Lecture 4: Computing systems of type «Grid»
### https://sites.google.com/site/clustergateorg/

- **Grid – before 2000 was used just for coomon electrical grid**
- **Basis:**
  - The load of computing facilities has stochastic nature (not even load during day, month, year, etc).
- **The idea is to create set of technical and administrative tools which might help users to get access to computing resources with current low load.**

Shevel.Andrey@gmail.com

CERN

- **Distributed Computing**
    - // synchronous processing
- **High-Throughput Computing**
    - // asynchronous processing
- **On-Demand Computing**
    - // dynamic resources
- **Data-Intensive Computing**
    - // databases
- **Collaborative Computing**
    - // scientists
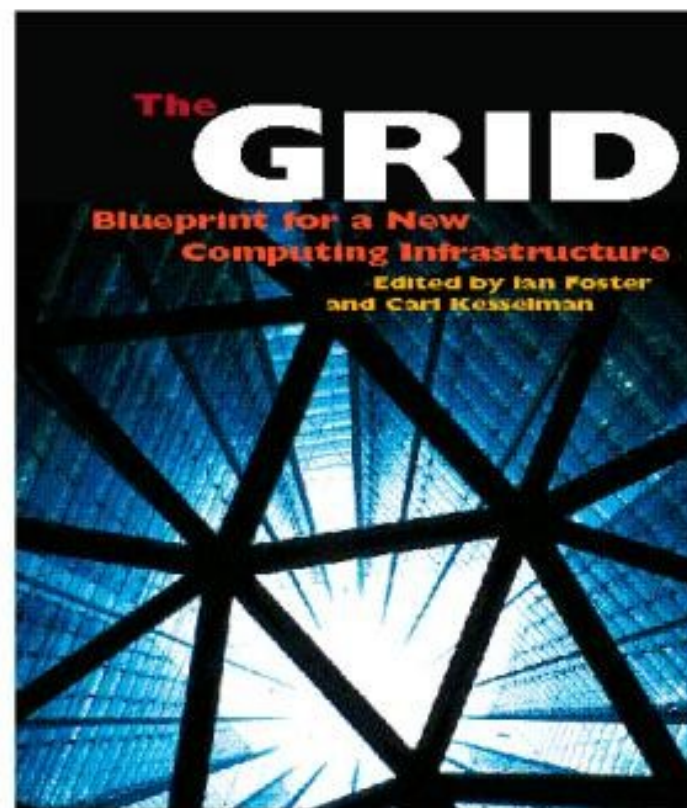
Ian Foster and Carl Kesselman, editors, "The Grid: Blueprint for a New Computing Infrastructure," Morgan Kaufmann, 1999, http://www.mkp.com/grids

Shevel.Andrey@gmail.com

Grid description

- http://en.wikipedia.org/wiki/Grid_computing
  - Grid – the set of computing resources from different organizations in general from different cities/countries/continents. All the resources are aimed to common goals.
  - Grid system can be considered as distributed computing network which is loaded with batch jobs.
  - In contrast to clusters Grid nodes are independent each other, more heterogeneous and distributed geographically.

Shevel.Andrey@gmail.com

Grid classification

- **Grid system for heavy computing (CPU intensive) but not required large volume of the data (examples):**
  - genomeathome.stanford.edu/
  - fightaidsathome.scripps.edu
  - http://milkyway.cs.rpi.edu/milkyway/
- **Grid system for CPU intensive tasks and for large volume of the data (so called Data Grid):**
 - wlcg.web.cern.ch/

Shevel.Andrey@gmail.com

Компоненты Грид систем

- **Grid Security Infrastructure (GSI)**;
- **Grid Resource Allocation Manager (GRAM**;
- **GridFTP – data transfer**;
- **Grid Resource Information Service (GRIS)** – information service about Grid resources
- **File Catalog (FC)**.
- *Usually the set of Grid software components is known as middleware.*

Globus Toolkit

http://www.globus.org/toolkit/about.html
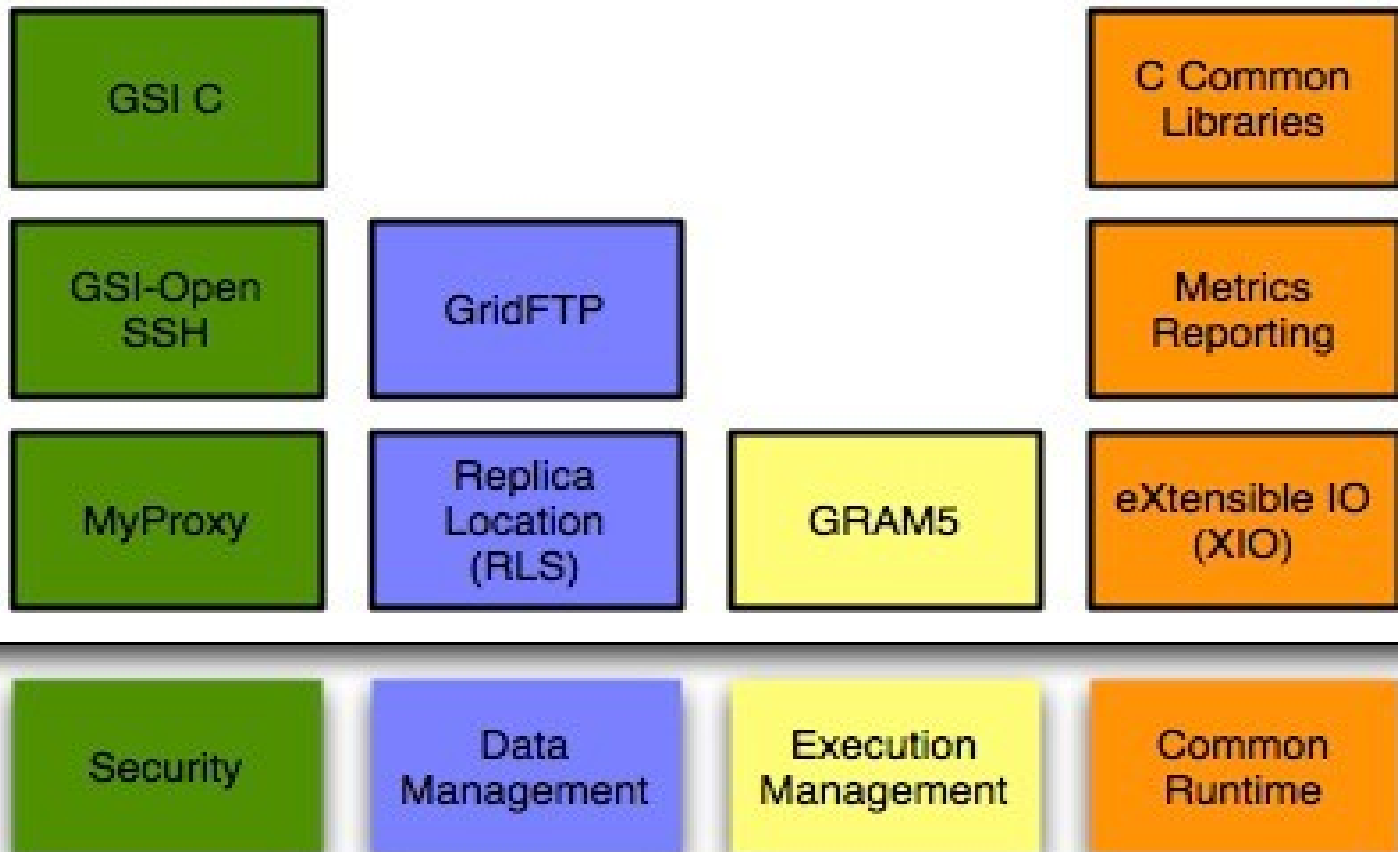
Some explanations

- *Protocol* – is precise set of rules how to transfer the data from one point to another point.

- *Service* – special ability, which permit to perform specific actions, for example, to transfer the data of determined type, to start processes on the computer, to give the access to some resources. In other words in addition to data transfer the actions not connected directly to the data transfer are assumed (protocol + behaviour).

Definitions

- Grid-Mapfile - /etc/grid-security/grid-mapfile
- DN – distinguished name
- CN – Common Name
- CA - Certificate Authority
- VO – Virtual Organization
- VOMS – Virtual Organization Management System
- PKI – Public Key Infrastructure

Shevel.Andrey@gmail.com

User certificate

- /DC=org/DC=doegrids/OU=People/CN=Andrey Shevel 438916 = DN

- User certificate – sertificate PKI in according to recommendations X509, developed International Telegraph Union in 1980-es

- Main PKI feature:
  - Sertificate consists of two parts: public key (PK) and secret key (SK), conversion PK~SK==1
  - usercert.pem & userkey.pem

Shevel.Andrey@gmail.com

Public-key infrastructure

Electronic business is picking up, and with it the need for secure electronic credentials is increasing. PKI is a way to prove identity in the online world. It also certifies that documents have not been tampered with.

❶ A document, such as a check, is digitally signed using hashing technology, the sender's private encryption key and the receiver's public key.
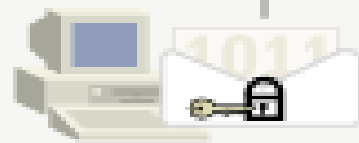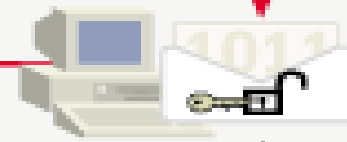
❷ The scrambled and encrypted document is sent.

❹ Using rehashing technology, the data from the received document is compared with that of the original document. This way, the document's authenticity can be assured.

❸ The document is decrypted using the receiver's private key and the sender's public key.

User certificate X509

- Certificate
  - Version
  - Serial Number
  - Algorithm ID
  - Issuer
  - Validity
    - Not Before
    - Not After
  - Subject
  - Subject Public Key Info
    - Public Key Algorithm
  - Subject Public Key
  - Issuer Unique Identifier (optional)
  - Subject Unique Identifier (optional)
  - Extensions (optional)
    - ...
  - Certificate Signature Algorithm
  - Certificate Signature

Shevel.Andrey@gmail.com

User access to Grid system

- To request the user cetificate from Certificate Authority (CA).

- To put the obtained certificate in the directory ~/.globus

- To add your certificate to the list of certificate of Virtual Organizations (VO), if it is required.

- To create temporary proxy (for 10 or 24 or 48 hours) to hold your certificate (proxy will answer with your certificate when outside Grid services ask you on your the certificate). During proxy creation you will be asked for pass phrase (relatively long phrase: 64 bytes or so).

Shevel.Andrey@gmail.com

Example globus-url-copy

- globus-url-copy -tcp-bs 8M -bs 8M -p 8 -vb file://`pwd`/bigfile.tar gsiftp://gridftp.ucar.edu/ptmp/user/
  - The copy of the file bigfile.tar on remote computer
  - Xx-bs и bs – block size
  - P – number of strems
  - Vb – подробная диагностика

Shevel.Andrey@gmail.com

How to organize the Grid system

- It is assumed that several organizations have significant computing clusters and have relatively close tasks (or same tasks).

- Is it assumed that authorities are convinced that it is very good if part of computing power will be used for common tasks.

- It has to be prepared and signed the document "Memorandum of understanding".

- After that you can start to install and tune the required Grid software to create Grid system.

Shevel.Andrey@gmail.com

Running Grid systems

- **Globus – http://www.globus.org**

- Grid: top ten questions -
  http://www.globus.org/alliance/publications/papers/topten.final.pdf

- Anatomy of the Grid -
  http://www.globus.org/alliance/publications/papers/anatomy.pdf

Shevel.Andrey@gmail.com

Additional info on Grid systems

- **Open Grid Forum - http://www.gridforum.org/**
- **The information page - https://sites.google.com/site/clustergateorg/grid**
- **European Grid Infrastructure - http://www.egi.eu/**
- **International Science Grid This Week http://www.isgtw.org/**
- **Open Science Grid - http://www.opensciencegrid.org/**
- **Wikipedia - http://en.wikipedia.org/wiki/Grid_computing**

End of Lecture

Shevel.Andrey@gmail.com