The Grid

A new paradigm in computing

Randall Sobie
Institute of Particle Physics
University of Victoria

Outline

Introduction

Potential applications

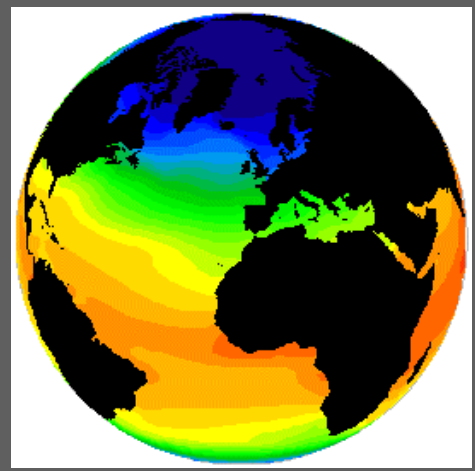
Grid architecture

Security

Canadian initiatives

Introduction

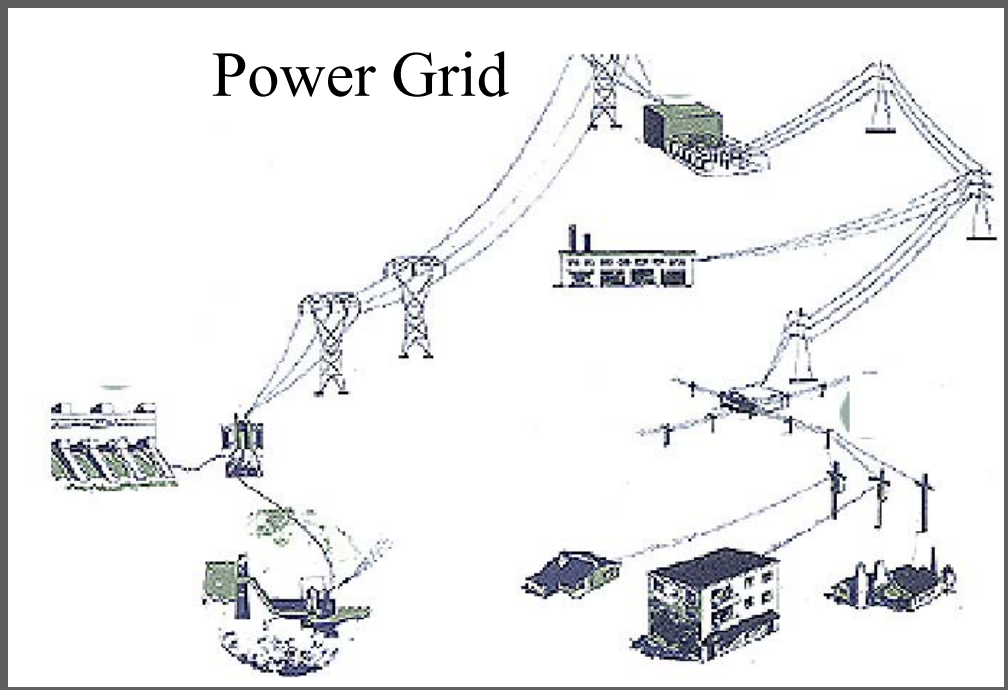
- We have seen remarkable advances in computer processing, storage and networking *(and the trend seems to be continuing)*
- Yet the demand for computing is growing faster than these technological advances
  - Bioinformatics
  - astronomical and particle physics images
  - satellite data
  - climate simulation
  - medical images



The Grid makes it possible for geographically distributed groups to access and share previously unavailable resources

Grids already exist

Sharing distributed resources is not a new concept.



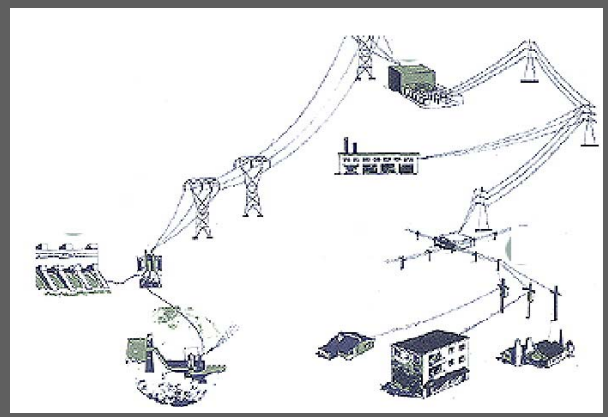
The use of power generated from distributed sources is more cost effective than building dedicated generators.

Similarly, the computing Grid will provide additional access and effective use of existing resources.

The significant advances in computer networking makes the idea of a computing grid a real possibility

Why are computing Grids possible now?

- The world is universally connected to the Internet
- Trend to collaborative research and sharing of resources
- Significant computing resources are pervasive due to the rapid increase in technology and decrease in cost
- Network capacity is increasing faster than storage capacity or processor speeds



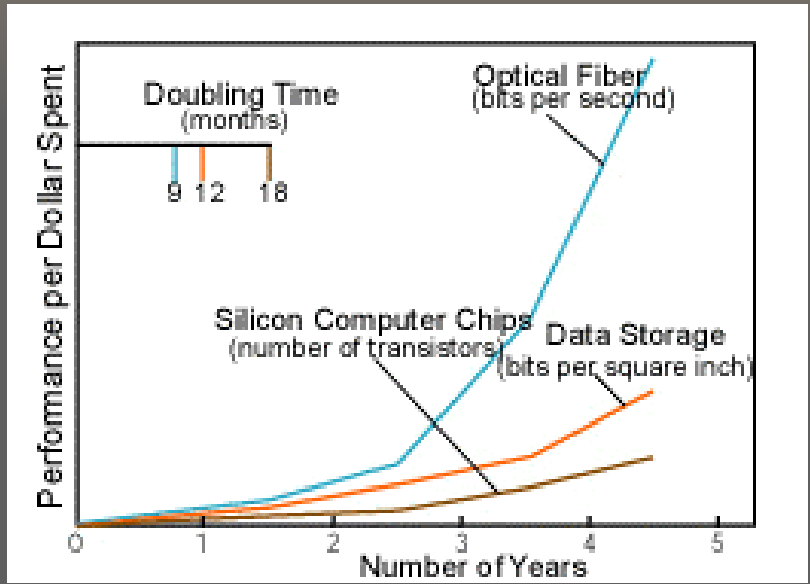
Moore's Law

Time to double…

Network	9 months
Storage	12 months
Processors	18 months

Over a 5 year period we expect

Network	x 100
Storage	x 30
Processors	x 10



If this trend continues, then communication becomes free

How do we exploit this network capacity?

We need to imagine applications that are communication intensive

- Pooling of distributed computational resources
- Streaming data to remote computers
- Linking sensors
- Connecting people in collaborative environments

For example,

- I no longer need to run code locally
- I can have my code access data from another remote site
- I can let my colleagues view the results with sophisticated visualization tools

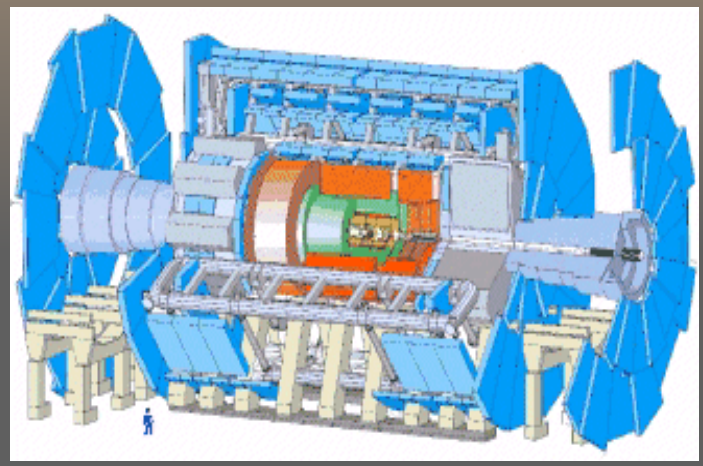
Potential applications

1. **Distributed computing**
   - Seti @ home
     - Screen savers searching for signatures for ET

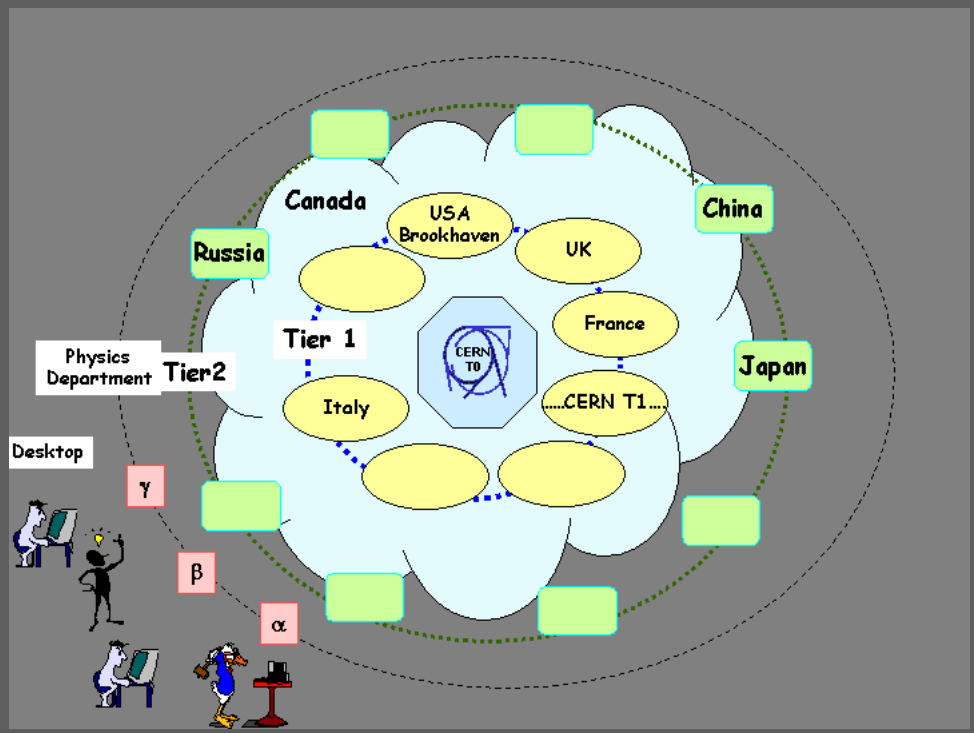
   - FightAIDS @ Home (Entropia Inc)
     - Thousands of home computers getting small pieces of data and performing calculations that model how drugs interact with various HIV virus mutations.

2. Large scale data analysis

- *Large Hadron Collider* at CERN will produce petabytes of data per year
- Analysis of the data will require a world wide Grid of processors and storage



ATLAS Detector



ATLAS Computing Model
Hierarchy of computing facilities distributed around the world

3. Real-time analysis

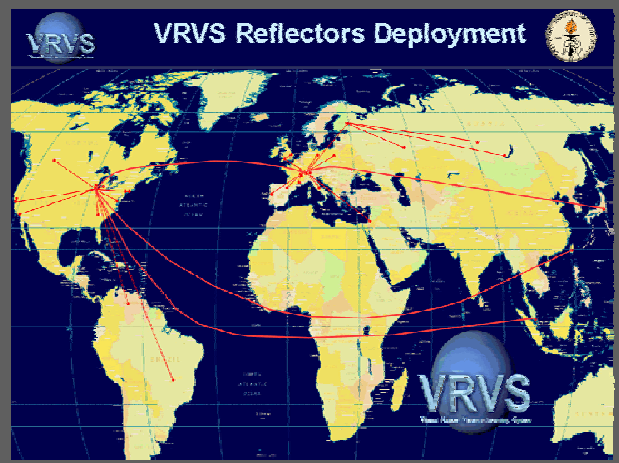
- The use of computers to analyze data from instruments in a real-time or quasi-real time environment

- The data is rapidly processed and the information is fed back to the operators.

Significantly enhances the capability of the instrument.

4. Collaborative work

- The Grid may facilitate long-distance communication and interaction



Global Grid Initiatives

Many countries have recognized the importance of the Grid

- UK E-Science project (200M pounds)
  - grid within the U.K. for particle physics research
- EU DataGrid project
  - grid for applications in high energy physics, environmental science, bioinformatics
- many projects in the US
  - DOE Science Grid

Grid Canada

- Initiated by CANARIE, NRC and C3.ca Association

Globus Tookit™

• No formal standards exist for Grids

• Most major Grid projects are being built on protocols and services provided by the Globus Toolkit ™

Globus Toolkit ™ (www.globus.org)

- Developed at Argonne National Laboratory aud University of Southern California
- Open-architecture, open-source set of services and libraries to support Grids and Grid applications
- Includes software for security, resource and data management, and fault detection

• In Canada, Platform Computing in Toronto, has the first commercially supported version of the Globus Toolkit (www.platform.com)

The goal of the Grid

What will the Grid do for the user who submits his job?

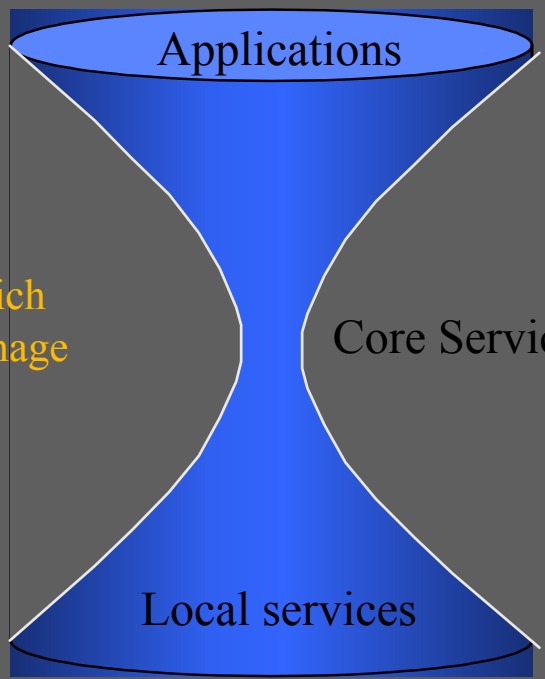
- Locate the remote resources that are available
- Optimizes the use of the widely distributed resources
- Organizes access to your data
- Deals with authentication to the different sites being used
- Interfaces to local resources
- Runs the job
- Monitors progress
- Recovers from problems

And tells the user when the job is complete!

Grid Architecture

The Architecture defines the fundamental components, specifies the purpose and function of the components, and how these components interact with one another.

Hourglass model:

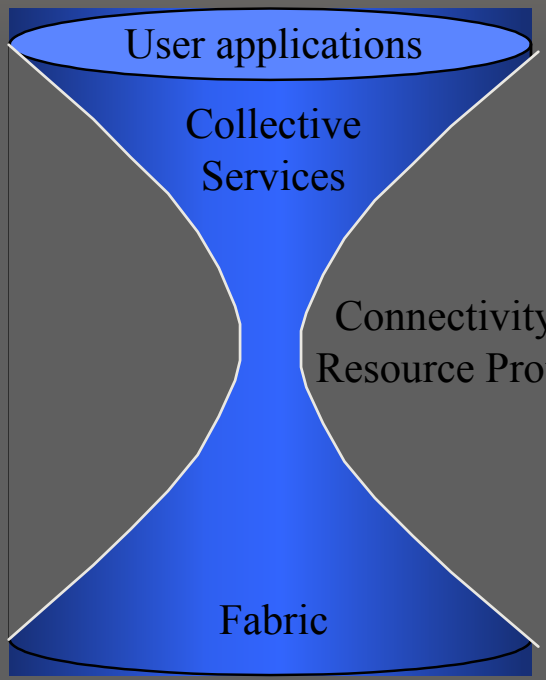


High level user applications

Protocols or sets of rules with which users and resources establish, manage and exploit relationships

Fabric and computing resources

Grid Architecture Overview



Collective layer – coordinates the use of multiple resources

Connectivity layer – controls how things talk to each other (eg Internet protocols)

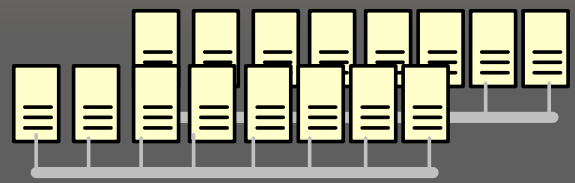
Resource layer – negotiates access to resources

Fabric layer - the resources available on the Grid (eg. computing resources)

Grid fabric layer

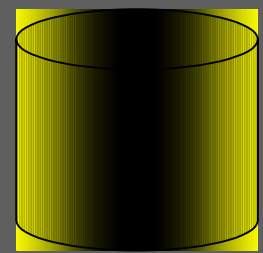
The fabric contains the resources which can be shared on the Grid

- Computational resources
- Storage systems
- Databases and catalogs
- Network resources
- Sensors



The fabric requires enquiry and resource management mechanisms

1. Computational resources
   - hardware and software configuration
   - dynamic information such as the load information
   - controlling the execution of jobs
2. Storage resources
   - Mechanisms to get and put files
   - Mechanism to control resources (space, bandwidth, …)



Connectivity and resource protocol layer

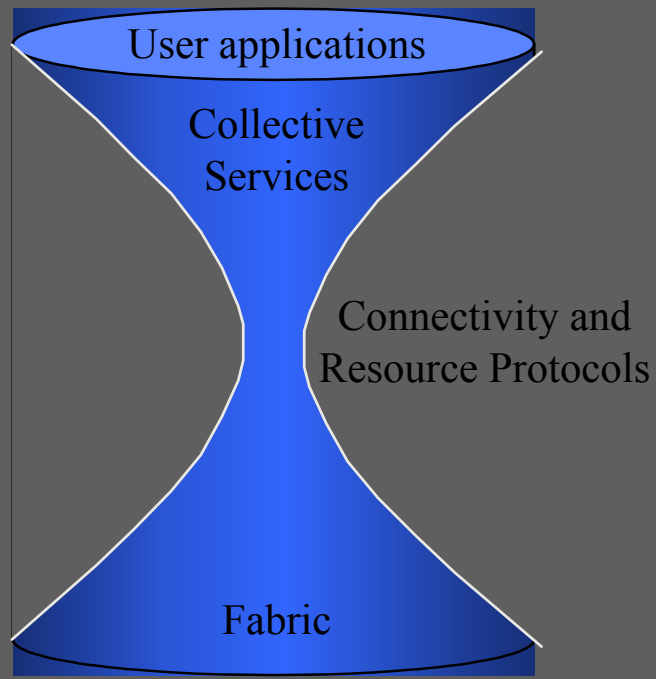
This layer has the set of protocols or rules with which users and resources establish, manage and exploit relationships.

Connectivity layer *("talking to things")*

- Communication protocols – exchange of resources between fabric layers or resources (eg TCP/IP)
- Authentication protocols – verifying and identifying users

Resource layer *("negotiating access to resources")*

- Information protocols - state and structure of a resource (configuration, load, usage policy, …)
- Management protocols - negotiating access to a resource



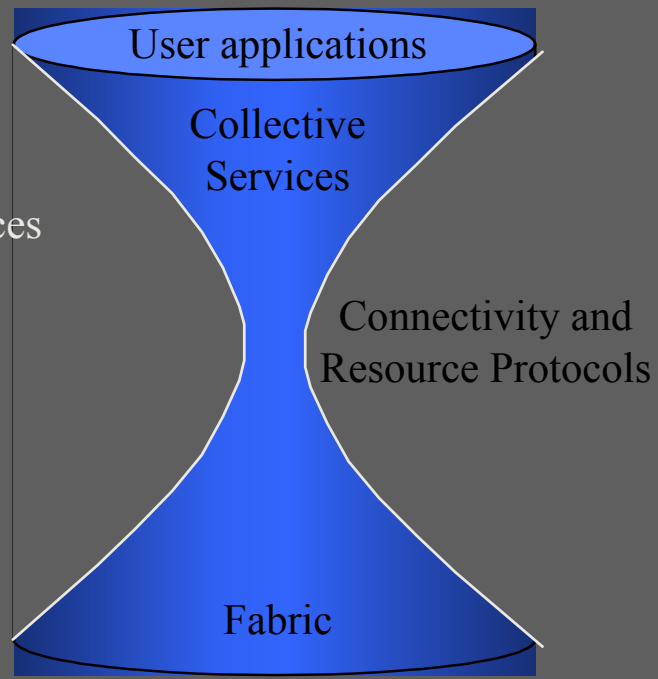
Collective layer

The Resource layer focus on interactions *with a single resource*.

The Collective layer contains protocols and services for interactions on *collections of resources*.

Examples:

- Services to discover the existence and properties of resources
- Monitoring and diagnostics
- Data replication services



Security

Traditional security is based upon the client - server model.

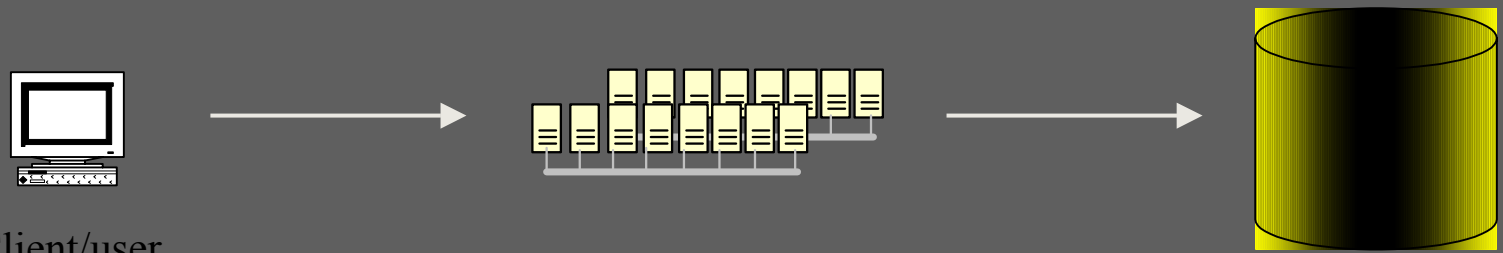
- Verification of each others identity
- Server determines whether to allow the client request

Client/user

Server

For the Grid, the distinction between clients and servers is blurred.

- A user submits a job to run on a resource and that resource needs to access data from another resource



Client/user

A resource can act as both a client and server.

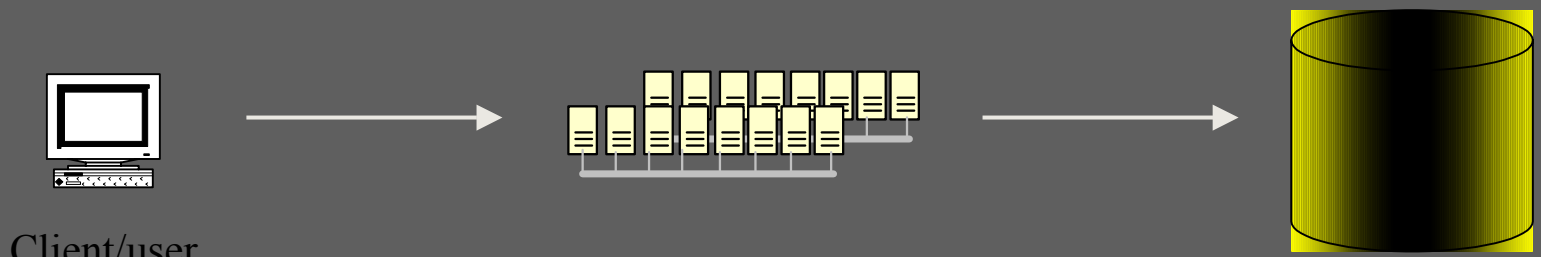
Security or Authentication requirements

Single sign-on

- Users login once and create a "grid proxy credential"
- The proxy credential gives the user's program the authorization to access resources

Interoperate with local security mechanisms.

- Grid security solutions must interoperate with local solutions
- Resource providers do not need to cooperate or interact on security issues



Client/user

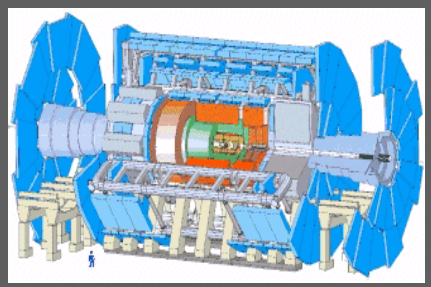
For example, a user has access rights to a processing and storage facility. There is no need for any security interaction between the two facilities.

Group or community authorization

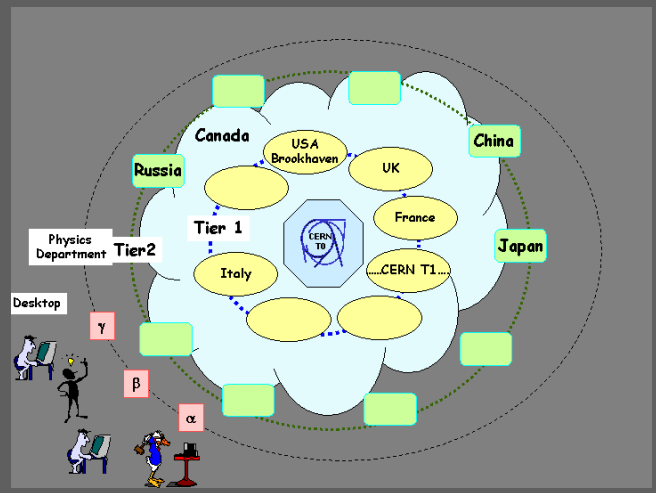
- There may be large, dynamic communities wishing to access resources over a Grid
- Each resource cannot keep track of individual memberships or privileges

There is the need for a trusted 3rd party to provide group certification.

- E.g.. ATLAS, Grid Canada, ….



2000 researchers



ATLAS Grid

Grid Activities in Canada

ATLAS experiment at CERN

- Establish a Canadian particle physics Grid
  - Utilize new CFI funded facilities
- Planned use
  - Production of simulated data
  - Analysis of physics data
- Current work
  - Linux testbed and large data storage
  - Participation in ATLAS "data challenges"
  - Large file transfer demonstration project using CANET4

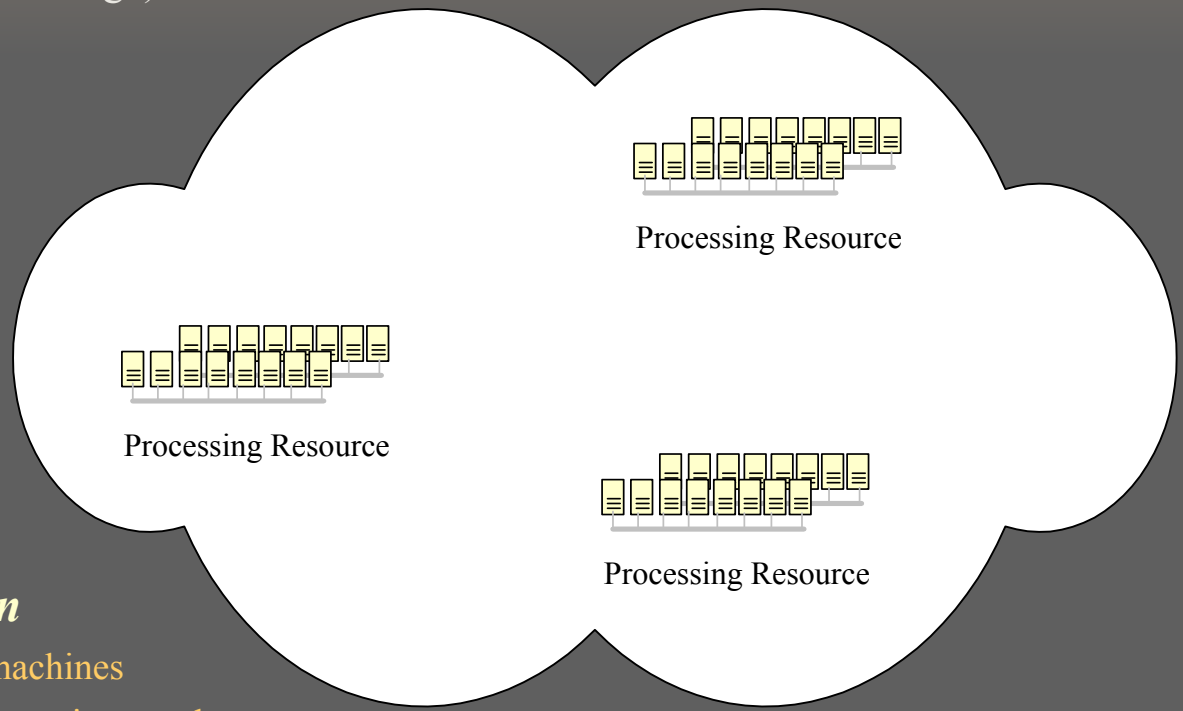
Alberta
SFU
UBC
TRIUMF
Victoria

Montreal
Carleton
Toronto
York

Many new CFI funded facilities

Canadian activities: How do we exploit our Grid?

We need access to specific information in order to effectively use a Grid ("Resource scheduling")



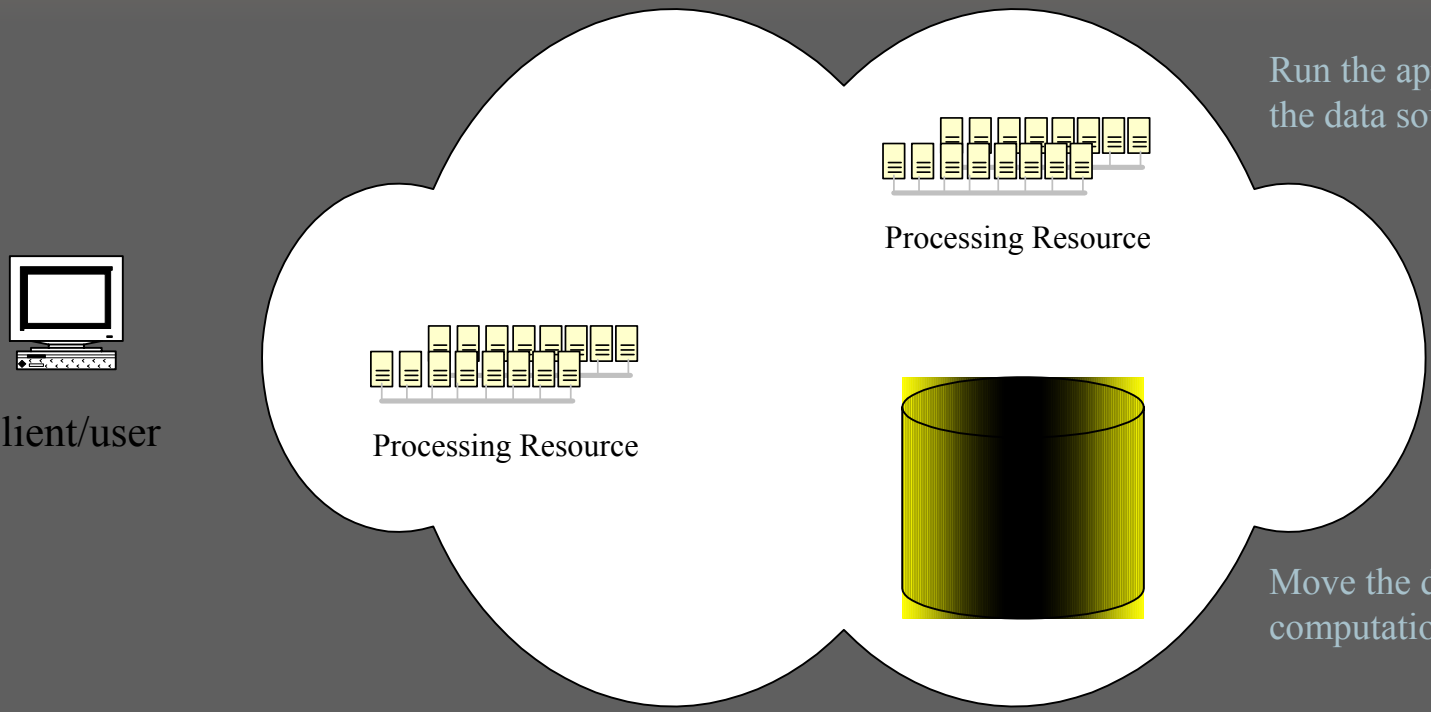
Static information

- type and number of machines
- operating system and version number
- range of compilers and software
- network bandwidth, topology

Dynamic information:

- current loading of the machines
- network loading and latency
- scheduling and charging strategies

Data – *the need to access to large amounts of data poses additional challenges*



We need to simultaneously identify processing and storage resources

Static + dynamic information => determine the optimal running configuration

Grid Canada

A partnership between CANARIE, the National Research Council, and C3.ca.

- Goal is to grid-enable compute, storage, and other resources across Canada
- Help these research communities deploy applications that can use the grid
- Build a Canadian Grid where researchers can access resources in an integrated and secure way.

- First meeting in November 2001 at the CANARIE Meeting (Toronto)
  - Identify interested communities (astronomy, climate, …)
  - Establish web site (www.gridcanada.ca)
- Second meeting in February 2002 at the Global Grid Forum (Toronto)
  - Establish short-term goal of a national Grid testbed
  - Discussion of long-term goals and plans
- Future meetings: Edmonton May 1st, HPCS 2002 in Moncton (June)

Grid-BC

BCNet Applications Advisory Committee – February 2002

Consider establishing Grid-BC:

- Educate the BC community on the potential of the Grid
  - talks, web information and link
- Identify interested groups – establish communication links
- Discuss local issues (e.g. the Grid depends on high speed network)
- Links with Grid Canada and international groups (e.g. Global Grid Forum)

Interested people should meet today

Grid BC
BC.NET

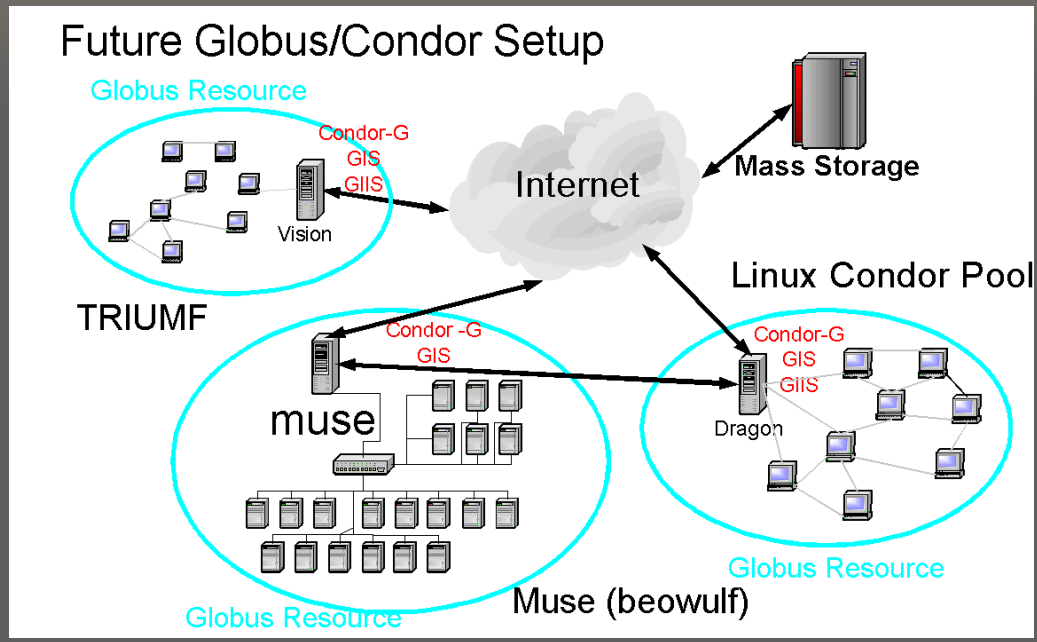
Summary

- The Grid is the next step after the world-wide Web
- It is made possible by the fantastic increase in network capacity
- It will allow scientists, industry to exploit the vast resource of computing capacity to address to new problems ("grand challenges")
- There is a large world-wide effort to develop the Grid
- Canadian researchers are starting to get involved
  - Grid Canada, Grid-BC and other organizations
- Exciting new area at the fore-front of computing

Linux testbed development at Victoria

MUSE cluster:

- 50 nodes + 3 TB disk
- Particle physics simulation and data analysis facility
- Cosmology groups (Beowulf)



Grid testbed

- User workstations as a private Globus resource
- Retired machines on MUSE have been setup as a public Globus resource
- Establish a national testbed for Linux
- Integrate new mass storage facility at Victoria into the testbed (Data Grid)