The Work Package 2 experience

Data Management on the Grid

Peter.Kunszt@cern.ch

DataGrid is a project funded by the European Commission
under contract IST-2000-25182

Outline

- Objectives and how they were met
- Achievements
- Lessons learned
- Future & Exploitation
- Questions

Objectives

DataGrid Technical Annex:

- ✓ Enable secure access to massive amounts of data in a universal global name space.
- ✓ Move and replicate data at high speed from one geographical site to another.
- ✓ Interface to heterogeneous mass storage management systems.
- × Manage synchronisation of ren Read-only write once versions.
- × Automate data caching and distribution according to dynamic usage patterns.

Network monitoring considerations.

Achievements

Delivering Middleware

- Many existing Grid components were included in the first release (like GDMP, Globus replica catalog)
- Based on first experience and on user feedback, the EDG-TB3 services have been designed and developed
- Pioneering role in the usage of J2EE-based web services (long time before OGSI)
- EDG-TB3 is a complete set of data management solution – but should still be considered first generation.
- **It's in use!**

Achievements

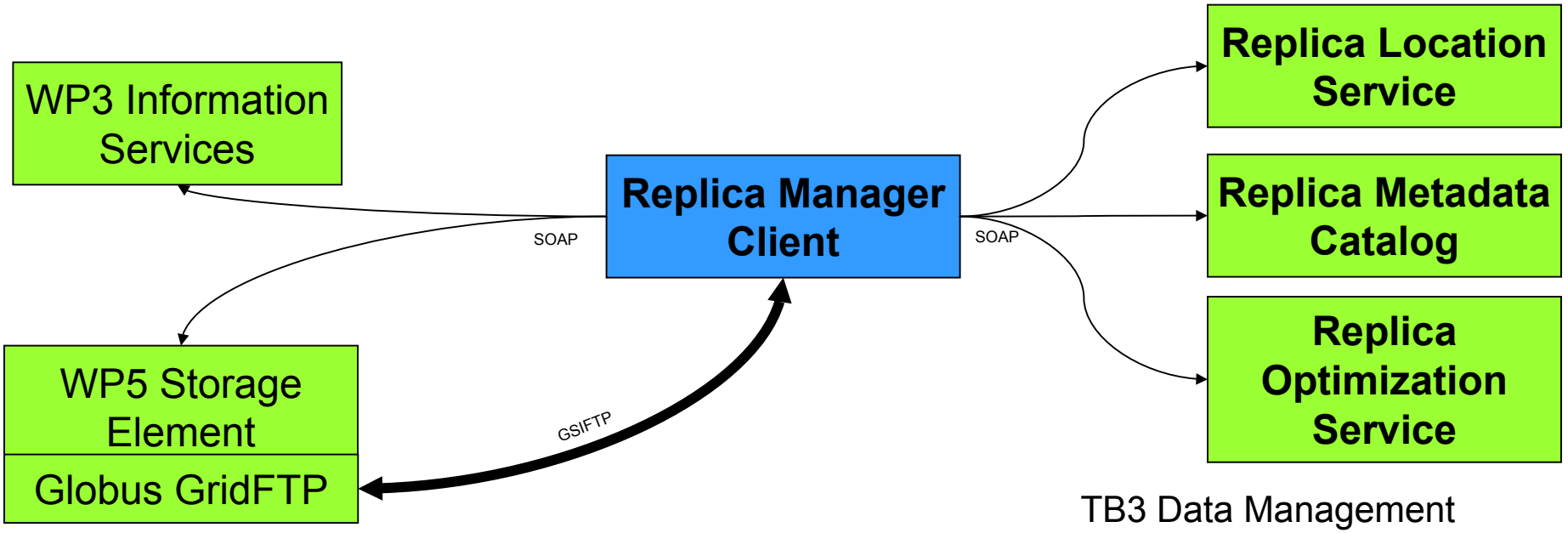
Collaboration

- Inside EDG: participation in all project groups and close collaboration with applications
  - Close coordination with WP5 in last year
  - Direct support of applications
- Collaboration with CrossGrid on storage resource metrics
- Collaboration with the Globus Project on Replica Location Service
- Strong participation in GGF and the various groups therein
- Participation in Storage Resource Management specification
- Interaction with many more groups, like DAIS project.

Achievements per Task

Data Replication Task

- Replica Manager (TB1+2+3)
- Replica Catalog (TB1+2) and Replica Location Service (TB3)
- Grid Data Mirroring Package (TB1+2)
- Replica Metadata Catalog (TB3)



TB3 Data Management

Achievements TB3

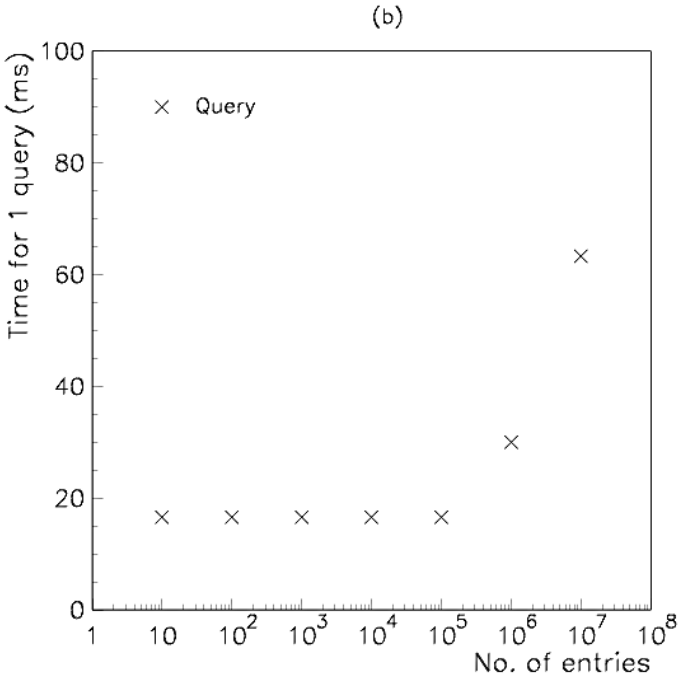
Important scalability improvements.

- **TB1+2 Replica Catalog** had serious limitations:
  - ~2000 entries maximum depending on file name length
  - not much concurrent usage possible.
- **TB3 Replica Location Service**
  - Limited only by underlying database (10^7 still ok for MySQL)
  - Good scalability with growing size
  - Very good scalability with number of concurrent users

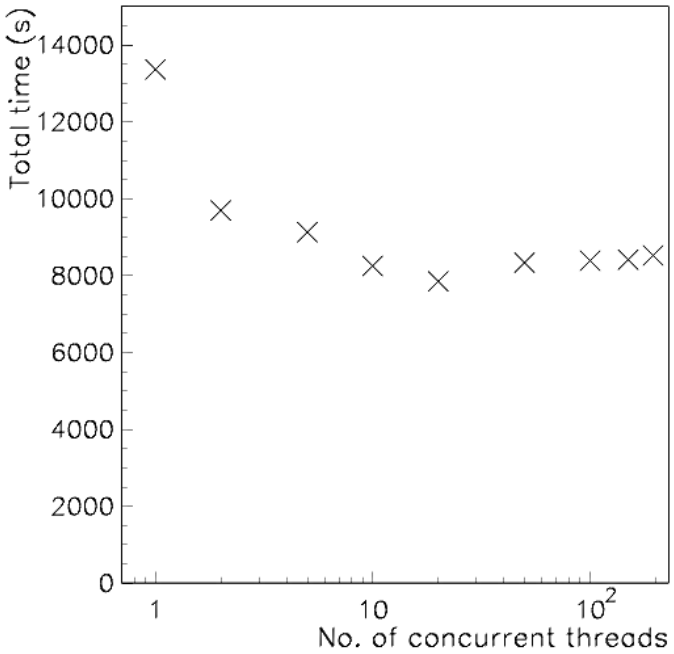
Important interoperability improvements

- Replica Manager is interoperable with various Storage interfaces: SE, SRM, GridFTP server (aka 'classic SE')
- Supporting both TB1+2 and TB3 information systems (MDS and R-GMA)

Scalability Figures



Time to run a lookup
(C++ API, MySQL)



Total time to add 500'000 entries
with concurrent threads,
one operation each (Java API)

See D2.6 for more details

Achievements per Task

Optimization Task

- Replica Optimization Service (TB3)
- Active research: OptorSim Simulation package
  - Simulation of optimal replica placement strategies
  - Includes simulation of network conditions

Metadata Access Task

- Spitfire – Grid Access to Relational DB (Demo)
  - Served as a prototype web service application for TB3 services
- Collaboration with DAIS group.
- Replica Location Service metadata
- Replica Metadata Catalog metadata

Achievements per Task

Security Task

- Secure java-based web services:
  - TrustManager for authentication (TB3). In use by WP3 and WP5 as well
  - AuthorizationManager for authorization (In operation). It can apply VOMS authorization information to the service.
- Secure clients in java and C++ to web services (TB3)
- Strong participation in EDG Security Group

Lessons Learned

Development Cycle

- In EDG it was not possible to do a proper requirements gathering, prototyping, testing, development fast enough for the lifetime of the project. A faster release cycle to the end-users will be possible from now on since future projects won't start from zero.
- We focussed on core services in EDG. The much needed end-to-end capabilities can now be added more easily since the users also know better what they want and how they want it.
- User interface and documentation are important and difficult to get right first time

Less is more

- Focusing on the basics: stability and usability paid off
- Extra features good, but should be pluggable because not all users want them

Lessons Learned

Security is key

- Can't 'add security later' – horizontal through all services
- Security mechanisms are deeply reflected in the design
- Lots of open issues: Performance, delegation, site buy-in…

Web Services work well

- Modular web service structure
- Pluggable QoS (deployable in open source or commercial environments)
- Based on standards: well supported by industry and open source community

Future & Exploitation

Products

- The LHC Computing Grid is running WP2 services (except Replica Optimization and Spitfire) and will maintain and support them for at least this year for their community.
- EGEE can benefit from existing WP2 services as a starting point
- Spitfire has served as an example for other projects already.
- The security infrastructure will serve as one of the bases for java-based web service infrastructures over SSL for the next projects.
- The optimization work has enriched the computer science community with many valuable insights through its many publications.
- The publications, documentation and tutorials serve as a reference for future projects.

Future & Exploitation

People

- All members of WP2 have gained valuable experience while working on EDG. Their expertise will be very useful to their future projects.

Processes

- The lessons learned in EDG will help improve the processes of the future projects that EDG members participate in.

A lot of work remains to be done

- Data sets and virtual data
- Application metadata bindings into the low-level services
- End-to-end integration with user applications
- …

BIG THANKS

To all people who have contributed to WP2.

CERN: Diana Bosio, Akos Frohner, Leanne Guy, Wolfgang Hoschek, Javier Jaen-Martinez, Marcin Kania, Arnaud Lacroix, Erwin Laure, Levi Lucio, Ben Segal, Heinz Stockinger, Kurt Stockinger

INFN: Giuseppe Andronico, Federico Di Carlo, Andrea Domenici, Flavia Donno, Livio Salconi, Marco Serra

PPARC: William Bell, David Cameron, Gavin McCance, Paul Millar, Caitriana Nicholson

HIP/CSC: Joni Hahkala, Niklas Karlsson, Juho Karppinen, Ville Nenonen, Marko Niinimäki, Tuomas Nissi, Henri Mikkonen, Olli Serimaa, Mika Silander, John White

KDC: Olle Mulmo, Bjorn Torkelsson, Gian Luca Volpato

ITC-IRST: Paolo Busetta, Luigi Capozza, Mark Carman, Ruben Carvajal-Schiaffino, Luciano Serafini, Floriano Zini

LCG: Itzhak Ben-Akiva, James Casey, Radovan Chytracek, Kálmán Kövári, Sophie Lemaitre

PPDG: Andrew Hanushevsky, Shahzad Muzaffar, Asad Samar

And: Brian Tierney(BNL), Aleksandr Konstantinov(NorduGrid/CrossGrid)