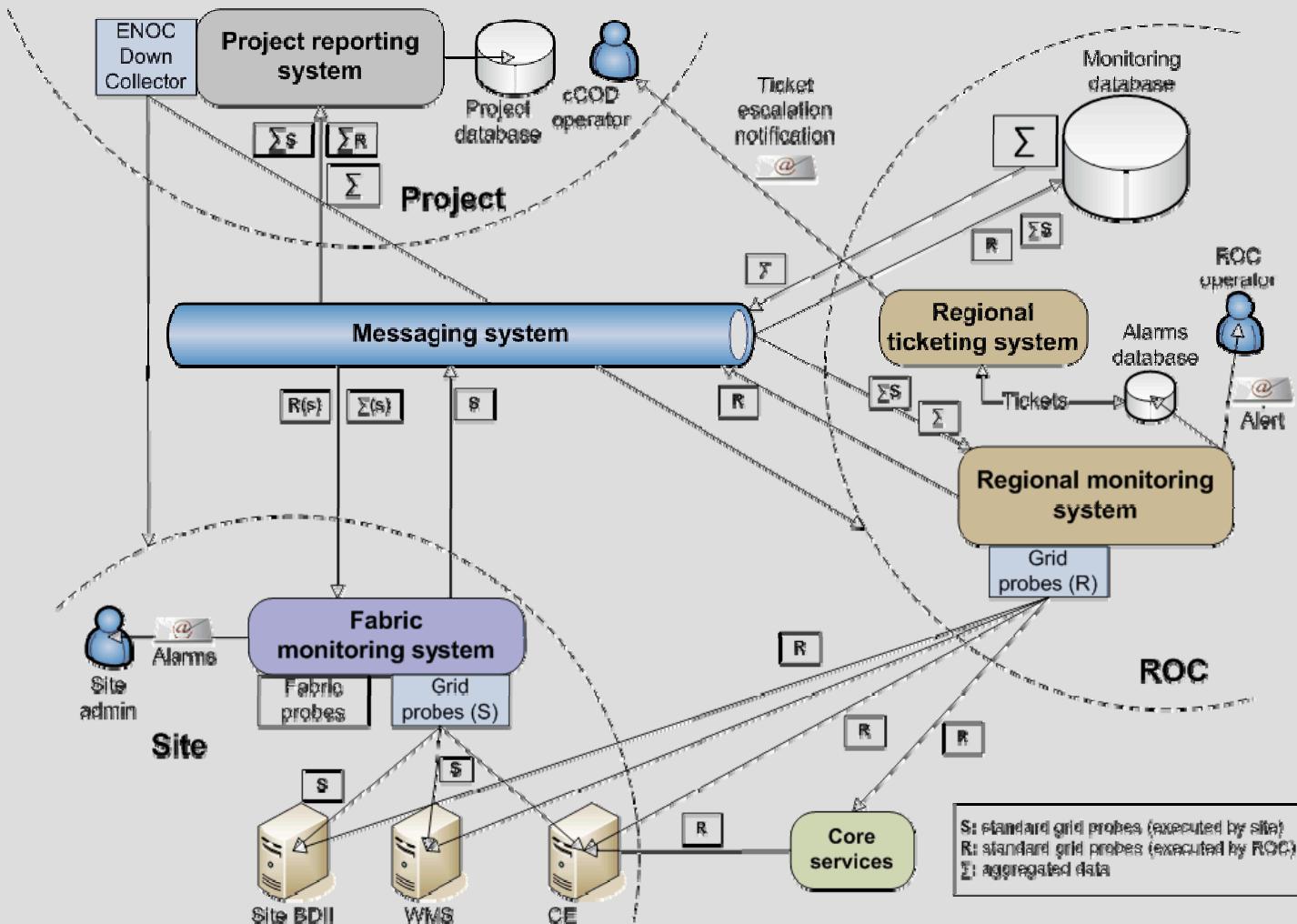
- **Monitoring and Discovery Service (MDS)**
- **Objava informacija i pretraživanje**
- **Upravljanje informacijama**
 - senzori i shema odvojeni od MDS-a
- **Zasnovan na WSRF-u**
- **Hijerarhijska arhitektura**



Nadzor – EGI nadzor servisa

- **Složeni sustav za nadzor stanja servisa**
 - obavještavanje administratora u slučaju problema
 - prikaz stanja sjedišta za operatore spletu
 - izrada izvještaja o dostupnosti i pouzdanosti
 - informacije za sustave za upravljanje poslovima
 - informacije o sjedištima za virtualne organizacije
- **Zasnovan na standardnim tehnologijama**
 - ActiveMQ messaging service, sustav za nadzor Nagios, JSON, XML

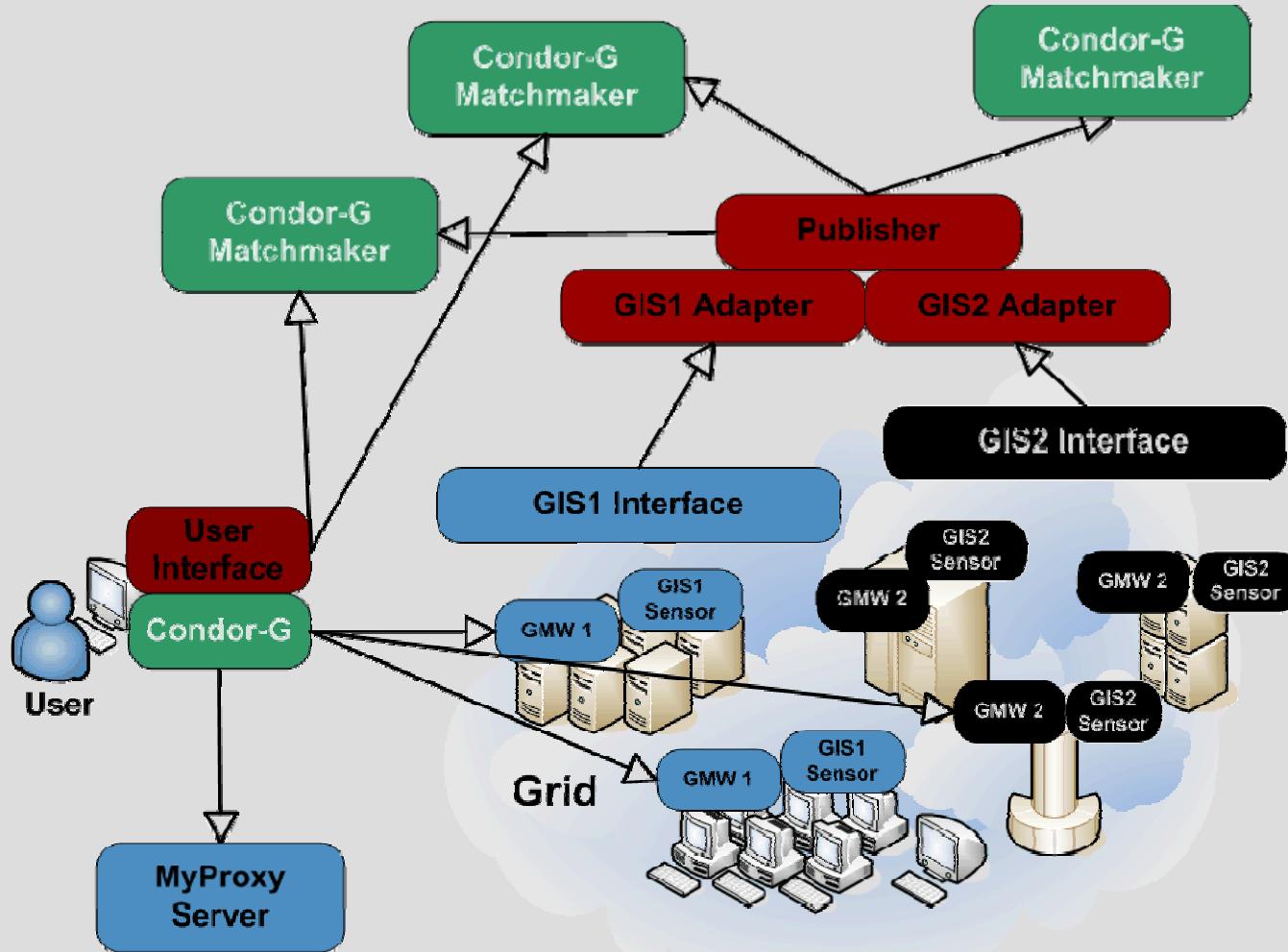
Nadzor – EGI nadzor servisa



Upravljanje poslovima

- **Prijenos datoteka na sjedišta**
- **Podržani posrednički sustavi**
 - GT2, GT4, UNICORE, Condor, ...
- **Otpornost na pogreške**
 - najčešće ponovno raspoređivanje posla (rescheduling)
- **Podržani tipovi poslova**
 - serijski, paralelni, paralelni na više sjedišta, workflow, polje poslova

Upravljanje poslovima – Condor-G



Ostvarenja spletja

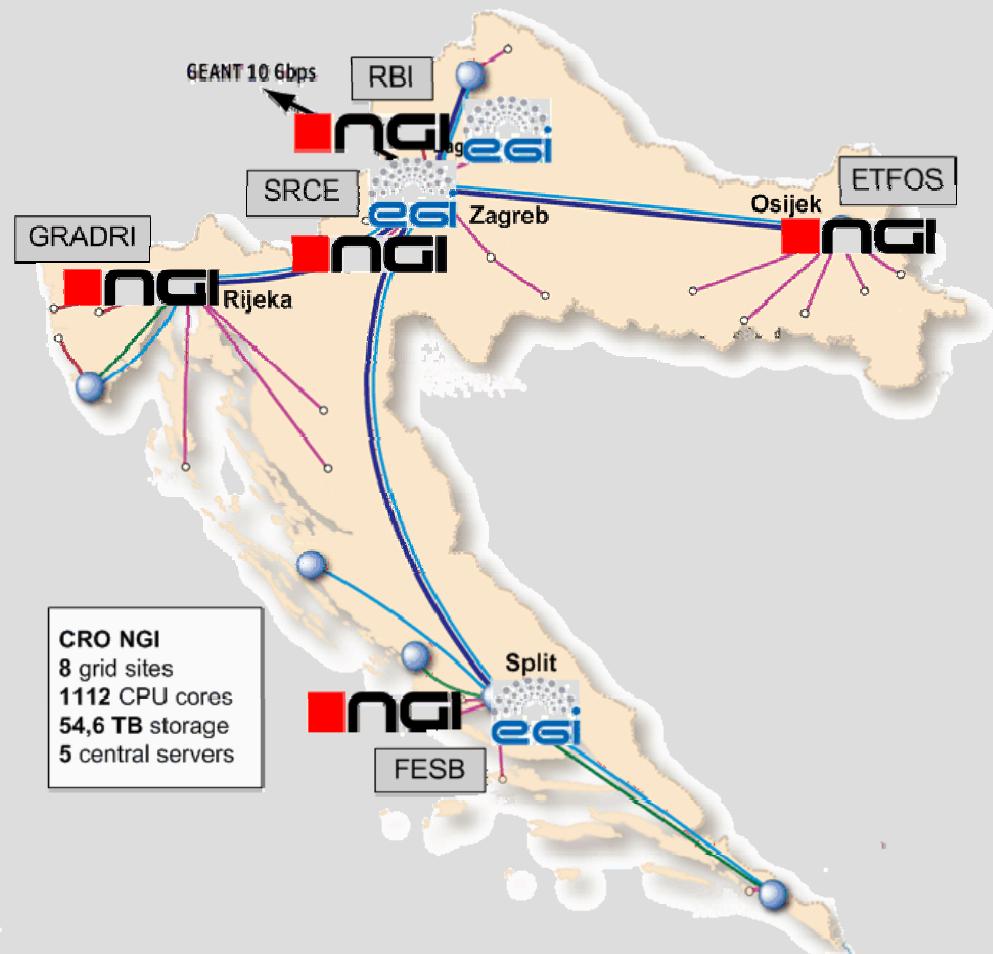
- **EGI**
 - engl. *European Grid Initiative*
 - 45 država (službeno povezanih)
 - 332 sjedišta
 - preko 200 000 procesorskih jezgri
 - preko 100 PB diskovnog prostora
 - posrednički sustavi gLite, ARC, UNICORE
- **DEISA**
 - superračunalni grid
 - 11 institucija
 - posrednički sustav UNICORE

Ostvarenja spletja

- **Open Science Grid**
 - bivši Grid3
 - GryPhyN, PPDG, iVDLG
- **TeraGrid**
 - ekvivalent DEISA-e
 - 11 institucija
 - više od 1 PFLOPs
 - više od 30 PBs prostora
 - 40Gb/s mrežne veze

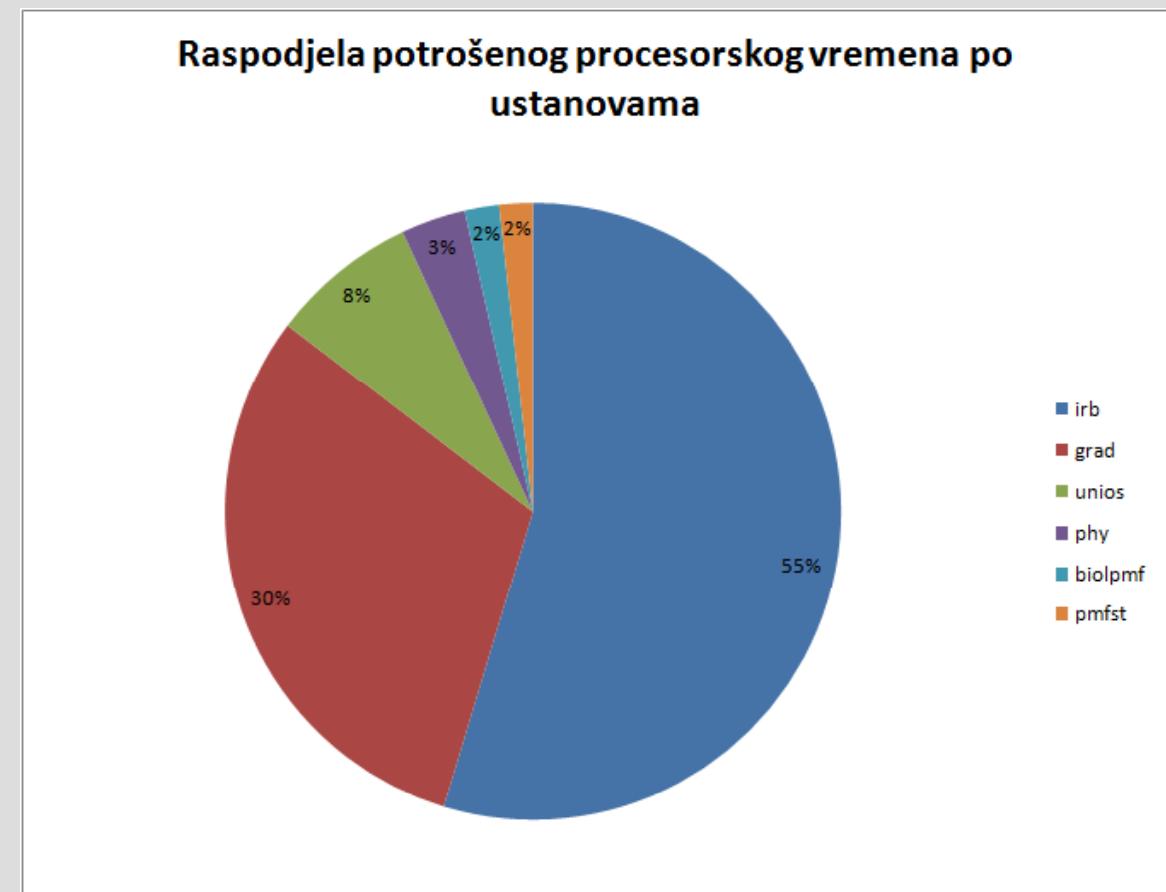
Ostvarenja spletta

- **CRO NGI**
 - splet u Hrvatskoj
 - dostupan akademskoj zajednici
 - 8 sjedišta
 - 1112 procesorske jezgre
 - 54,6 TB diskovnog prostora



Ostvarenja spletu

- **CRO NGI**
 - 8 ustanova partnera
 - 25 ustanova korisnica
 - 106 korisnika



Literatura

- M. Baker, et al, “**Cluster Computing White Paper**”, 2001
- R. Buyya (editor), “**High Performance Cluster Computing: Architectures and Systems, Volume 1**”, Prentice Hall, NJ, USA, 1999
- R. Buyya (editor), “**High Performance Cluster Computing: Programming and Applications, Volume 2**”, Prentice Hall, NJ, USA, 1999
- W. Gropp, K. Kennedy, L. Torczon, A. White, J. Dongarra, I. Foster, and G. C. Fox, “**The Sourcebook of Parallel Computing**”, Morgan Kaufmann, 2002

Literatura

- I. Foster and C. Kesselman, “**The Grid 2: Blueprint for a New Computing Infrastructure**”, Morgan Kaufmann, 2003
- F. Berman, G. Fox, and A. J. G. Hey, “**Grid Computing: Making the Global Infrastructure a Reality**”, Wiley, 2003
- I. Foster, “**What is the Grid? A Three Point Checklist**”, GRIDToday, 2002
- I. Foster, Y. Zhao, I. Raicu, and S. Lu, “**Cloud computing and grid computing 360-degree compared**”, 2008

Literatura

- W. Gentzsch, “**Sun Grid Engine: towards creating a compute power grid**”, Proceedings of the first IEEE/ACM International Symposium on Cluster Computing and the Grid (2001), pp 35-36, 2001
- I. Foster, “**Globus Toolkit Version 4: Software for Service-Oriented Systems**”, IFIP International Conference on Network and Parallel Computing, Springer-Verlag LNCS 3779, pp 2-13, 2006

Hvala na pažnji!

