UNIT - II

Grid Services.

OGSA:

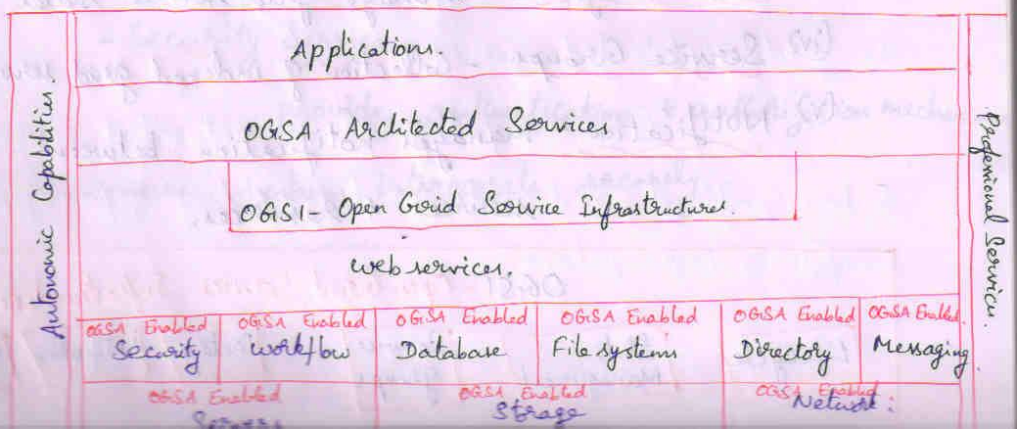
Open Grid Service Architecture. It defines how different components will interact each other in grid environment. It is a set of standards defining the way in which information is shared among diverse components of large, heterogeneous grid systems.

A grid system is a scalable WAN that supports resource sharing and distribution.

Architecture of OGSA:

The OGSA architecture consists of four layers. They are

(i) Physical and logical Resources layer.
(ii) Web Service layer.
(iii) OGSA architected Grid Services layer.
(iv) Grid Applications layer.



a) Physical and Logical Resources layer:-

Servers, storage, networks are the physical resources. Database managers, workflow managers are logical resources. The logical resources manage physical resources. Both logical and physical resources are OGSI enabled services.

b) web Services layer:-

web service is software available on that could interact with other software using XML. It consists of Open Grid service Infrastructure (OGSI) sublayer which specifies grid services and provide consistent way to interact with grid services. It also extends web service capabilities.

It consists of 5 interfaces:-

(i) Factory - provide way for creation of new grid s
(ii) Lifecycle - Manages grid service life cycles.
(iii) State Management - Manage grid service states.
(iv) Service Groups - Collection of indexed grid servi
(v) Notification - manages notification between services & resources.

OGSI - Open Grid Service Infrastructure

c) OGSA Architected Services Layer:

This layer is mainly classified into three service categories. They are.

(i) Grid Core Services.

(ii) Grid Program Execution Services.

(iii) Grid Data Services.

(i) Grid Core Services:

It composed of 4 main types of services.

- Service Management

- assist in installation, maintenance & troubleshooting tasks in grid system.

- Service Communication.

- it includes functions that allow grid services to communicate.

- policy Services

- provides framework for creation, administration, & management of policies for system Operation.

- Security Services.

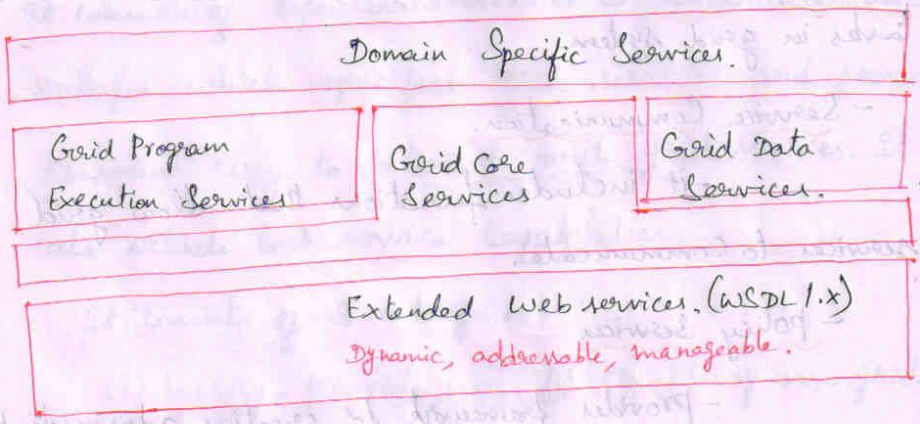
- provide authentication & authorization mechanisms to ensure systems interoperate securely.

(ii) Grid Program Execution Services:

It supports unique grid system high performance computing, collaboration, parallelism. It also supports Virtualization of resource processing.

(iii) Grid Data Services:

It supports data virtualization, and provide mechanism for access to distributed resources such as databases and files.



d) Application layer:

This layer comprises of applications that use the grid architected services.

Grid computing allows networked resources to be combined and used. Grid computing offers great benefit to organizations.

OGSI - Open Grid Services Infrastructure.

OGSI provides detailed description about grid services in a formal and technical specification manner. It also defines the working of grid services in a described way. GT3 (Globus Toolkit) includes a complete implementation of OGSI.

Other implementations are

a) OGSI::Lite (Perl)1

b) UNICORE OGSA Demonstrator 2.

The OGSI specification defines grid services and build upon web services. OGSI creates an extension model for web Services Definition Language (WSDL) called GWSDL (Grid WSDL) due to interface inheritance and service data for expressing state information. 2 marks

The Components of OGSI are.

a) Lifecycle.

b) State Management.

c) Service Groups.

d) Factory

e) Notification

f) Handle Map.

Data Intensive Grid Service Models:

The grid applications are normally grouped into two categories.

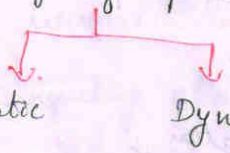
a) Computation Intensive.

b) Data Intensive.

The data intensive applications deal with massive amounts of data. The grid system must specially designed to discover, transfer and manipulate the massive data sets. Transferring the massive data is a time consuming task.

Data access method is also known as caching, which is often applied to enhance data efficiency in a grid environment. The Replication strategies determine when and where to create a replica of data.

Strategies of Replication.



a) Static Method:

The locations and number of replicas predefined and it cannot be modified. Replication

Operation require little overhead. It doesnot adopt ondemand changes, bandwidth and storage variability. Optimization is required to determine the location and number of data replicas.

b) Dynamic Method:-

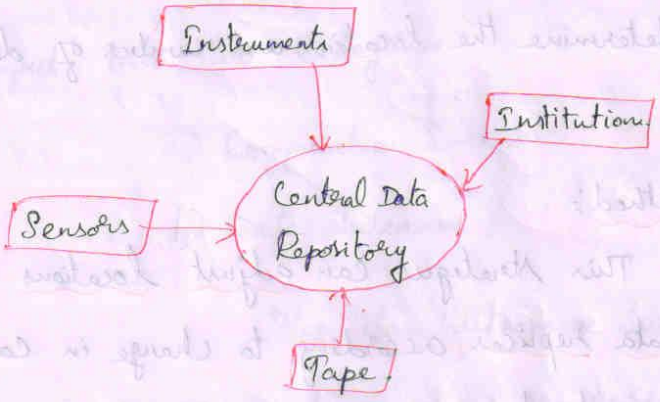
This strategies can adjust locations and number of data replicas according to change in conditions. There is frequent data moving operations which result in more overhead. Optimization may be determined based on whether the data replica is being created, deleted or moved. The most common replication include preserving locality, minimizing update costs and maximizing profits.

Grid Data Access Models:-

The Grid Data Access Models consists of four access models for organizing a data grid. They are

a) Monadic Method.
b) Hierarchial Model.
c) Federation Model.
d) Hybrid Model.

a) Monadic Method:

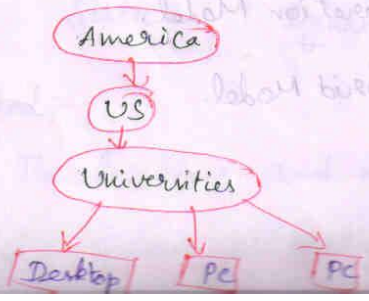


This is a Centralized data repository model. All datas are saved in it. When user wants to access some data they have no submit request directly to Central repository. No data is replicated for preserving data locality.

Disadv:- (i) For larger grid, this model is not ~~used~~ good interms of performance and availability.

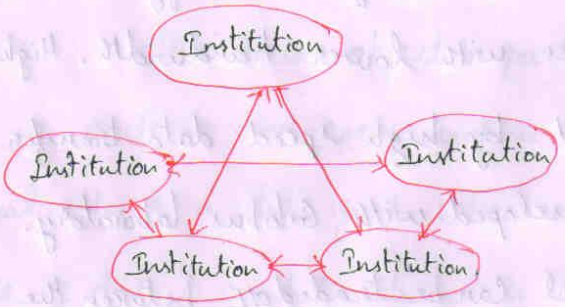
(ii) Data Replication is permitted only when fault tolerance is demanded.

b) Hierarchial Model:-



It is suitable for building a large data grid which has only one large data access directory. Data may be transferred from the source to second level. After being forwarded several times specific data objects are accessed directly by users. Higher level data center has a wide coverage area. Security services are easier to implement in this model.

c) Federation Model :- (Mesh Model)

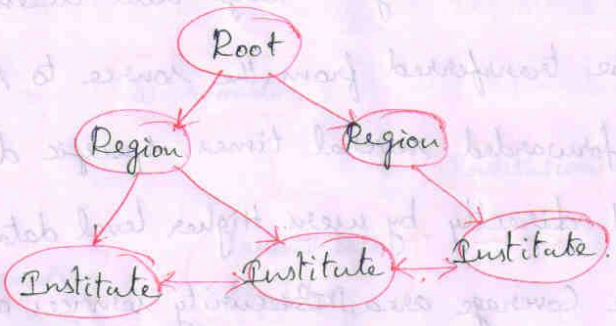


The federation model is suited for designing a data grid with multiple source of data supplies. It is also known as mesh model. The data is shared and items are owned and controlled by their original owners. Only authenticated users are authorized to request data from any data source.

Disadv :-

This mesh model cost the most when the number of grid institutions becomes very large.

e) Hybrid Model:



The hybrid model combines the best features of hierarchial and mesh models. Traditional data transfer technology such as FTP app for networks with lower bandwidth. Higher bandw are exploited by high speed data transfer tools suc GridFTP developed with Globus laboratory. The cost hybrid model can be traded off between the two extr models of hierarchial and mesh-connected grid

Parallel vs Striped Data Transfer:-

Parallel Data Transfer:-

It opens multiple data streams for passing subdivided segments of a file simultaneously. Although the speed of each stream is same as in sequential streaming, the total time to move data in all streams can be significantly reduced compared to FTP transfer.

Striped Data Transfer:-

The data object is partitioned into a number of sections and each section is placed in an individual site in a data grid. When a user requests a piece of data, a data stream is created for each site in a data grid. and all the sections of data objects are transferred simultaneously.