Grid Training for Power Users
Institute of Physics Belgrade

e-infrastructure

Grid Training for Power Users
Institute of Physics Belgrade

Overview of the Training Agenda

Dusan Vudragovic

SCL, Institute of Physics Belgrade

Serbia

dusan@ipb.ac.rs

Training Agenda

HANDS-ON

- Overview of Grid Computing
- Setting Up User Account
- Job Management
  - Single Job Submission
  - Job Collections Submission
  - Parametric Job Submission
  - Direct Acyclic Graphs (DAG) Job Submission
  - Message Passing Interface (MPI) Job Submission
- Data Management
- Advanced Job Submission

http://indico.ipb.ac.rs/conferenceDisplay.py?confId=280

Grid Training for Power Users
Institute of Physics Belgrade

Overview of Grid Computing
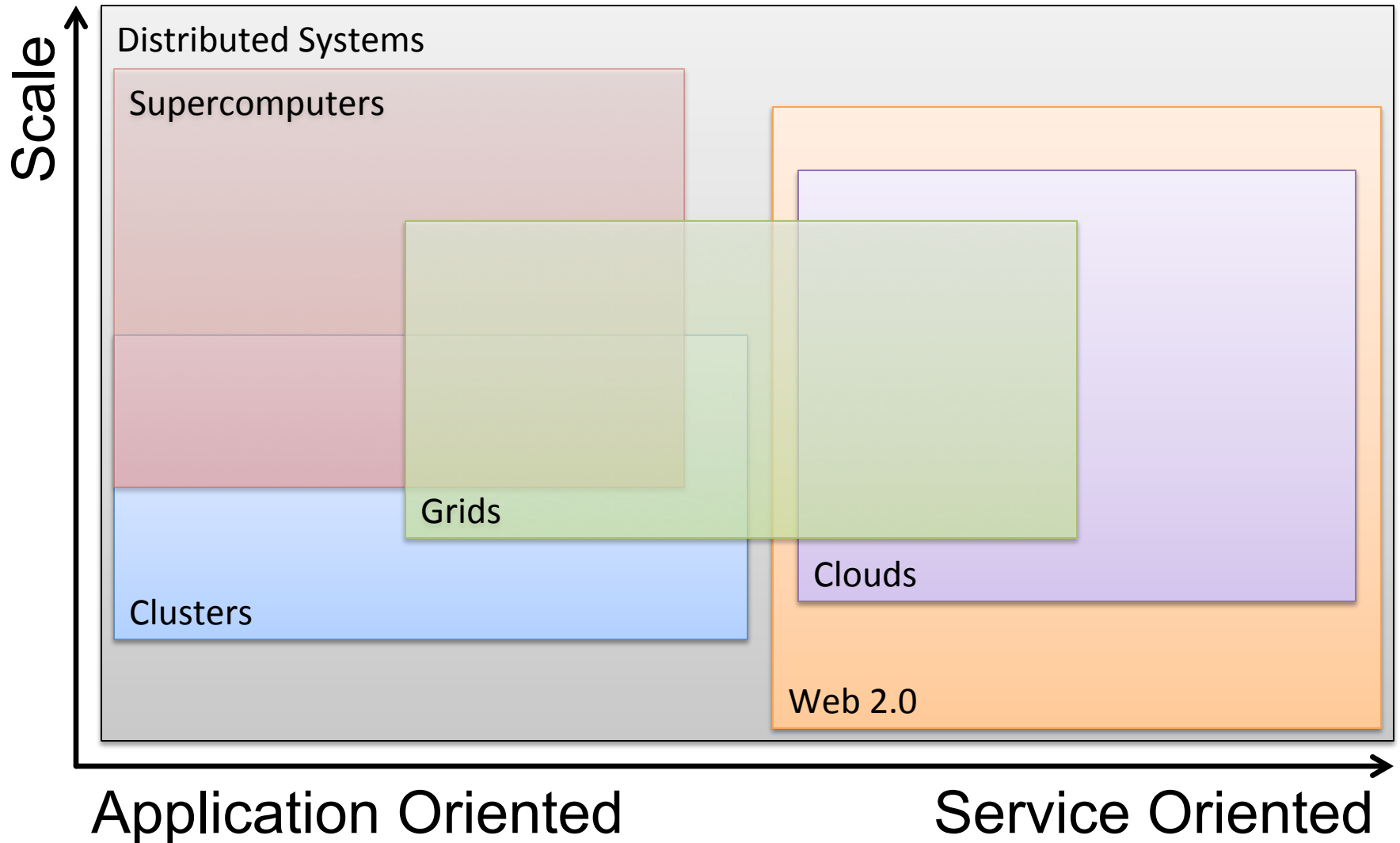
Dusan Vudragovic

SCL, Institute of Physics Belgrade

Serbia

dusan@ipb.ac.rs

28/05/2012

- Distributed Computing Systems
- AEGIS - Serbian Grid Infrastructure
- EGI-InSPIRE
- gLite Architecture

- Distributed system consists of multiple autonomous computers that communicate through a computer network

- Clusters, Supercomputers, Clouds, Grids
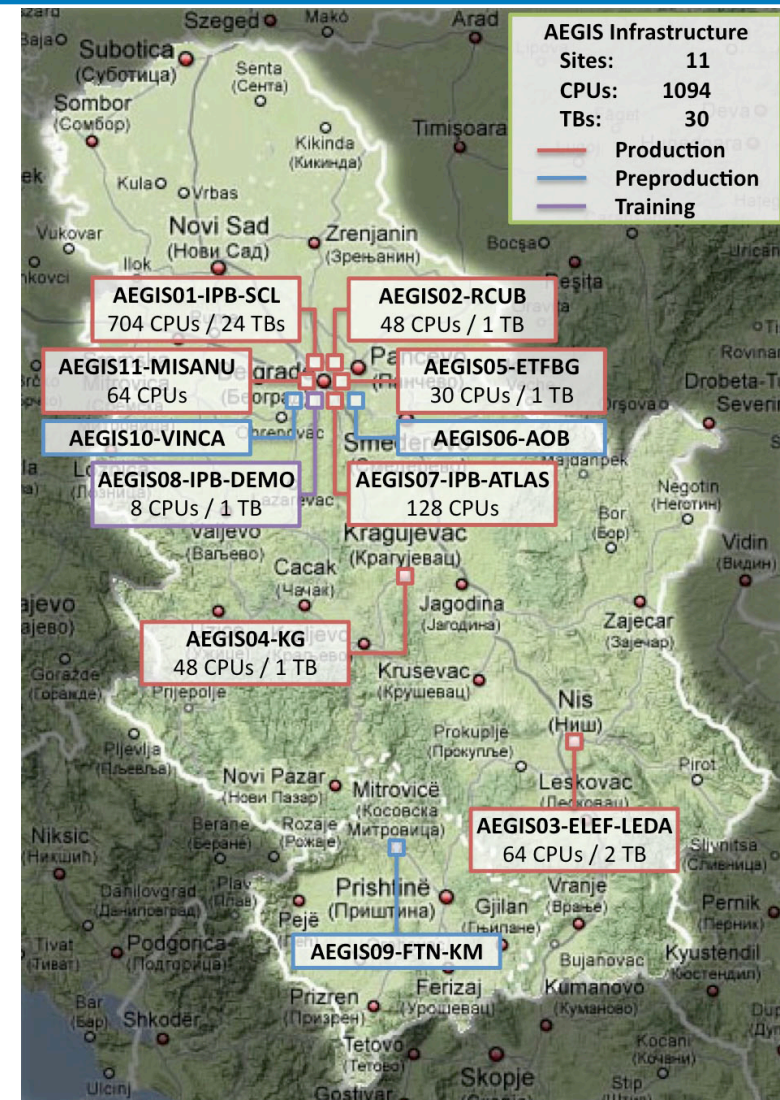
Grid Characteristics

- Federated yet separately administered resources, spanning multiple sites, countries and continents
- Heterogeneous resources (e.g. hardware architectures, operating systems, storage back-ends, network setups)
- Distributed, multiple research user communities (including users accessing resources from varied administration domains) grouped in Virtual Organizations (VO)
- Mostly publicly funded (both resources and engineering, but not necessarily from the same funding source), at local, national and international levels
- Range of data models, ranging from massive data sources, hard to replicate (e.g. medical data only accessible at hospital premises), to transient datasets composed of varied file sizes

- Grid middleware provides the user with high level services for:
  - Identification of entities (i.e. users, systems and services)
  - Scheduling and running computational jobs
  - Accessing and moving data
  - Obtaining information on the grid infrastructure, applications

AEGIS

- AEGIS = **A**cademic and **E**ducational **G**rid **I**nitiative of **S**erbia
  http://www.aegis.rs/
- Founded in April 2005
- Mission:
  to provide Serbian research and development community with reliable and sustainable grid infrastructure
- Specific goals:
  - Coordinate efforts
    - to further develop academic and educational high performance computing facilities
    - to help integration into AEGIS infrastructure
  - Organize dissemination and training activities
  - Promote development and deployment of applications on AEGIS infrastructure
  - Coordinate fund raising efforts (NIP) to improve AEGIS infrastructure and human resources
  - Facilitate wider participation of AEGIS members in FP7, and other international GRID projects
  - Create a national GRID development policy
  - Lobby for its position within the national research development policy

- Coordinating institution:
  - Institute of Physics, Belgrade
- University computer centres: (4)
  - AMRES – RCUB - University of Belgrade, NREN of Serbia
  - ARMUK - University of Kragujevac
  - ARMUNS - University of Novi Sad
  - JUNIS - University of Nis
- Public research organisations: (18)
  - Institute of Nuclear Sciences, Belgrade
  - Institute for Chemistry, Technology and Metallurgy, Belgrade
  - Institute of Mathematics, Serbian Academy of Sciences and Arts
  - Astronomical Observatory, Belgrade
  - Faculty of Electrical Engineering, University of Belgrade
  - Faculty of Mechanical Engineering, University of Belgrade
  - Faculty of Organizational Sciences, University of Belgrade

  - Faculty of Mathematics, University of Belgrade
  - Faculty of Physics, University of Belgrade
  - Faculty of Technology and Metallurgy, University of Belgrade
  - Faculty of Sciences, University of Novi Sad
  - Faculty of Technical Sciences, University of Novi Sad
  - Faculty of Agriculture, University of Novi Sad
  - Faculty of Sciences, University of Nis
  - Faculty of Electronic Engineering, University of Nis
  - Faculty of Mechanical Engineering, University of Nis
  - Faculty of Natural Sciences, University of Kragujevac
- International collaborations   (2)
  - ATLAS group in Serbia
  - CMS group in Serbia
- SMEs: (2)
  - Irvas International d. o. o., Nis
  - South European Weather Agency, Belgrade

AEGIS Infrastructure

- Production:
  - AEGIS01-IPB-SCL
    (704 CPUs, 26 TB)
  - AEGIS02-RCUB
    (48 CPUs, 113 GB)
  - AEGIS03-ELEF-LEDA
    (64 CPUs, 1.5 TB)
  - AEGIS04-KG
    (48 CPUs, 480 GB)
  - AEGIS07-IPB-ATLAS
    (128 CPUs)
  - AEGIS11-MISANU
    (64 CPUs)
- Certification:
  - AEGIS05-ETFBG
  - AEGIS09-FTN-KM
- Demo/training:
  - AEGIS08-IPB-DEMO
- New:
  - UOB Faculty of Physics

AEGIS Infrastructure
Sites: 11
CPUs: 1094
TBs: 30
Production
Preproduction
Training

AEGIS01-IPB-SCL
704 CPUs / 24 TBs

AEGIS02-RCUB
48 CPUs / 1 TB

AEGIS11-MISANU
64 CPUs

AEGIS05-ETFBG
30 CPUs / 1 TB

AEGIS10-VINCA

AEGIS06-AOB

AEGIS08-IPB-DEMO
8 CPUs / 1 TB

AEGIS07-IPB-ATLAS
128 CPUs

AEGIS04-KG
48 CPUs / 1 TB

AEGIS03-ELEF-LEDA
64 CPUs / 2 TB

AEGIS09-FTN-KM

AEGIS and EGI

- EGI.eu created in February 2010
- Established as an international consortium based in Amsterdam
- Serbia represented in the EGI Council and other bodies by IPB
- Coordinates EGI-InSPIRE project, May 2010 – April 2014
- IPB represents Serbia as a partner

- FP7 RI-261323 project, ESFRI
(European Strategy Forum on Research Infrastructures)
http://www.egi.eu/

    - WP1 (NA1): Management
    - WP2 (NA2): External relation
    - WP3 (NA3): User community coordination
    - WP4 (SA1): Operations
    - WP5 (SA2): Provisioning the software infrastructure
    - WP6 (SA3): Services for HUC
                        (Heavy User Communities)
    - WP7 (JRA1): Operational tools

        * IPB involved

EGI and EMI

- EMI is a collaboration of four major European middleware providers:
  - ARC
  - dCache
  - gLite
  - UNICORE
- EMI aims to deliver a consolidated set of middleware components for deployment in EGI, as part of the Unified Middleware Distribution (UMD)

http://www.eu-emi.eu/

- Comes from a number of current and past Grid projects, like DataGrid, DataTag, Globus, EGEE, EMI, and WLCG.
- Security
  - Based on the Grid Security Infrastructure (GSI)
    - Public key encryption
    - X.509 certificates
    - Secure Sockets Layer (SSL) communication protocol
  - Enables secure authentication and communication over an open network
  - EUGridPMA (European Policy Management Authority for Grid Authentication)
  - Certification Authority (CA)
  - Virtual Organization Membership Service (VOMS)
  - MyProxy server (PX)

- ## User Interface
  - ### CLI tools to perform some basic Grid operations:
    - list all the resources suitable to execute a given job
    - submit jobs for execution
    - cancel jobs
    - query the status of jobs and retrieve their output
    - copy, replicate and delete files from the Grid
    - submit and manage file transfer jobs
    - retrieve the status of different resources from the Information System

- ## Computing Element
  - Grid Gate (GG) acts as a generic interface to the cluster
  - Local Resource Management System (batch system)
  - LCG and CREAM

- ## Storage Element
  - Control simple disk servers, large disk arrays or tape-based Mass Storage Systems
  - GSIFTP (a GSI-secure FTP), RFIO
  - Disk Pool Manager (DPM), CASTOR, dCache

- Information Service
  - Lightweight Directory Access Protocol (LDAP)
  - resource-level BDII, site-level BDII, top-level BDII
- Workload Management System (WMS)
  - Job Description Language (JDL)
- Logging and Bookkeeping service (LB)