Comparing GT2, GT3 and GT4

- Based on WSRF (*Web service Resource Framework*)
- Grid standards now compatible with WS
- Many pre-WS protocols and standards still supported
- June 2008 – version: 4.2.1

[Visit Globus website](http://www.globus.org)
Possible distributed application *without* using Globus Toolkit:
Possible distributed application with using Globus Toolkit:

(http://www.globus.org)
Based on web services

Considerable improvements in:

- performance,
- usability,
- documentation,
- following standards
- functionality
- robustness
• New installer for specifying the set of components for installation
  
  – possibility to omit components
  
  – avoid unnecessary configuration
  
  – automatic configuration and autoconf for more platforms

• Added hundreds of tests

• Updates for following standards

• New GridFTP server.
  
  – wuftpdp code replaced with high performance Globus XIO library
  
  – Supports distributed storage, HPSS (High Performance Storage System) support.

• Better support for Windows (XIO incl.)
- Possibility for user-provided scheduler (*Community Scheduler Framework*)

- Client and server program support also for C++ and Python (in addition to C and Java)

- New possibilities in security. e.g. message level security for SOAP messages based on *WSSecurity* and *SecureConversation* specifications)

- Support for multiple authorisation frameworks (like *grid-mapfile* (ACL)); user-provided authorisation manager SAML (*Security Assertion Markup Language*) protocol; MyProxy certificate repositories; CAS (*Community Authorization Service*)

- GT4 can be integrated with VOMS (*Virtual Organization Management Service*) and PERMIS.

- MDS4 (*Monitoring and Discovery Services*) gives more functionality (each GT4 web service container is preconfigured with *MDS-Registry* service, which
manages information, which services implemented in the given container (using XSLT templates).

- Big effort in performance:
  - GRAM subjobs written in C (winning 80% in performance (due to JVM))
  - GridFTP able to use 80% of raw iperf; Upto 1800 clients

In addition, quite a few tools:

- Nimrod-G (parameter learning system), Condor-G jobscheduler; Ninf-G (remote procedure call system); GRMS (Grid Resource Management System) (metascheduler); OGCE (Open Grid Computing Environment) – software for portals.

- GRAM interface to Condor; Platform’s Load Sharing Facility (LSF); Sun Grid Engine; Portable Batch System; etc.
7.2 WSRF

Web Service Resource Framework
OGSI (Open Grid Services Infrastructure), (1.0 released in July 2003)
OGSI was extension to WSDL and XML, which:

• introduced stateful Web services;

• defined

  – mechanisms for creation of web service instances, naming conventions, time-to-live concept;
  
  – service state data and mechanisms for corresponding queries;
  
  – asynchronous notification mechanisms for service state data changes;
  
  – service instance grouping and management
  
  – service runtime error general management
Shortcomes of OGSI:

• Too many things for one specification – called as “composed specification”

• Does not work well with existing web service tools. Using mainly XML Schema

• Too object-oriented. (The follower, WSRF makes clear difference between service and resource for a given service instance.)

• WSRF has broader support for programming languages
WSRF (Web Service Resource Framework)

- keeps actually all OGSI functionality
- changes syntax (e.g. WS-Addressing) using different terminology to represent it
- breaks OGSI functionality into 5 different sub-specifications + notification specifications

Current version: WSRF v1.2

03.04.06: WSRFv1.2 specification accepted as OASIS standard (OASIS – not-for-profit global consortium that drives the development, convergence and adoption of e-business standards OASIS (http://www.oasis-open.org)
WSRF sub-specifications

   - also defines the pattern by which resources are accessed through Web services
   - the means by which WS-Resources are referenced

2. **WS-ResourceProperties** – standardises the means by which the definition of the properties of a WS-Resource may be declared as part of a Web service interface
   - properties represents a projection of (or a view) on the WS-Resource’s state
   - projection is defined in terms of a resource properties document (which defines a basis for access to the resource properties through Web service interfaces)
   - defines a standard set of message exchanges that allow a requestor to query or update the property values of the WS-Resource
3. **WS-ResourceLifetime** – standardised messages to destroy resource (immediate or scheduled)
   - resource properties [WS-ResourceProperties] that may be used to: inspect and monitor the lifetime of a WS-Resource

4. **WS-ServiceGroup** – heterogeneous by-reference collection of Web services
   - wide variety of collections of services or WS-Resources, including registries of services and associated WS-Resources
   - Members of a ServiceGroup are represented using components called *entries*
   - A ServiceGroup entry is a WS-Resource.
   - The Web service associated with a ServiceGroup entry can be composed from a variety of Web services standards including **WS-ResourceLifetime** which defines standard patterns by which resources can be destroyed, **WS-BaseNotification** which defines how third parties may subscribe to be informed of changes to the ServiceGroup and **WS-ResourceProperties** which defines how the properties of a ServiceGroup and its entries are made accessible through a Web service interface.
5. **WS-BaseFaults** – Problem determination in a Web services setting is simplified by standardising a base set of information that may appear in fault messages.

– defines an XML Schema type for base faults, along with rules for how this base fault type is used and extended by Web services.

**+ WS-Notification-specifications** – standardised approach for notification services

– Web-services approach to notification

– topic-based publish/subscribe pattern

**OGSI → WSRF** – transform:

**Grid Service Reference → WS-Addressing Endpoint Reference** – Uses the endpoint reference properties of WS-Addressing to identify a stateful resource associated with the Web service at the designated endpoint.

**Grid Service Handle → WS-Addressing Endpoint Reference & WS-RenewableReference**

– WS-RenewableReferences introduces policy annotations to the WS-Addressing end-
point reference that allow ”handles” and “handle Resolvers” to be an integrated part of the endpoint reference. Use of the policy annotations provides for additional endpoint reference stability.

HandleResolver portType –> WS-RenewableReferences  – Integration of Handle Resolution service references in the endpoint reference.

GridService portType service data access –> WS-ResourceProperties – Multiple operations instead of one extensible operation, supporting simpler binding to existing programming models.

GridService portType lifetime management –> WS-ResourceLifetime – Removes the superfluous “terminate before” operation. Cosmetic changes to others.

Notification portTypes –> WS-Notification – Generalises notification to hierarchical topic-based pub/sub mechanism.

Factory portType – Now treated as a pattern; thus, no specific operation.

ServiceGroup portTypes –> WS-ServiceGroup – Only minor changes

Base fault type –> WS-BaseFault – Only minor changes.
7.3 Data Management) - RFT

- Fire-and-forget transfer
- Web services interface
- Many files & directories
- Integrated failure recovery
- Has transferred 900K files
GridFTP

- 100% Globus code
  - No licensing issues
  - Stable, extensible

- IPv6 Support

- XIO for different transports

- Striping → multi-Gb/sec wide area transport
  - 27 Gbit/s on 30 Gbit/s link

- Pluggable
Replica Location Service (RLS)

• Identify location of files via logical to physical name map

• Distributed indexing of names, fault tolerant update protocols

• GT4 version scalable & stable

• Managing ~40 million files across ~10 sites
OGSA-DAI (*Data Access and Integration*)

- Provide service-based access to structured data resources as part of Globus
- Specify a selection of interfaces tailored to various styles of data access?starting with relational and XML
- Supports data access, insert and update
  - Relational: MySQL, Oracle, DB2, SQL Server, Postgres
  - Supports data delivery
    * SOAP over HTTP
    * FTP; GridFTP
    * E-mail
* Inter-service

- Supports data transformation XSLT ZIP; GZIP
- Supports security
  
  * X.509 certificate based security
7.4 Information Services

Monitoring and Discovery

- Each service should be followed and discovered using common mechanisms;
  - These mechanisms are: WSRF/WSN

- A common aggregator framework for collecting information from services, thus:
  - **MDS-Index**: Xpath queries, with caching
  - **MDS-Trigger**: perform action on condition
  - **(MDS-Archiver)**: Xpath on historical data

- Deep integration with *Globus containers & services*: every GT4 service is discoverable
  - GRAM, RFT, GridFTP, CAS, ...
WS GRAM architecture
WS GRAM architecture

Delegated credential can be: Made available to the application
WS GRAM architecture

Delegated credential can be:
Used to authenticate with RFT
WS GRAM architecture

Delegated credential can be: Used to authenticate with GridFTP
7.5 Globus Toolkit 5

Version 5.0.0 – January 2010

Latest: 5.2.0 – 15 December 2011

• Some components of GT4 are not included in GT5 (e.g. GT4 Java Core, WS-GRAM4, RFT), to be replaced by new software under development (e.g. Crux, Globus.org Service)

• GRAM5
  – backward compatible with GRAM2
  – WS-GRAM (GT4) dropped
  – RSL language identical to GRAM2

• File transfer
  – GridFTP new functionality (eg. stall detection, recovery from server or network failures, synchronization capabilities...)

- RFT service discontinued, replaced with
- Globus Online
  - service provided by Globus
  - transfers between endpoints

- MDS

- about to come: IIS ([Integrated Information Services](http://www.ci.uchicago.edu/wiki/bin/view/IIS/WebHome)) (currently alpha)
7.6 Globus ecosystem

- Condor-G
- MPICH-G2
- GRMS
- Nimrod-G
- Ninf-G
- Open Grid Computing Env.
- Commodity Grid Toolkit
- GriPhyN Virtual Data System
- Virtual Data Toolkit
- GridXpert Synergy
- Platform Globus Toolkit
- Oracle (Sun) Grid Engine
- PBS scheduler
- LSF scheduler
- GridBus
- TeraGrid CTSS
- NEES
- IBM Grid Toolbox
- ...
8 Other Grid Implementations

8.1 UNICORE

(Uniform Interface to Computing Resources)

UNICORE documentation

Videos & presentations there – see

• Introduction to UNICORE ← The slides we were looking at the lecture...

• Rich Client Tutorial 1: First start

• Rich Client Tutorial 2: Single jobs

• Rich Client Tutorial 3: Basic Workflows

• + tutorial videos 6 and 7 at the same place