UNICORE in XSEDE

Bernd Schuller, Bastian Demuth
Institute for Advanced Simulation (IAS)
Jülich Supercomputing Centre (JSC)
Forschungszentrum Jülich
Forschungszentrum Jülich

Founded in 1956 as nuclear research facility
“Public non profit” Company: 90% Federal Republic of Germany
10% Federal State of North Rhine-Westphalia

Research Areas

Health

Energy & Environment

Information

Key Competency (scientific computing, physics (Nobel prize for Peter Grünberg in 2007))

Annual Budget: ~ 450 million €
Staff: ~ 4700 (~ 1600 scientists, ~ 400 Ph.D. students)
more than 800 visiting scientists per year from 50 countries
Juelich Supercomputing Centre (JSC)
Jülich Supercomputing Centre (JSC)

Staff: ~ 170 including
  15 Ph.D. students, 8 full professors

Areas of research, development, and service
  National High Performance Computing Center (operation of supercomputers, user support)
  Computational Science (computer simulation, numerical and stochastic methods, cluster computing)
  Central IT-infrastructure for FZJ (network, backup)
  Education & Training (mathematical-technical software developers, students, young scientists, and users)
  Federated Systems & Data (UNICORE development, operation & support, application enabling, research)
Compute resources at JSC - JUQUEEN

Blue Gene/Q, 131072 processors (PowerPC A2, 1.6 GHz), 1.6 PetaFlop/s, 131 TeraByte Memory, currently rank 8 in top500 list
A brief history ...

- **UNiform Interface to COmputing Resources**
- Initial development started in 1997 (German projects)
- Access the German High Performance Computing centres
- Primary goals: seamless, secure, and intuitive
- Since 2002: development in national and EU projects
- UNICORE 6.0 released in 2007
  - Completely based on HTTP and Web Services
  - Improved incrementally ever since
Selected Projects

UNICORE

UNICORE Plus

EUROGRID

GRIDSTART

GRIP

OpenMolGRID

UniGrids

VIOLA

DEISA

OII-Europe

EGEE-II

PRACE

SMARTLM

Fit4Green

SLA4DGRID

EMI

UNICORE

Swing GRID

BEgrid

BiG

NIC

AERO GRID

WisNet GRID

UIMA-HPC

2001

2002

2003

2004

2005

2006

2007

2008

2009

2010

2011

2012

2013

2014

17th July, 2012
Some heads might already be in the clouds...

• ... ours are still “in the Grids”

• Enough unresolved problems to keep us busy
• XSEDE/PRACE model: give computational power to the people with the best ideas, basically for free
• High-end HPC cannot be easily outsourced to cloud providers
• HPC requires devoted support, know-how and active research
What can we do for you?

- Unified submission of compute jobs
- Independent of batch system
- System peculiarities are hidden
- Software-As-A-Service, pre-configured applications
- Graphical user interface

- Data movement and management
  - Client → Server, Server → Client and Server → Server
  - Scheduled transfers, progress monitoring
  - Support for different transfer protocols
  - Metadata & Queries

- Workflows and automation
  - If-Statements, loops, sweeps
  - Graphical editing and monitoring
Basic Architecture

Client Layer

Web Service Layer

Execution Management Layer

Resource Layer

Graphical Client

Command Line Client

Gateway

UNICORE Services

OGSA-*

UNICORE/X Service Container

XNJS - Backend

Target System Interface

Local RMS (e.g. Torque, LL, LSF, etc.)

AuthZ Attributes

AuthZ Policies

IDB (Resource description)
Clients
UNICORE commandline client (UCC)

- Provides access to all UNICORE features
- “Batch” function for high-throughput job processing
- Scriptable, extensible
- Linux: bash autocomplete, Emacs integration
UNICORE Rich Client (URC)

- Building, submitting and monitoring jobs and workflows
- Integrated data and storage management
- Pluggable credentials
- Help system, getting started pages
- Extensible
- Based on the Eclipse framework
Example: MOPAC - 5

Application-specific plugins
Job submission

- Support for common batch systems
  - Torque, Slurm, LoadLeveler, LSF, ...

- Extensible job resource requirements
  - Cater for specialities of HPC systems

- Execution environments
  - Simplify creation of parallel jobs for end-users
  - Hide “details” of site-specific parallel environments
Data Movement

1. Import/Export file
   - Gateway – Site 1
   - SMS
   - FTS
   - UNICORE/X Site2

2. Move data
   - Target System Interface – Site 1
     - Local RMS (e.g. Torque, LL, LSF, etc.)

1. Send/Rcv file (@ Date/Time)
   - Gateway – Site 2
   - SMS
   - FTS
   - UNICORE/X Site1

2. Move data
   - Target System Interface – Site 2
     - Local RMS (e.g. Torque, LL, LSF, etc.)
File transfer options

- Extensible set of transfer protocols
  - *Built-in*: BFT, OGSA-ByteIO
  - *Optional*: High-performance UFTP solution

- Data staging from/to job directory has additionally
  - *GridFTP*
  - *http(s), file://, ftp, scp, mailto*
UNICORE/X Metadata system

- Client
- Metadata Service
- Storage Service
- Gateway
- Local RMS (e.g. Torque, LL, LSF, etc.)

UNICORE/X Service Container

- Metadata Manager
- Indexer
- Metadata Extractor
  - Apache Lucene
  - Apache Tika

Target System Interface
Metadata client tools

- Client tools currently only available in UCC
- Create, read, update, delete metadata
- Trigger server-side metadata extraction and indexing
- Search (fulltext, fuzzy, wildcards)
- Example: list file properties including metadata
Workflow services

Client

Gateway

Service Registry

Workflow Engine

Service Orchestrator

Gateway – Site 1

UNICORE Services

XNJS

UNICORE/X Site 1

Target System Interface – Site 1

Local RMS

Gateway – Site 2

UNICORE Services

XNJS

UNICORE/X Site 2

Target System Interface – Site 2

Local RMS
Workflow features

- Simple graphs (DAGs)
- Workflow variables
- Loops and control constructs
  - while, repeat, for-each
  - if-else
- Conditions
  - Exit code, file existence, file size, workflow variables
UNICORE in the XSEDE Architecture

**Extended Architecture**

Access Layer

- Applications, Portals and Gateways
- Transparent access via the file system
- APIs and CLIs

Services & Web Services Infrastructure

- XSEDE Enterprise Services
  - JSDL/BES
  - HPC-BP
  - GridFTP
  - WSI-BSP
  - RNS/ByteIO

- Community Provided Services
  - GRAMS
  - REST/RMI
  - Amazon EC2

- Application Development

Resources

- Core Enterprise Resources, e.g., RP resources
- Other Resources, e.g., Campus centers, Amazon, Research Group Data

17th July, 2012
Standards ...
... are harder than you'd think
Some standards in the service layer

- **JSDL**
  - Job Submission Description Language
  - Application or Executable
  - File staging
  - Extensible (param. sweep, parallel apps, ...)

- **OGSA-BES**
  - Job submission
  - Job management
  - Bulk operations

- **RNS**
  - Tree of resources

- **ByteIO**
  - File access

- **GridFTP**
  - High perf. data movement

- **SRM**
  - Storage resource management

- **GLUE 2**
  - Information model
  - Describes Grid entities (Services, Resources, ...)
  - Adopted in XSEDE and EMI
How users will benefit from standards

- Interoperation of Genesis II and UNICORE
  - *Genesis II client*
    - Login via username/password, certificates hidden from user
    - Submit jobs to UNICORE or Genesis via BES
  - *GFFS support:*
    - User home
      - Submission and monitoring with FUSE driver
  - *Campus Bridging use case*

- Use existing data from GridFTP (and/or SRM) servers
- Common way of describing resources in information systems
Achievements during Year 1

- Interoperation with Genesis II
  - Genesis II client can run jobs through UNICORE
  - Support for trust delegation
  - Staging from and to the GFFS
- Support for Globus Grid-map-files
- XSEDE-specific installation package
- Simplified core service installation procedure
- Improved documentation
- Systematic and comprehensive testing
UNICORE Beta Deployment in XSEDE (1)
UNICORE Beta Deployment in XSEDE (2)

• Services can be accessed via unmodified UNICORE clients
• Interoperation with PRACE => global HPC Grid

• Comprehensive testing in XSEDE generates functional requirements, bugs, improvement ideas
• Workflow system under evaluation now
• URC under evaluation soon
Current development and Outlook

Integration into XSEDE

- Improved GridFTP support
- Integration with XSEDE authN (e.g. MyProxy logon)
- SRM support?

Standards

- RNS implementation in UNICORE
- Work on JSDL and BES evolution (with our Genesis II colleagues)

Enhanced capabilities

- JSDL parameter sweep support
- Re-activate Condor TSI
- BES support in the workflow system