

Grid računarstvo i računarstvo visokih performansi

Vladimir Slavnić

Laboratorija za primenu računara u nauci
Institut za fiziku u Beogradu
Univerzitet u Beogradu

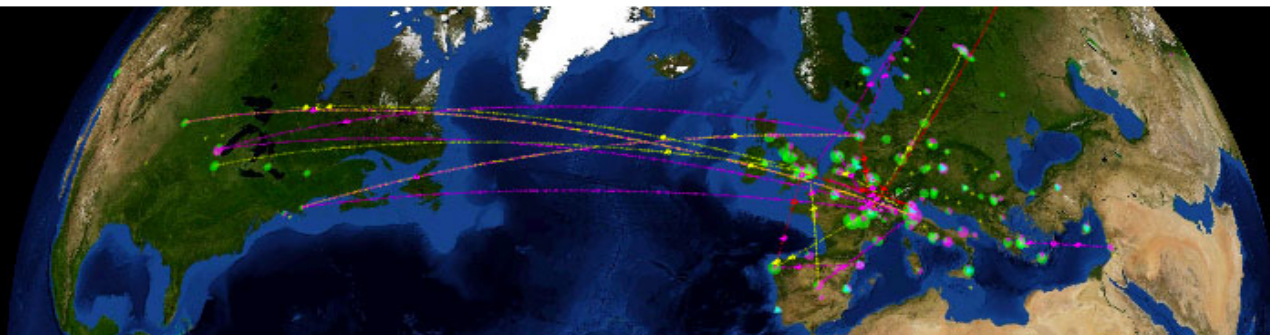


13th International Masterclasses 2017
Fizički fakultet
Univerzitet u Beogradu
06. mart 2017.

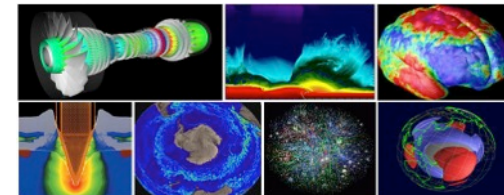
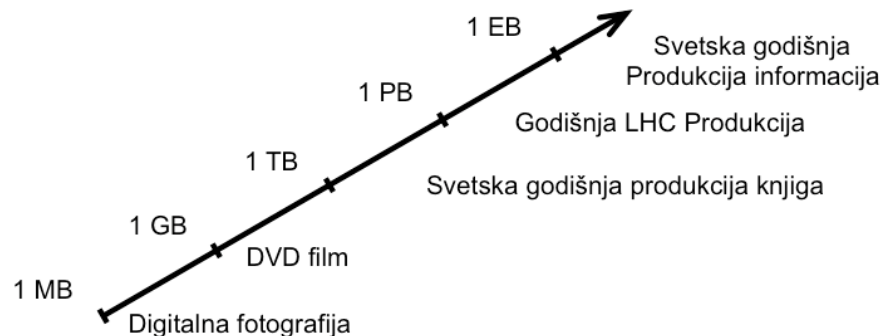
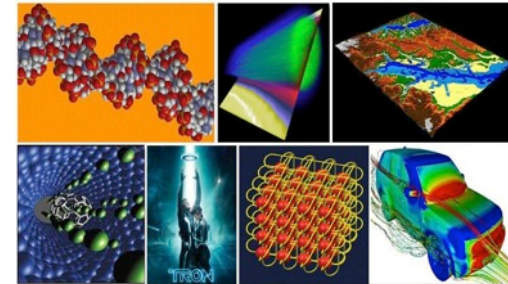
INSTITUTE OF PHYSICS
BELGRADE



- Motivacija
- Mogućnosti i Grid kao rešenje
- Grid arhitektura i middleware
- Računarstvo visokih performansi
- Grid i HPC inicijative
- Nacionalna infrastruktura



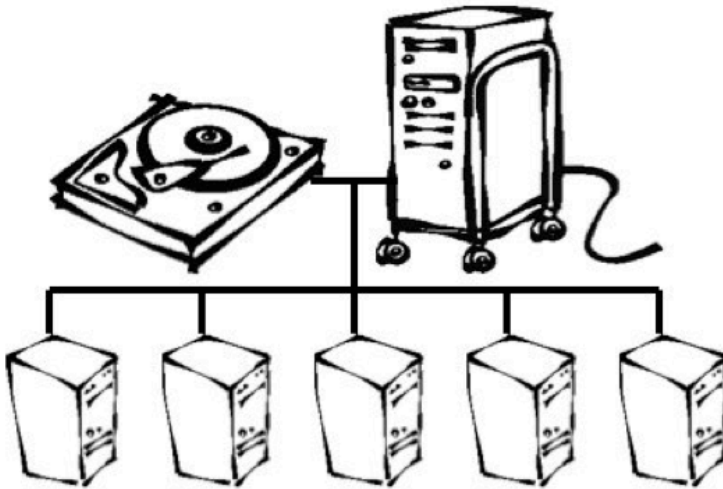
- Nauka postaje sve više digitalna
 - Numeričke simulacije
- Fizika elementarnih čestica je jedna od važnih naučnih disciplina koja koristi velike količine računarskih resursa
 - Velike međunarodne kolaboracije
 - Velike količine podataka iz eksperimenata
 - Large Hadron Collider (LHC) u CERN-u
 - Više od 600 miliona sudara u sekundi
 - ~50 PB/god (čak i kad se isfiltrira 99% podataka)



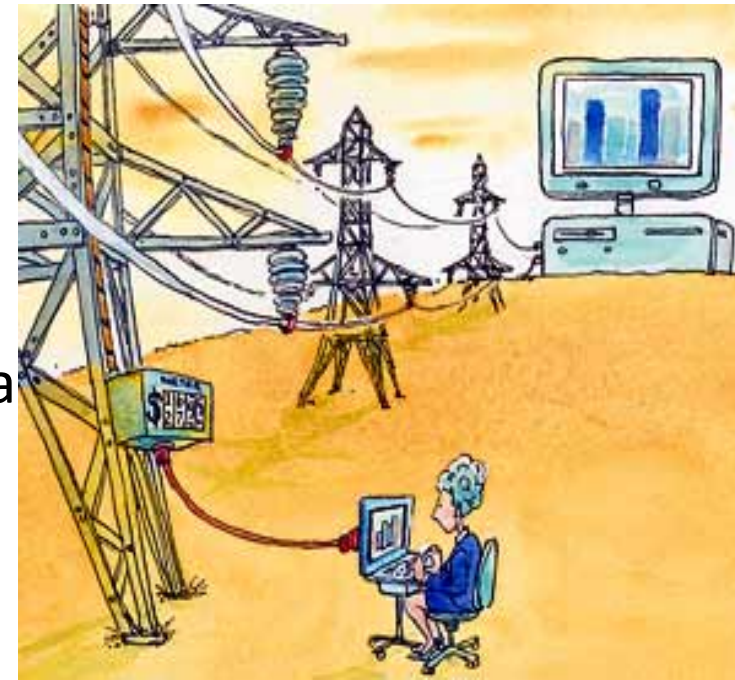
- Za simulaciju jednog sudara (događaja) u nekom od LHC detektora potrebno je ~15 min. na današnjim PC procesorima
- Za dobru analizu jednog raspada potrebno je u proseku 100,000 događaja
- Za potpunu simulaciju samo jednog raspada potrebno je u proseku 3 godine procesorskog vremena
- Analiza podataka na LHC zahteva kompjutersku snagu ekvivalentnu 100,000 današnjih PC procesora

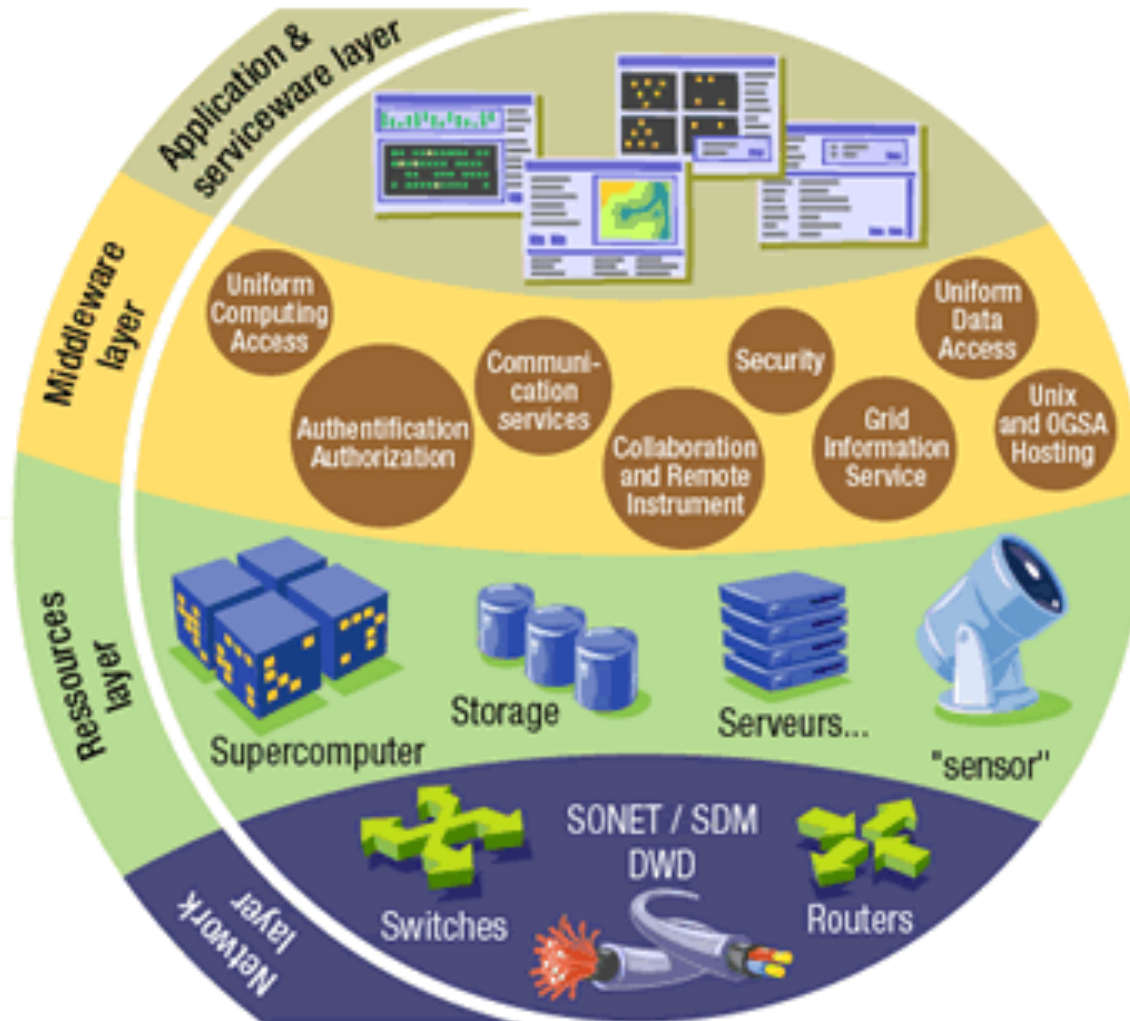


- World Wide Web – jednostavna komunikacija; omogućeno jednostavno predstavljanje i povezivanje dokumenata na svetskom nivou pomoću hiperlinkova
- Linux – pouzdan operativni sistem
- Switch tehnologija – brza komunikacija između računara
- Klaster – mnoštvo računara povezanih u jedinstvenu celinu

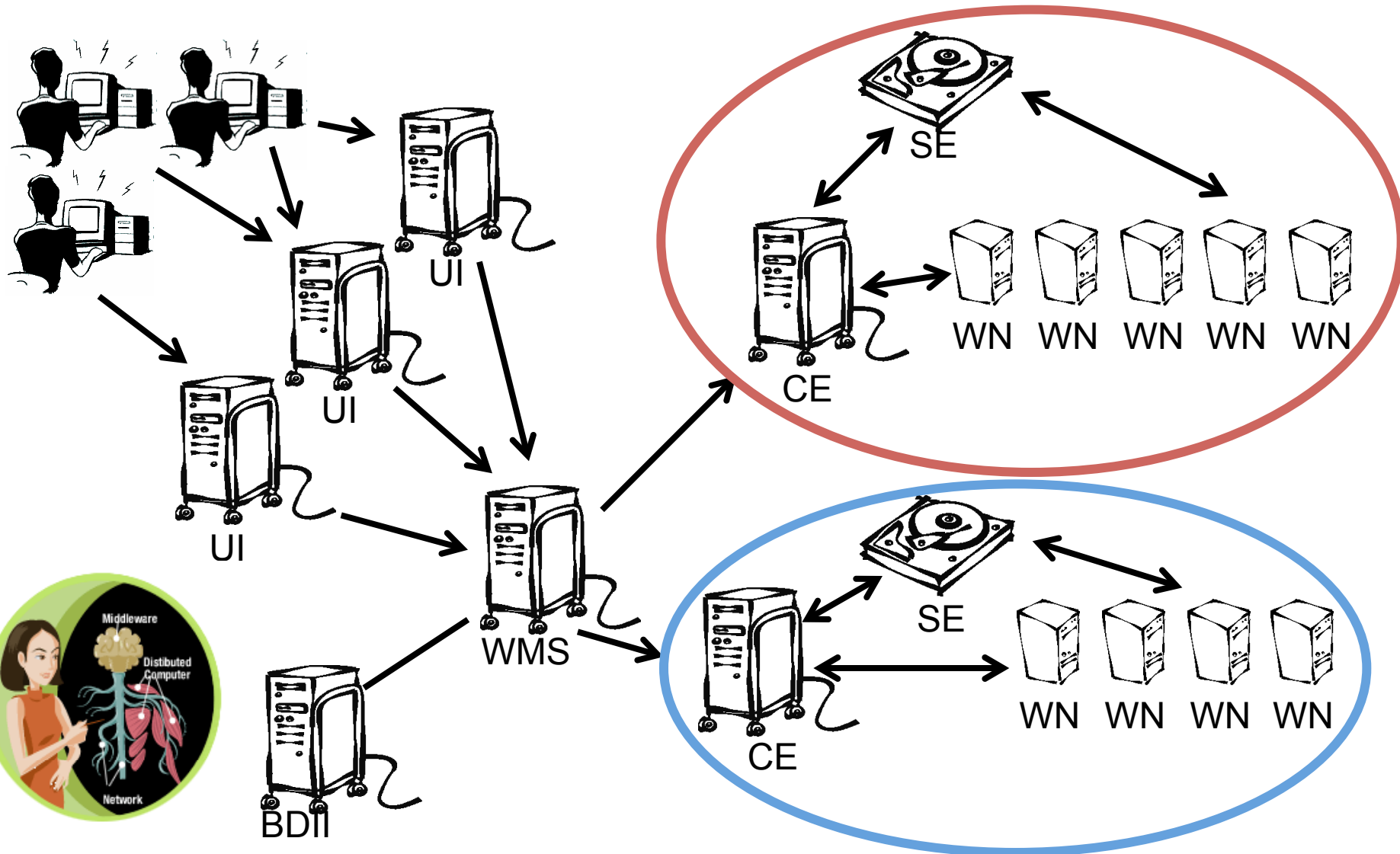


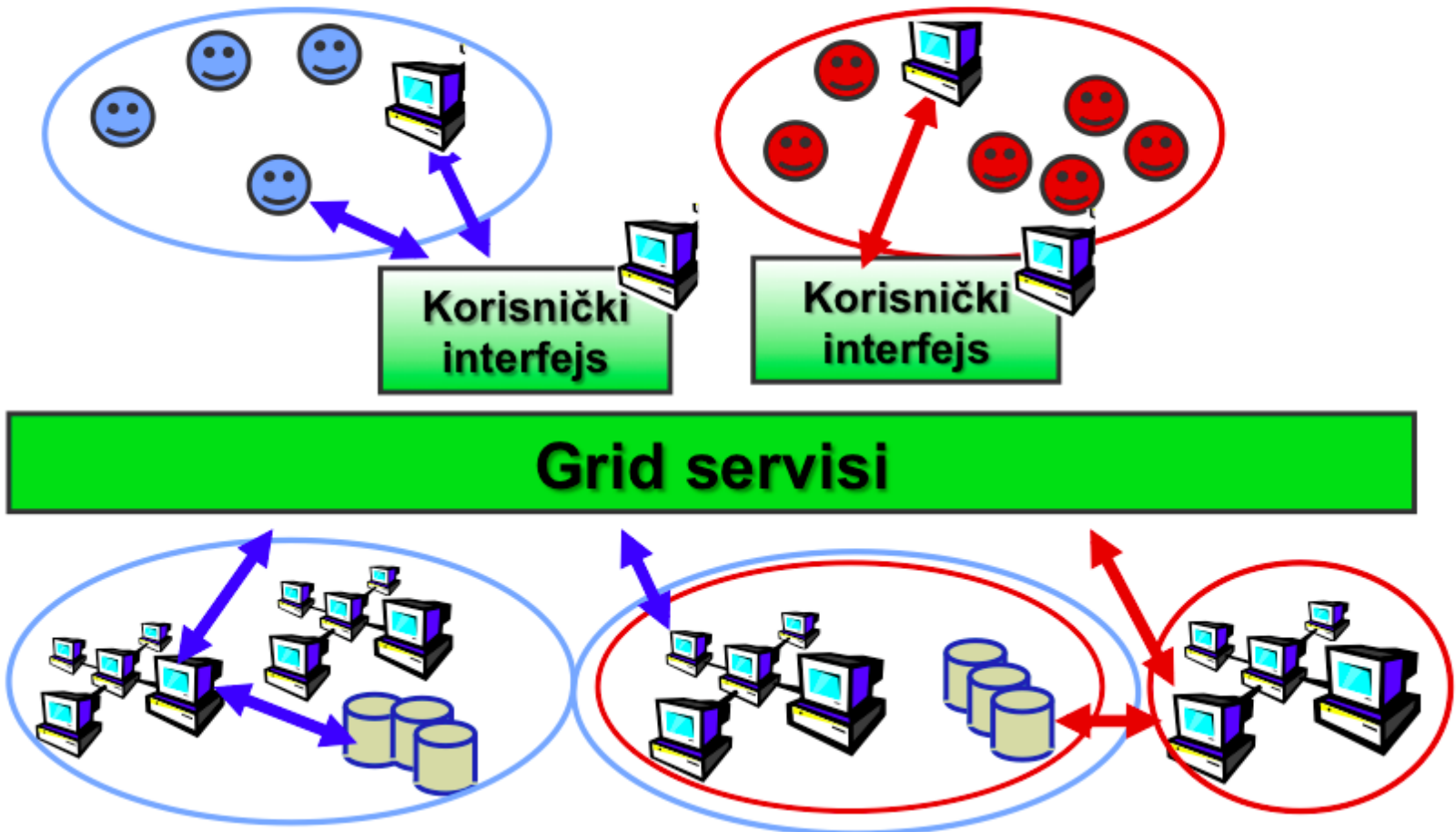
- Grid računarstvo
 - Geografski distribuiran kompjuterski sistem
 - Baziran na deljenju procesorskih i skladišnih resursa
 - Analogon električnoj mreži
- Zajednička upotreba distribuiranih računarskih resursa (CPU, podaci, itd.) na siguran način, tako da korisnici mogu da saraduju u okviru virtuelnih organizacija (VO)

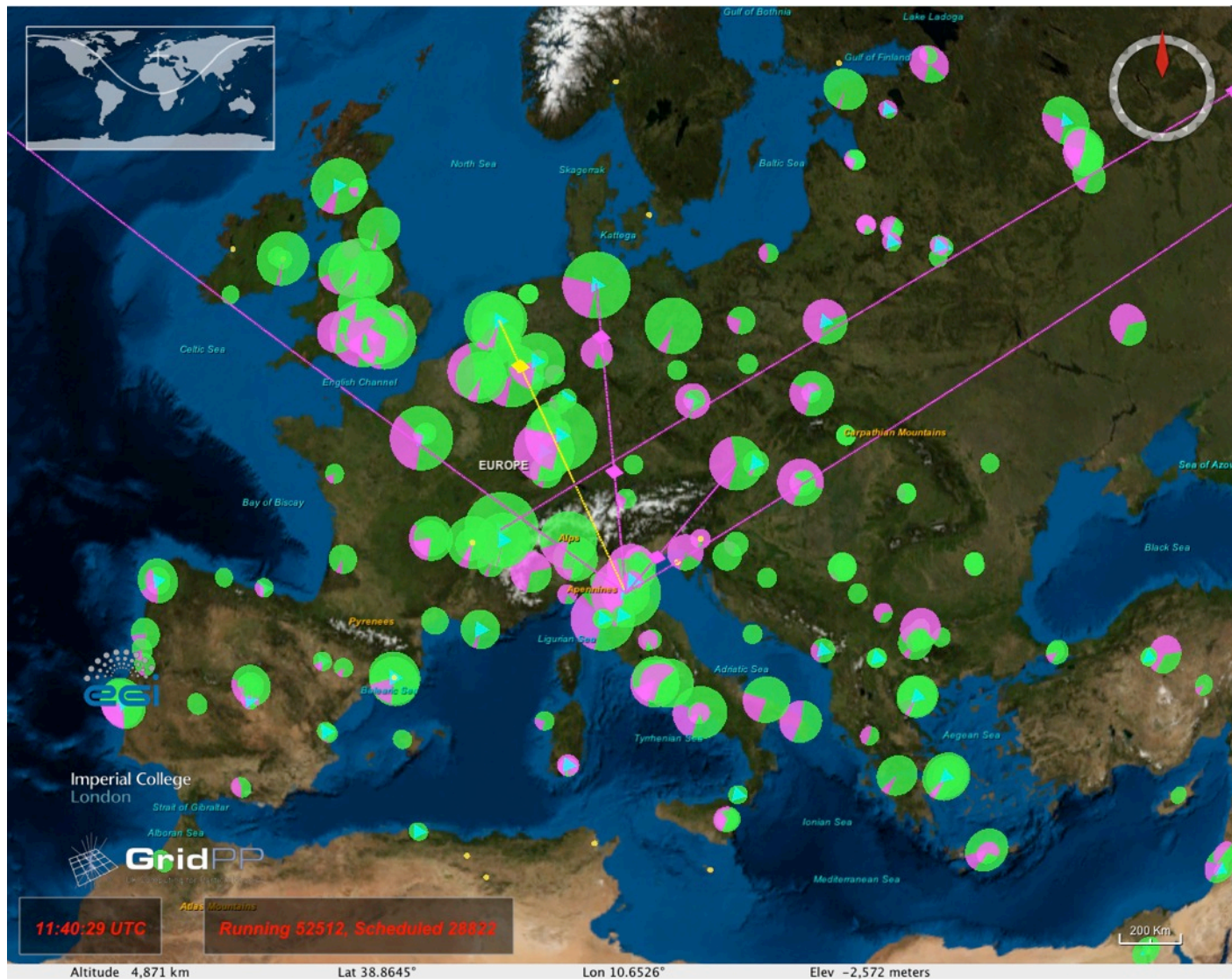




- Grid počiva na naprednom softveru, koji zovemo middleware I koji povezuje resurse i aplikacije
- Grid middleware
 - Nalazi odgovarajuće klastere na kojima aplikacije mogu da budu izvršene
 - Optimizuje upotrebu resursa
 - Organizuje efikasan pristup podacima
 - Vodi računa o autentifikaciji korisnika na različitim klasterima
 - Obezbeđuje izvršavanje simulacija i nadgledanje njihovog statusa
 - Isporučuje rezultate simulacija naučnicima
- Najrasprostranjeniji je gLite middleware







- Capacity vs. Capability computing
- Uvek nam je potrebno više računarske snage
- Želimo da smanjimo vreme izvršavanja naših aplikacija
- HPC sistemi pokušavaju da simultano maksimizuju performanse u tri važna aspekta:
 - Računanje (number crunching) – broj operacija u sekundi
 - Pristup podacima – brzine prenosa podataka unutar ili između računara
 - Skladištenje podatka – kapaciteti uređaja za skladištenje: keš, RAM, diskovi...
- Smanjivanje vremena izvršavanja programa se dobija paralelizacijom
- HPC zajedno čine:
 - Brzi računarski sistemi
 - Brze računarske mreže (interconnect)
 - Softverski alati
 - Paralelne programske paradigmekoji omogućavaju ranije nezamisliva računanja mogućim



| Rank | Site | System | Cores | Rmax (TFlop/s) | Rpeak (TFlop/s) | Power (kW) |
|------|--|--|------------|-------------------|--------------------|---------------|
| 1 | National Supercomputing Center in Wuxi China | Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway NRCPC | 10,649,600 | 93,014.6 | 125,435.9 | 15,371 |
| 2 | National Super Computer Center in Guangzhou China | Tianhe-2 (MilkyWay-2) - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P NUDT | 3,120,000 | 33,862.7 | 54,902.4 | 17,808 |
| 3 | DOE/SC/Oak Ridge National Laboratory United States | Titan - Cray XK7 , Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x Cray Inc. | 560,640 | 17,590.0 | 27,112.5 | 8,209 |
| 4 | DOE/NNSA/LLNL United States | Sequoia - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom IBM | 1,572,864 | 17,173.2 | 20,132.7 | 7,890 |
| 5 | DOE/SC/LBNL/NERSC United States | Cori - Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect Cray Inc. | 622,336 | 14,014.7 | 27,880.7 | 3,939 |
| 6 | Joint Center for Advanced High Performance Computing Japan | Oakforest-PACS - PRIMERGY CX1640 M1, Intel Xeon Phi 7250 68C 1.4GHz, Intel Omni-Path Fujitsu | 556,104 | 13,554.6 | 24,913.5 | 2,719 |
| 7 | RIKEN Advanced Institute for Computational Science (AICS) Japan | K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect Fujitsu | 705,024 | 10,510.0 | 11,280.4 | 12,660 |
| 8 | Swiss National Supercomputing Centre (CSCS) Switzerland | Piz Daint - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100 Cray Inc. | 206,720 | 9,779.0 | 15,988.0 | 1,312 |
| 9 | DOE/SC/Argonne National Laboratory United States | Mira - BlueGene/Q, Power BQC 16C 1.60GHz, Custom IBM | 786,432 | 8,586.6 | 10,066.3 | 3,945 |
| 10 | DOE/NNSA/LANL/SNL United States | Trinity - Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Aries interconnect Cray Inc. | 301,056 | 8,100.9 | 11,078.9 | 4,233 |

- Grid:
 - Worldwide LHC Computing Grid (WLCG)
 - Preko 40 država, ~170 sajtova
 - European Grid Infrastructure (EGI)
 - ~300 sajtova, ~730,000 CPU, ~300 PB
 - EGI Federated Cloud
 - Grid akademskih cloud infrastruktura i virtualizovanih resursa
- Superračunari:
 - PRACE - Evropska HPC inicijativa
 - PRACE Tier-0 superračunari (Rmax)
 - CURIE (1.7 PFlop/s), MARCONI (13 PFlop/s), Hazel Hen (7.4 PFlop/s), JUQUEEN (5.9 PFlop/s), MareNostrum (1.0 PFlop/s), Piz Daint (9.8 PFlop/s), SuperMUC (3.2 PFlop/s)
- Regionalne inicijative
 - VI-SEEM projekat
 - Dosadašnji projekti
 - SEE-GRID serija projekata
 - HP-SEE - HPC inicijativa u jugoistočnoj evropi
- Institut za fiziku Beograd koordinira Grid i HPC aktivnostima u Srbiji

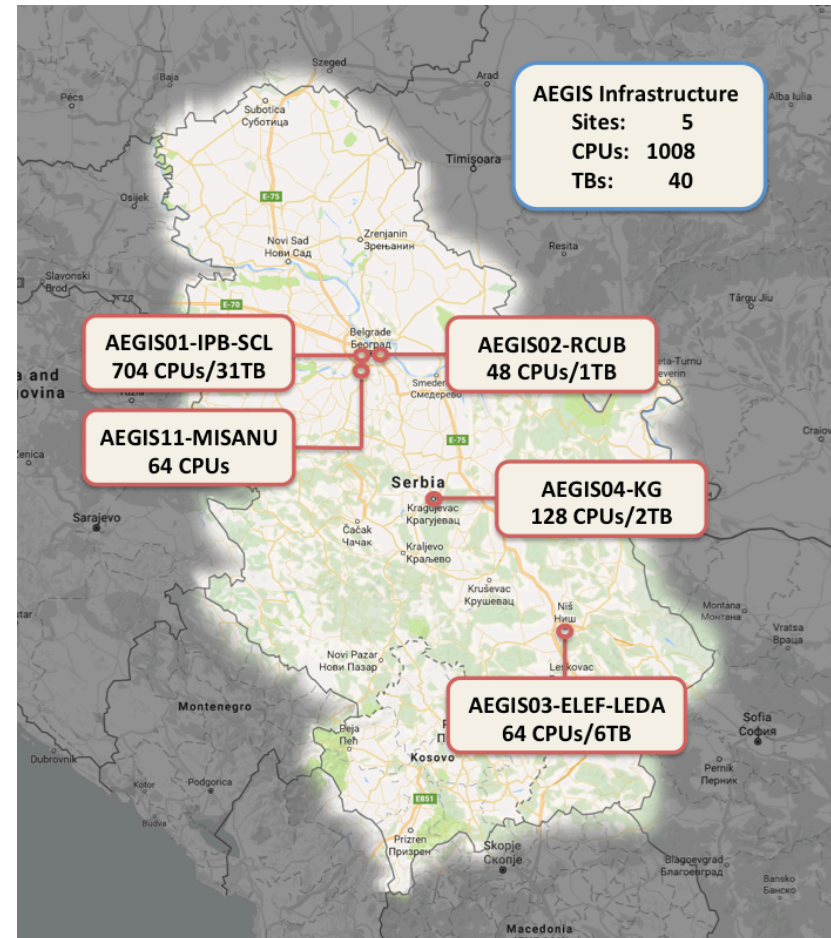


WLCG
Worldwide LHC Computing Grid



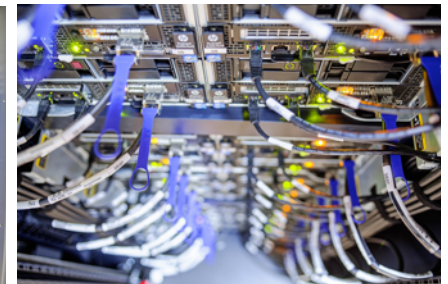
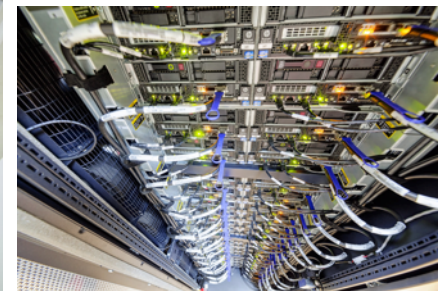
Nacionalna Grid infrastruktura

- AEGIS - Akademska i obrazovna Grid inicijativa Srbije
 - Formirana 2005. godine (~1000 CPUs, 40 TB)
- Institut za fiziku:
 - PARADOX III (AEGIS01-IPB-SCL)
 - Core Grid servisi



Nacionalna HPC infrastruktura [1/2]

- PARADOX IV instalacija
 - Rekovi sa vodenim hladjenjem
 - 106 mašina, 1696 CPUs
 - 2x8-core Intel Sandy Bridge E5-2670 @ 2.6 GHz
 - 32 GB of RAM
 - 106 GPU NVIDIA Tesla M2090
 - 96 TB skladišnog prostora
 - Lustre paralelni fajl sistem
 - QDR Infiniband 40 Gbps



- BioIRC HPC sistem
 - Lociran u BioIRC - Bioengineering R&D Centru u Kragujevcu
 - 2 management/storage servera - 12TB skladišnog prostora, Lustre FS
 - 22 WNs
 - 2 x Intel Xeon E5-2670, 16 cores, 64GB RAM
 - 4 Tesla M2090
 - Infiniband QDR interconnect



