

ILEADUE



eGEE Enabling Grids for E-scienceE

Grid Services

Heinz Stockinger
Swiss Institute of Bioinformatics



www.eu-egee.org



EGEE-II INFSO-RI-031688 EGEE and gLite are registered trademarks

eGEE Enabling Grids for E-scienceE

Outline

- **Move from “physicist's view to computer scientist's view”**
- **What is “Grid Technology”?**
 - Academia / Industry
- **General Grid Architecture**
 - OGSA Basics
- **Building blocks for a Grid Service**
 - Web-services
 - WSRF
- **Grid Services and Globus Toolkit 4 (GT4)**

EGEE-II INFSO-RI-031688 CSC 2007, Grid Track, Dubrovnik, Croatia 2

- **Some slides are based on material included in**
 - The Globus Toolkit 4 Programmer's Tutorial (<http://gdp.globus.org/gt4-tutorial/>)
 - © **Borja Sotomayor** and available for use and redistribution under the terms of the Globus Toolkit Public License (<http://www-unix.globus.org/toolkit/license.html>)."
- **Some of these slides have been taken from the Globus Alliance and**
 - in particular **Ian Foster, Lisa Childers**

- **Grid idea and first projects come from the academic environment**
 - Globus Toolkit
 - Condor
 - Legion, Avaki
 - EGEE: gLite
 - UNICORE
 - NAREGI
 - ...
- **Several companies currently “sell Grid technology”**
 - Sun Grid Engine
 - Oracle 10g Real Application Cluster (RAC)
 - IBM websphere
 - ...

- **Is there such a thing as “pure Grid technology”?**
 - Globus? Web services?
- **Is there something that is commonly agreed as “Grid technology”?**
 - Is there an added value?
- **There might not be some “really new stuff”**
- **There is not even an agreed definition on what a Grid is**
- **However, key standards and implementations for certain areas have been created to:**
 - Share resource
 - Allow transparent access
 - GSI (Grid Security Infrastructure)
 - GridFTP for file transfer
 - GRAM (or related protocols) for access to computing power
 - SRM for access to storage

- **Globus Toolkit (GT) is the most known and commonly used Grid middleware**
- **Several projects world-wide build on it**
 - Globus Toolkit is also used in gLite
- **Globus 2 provided many communities with a base toolkit to create prototype and production Grids**
 - Lots of experience has been gained by many communities such as High Energy Physics, Climate Modelling, ...

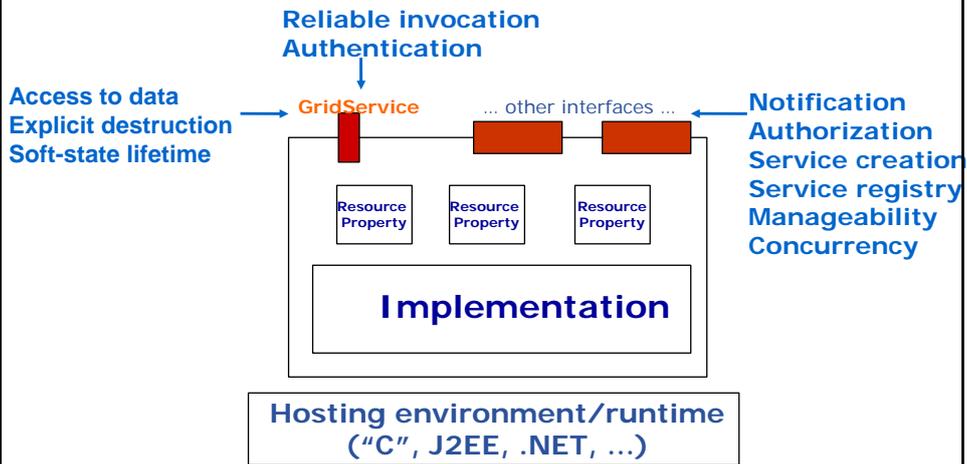
- **Condor, Globus, and Legion** have been among the most dominant Grid systems since the late 90ies
- **The Grid community needed to be organized**
 - Open Grid Forum (OGF)
 - Connects Grid researchers and industries all over the planet
- **Open Grid Forum:**
 - Standardisation body
 - Does not promote single technologies but Grid standards
 - Make Grids interoperable
 - There is no such thing as a “single” Grid
 - Define a standard Grid Architecture
 - Open Grid Service Architecture (OGSA)
 - Lots of involvement from industry
 - *Which also exposes the different views!*

- **Standard Architecture**
- **Service orientation to virtualise resources**
 - Everything is a service
- **Web services are the underlying technology**
 - Standard interface definition mechanisms
 - Evolving set of other standards: security, etc.
- **From Grids (Globus Toolkit)**
 - Service semantics, reliability & security models
 - Lifecycle management, discovery, other services
- ➔ **A framework for the definition & management of composable, interoperable services**

“The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration”, Foster, Kesselman, Nick, Tuecke, 2002

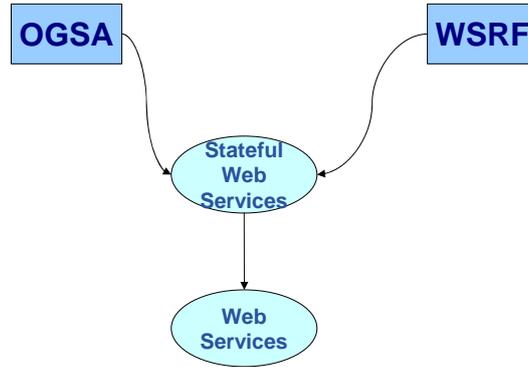
- **Grid service**
 - Standard interfaces and behaviours that address key distributed system issues: naming, service state, lifetime, notification
 - A Grid service is a Web service
 - **Stateful resource** is added
- **... supports standard service specifications**
 - Agreement, data access & integration, workflow, security, policy, diagnostics, etc.
 - Target of current & planned GGF efforts
- **... and arbitrary application-specific services based on these & other definitions**

- **Stateful information is stored in a Resource**
 - More specifically, in a **Resource Property**
- **A single Resource Property can be regarded as a single value (variable)**
- **Definitions (WS-ResourceProperties):**
 - A resource property is a piece of information defined as part of the state model of a WS- Resource.
 - A resource property may reflect a part of the resource's state, meta-data, manageability information, etc.



- Based on OGSA
- Standard defined by Global Grid Forum and OASIS
- Specifies how to make web services stateful

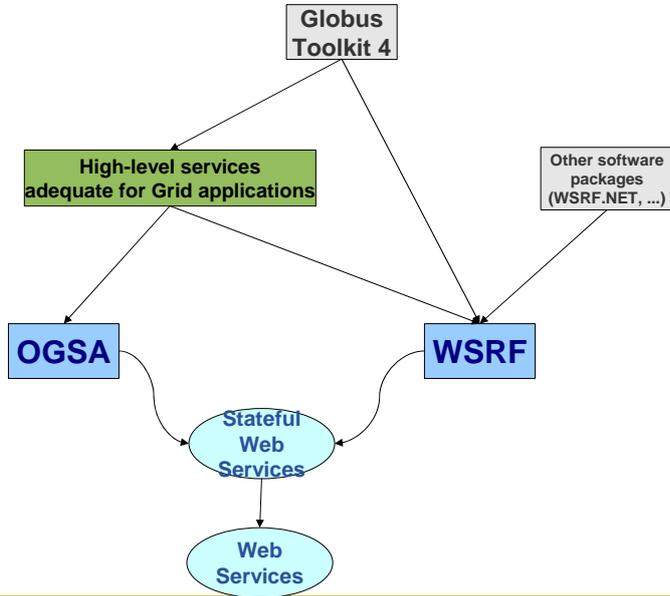
- **OGSA = Architecture**
- **WSRF = Infrastructure on which the architecture is built**



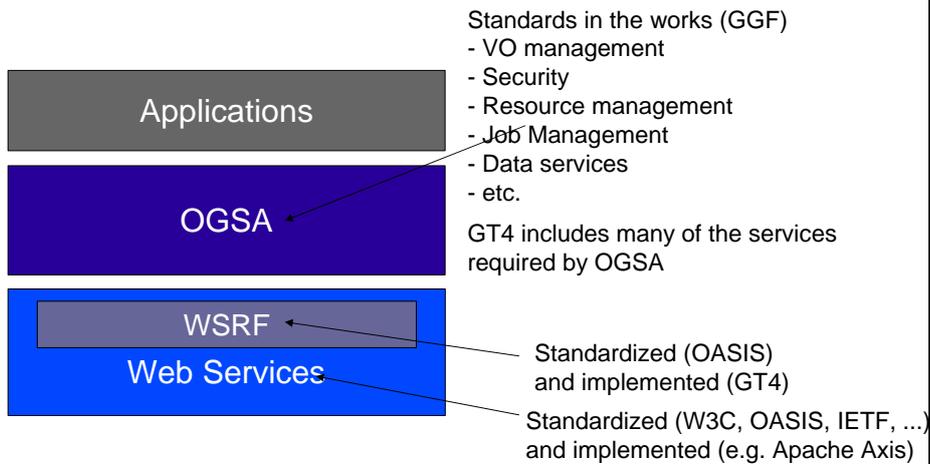
The definition of WSRF means that Grid and Web communities can move forward on a common base*

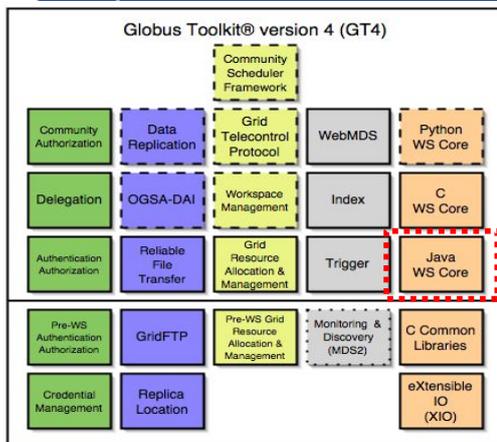


*WSRF evolutions, changes are now being lively discussed – mainly industry driven



- **GT4 includes a complete implementation of WSRF specification**



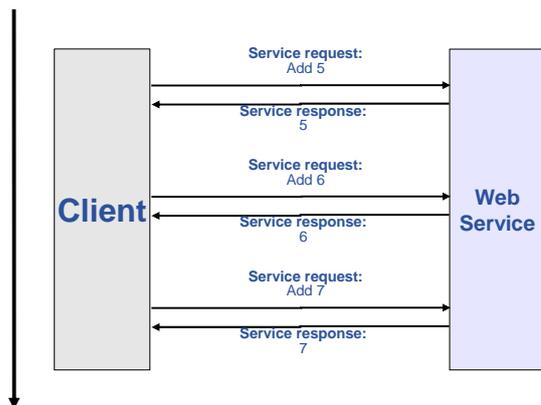


Covered today

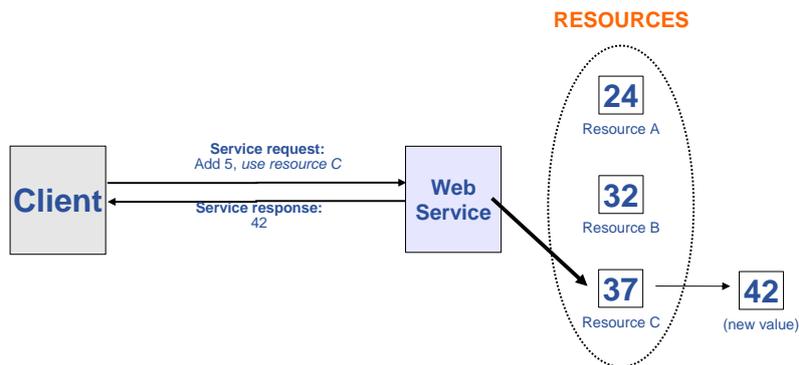
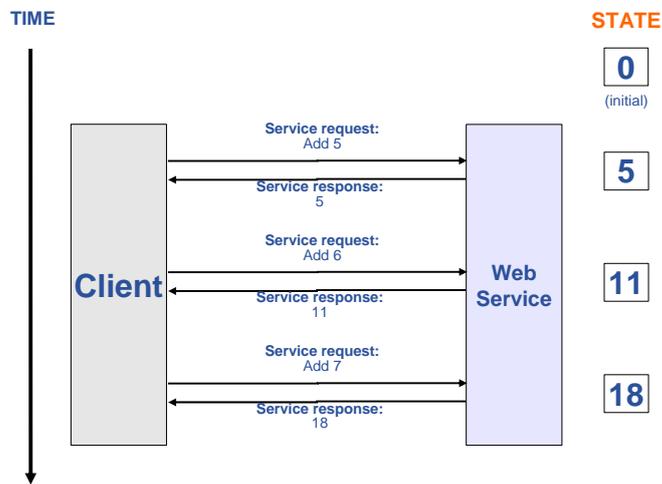


- Core GT Component: public interfaces frozen between incremental releases; best effort support
- Contribution/Tech Preview: public interfaces may change between incremental releases
- Deprecated Component: not supported; will be dropped in a future release

TIME

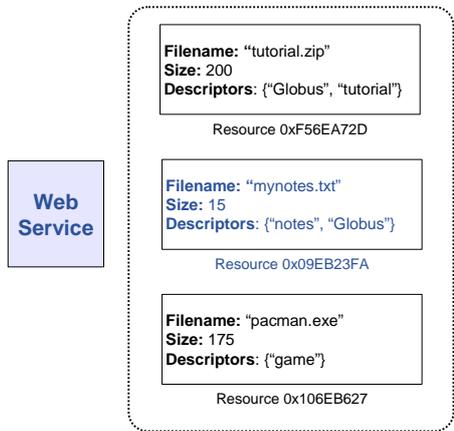


No state information is kept!



Remember: it's called
“Web Services Resource Framework” (WSRF)

RESOURCES

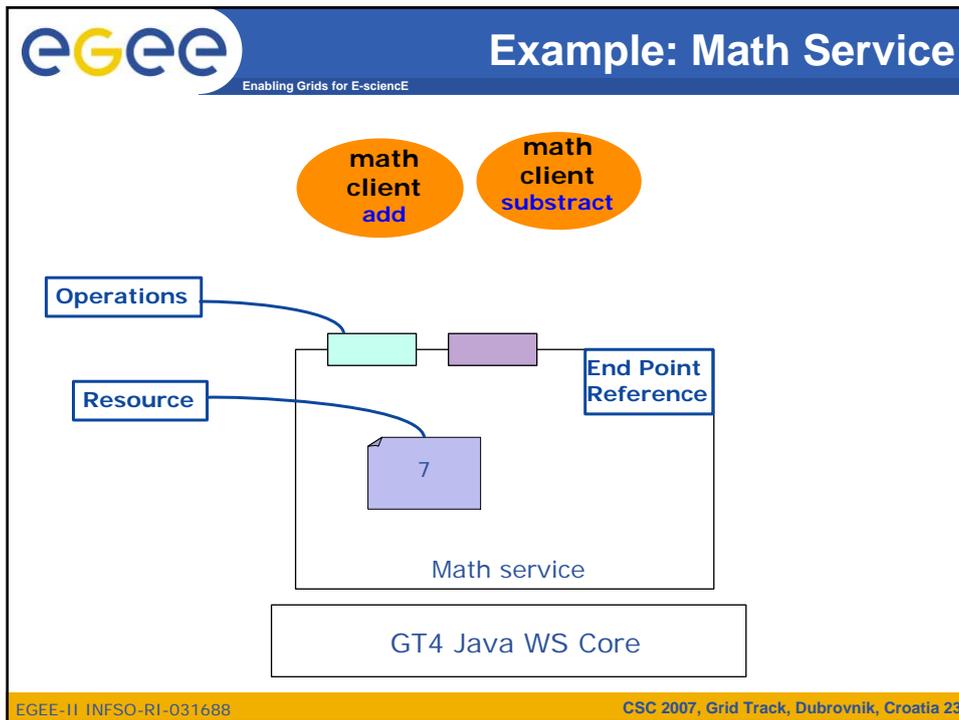


Each resource gets a unique ID

WS + Resource = **WS-Resource**

Address of WS-Resource is called **endpoint reference (EPR)**

- **WS-ResourceProperties**
 - A resource is composed of 0 or more resource properties (see previous slide)
- **WS-ResourceLifetime**
 - Resources can be started (created) and stopped (destroyed) at any time when the web service is running
- **WS-RenewableReferences**
 - End point references can be renewed if it becomes invalid
- **WS-ServiceGroup**
 - Web services and WS-Resources can be grouped together
- **WS-BaseFaults**
 - Faults are the equivalent to exceptions in Java or C++ to handle problems that occur at run-time
- **Related: WS-Notification, WS-Addressing**



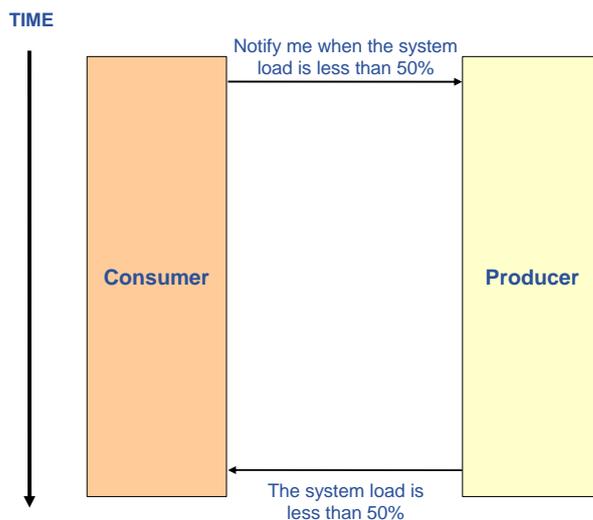
EGEE Enabling Grids for E-scienceE

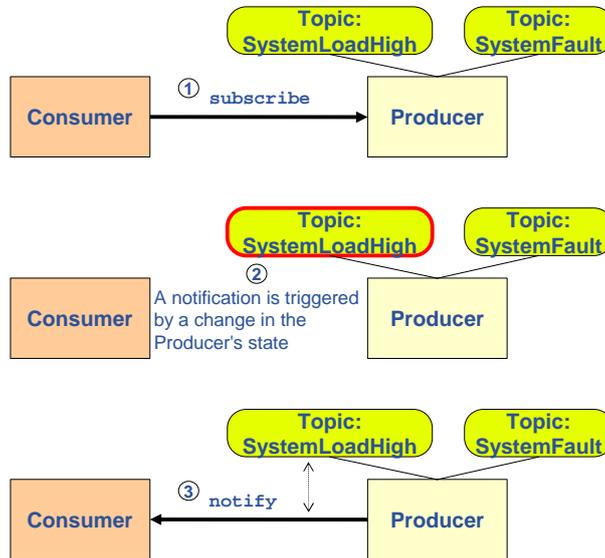
WS-ResourceLifetime

- **It's all about lifetime management**
 - A Grid service needs to be deployed (started) and destroyed (stopped)
 - “How long is a resource available?”
 - The lifetime is not equivalent to the availability of the service container
- **End Point Reference (EPR) can be used to destroy**
- **WS-ResourceLifetime provides different lifetimes:**
 - Immediate destruction
 - The resource gets destructed immediately after a client calls the termination method
 - Scheduled
 - A resource is available for a certain amount of time (e.g. 10 minutes)
 - Called “lease-based” model
 - Interested parties must renew a lease

EGEE-II INFISO-RI-031688 CSC 2007, Grid Track, Dubrovnik, Croatia 24

- **Based on a common design pattern**
 - Observer/Observable, Model-View-Controller
- **One can get a notification if a Resource Property changes**
- **WS-Notification consists of:**
 - WS-Topics
 - “Items of interest for **subscription**”
 - WS-BaseNotification
 - Defines standard interfaces of notification **consumers** and **producers**
 - Producers need to expose a **subscription operation**
 - Consumers need to expose a **notification operation**





- **OGSA is relatively new**
- **OGSA basis changed significantly:**
 - 2003: OGSi
 - 2004: WSRF
 - Major incompatible change!
- **Production Grids (EGEE, TeraGrid, OSG, DEISA, ...) did not follow this**
 - Infrastructure needed to be kept up running
 - Incompatible changes hard to deploy on systems encompassing several hundreds of sites
 - Hence production Grids are typically quite conservative and wait for **established** standards (e.g. gLite tries to follow WS-I)






www.eu-egee.org

EGEE-II INFSO-RI-031688

EGEE and gLite are registered trademarks

Information on GT4 Exercises



GT4 Web Service Core

- Will be used in exercises
- Build web services in Java
- Follow the tutorial by Borja Sotomayor
 - <http://gdp.globus.org/gt4-tutorial>
- Only one chapter is compulsory
 - Chapter 3. Writing Your First Stateful Web Service in 5 Simple Steps
 - Step 1: Defining the interface in **WSDL**
 - Step 2: **Implementing** the service in **Java**
 - Step 3: Configuring the deployment in **WSDD** (and JNDI)
 - Step 4: Create a **GAR** file with Ant
 - Step 5: **Deploy** the service into a **Web Services container**
 - It is important that you understand the steps since they are also required in the mini project.

EGEE-II INFSO-RI-031688

CSC 2007, Grid Track, Dubrovnik, Croatia 30

- **WSDL file – additions**

- xmlns:wsrp
- xmlns:wsrpw

We use XML Schema for this purpose

- **<!-- RESOURCE PROPERTIES -->**

```
<xsd:element name="Value" type="xsd:int"/>
<xsd:element name="LastOp" type="xsd:string"/>
<xsd:element name="MathResourceProperties">
  <xsd:complexType>
    <xsd:sequence>
      <xsd:element ref="tns:Value" minOccurs="1"maxOccurs="1"/>
      <xsd:element ref="tns:LastOp" minOccurs="1" maxOccurs="1"/>
    </xsd:sequence>
  </xsd:complexType>
</xsd:element>
```

- **<!-- PORTTYPE -->**

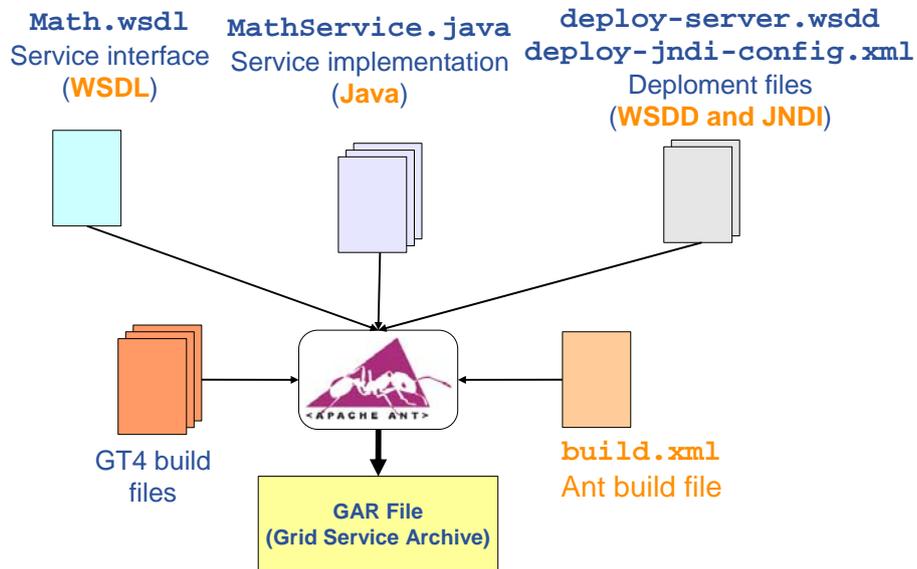
```
<portType name="MathPortType"
  wsdlpp:extends="wsrpw:GetResourceProperty"
  wsrp:ResourceProperties="tns:MathResourceProperties">
```

- **Web service container is provided by Globus Toolkit**

- We use the Globus container rather than Tomcat, etc.
- Start the container **without** using **security**
 - \$GLOBUS_LOCATION/bin/globus-start-container -nosec

- **The build process uses Ant**

- For people familiar with Makefiles: Ant is equivalent to makefiles
- Uses XML to express build properties



- **GT4 Tutorial links:**
 - <http://gdp.globus.org/gt4-tutorial/>
 - <http://www-unix.mcs.anl.gov/~childers/tutorials/BAS/GPN/>
 - <http://www.globus.org/toolkit/tutorials/BAS/SDSC/index.html>

- We have seen what **Grid services** are and how they can be implemented with **Globus Toolkit 4**
- **WSRF is a very new standard that is not yet fully finished**
 - Will be included into IBM's websphere
 - Not fully yet accepted as an industry standard
 - Industry continues to discuss certain aspects of it
- **Many Grid projects are still in the transition phase between:**
 - Conventional Web Services
 - Grid Services based on WSRF
- **However, Service Oriented Architectures (SOA) and Web services are full industry standards**
- **Caveat: building new services on OGSA is easy – making them scalable and performant for a large scale infrastructure is a major task!**