Grid Computing 101

Karthik Arunachalam
IT Professional
Dept. of Physics & Astronomy
University of Oklahoma

Outline

- The Internet analogy
- The Grid overview
- Grid computing essentials
- Virtual Organization
- ATLAS project
- Demonstration of a grid job

The Internet

- Data is shared
- Data could be stored in flat files, databases, generated dynamically on the fly. It could be in different formats, languages etc.
- Data is shared at will by the owners
- Policies are defined for data sharing (who, when, how, how much, etc.)
- Data is stored (distributed) across computers (web servers)
- Sharing of data is possible because servers and clients are glued together by the network and protocols
- Sharing is only one side of the story

The Internet…

- Shared data is accessed by clients
- Clients are removed from the complexity of the server side
- Clients with common interests could form virtual groups (social networking!)
- Server side could keep track of client's activity and account for it
- The internet is unreliable
- The internet as a utility (web services)

The Grid

- Resources are shared
- Resources could be CPUs, Storage, Sensors etc. (possibly anything that could be identified using an IP address)
- Resources are shared at will by the owners
- Policies are defined for resource sharing (who, when, how, how much etc.)
- Resources are housed (distributed) across organizations and individuals
- Sharing possible because resources are glued together by the network and the **Middleware**
- Based on open standards which continuously evolve (The Open Grid Forum)

The Grid…

- Shared resources are accessed by clients
- Clients are removed from the complexity of the Grid through middleware
- Clients with common interests could form Virtual Organizations (VO)
- Server side could keep track of client’s activity and account for it
- The grid is unreliable too ☹
- The **Grid** as a utility

Why Grid computing?

- Answer: The distributed computing model
  - Geographically dispersed communities
  - Talent is distributed
    - Harnessing local expertise and creativity
  - Financial constraints
    - Distributed funding model
  - Risk mitigation
    - No single point of failure; Agility to adapt to changes
  - Round the clock support
  - Keeps expertise where it is and avoids brain drain ☺
  - High speed networks and robust middleware make this possible

Who uses the grid?

- Primarily scientists and Researchers
- Various fields: Physics, Chemistry, Biology, Medicine, Meteorology and more
- For what: To solve complex problems
- No single centralized resource is powerful enough to model/simulate/run/solve these
- Virtual Organizations are at the core
  - Individuals with common goals/interests. Example: ATLAS, CMS, DOSAR etc.
- Somewhat removed from complexity of the grid using Middleware

User expectations

- Single sign-on (using grid proxy) authentication procedure
  - Sign-on once and use the grid for extended periods of time
- Methods to submit jobs, verify status, retrieve output, control jobs, view logs etc.
- Fast, reliable and secure data transfer, storage and retrieval using protocols that are easy to use and robust
- Reasonably quick completion of jobs
- Additional troubleshooting if they need more information
- Good accounting information
- Robust grid infrastructure that seamlessly provides them with the grid services they need anytime, anywhere

Virtual Organizations (VOs)

- What are VOs?: Groups of people who are distributed geographically, wanting to achieve a common goal
- How are they implemented? In software as a set of grid identities, organized into groups, with roles assigned to individuals
- VOs have an agreement with collaborating universities, institutes, and national labs to use their computing resources
- To be able to use the grid resources for a specific purpose, one should join a specific VO (Example: The ATLAS VO)
- How to join a VO? Obtain a grid certificate from a trusted Certificate Authority (CA),like DOE and request to become part of a particular VO (corresponding to the experiment which you are part of)

Virtual Organizations (VOs)…

- Grid certificates are like passports and becoming part of a VO is like obtaining a visa on your passport.
- Grid certificates identifies an individual uniquely using a Distinguished Name (DN)
- Once approved by the a representative of your experiment your Distinguished Name (DN) will be added to the list of DNs that are part of the VO
- Now you will be recognized by all collaborating labs and institutes as part of the VO and you will be allowed to use the grid resources, subject to policy guidelines
- Grid certificates have a limited validity time (usually 1 year) and they have to be renewed to stay valid
- Create a grid proxy (X509 certificate) on your localhost and use it as your single sign-on mechanism to submit jobs to the grid

The Ideal Grid

- Ideal Grid would function like a **utility**
- Similar to Electricity, internet, water, gas, telephone…
- Pay as you use - similar to any other utility
- Plug the client into the grid and harness the power of its resources
- Shouldn’t matter where the resource is, who maintains it, what type of hardware, software etc.
  - High speed networks, grid middleware make this possible
- Focus on the science rather than setting up, maintaining/operating the computing infrastructure behind it.
- The grid is NOT ideal yet – this means more work needs to be done

The Grid Architecture

- Describes the design of the grid
- Layered model
  - Hardware centric lower level layers
    - Network layer that connects the grid resources. High speed networks enable seamless sharing of resources and data.
    - Resource layer – the actual hardware like the computers, storage etc. that are connected to the network.
  - User centric upper level layers
    - Middleware that provides the essential software (brains) for the resource to be “Grid enabled”
    - The application layer containing applications that the grid users see and interact with.
- Helps end users to focus on their science and not worry about setting up the computing infrastructure

Grid Resources

- CPUs (from PCs to HPCs), Storage, Bandwidth, software…
  - Who provides these and why?
  - Common interests and goals – remember the Virtual Organization (VO)
- Dedicated resources
  - Completely dedicated to be used by a VO
- Opportunistic resources
  - Harvesting ideal computing cycles
  - You can donate your ideal cycles!
- Set of resources connected to form a specific grid (Eg: Open Science Grid). Individual grids connected to form one single global grid

Grid Resources…

- Sharing of resources is based on trust and policies
- The car pooling analogy
- VO plays an important role in ‘trust’ – become part of the VO
- Policies at grid and site level: Regarding usage, security, authentication, priorities, quota etc.
- Generally expect grid users to abide by policies. Policies could also be enforced.
- Authentication done using grid proxy certificate issued by a trusted authority.
- Usage of resources could be accounted for

Grid's glue - Middleware

- OK, I have the resource and want to share it – The question is how do I do it?
- The network is essential. But simply hooking the resource to the network doesn't enable sharing
- Grid Middleware provides the essential components for my resource to become part of the grid
- The grid software contains the grid middleware. For example the OSG software stack contains the Globus toolkit
- Made up of software programs containing hundreds of thousands of lines of code
- Installing the grid software is the first step toward making your resource grid enabled

The ATLAS project

- ATLAS – Particle Physics Experiment at Large Hadron Collider (LHC) at CERN, Geneva, Switzerland.
- LHC is the largest scientific instrument on the planet!
  - Scientists trying to re-create the moment after the big bang happened
  - ATLAS detector will observe/collect the collision data to be analyzed for new discoveries
  - Origin of mass, discovery of new particles, extra dimensions of space, microscopic black holes etc.
- Late 2009 to early 2010 – startup of LHC and first event collisions expected
- 10 to 11 months of intensive data collection expected
- Experiment is expected to last for 15 years

The ATLAS project…

- LHC will produce 15 Petabytes (15 million GBs) of data annually.
- ATLAS designed to observe one billion proton collisions per second – combined data volume of 60 million megabytes per second
- Lots of junk data. Only some interesting events.
- Atlas Trigger system helps filter interesting events for analysis
- ATLAS will collect only fraction of all the data produced – around 1 petabyte (1 million gigabytes) per year
- This data needs to be accessed and analyzed by Physicists

Storing & Analyzing ATLAS data

- 1 petabyte of data per year to be analyzed
- Enormous computing power, storage and data transfer rates needed
- No single facility, organization or funding source capable of meeting these challenges
- One of the largest collaborative efforts attempted in physical science
- Thousands of physicists and from 37 countries, more than 169 universities & laboratories involved

Storing & Analyzing ATLAS data…

- Grid computing to the rescue
- Computing power and storage distributed geographically across Universities & laboratories – all connected with high speed networks
- Physicists are collaborating together as the ATLAS Virtual Organization (VO)!
- To become part of ATLAS: Obtain a grid certificate and apply to become a member of ATLAS VO
- ATLAS jobs are embarrassingly parallel i.e. each sub-calculation is independent of all the other calculations – hence suitable for High Throughput Computing
- Hierarchical model of data distribution
  - Single Tier 0 at CERN
  - 10 Tier 1 centers spread across the globe
  - Several Tier 2 centers under each Tier 1

OU's ATLAS Tier2 center - Hardware

- OU's OCHEP tier2 cluster is part of the US-SWT2 center (along with UTA)
- 260 core Intel(R) Xeon(R) CPU E5345 @ 2.33GHz
- 2 GB of RAM per core (16 GB per node)
- 12 TB of storage (to be increased to 100 TB soon)
- 5 head nodes (1 CE + 1 SE + other management nodes)
- 10 Gbps network connection from head nodes
- Connected to NLR via OneNet

OU's ATLAS Tier2 center - software

- US ATLAS is part of the Open Science Grid (OSG)
- OSG (0.8) software stack is installed as the grid software on the OCHEP cluster head nodes. This provides the grid's middleware glue. ROCKs is used as cluster software
- Condor is used as the local batch system for queuing, scheduling, prioritizing, monitoring and managing jobs at the site level
- The Compute Element is the gatekeeper for the cluster. This is where the jobs get submitted to the cluster

OU's ATLAS Tier2 center – software…

- ATLAS jobs are managed through the Panda (Production and Distributed Analysis) distributed software system
- Distributed Data Management (DDM) system (DQ2 software) is used to manage and distribute data
- Network performance is tested and tuned continuously using the PerfSonar software toolkit from Internet2
- Monitoring and managing of the cluster has been completely automated using a collection of scripts that could provide alerts and take actions
- Opportunistic resources: OU Condor pool (> 700 lab PCs), OSCER's Sooner HPC cluster

Demonstration

- Basics
  - Initiate the grid proxy
  - Running a job on the grid
    - Run the job
  - Submitting a job on the grid
    - Submit the job
    - Check the status
    - Retrieve the output
  - Condor batch system information

References

- Acknowledgements & useful links
  - http://www.cern.ch
  - http://www.atlas.ch
  - http://www.isgtw.org
  - https://pki1.doegrids.org/ca/
  - http://www.opensciencegrid.org/
  - http://www.cs.wisc.edu/condor/overview
  - http://www.nhn.ou.edu/~atlas/atlas/grid/