Grid Research Centre

Rob Simmonds
Dept. Computer Science,
University of Calgary

GRID
RESEARCH CENTRE

Grid Research Centre

- Location

  Department of Computer Science
  ICT Building,
  University of Calgary

- Main sponsors
  - Hewlett Packard
  - Alberta Innovation and Science
  - Western Economic Diversification

Distributed Systems Lab

- TeleSim group
  - Parallel discrete event simulation
  - Network simulation and emulation
  - Network performance evaluation
  - Wireless network protocol development
  - Modelling languages
- Grid Research Centre
  - Grid services evaluation for WestGrid
  - Grid OS project
  - Automated grid monitoring
  - Scientific data management

GRC People

- Rob Simmonds
- Martin Arlitt (HP Labs)
- Cameron Kiddle
- Diwakar Krishnamurthy
- Roger Curry
- Mark Fox
- Andrey Mirtchovski
- Phil Rizk
- Id Adewale
- Abhishek Gaurav
- Nayden Markatchev

GRID
RESEARCH CENTRE

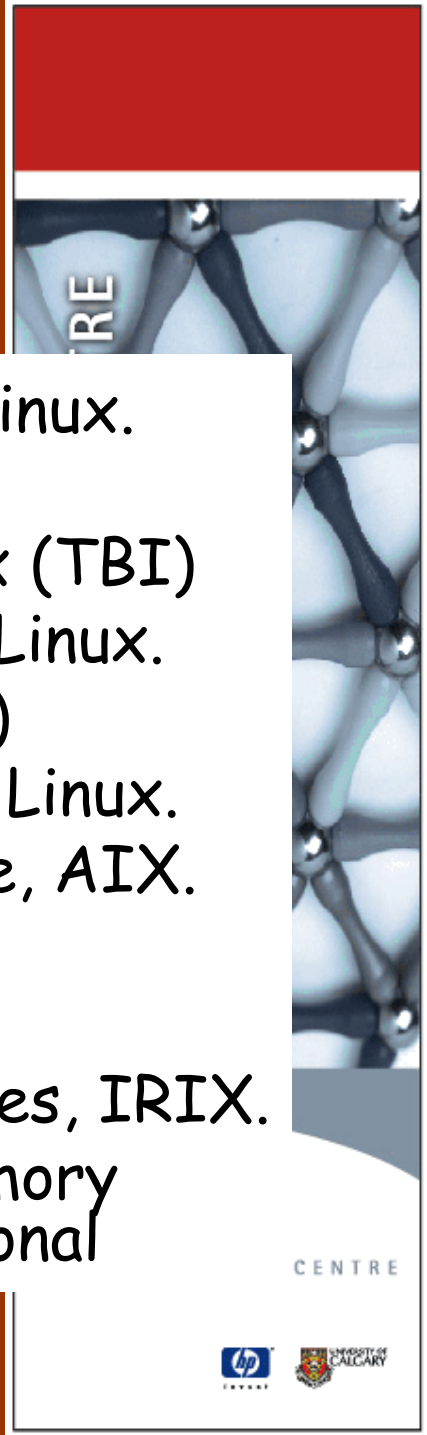
WestGrid

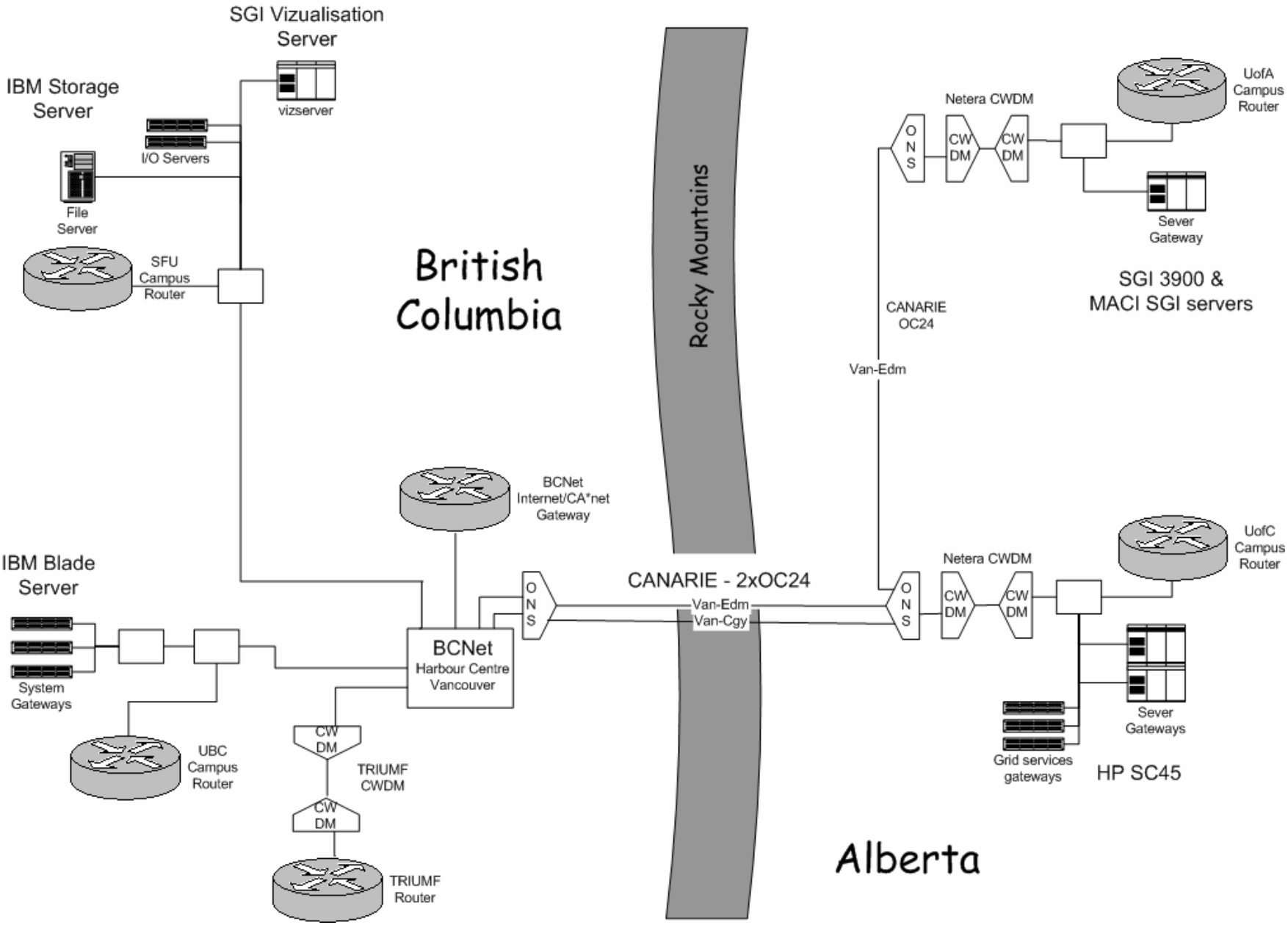
- WestGrid provides grid-enabled HPC and collab/viz resources to Canadian researchers
- Close links between GRC and WestGrid
  - WestGrid CTO is co-PI of GRC
  - GRC recommends grid technologies to WestGrid

GRID
RESEARCH CENTRE

WestGrid Resources

- Cray XD1 – 6 FPGA and 12 Opteron procs, Linux.
- HP SC45 – 144 Alpha procs, Tru64.
- HP XC Infiniband cluster – 256 procs, Linux (TBI)
- IBM blade centre – 1680 Xeon processors, Linux.
- 2x IBM p695 – 64 Power5 procs, AIX (TBI)
- IBM blade centre – 54 PowerPC-970 procs, Linux.
- IBM storage node – 35TB disk, 200 TB tape, AIX.
- Parasol Genematcher-II – 28800 procs.
- SGI Origin – 256 MIPS procs, IRIX.
- SGI Onyx – 20 MIPS procs, 12 graphics pipes, IRIX.
- + 4 older clusters, 6 older large shared memory systems, service provision nodes and additional gateway nodes
- + Access grid facilities at all sites




SGI Vizualisation Server
vizserver
IBM Storage Server
I/O Servers
File Server
SFU Campus Router
British Columbia
Rocky Mountains
Netera CWDM
ONS
CWDM
CWDM
UofA Campus Router
Sever Gateway
SGI 3900 & MACI SGI servers
CANARIE OC24
Van-Edm
BCNet Internet/CA*net Gateway
IBM Blade Server
System Gateways
BCNet Harbour Centre Vancouver
CANARIE - 2xOC24
Van-Edm
Van-Cgy
Netera CWDM
UofC Campus Router
Sever Gateways
Grid services gateways
HP SC45
UBC Campus Router
CWDM
TRIUMF CWDM
CWDM
TRIUMF Router
Alberta

Grid Monitoring

GRID
RESEARCH CENTRE

Grid Monitoring

- Joint project with HP Labs
- Researching
  - What information is required to understand operation of grid environments?
  - How to collect this information?
  - How to store this information?
  - How to display this information?
- Creating a meta-model for monitoring components

GRID
RESEARCH CENTRE

Cluster monitoring

- Current cluster monitoring tools show lots of information
  - Too much information can make it hard to see what is important
  - Correlation done by eye
  - Information not easily exchanged between tools

GRID
RESEARCH CENTRE

Grid complications

- Heterogeneous environment
- Mix of security policies and network/firewall configurations
- Huge number of available resources

GRID
RESEARCH CENTRE

System description

- To automate deployment of monitoring tools, structural and configuration information needs to be available
  - Need to know dependencies to quickly locate problems, or to predict future problems
  - Need to discover this information and track changes

GRID
RESEARCH CENTRE

Current representations

- Have been examining existing ways representing information
  - Common Information Model (CIM)
    - DMTF standard
    - Used by many equipment vendors
    - Adopted by GGF
  - Grid Laboratory Uniform Environment (GLUE)
    - Used by Globus MDS, so widely deployed in grid environments
- Neither of these classify data items clearly

Data model

- Currently working with relational system description
- Separate system used to archive changes and provide historical information
- Aim to understand how to apply what we are learning to existing standards

GRID
RESEARCH CENTRE

Statistics

- Need to combining information on different time scales
- Correlating data sources to locate related issues
  - Jobs with property X always fail
  - Jobs with prop. X fail on node Z
  - Resource R unresponsive when job with property P is run

GRID
RESEARCH CENTRE

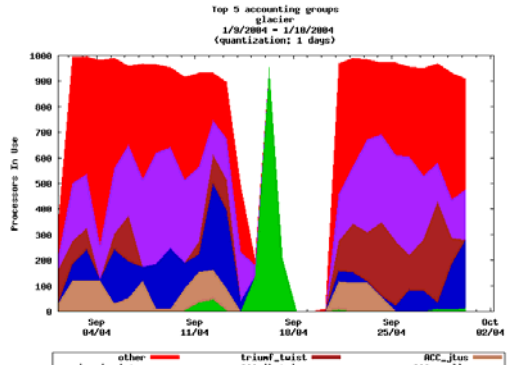
Use cases

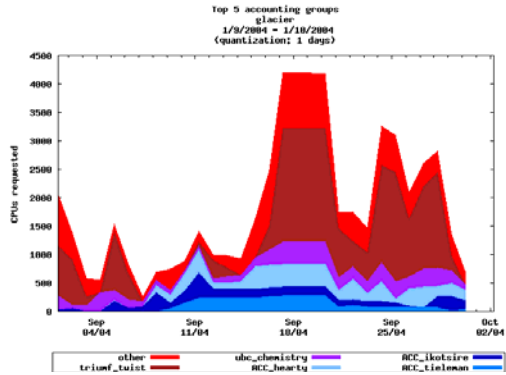
- Use cases developed to find common monitoring requirements
  - Why is node idle?
  - Why did job fail?
  - Why is job not starting?
  - Where could job run?
  - What type of job could run now?
  - Why is job pattern not matching what is defined by policy
  - … and that is just a few job related issues.

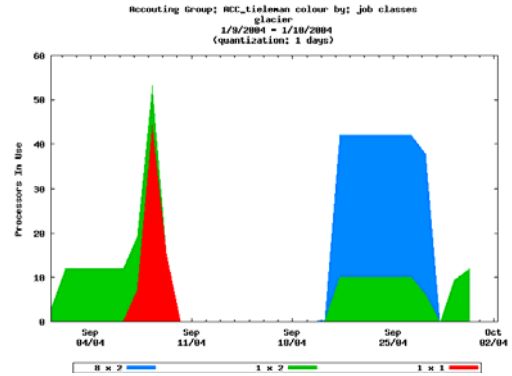
GRID
RESEARCH CENTRE

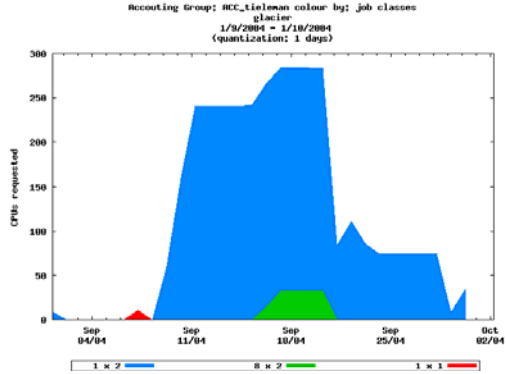
Data analysis

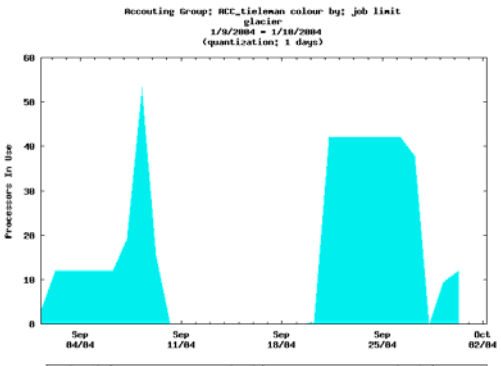
- Find minimum information to explain system behaviour
- Need to exploit data-mining techniques

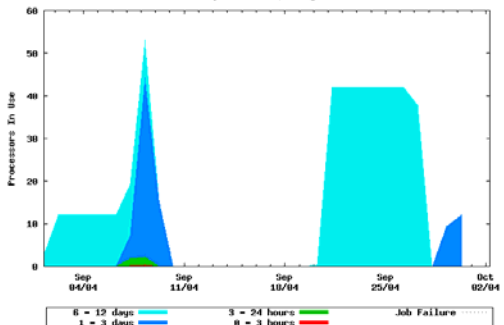


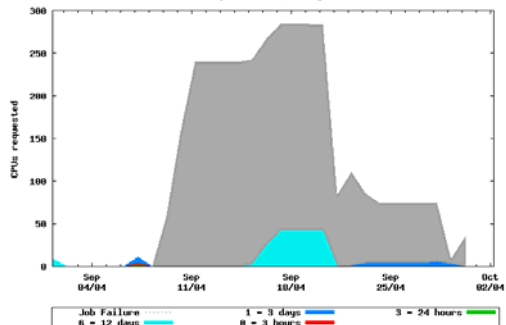












Data Management

GRID
RESEARCH CENTRE

Data Management

- Creating Proactive Data Management Service (PDMS)
- Primarily focused on supporting big science problems with large amounts of data
- Supported by CANARIE
- Working with two groups of scientists to determine needs

GRID
RESEARCH CENTRE

PDMS

- Provides a data management and replication service
- Aim to provide interoperability with user-controlled lightpath networks
- Being written using new GGF WSRF standards
- Uses Globus RLS and MCS to enable interoperability with execution planning tools

GRID
RESEARCH CENTRE

PDMS (II)

- Allows users to define schema to uniquely identify each data item in their collections
- Enables data in specified views to be replicated or downloaded
- Designed to interoperate with workflow managers performing co-allocation of resources

GRID
RESEARCH CENTRE

Research Issues

- Considering many data replication and caching issues
- Particular interest in optimizing for network conditions
  - Take advantage of schedulable high bandwidth links
  - Dealing with networks with bad segments

GRID
RESEARCH CENTRE

Collaborations

- Grid monitoring project
  - HP Labs
- Data management project
  - University of Bath (UK), Carlton University
- Distributed OS for grid
  - LANL
- Simulation/Emulation projects
  - Nortel, Siemens, Telus Advanced Communications, Axia Supernet, Electronic Arts, Netera and others

GRID
RESEARCH CENTRE

Summary

- Grid Research Centre based at University of Calgary
- Working on some key problems faced in grid environments
  - How to monitor systems and jobs
  - How to manage large data sets
- Have established collaborations and aim to contribute towards standards

GRID
RESEARCH CENTRE

http://grid.ucalgary.ca