Grids and Security

Ian Neilson
Grid Deployment Group
CERN
• Background
  • Grids
  • Grid Projects
• Some Technical Aspects
  • The three or four A’s
• Some Operational Aspects
  • Security Coordination and Incident Response
• Other stuff
Grids

LCG – LHC Computing Grid

EGEE – Enabling Grids for e-Science in Europe

"[Grids] enable the sharing, exchange, discovery, and aggregation of resources distributed across multiple administrative domains ..."

Sun Microsystems

"A VO is a participating organization in a grid to which grid users must be registered and authenticated in order to gain access to the grid’s resources. A VO must establish resource-usage agreements with grid resource providers. Members of a VO may come from many different home institutions, may have in common only a general interest or goal (e.g., CMS physics analysis), and may communicate and coordinate their work solely through information technology (hence the term virtual). An organization like an HEP experiment can be regarded as one VO. A more comprehensive definition can be found at ...." Glossary - http://www.opensciencegrid.org/home/terminology.html
LCG in one slide

- Computing fabric for the Large Hadron Collider experiments
- Operating 2007+
- 95 sites
- 31 countries
- 9000 CPUs
- 6 TB storage

http://goc.grid-support.ac.uk/gppmonWorld/gppmon_maps/lcg2.html
EGEE in one slide

- 70 institutions in 28 countries, federated in regional clusters

- 32MEUR for first 2 years (plans for another 2 years)

- Deployment and reengineering project

- 50% operations & support, 25% training & appl. support, 25% reengineering
Scaling up, surely...

<table>
<thead>
<tr>
<th>Year</th>
<th>Applications</th>
<th>Resource Centers</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>High Energy Physics, Bioinformatics</td>
<td>10 Sites, 1K CPUs, 10 TB</td>
</tr>
<tr>
<td>2004</td>
<td>High Energy Physics, Bioinformatics, Chemistry, Astronomy</td>
<td>20 Sites, 3K CPUs, 50 TB</td>
</tr>
<tr>
<td>2005</td>
<td>Chemistry, Astronomy, Earth Observation</td>
<td>50 Sites, 20K CPUs, 500 TB</td>
</tr>
<tr>
<td>2006</td>
<td>Chemistry, Astronomy, Earth Observation, Industry, Biodiversity</td>
<td>100 Sites (Global), 50K CPUs, 1 PB</td>
</tr>
<tr>
<td>2007</td>
<td>Chemistry, Astronomy, Earth Observation, Industry, Biodiversity, Climate Modeling, Nanotechnology</td>
<td>100 Sites (Global), 50K CPUs, 1 PB</td>
</tr>
</tbody>
</table>
LCG/EGEE Security environment

- The players

  Personal data
  Roles
  Usage patterns
  ...

  Users

  Grid

  Sites

  Resources
  Availability
  Accountability
  ...

  VOs

  Experiment data
  Access patterns
  Membership
  ...

The Risks

- Top risks from Security Risk Analysis
  - Launch attacks on other sites
    - Large distributed farms of machines
  - Illegal or inappropriate distribution or sharing of data
    - Massive distributed storage capacity
  - Disruption by exploit of security holes
    - Complex, heterogeneous and dynamic environment
  - Damage caused by viruses, worms etc.
    - Highly connected and novel infrastructure
Policy – the Joint Security Group

Certification Authorities

Usage Rules

Incident Response

Audit Requirements

Security & Availability Policy

GOC Guides

User Registration

Application Development & Network Admin Guide

http://cern.ch/proj-lcg-security/documents.html
The goal

Authorization Policy Architecture

- Delegation
- PKI Identity
- User's behalf
- PKI/Kerberos Identity
- Translation Service
- Site Policy
- Resource Attributes
- Site Policy
- Resource Attributes
- PKI/Kerberos Identity
- Local Site
- Kerberos Identity
- Policy and attributes
- Allow or Deny
- Site/Resource Owner
- Authorization Service/PDP
- Policy Enforcement Point
- Standardize
- Other Stakeholders
Authentication Infrastructure

- Users and Services own long-lived (1yr) credentials
  - Digital certificates (X.509 PKI)
  - European Grid Policy Management Authority
    - “... is a body to establish requirements and best practices for grid identity providers to enable a common trust domain applicable to authentication of end-entities in inter-organisational access to distributed resources. ...”
  - www.eugridpma.org covers EU (+ USA + Asia)
  - Shared infrastructure between all EU FP6 Grid projects (and others)
  - Establish and “audit” common minimum operational requirement
Authentication Issues

- Do trust mechanisms scale up?
  - Lots of CAs.....
- Can users keep the secret?
- "On-line" certification authorities & Certificate Stores
  - Kerberized CA
  - MyProxy certificate store
  - Virtual SmartCard
AAA : Authorization 1: VO-based

- User Registers
  - Accepts Usage Rules
  - Provides personal/contact data
  - Request to join VO
    - VO managers add to VO database
  - Certificate Identity (DN) captured
- Submits job…
  - Creates short-lived proxy using long-lived certificate
  - Proxy ‘travels’ with the job
- …jobs arrive at resource
  - Checks certificate validity
    - Trusted CAs and revocation lists
  - Checks user is authorized – ‘whitelist’
    - Downloaded from Registration/VO database
  - Maps certificate DN to a local account
  - Runs job
- Currently in use by Nordugrid ARC, LCG/EGEE
AAA : Authorization 2 : VO/Role-based

- User Registers
  - .... As above but may be assigned a role by VO
- Creates proxy
  - Contacts VO server to sign user’s attributes into proxy
- Submits jobs
  - Proxy ‘travels’ with the job
- Resources authorize access
  - Checks certificate validity
    - Trusted CAs and revocation lists
  - Checks user authorization – from attributes in the proxy
    - Allows for one user, multiple VOs, and multiple roles
  - Maps to a local account
  - Runs job
- Being deployed by LCG/EGEE
AAA: Accounting

- Accountability
  - Little thought given here 😞
  - Retention of logs at sites
    - Dispersed information
    - No standard formatting
    - ‘Debug’ information
  - Usual concerns of privacy

- Billing
  - By user, by VO, per site, aggregated?
    - Need to sort out local from grid usage
So, ... but scaling up securely?

- Diverse audience
  - Operations, Middleware, Applications, End users
  - Regional differences
- Impossible and Contradicting requirements
  - Traceability and Anonymity tradeoffs
  - Performance and Security tradeoffs
- Lucky nothing has happened so far
  - Grid is on the hackers radar
- User certificates and keys are spread all over
  - Private key file scrambled - password?
  - File protections?
  - Similar to the SSH key problem
- Proxy certs way too unrestricted
- Too many services operate as root
- VERY hard to audit what’s going on.
Global Grid Forum activities

- Information exchange
- Grid Security Infrastructure (GSI)
- Proxy certificates (now RFC3820)
- CA operations recommendations
- Site operations recommendations
- Authorization
- Firewall issues
- Application domain interests
- ...
- Workshop on Operational Security
  - [http://grid.ncsa.uiuc.edu/ggf12-sec-wkshp/](http://grid.ncsa.uiuc.edu/ggf12-sec-wkshp/)
Security Coordination Activities in LCG/EGEE

- Updates to Incident Response Agreement
  - In collaboration with Open Science Grid
    - “To guide the development and maintenance of a common capability for handling and response to cyber security incidents on Grids.”
  - The capability will be established through
    - common policies and processes, organizational structures,
    - cross-organizational relationships,
    - common communications methods, and
    - a modicum of centrally-provided services and processes.
  - Managed contacts lists
  - Links with policy, development and deployment activities
Security Coordination Activities in LCG/EGEE

- Security Service Challenges
  - Exercise response procedures in controlled manner
    - Non-intrusive
    - Compute resource usage trace to owner
      - E.g. Run a job to send an email
    - Storage resource trace to owner
      - E.g. Run a job to store a file
    - Disruptive
      - Disrupt a service and map the effects on the service and grid
Operational Security Coordination Team

- EGEE operational channels still being established.
- Does not have central authority over sites
Other issues

• Software and standards immaturity
  • “Production quality” by academia standards
  • 80/20 rule often applied
  • Standards either far ahead or far behind

• Firewalls
  • Need to regard network connectivity as another resource
Thank You

• Thanks to UK PPARC for my funding in LCG
• Acknowledgement to
  • Olle Mullmo who took several slides stolen from older presentations made by Steve Tuecke, Von Welch, Frank Siebenlist, and others
  • Dave Kelsey