Grid Site: How-To

Andrey Kiryanov, NRC KI – PNPI

SPD Collaboration Meeting, 23-27 October 2023

Introduction

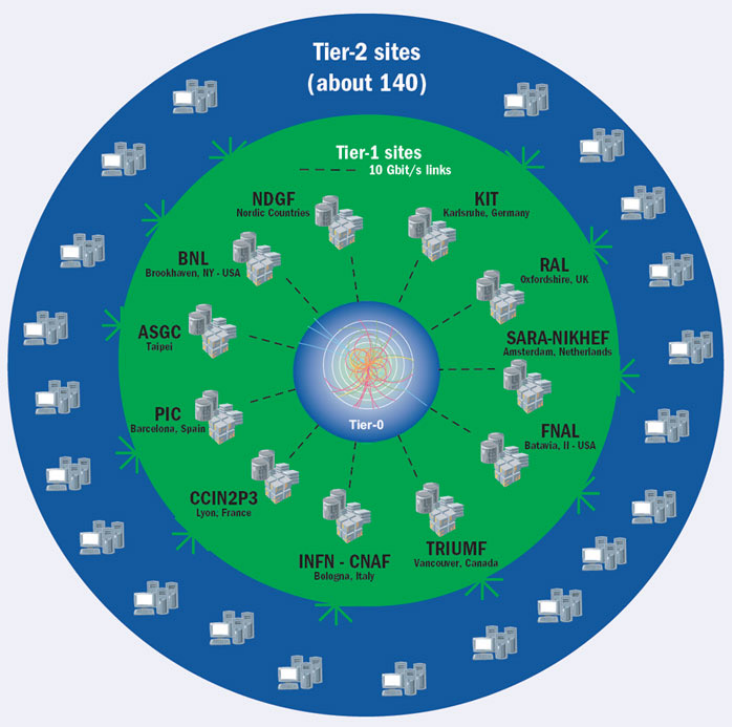
- What is “the Grid”?

  The Grid: Blueprint for a New Computing Infrastructure

  by Ian Foster & Carl Kesselman published in 1998

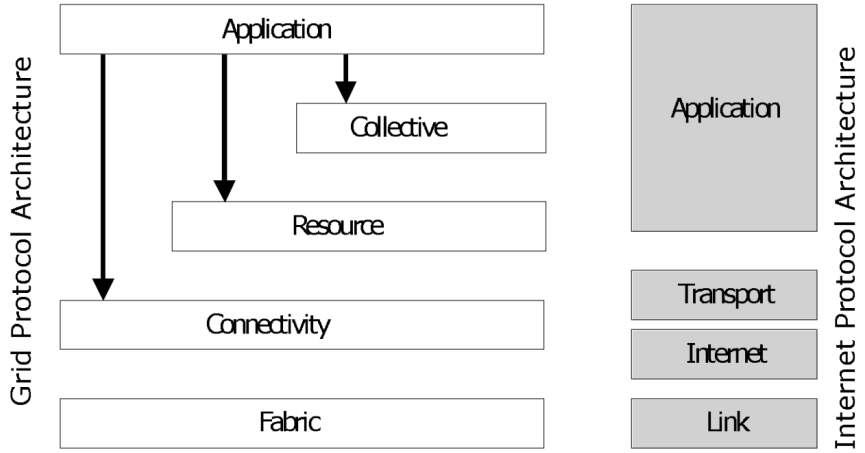
  — A highly theoretical book with ideas on new ways of interconnecting information resources
  — Was an inspiration for CERN and LCG (LHC Computing Grid)
  — At the time LHC was expected to produce an unprecedented amount of data (10 PB per year) which needed to be stored **and processed**

Proposed architecture



The MONARC Project
is an acronym for
"Models of Networked Analysis at Regional Centres for LHC Experiments"

The Grid architecture



Short history of LCG

- First software release in 2003
  - Funding projects worked side-by-side: EDG, EGEE, EMI
- Renamed to WLCG (Worldwide LHC Computing Grid) around 2006
- Software was initially heavily based on the Globus toolkit
  - Discontinued in 2018
- Real operation started in 2010 with LHC start and generally was a success, but lessons were learned
- Over time, what was initially a unified software distribution built by CERN turned into a couple of independent components, of which only disk storage (EOS) is actively developed at CERN
  - Everything else is either abandoned or maintained by other institutes
  - Job management moved from central infrastructure (Resource Brokers) to experiment's responsibility
  - Data movement service (FTS) is still available centrally at CERN

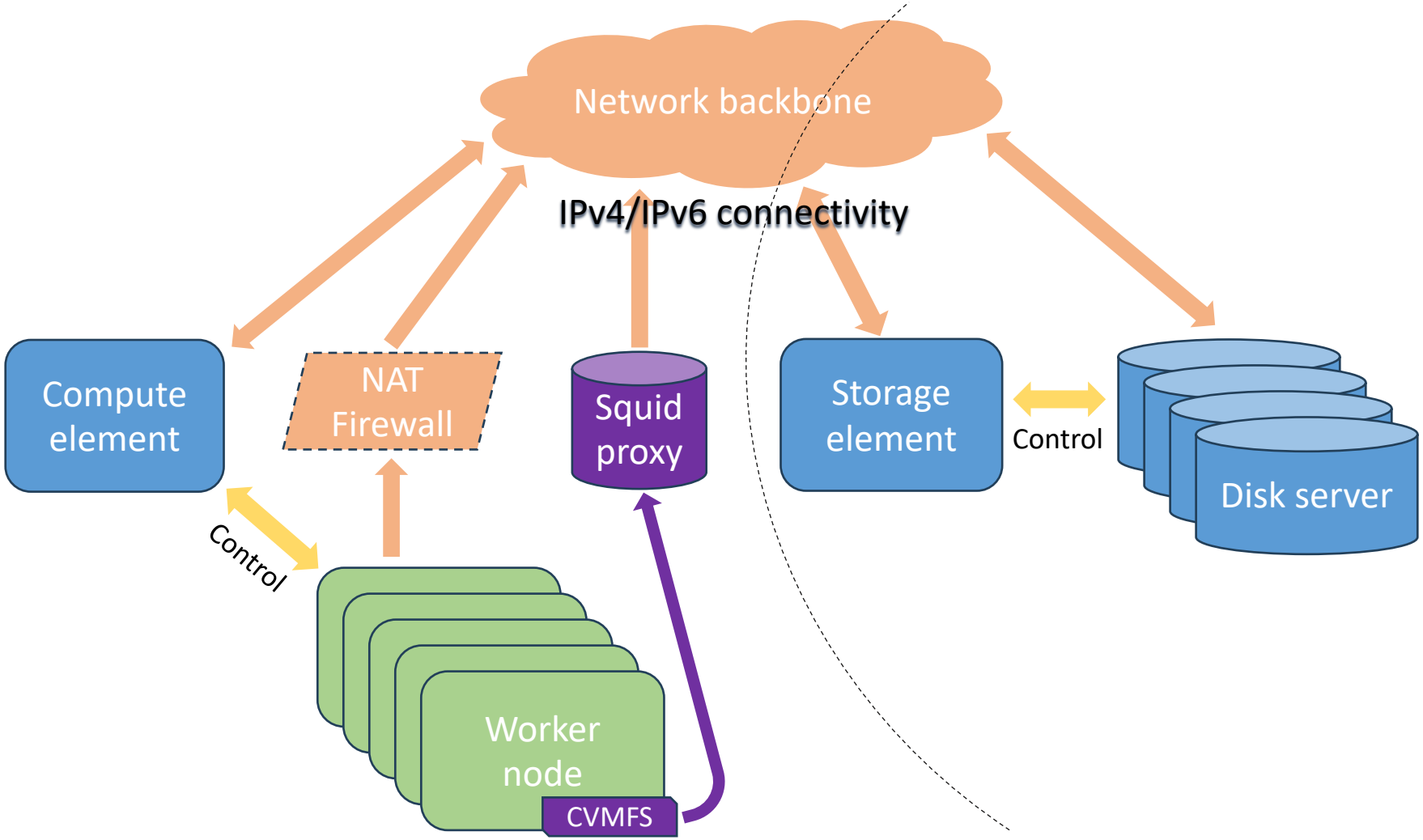
Main central building blocks

- GOCDB as a central (static) registry of resources
- LDAP-based information system with Glue1 and Glue2 schemas for dynamic resource information (mostly only compute)
- APEL-based accounting
- Centralized testing based on SAM and HammerCloud
- X.509 certificates and a chain of trust managed by EUGridPMA (with national CAs)
- VOMS and attribute certificates for VO management

VO-specific central building blocks

- *(yet another)* Information system
  - CRIC
- Workflow management system
  - PanDA
- Data management system
  - Rucio, FTS
- VO-specific tests
  - HammerCloud

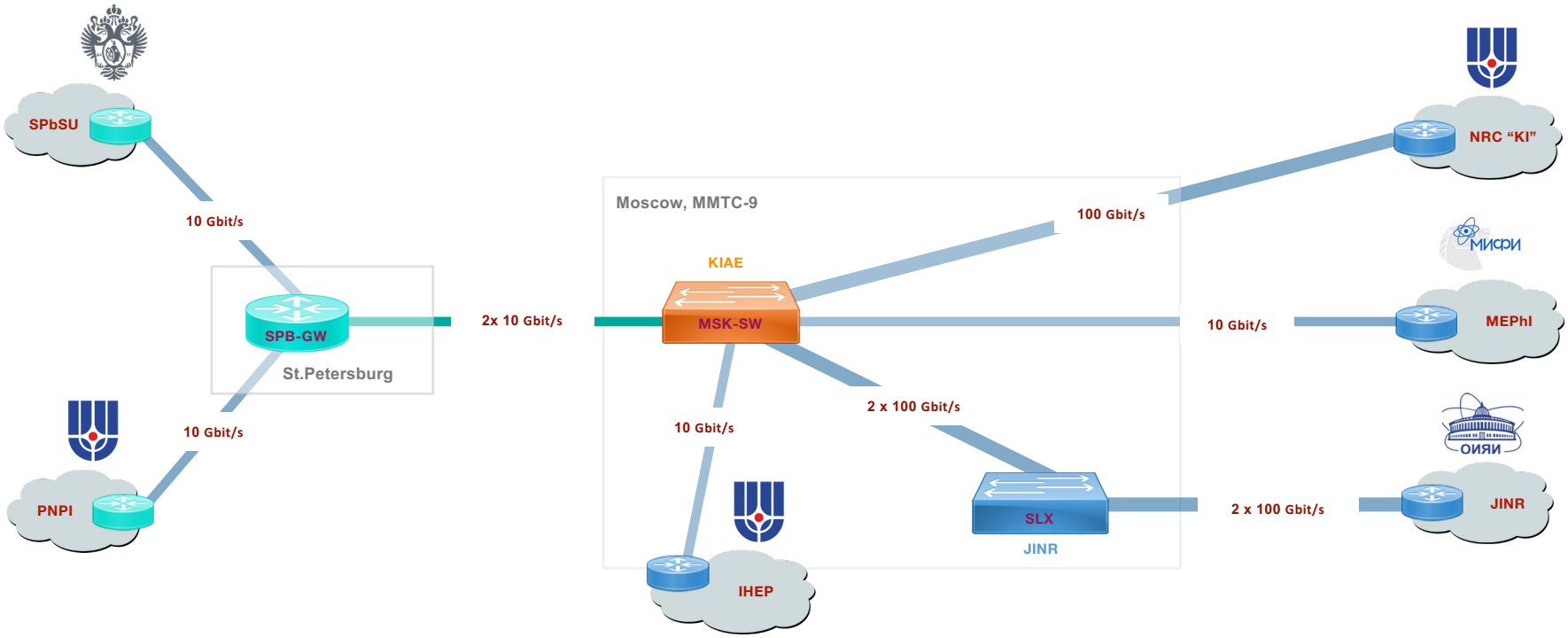
LCG Site Structure



LCG Site Components

- Network
  - 10 Gbps network connection is a must nowadays
  - WLCG requires both IPv4 and IPv6
- Compute Element (CE)
  - LCG CE and CREAM discontinued
  - ARC CE (Nordugrid) and HTCondor-CE available
  - Worker nodes need outgoing connectivity
  - CVMFS needs Squid proxy
- Storage Element (SE)
  - DPM / DOME discontinued
  - EOS and dCache available
  - Disk nodes need incoming connectivity

Russian scientific backbone



LCG Site Installation Steps (1)

- Talk to your ISP, request an IPv6 subnet is you don't have any yet
- Figure out how to connect to the Russian scientific backbone (RDIG-M), which is jointly managed by NRC KI and JINR
- Choose you domain and host names
  - It will be hard to change them later
  - Worker nodes can stay in the intranet, just do not route all of them through a single 1 Gbps bottleneck
- Request host certificates from a national CA for CE, SE and disk servers
  - In Russia this is RDIG CA operated by NRC KI
  - Nothing can stop a large experiment from deploying and managing its own CA
- Decide on a type of CE and deploy a batch system for the worker nodes
  - Slurm is a strong choice nowadays
- Deploy a CVMFS on the worker nodes

LCG Site Installation Steps (2)

- Decide on a storage system and whether you will manage tapes
- Deploy disk nodes
- Configure your experiments' VOs
  - Ask your experiment for a VOMS configuration package
  - Configure per-VO batch queues (partitions)
- Test everything thoroughly
- Register your endpoints in the experiments' resource registries
- …
- Profit!

LCG Site Support

- Small and medium-sized sites can be easily managed by one or two people
  - Depends on the level of automation for the time-consuming tasks like node re-installation in the case of a disk failure
  - Make use of cluster management and monitoring tools
    - Ansible, Puppet, Zabbix, etc.
- Pledges
  - SLA between a VO and a site
  - Covers compute resources (normalized cores), storage resources and **reliability**
  - Declare a downtime before changing things!

LCG Site Reliability

- 97% reliability means ~11 days of downtime per year

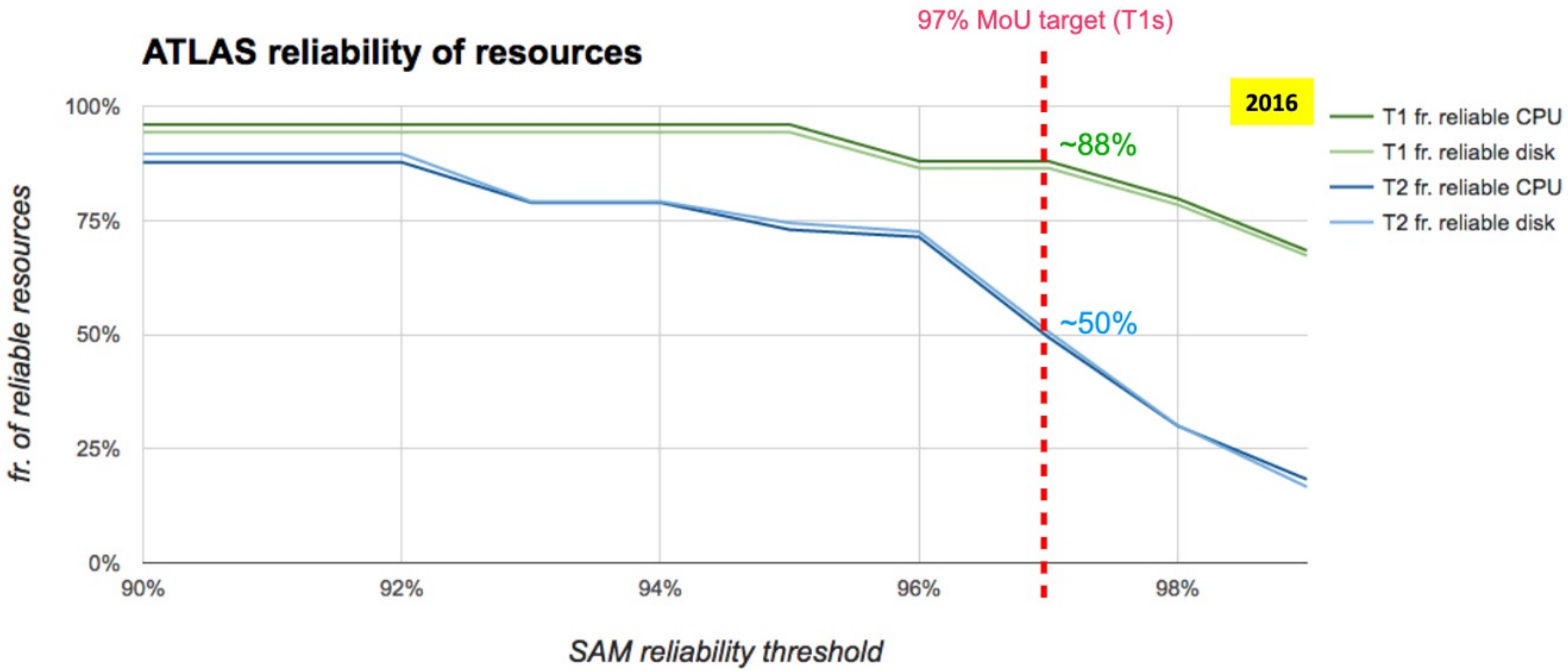


Chart by J. Flix – PIC/CIEMAT – jfix@pic.es
March 2017 GDB – ISGC2017 - Taipei

Documentation is the key

- Slurm
  - https://slurm.schedmd.com/documentation.html
- ARC CE
  - https://www.nordugrid.org/arc/arc6/index.html
- EOS
  - https://eos-docs.web.cern.ch/diopside/index.html
- dCache
  - https://www.dcache.org/documentation/
- HTCondor-CE
  - https://htcondor.org/htcondor-ce/documentation/

Thank you!