Authorization for LCG (VOMS)

David Kelsey, CCLRC/RAL, UK
d.p.kelsey@rl.ac.uk

LCG GDB Meeting,
CERN, 8 February 2006

Overview

- **Report on LCG Authorization workshop (Sep 2005)**
  - Just one, not a series
- **LHC experiment requirements (for AuthZ)**
- **What happened since then**
- **Some other comments**

note

- **not many details**
  - The other talks will (hopefully!) tell the whole story

History

A quote on VOMS…

- **All application communities require**
  - Groups/roles in VOMS and file ACL's
    - The highest priority
- **Ready, being tested (INFN, NIKHEF, CERN…)**
- **But dropped off end of list for 2.0 release!**
- **High priority item**
  - integration immediately after 2.0

(D. Kelsey in Security Summary, EU DataGrid project conference – Barcelona 15 May 2003)

LCG Authorization Workshop

- 13th September 2005 (CERN)
- See *http://agenda.cern.ch/fullAgenda.php?ida=a054503*
  - *A workshop to plan use of current Grid Authorization technology by the LHC experiments over the next year (or two), aiming in particular for Service Challenge 4 (April 2006). It is essential to achieve interoperability between the different Grids providing resources to the LHC experiments.*
- Attended by ~40 people
  - Experiments, Security experts/developers, deployment, …
- Date was fixed next to EGEE Middleware Security Group meeting to encourage US participation
  - To tackle interoperability issues
  - But bad date for LHCb (apologies again!)

Workshop agenda

- **Experiment plans/ requirements**
  - All presented except LHCb (clash with collaboration meeting)
- **Current middleware components**
  - EGEE VOMS, LCAS, LCMAPS
  - OSG PRIMA, GUMS
  - LCG Grid services (Data Management, Workload Management)
- **VOMS deployment plans (CERN for LHC expts)**
- **Discussion on groups, roles and capabilities**
- **Future plans**
  - G-PBox
  - EGEE
  - Globus

Experiment requirements

- **The general need for VOMS and fine-grained access control has been known for a long time**
- **Presented in many places (not just the AuthZ ws)**
  - EDG, EGEE
  - LCG Baseline services report
  - Computing TDRs
  - Ongoing discussions in EGEE TCG
- **ALICE**
  - Presented **Efficient data access authorization using catalogue based authorization tokens**
    - GSI, xrootd
  - Access policy set centrally – in catalogue
  - roles/groups 5 to 10 during next 12 months
- **ATLAS (Alessandro de Salvo) & CMS (Stefano Belforte)**
  - See next slides

Groups and roles [1]

- Needed for
  - Resource allocation
  - Data and space management

- Current VO implementation
  - 2 groups
    - “lcg1” [ATLAS users]
    - “usatlas” [OSG users only]
      - NB: all lcg1 users are allowed on USATLAS sites
  - 4 roles (currently implemented as LDAP groups)
    - “admin” [the VO administrators]
    - “lcgadmin” [the LCG VO software managers or SGMs]
    - “usprod” [the production managers for OSG]
    - “ussoft” [the OSG software managers]

- Current implementation in OSG (VOMS based)
  - 4 groups/roles
    - /atlas/usatlas/Role=production:
      data production coordinators
    - /atlas/usatlas/Role=software
      people that need to install remove software and debug applications
    - /atlas/usatlas
      USATLAS users
    - /atlas/lcg1
      the rest of ATLAS

Groups and roles [2]

- *Migration to the VOMS implementation*
  - *Migration of all the current groups and roles to the new system*
  - *Introduction of a new set of groups and roles*
    - *For Data Management*
    - *For Workload Management*

Workload Management groups and roles

- **Workload Management roles (3)**
  - *Grid software administrator*
    - *Responsible of the installation of the experiment software.*
  - *Production manager*
    - *Production user, will have higher priority than normal users for official group productions and will be able to place files in commonly accessible areas*
  - *User*
    - *Any normal user*

- **Workload Management groups (~20)**
  - *Physics and Combined Performance working groups*
    - *One group for each Physics Working Group and Combined Performance Group*
  - *Testing, validation and central production activities groups*

Database and Data Management groups and roles

- **Database access roles (5)**
  - **Administrator**
    - *Administrators manage the installation of database servers and give access rights to other users.*
  - **Developer**
    - *Database applications developers for particular software domains (full access right to particular databases)*
  - **Editor**
    - *People having UPDATE or DELETE rights*
  - **Writer**
    - *People having INSERT or SELECT rights*
  - **Reader**
    - *People having only the SELECT privileges*
- **Data Management groups and roles**
  - *The same groups and roles as for the Workload Management*

The need

- We take authorization to mean :
  Policies and Resource Management
- I.e. we (CMS) needs those and expect Grid to provide them using authorization tools (and other tools)
- Policies:
  - Who can use given resources (disk, CPU, network)
  - Who can decide the former
- Resource Management
  - How much resources are allocated to different activities

CMS organization

- CMS is a large collaboration, unmanageable as such
- Will be structured with groups and subgroups
- Groups by physics interest
  - Higgs, Higgs → leptons, Higgs → 4muons, H → 4mu trigger …
- Groups by mundane affinity
  - By site: The folks at CMS-Tier2 in Rome
  - By resource: The people using an Analysis Facility next to a T1
  - By funding: The INFN physicists
- Groups by service tasks
  - Calibration → Detector → Sub-detector → specific variable
  - Reconstruction
  - Monte Carlo production

- Policies and resource management need to match this granularity, possibly down to individual physicists

Wider scope

- The work of a physics group may require changes beyond the resource they "directly control":
  - Request for urgent MonteCarlo samples
  - Request for dedicated (re)reconstruction of data hosted at Tier1's
  - Replica of data from Tier1 (tape) to Tier2
  - Etc.
- Times 10 (?) major physics groups
  - Policies at Tier1 change very often !

Yes, VOMS

- So it is pretty clear we need VOMS roles and groups

- We need them to match the granularity of the CMS VO
  - Which will change
  - Which we do not know exactly
- Imagine two thousand physicists
  - Usually difficult to work by gentlemans agreement with groups of more then a handful of people
  - So policies and resource management may have to reach down to very small groups

- Yes, will try to limit the groups, but can not commit to any number of them, especially in the long term

No, VOMS is not enough

- Then we need the grid tools to be able to use VOMS roles and groups to control allocation of resources

- Mapping VOMS groups to unix groups ?
- Using VOMS directly in ACL's ?
- Using VOMS group to select CE ?
- Using VOMS groups to reorder global task queue ?

- But we need this "now"

- CMS goal for summer 2006: allocate resources at Tier1's separately for Monte-Carlo and Analysis
  - Should we be ashamed of asking so little ?

Discussions during workshop

- **Cannot today implement large numbers of groups/roles**
  - As need a unix gid for every combination (too many)
- **Agreed to limit the number**
  - For Jan 2006 – the aim is …
    - (Per VO) 2 to 4 groups and 2 to 4 roles (with a max sum of 6)
- **Useful to use similar names (e.g. role lcgadmin)**
- **Agreements:**
  - Create a mail list to continue discussion
    - project-lcg-authz at cern.ch (with archive)
  - Maarten Litmath to write document for experiments to consider

Since the workshop

- **Debate on VOMS "capabilities"**
  - Attractive for batch system priorities, but
  - VOMRS can not handle these (agreed before not to use them)
  - I think these have now gone away (yes?)
- **Proposal for Groups/Roles**
  - Prepared by Maarten Litmaath and circulated in October
  - Discussion followed (particularly CMS)
- **Should batch priorities be handled by a role?**
  - E.g. Role = prio-high
  - Experiments want central control
    - Sites need do no config changes
    - But difficult to manage
- **LCG 2.7 contains new default groups.conf**
  - With standard roles (lcgadmin and production)
  - 4 CMS physics groups

Other comments

- **Users will need to be registered in the CERN HR database**
  - Will this cause problems? Are they fore-warned?
- **Accounting**
  - Required at group/role level
  - Is this possible?
- **EGEE, G-PBox, GGF, OGSA-AuthZ, SAML, all being worked on – everything solved in the future**
  - But the future is always a long time ahead!
- **SC4 is a chance to test reasonably simple implementations**
  - Work on any interoperability issues
  - Lets get it working and learn!