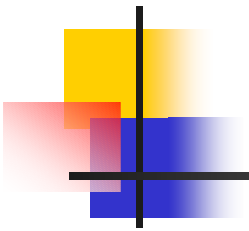
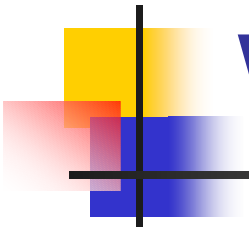
The Open Grid Services Architecture and Data Grids

Karthik Ram Venkataramani
kv8@cse.buffalo.edu
University at Buffalo



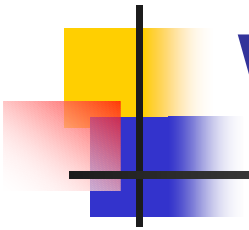
Introduction

- Sharing and Coordinated use of diverse resources in dynamic VO's.
- Need for explicit structure for Grid Components
- OGSA defines an Architecture for the Grid.
- Defines a set of rules that make up a Grid Service.



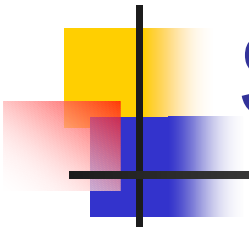
Web Services

- Technique for describing software components to be accessed
- Define methods for accessing and discovering the components
- Serve as a programming language, programming model and system software



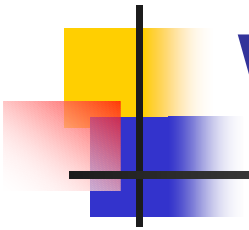
Web Service Standards

- Simple Object Access Protocol (SOAP)
  - Messaging between service provider and requester.
  - Follows http Request-Response model.
  - XML envelope to over the message to describe what is in, and how to process it.
  - Defines an XML based convention for the Request-Response.
  - SOAP payloads can also be carried in FTP, JMS.



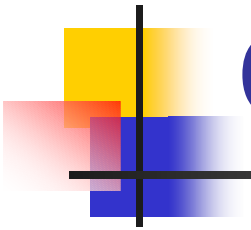
Standards ...

- Web Services Description Language (WSDL)
  - Standardized XML format for describing network services
  - Describes the name, location and ways to communicate.
  - Defines web services as a set of endpoints operating on messages (document oriented or RPC payloads)



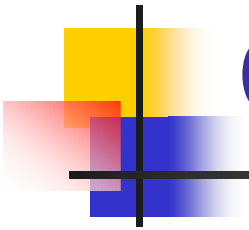
Why Web Services ??

- WSDL supports dynamic discovery and composition of services in heterogeneous environments
- A framework based on Web Services can exploit numerous tools and extant services.



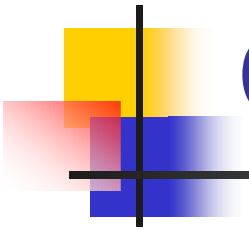
Grid Services and OGSA

- Defined as a web service that provides a set of well defined interfaces and conventions
- Interfaces address
  - Discovery, Dynamic Service Creation
  - Lifetime management, notification
- Conventions include
  - Naming services and upgradeability



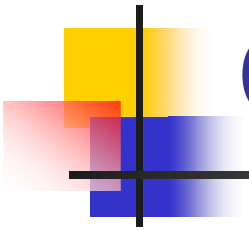
OGSA ...

- OGSA defines the semantics of a Grid Service Instance
- Defines basic behavior and does not specify the what a service does and how it does it.
- Doesn't address issues of implementation programming model, language, tools and execution environment.



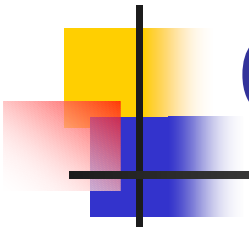
OGSA Technical Details

- The OGSA Service Model
  - Basic Premise: Everything is represented by a Service
  - Uniform Service Oriented Model
  - Specifies core set of consistent interfaces from which all Grids are implemented
  - Grid Services maintain internal state for the lifetime of the Service.



OGSA Service Model ...

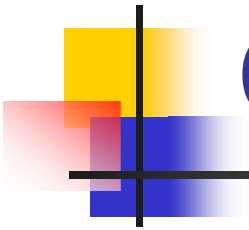
- Grid Service Instance refers to a particular instantiation of a service.
- State oriented service facilitates failure recovery mechanisms.
- OGSA services can be created and destroyed dynamically.



Grid Service Handle (GSH)

- Every Grid service instance is assigned a globally unique GSH.
- GSH carries no protocol specific or instance specific information.
- All instance specific information are encapsulated in a single abstraction called Grid Service Reference (GSR)

Grid Service Reference (GSR)

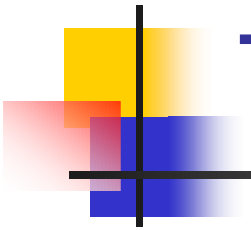


- Unlike GSH, GSRs change during the Service's lifetime.
- GSR has an explicit expiration time
- OGSA mappings define mechanisms for obtaining an updated GSR.
- A GSR does not guarantee access to a Grid service, local grid policies may enforce their constraints.

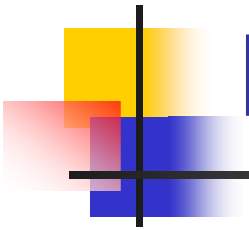
Grid Service Interface

- OGSA doesn't define a specific set of services but defines a set of interfaces for manipulating service models.
- A mandatory interface which must be supported by all Grid Services.
- Interface for manipulating service handles, reference abstractions.

Transient Services: Factories



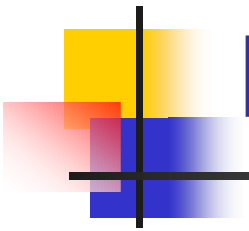
- Services implement this interface to create new Grid service instances.
- This service is called a *factory*.
- *CreateService()* operation creates a requested service and returns the GSH and initial GSR.
- Again OGSA does not specify how the instance is created.



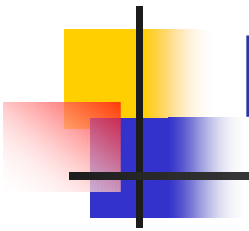
Lifetime Management

- Soft-state approach where every instance is created with a specific lifetime
- Initial lifetime can be extended by explicit client request.
- Client can always know when the Grid service will terminate.
- Resource consumption at hosting environment is always bounded

Managing Handles and References

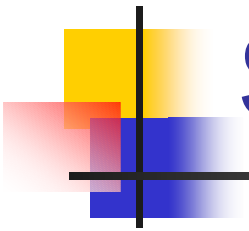


- GSH lives forever, but GSR expires
- A handle-to-reference mapping interface takes a GSH and returns a valid GSR.
- Mapping operations may be controlled and requests denied.
- Again a valid GSR doesn't promise access to the grid service



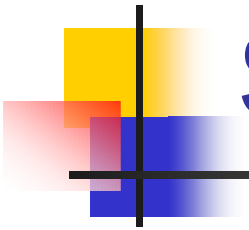
HandleMaps

- Every Grid service instance is always registered with at least one home handleMap.
- GSH includes the handleMap's identity
- All handleMap services are also identified by a URL
- Mapping operation is bound to a single protocol like HTTP.



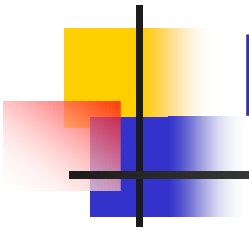
Service data and discovery

- Each Grid service instance has an associated service data (usually a collection of XML elements)
- The mandatory GridService interface defines the operation for querying and retrieving service data
- GSH, GSR, primary key, handleMap etc are service data members



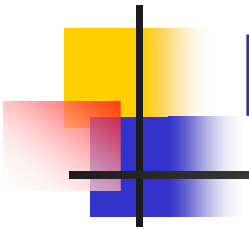
Service discovery

- Process of identifying a subset of GSH's.
- Attributes used include interfaces provided, number of requests serviced, load, policy statements etc.
- A registry interface provides operations by which GSH's can be registered with the registry service.
- Registration is soft-state operation and must be periodically refreshed.



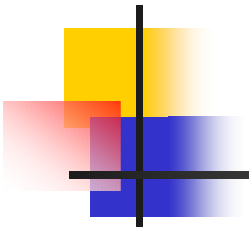
Notification

- OGSA notification framework allows clients to be notified by messages.
- Interface also includes framework for asynchronous one way delivery of notifications.
- The source must support the OGSA *notificationsource* interface.
- Clients wishing to receive notifications must implement *notificationsink* interface.



Data Grids

- Address computational and data intensive applications
- Combine huge amounts of data and resources which are geographically distributed
- Provide very high availability and reliability



Issues in Data Grids

- Resource scheduling
- Data access optimization
- Network storage
- High speed data transfers
- Data Management
- Security

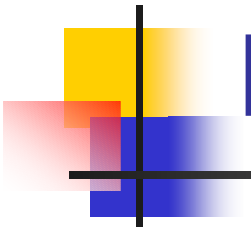
Data Grid Services

- In most cases Data Management involves file handling
- Typically Data Grids need to handle various types of data
  - Files
  - File Collections
  - Relational databases
  - XML Databases
  - Virtual Data Sets
- Data must be identified using a mechanism –Grid Data Handle (GDH)

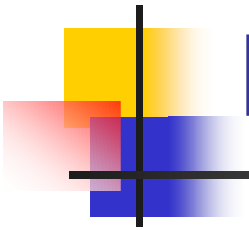
Grid data Handle (GDH)

- Similar to a GSH in OGSA.
- Unique identifier to locate and retrieve data
- GDH is immutable, facilitates location, tracking, transmission etc

Grid Data Reference and Data Registry

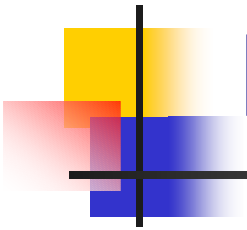


- GDH carries no protocol specific or instance specific information.
- The Grid Data Reference holds all protocol and instance specific information
- GDR includes data location, access protocols supported, data lifetime and other meta data.
- Data Registry holds the mapping from GDH to GDR



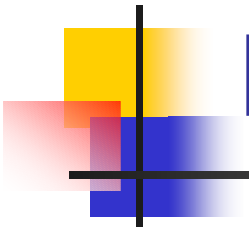
Data Grid and OGSA

- Factories
  - Increase robustness and availability of services
  - Transactions are made easy by building on functionality of factories
  - Each Component can have a dedicated factory and be instantiated.



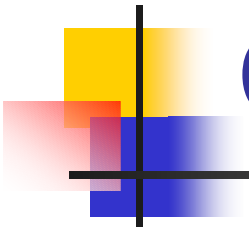
Data Grids and OGSA ..

- Registries
  - Where should registry interfaces be deployed?
  - Scalability issues restrict registries to be kept in as few nodes as possible
  - Must be kept in nodes which have high QOS.



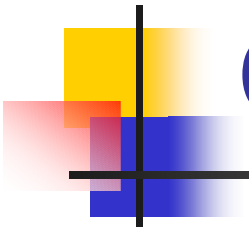
Data Grids and OGSA ..

- Service Lifetime Management
  - Keep a set of factories for very long lifetime and keep them alive using OGSA lifetime extension mechanisms
  - Setup factories which create new services on demand. (Applications are responsible for keeping services alive)
  - Redirect incoming applications to existing services based on loads.



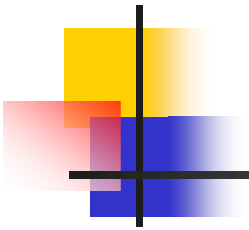
OGSA Summarized .

- OGSA provides an architecture for the Grid
- Provides a generic framework for interoperability among heterogeneous service implementations
- Uses the web service model for building a Grid service architecture



OGSA Summarized ...

- Specifies GridService as a mandatory interface to be included. The GridService port has three operations
  - FindServiceData()
  - Destory()
  - SetTerminationTime()
- OGSA standard service ports are
  - HandleMap
  - Registy
  - Factory



ISSUES ..

- Dealing with service overloads, VO partitioning, factory/registry unavailability
- Elaboration on QOS metrics, probably make them as a separate namespace that can be queried
- Enforcing local and VO wide security policies, local or global authorization?

References

- Grid Computing-
  Making the Global Infrastructure a Reality.
  Fran Berman, Geoffrey Fox, Anthony Hey

- http://www-fp.globus.org/ogsa/

- An Analysis of The Open Grid Services Architecture
  Dennis Gannon, Kenneth Chiu, Madhusudhan Govindaraju, Aleksander Slominski
  Department of Computer Science, Indiana University, IN