Portals and Credentials

David Groep

Physics Data Processing group NIKHEF

- ## Portals all around
  - – EGEE TCG Portal working group
  - – Dutch BiG Grid portals

- Started in 2007 in order to …

 *"propose "best-practice" rules for the access of portals to the grid. […] To do so, a portal responsible should […] then to be able to register this portal certificate to a VO allowed on the grid. Once the portal have been accepted into the concerned VO, it should be able to store and access data inside the VO area, and also to run job on site accepting this VO. […]"*

- Lead by Christophe Blanchet with others
- Identified a set of 5 portal scenarios, ranging from simple queries to complex workflow execution.

Example: Christophe's GPS@ portal

- ## BLAST searches on the grid

- ## Provide Biologists with an usual Web interface: NPS@
    - NPS@ Web portal online since 1998
    - 46 tools & 12 updated databases
    - + 9,000,000 jobs & 5,000 jobs/day
- ## Ease the access to updated databases and algorithms.
    - Protein databases are stored on the grid storage as flat files, encrypted if needed.
    - Wrapping legacy bioinformatics applications
    - Transparent remote access through local file-system accesses
- ## Display results in graphical Web interface.

- ## Has to complete with 'free' portals in the genomics community
- ## Virtually anonymous access

- KNMI/SRON/SARA/Nikhef effort

- Processing Sciamachy data
  - Predefined workflow
  - Large input data sets
  - Access limited to identified researchers
  - Raw data is actually protected as well

- Portal controls access through GUI
  - User identify use username/password
  - NADC processing created the workflow
  - Upload output data to dedicated system

- Jobs submitted to the grid identify themselves as a Robot
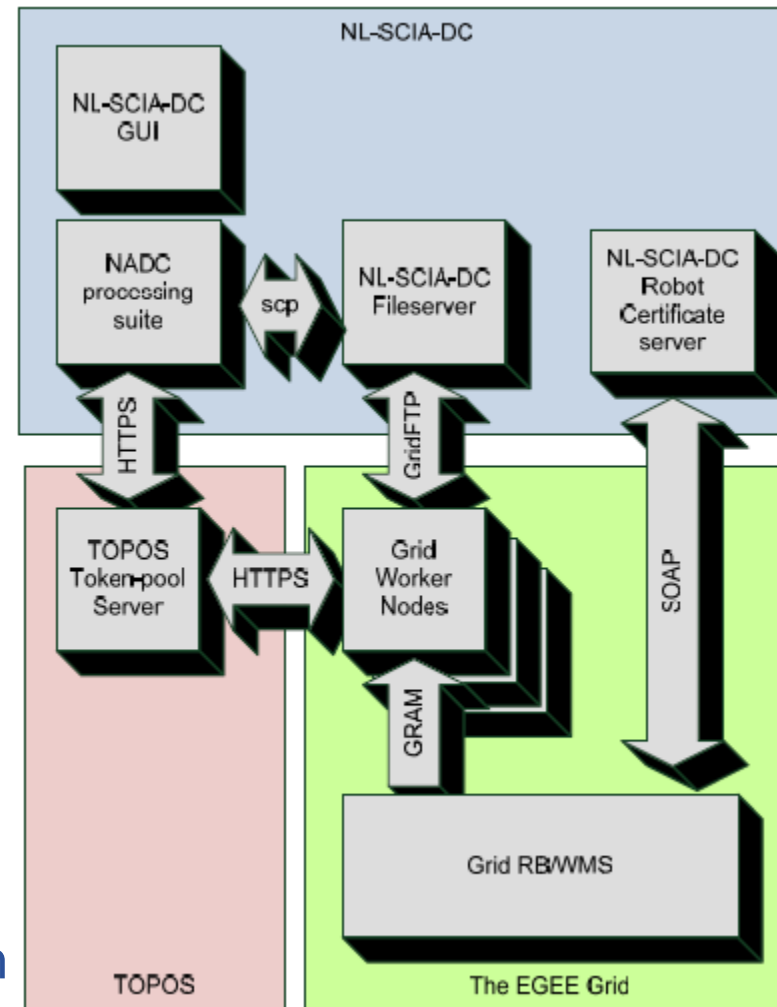
A Robot *What*? A Robot Certificate:

- 'Automated Client' (see the old OGF document)
  - Identified as such in the CN
    "Robot: <what-i-am>" plus name of a human responsible

- With private key held on a secured hardware device

- As per boiler-plate text from the UK, NL and IT CP/CPSs

- Questions to ask

- **Categories of users:**
    - 1) End user is "anonymous" (i.e. neither grid credential nor registration).
    - 2) Pseudo-anonymous users (portal registration)
    - 3) Identified users w/o grid credentials (maybe other certificates).
    - 4) Identified users with grid credentials; portal credentials used.
    - 5) End user has his/her own grid credentials.

From: Christophe Blanchet and TCG Portal WG

- More Questions

- Per use case
    - Is Grid job well-known (canned job) or free?
    - Is Grid job executed under Portal ID or User ID?
    - Is Grid data well-known (canned data) or free?
    - Is Grid data owned by Portal ID or User ID?
- For User ID
    - Is user well identified (photo ID)?
    - Or just username/email-address and password?
    - Does user ID need to be transferred to the Grid
- Need to address accounting issues (as discussed earlier)
    - And data privacy issues

From: Date Kelsey, TCG Portal WG

- Classify by auth method or function? BiG Grid tried function:

1. The Web User invokes functionality on the Portal where jobs submitted to the Grid use executable code that is provided by the Portal to the Grid as part of the job submission process. All parameters and input data are defined exclusively by the Portal and cannot be influenced by the user.

2. The Web User invokes functionality on the Portal where jobs submitted to the Grid use executable code that is provided by the Portal to the Grid as part of the job submission process. The Web User may only provide run-time parameter settings from an enumerable and limitative set, and may select data files from a enumerable repository of data files that are pre-vetted for use by the Portal.

3. The Web User invokes functionality on the Portal where jobs submitted to the Grid use executable code that is provided by the Portal to the Grid as part of the job submission process. The Web User may provide run-time parameter settings from an enumerable and limitative set, and may provide non-validated input data to the executable code.

4. The Web User invokes functionality on the Portal where jobs submitted to the Grid use executable code that is provided by the Web User. Whether this code is passed through unmodified by the Portal and is submitted to the Grid as-is, or whether this code is inspected and analysed on the Portal does not change the classification of this Portal.

- ## Common elements
  - Should fit in the JSPG "Security and Availability Policy"

In addition to all other Policies, the following conditions apply to all Portals:

1. The Portal, the VO to which the Portal is associated, the Portal manager are all individually and collectively responsible and accountable for all interactions with the Grid, unless a credential of a Strongly Identified Web User is used to interact with the Grid.
2. The Portal must be capable of limiting the job submission rate.
3. The Portal must keep audit log to associate any interactions with the Grid as defined the audit policy, within the scope of the service.
4. The Portal manager and operators must assist in security incident investigations by the Grid and by any Resource Provider who can justifiably claim to provide or to have provided resources to the Portal.
5. Where relevant, private keys associated with (proxy) certificates must not be transferred across a network, not even in encrypted form. Other re-useable private data should not be transferred across a network, and if transferred must be encrypted when sent across a network.
6. A Portal must not store or obtain long-lived reusable authentication information for its end-users that is valid for more than 1 Ms.

2.1 Class-1 "Web Rendering Automaton" Portals

By registering a Class-1 Portal in a Virtual Organisation, you agree to the conditions laid down in this section and documents references therein.

1. The Portal may offer services to all Web Users.
2. The Portal must use a Robot Certificate to interact with the Grid.
3. No data may be stored on the Grid as a result of any action by a Web User, except for transient data stored in designated scratch areas on the computational resources provided to the running jobs.
4. Maximum submission rate must be specifically agreed between Portal and Grid
5. The Portal must keep enough information to associate any interactions with the Grid with a particular Internet address and (tcp) port used by the requester.

for example: render latest forecast, update a picture)

2.2 Class-2 "Parameter" Portals

By registering a Class-2 Portal in a Virtual Organisation, you agree to the conditions laid down in this section and documents references therein.

1. The Portal may offer services to Pseudonymous, Identified and Strongly Identified Web Users.
1. The Portal may either use a Robot Certificate or authentication information provided to obtain a User credential specific to the Web User and use these for interactions with the Grid.
2. No data may be stored on the Grid as a result of any action by a Web User, except for transient data stored in designated scratch areas on the computational resources provided to the running jobs.
3. The job submission rate may be limited differently for Pseudonymous, Identified and Strongly Identified Web Users, and the maximum submission rate by the Portal induced by Pseudonymous and Identified Web Users must be specifically agreed between you and the Grid
4. The Portal must keep enough information to associate any interactions with the Grid with a particular user. If the user was Identified or Strongly Identified, relevant authentication information must be recorded and archived.

2.3 Class-3 "Data Processing" Portals

By registering a Class-3 Portal in a Virtual Organisation, or by connecting a Class-3 Portal to the Grid infrastructure, you agree to the conditions laid down in this section and documents references therein.

1. The Portal may offer services to Identified and Strongly Identified Web Users.
2. The Portal may use a Robot Certificate, or alternatively may use the authentication information provided to obtain a User credential specific to the Web User and use these for interactions with the Grid.
3. When a Robot Certificate is used to store data on the Grid as a result of an action by a Web User, it may only be stored in locations that have been specifically agreed between you and designated Resource Providers, except for transient data stored in designated scratch areas on the computational resources provided to the running jobs. When a User Credential is used, data may be stored in all Grid locations where the User has permission to store such data.
4. The Portal must keep enough information to associate any interactions with the Grid with a particular Web User.
5. The system used to authenticate Identified Users must be adequately secured. In particular additional requirements apply:
   a. Web Users must be notified of all registrations, modifications and of removal of their data in the authentication database.
   b. The authentication database must contain enough information to contact the Web User for as long as authentication is possible.
   c. Entering authenticating information in the database, including resets of such information, must be appropriately authenticated.

2.4 Class-4 "Job Management" Portals

By connecting a Class-4 Portal to the Grid, you agree to the conditions laid down in this section and documents references therein.

2. The Portal may offer services only to Strongly Identified Web Users, or to Identified Web Users where both the Portal system itself and the authentication of Identified Web Users meets the requirements of either the SLCS or MICS IGTF Authentication Profile.
3. The Portal must use User credentials specific to the Web User and use these for all interactions with the Grid.
4. The Portal operations must comply with the Site Operations Policy.

- Based on this *interim* policy,
  BiG Grid allows registration of Robot certificates in its Vos

- Two portals with robot certs now in production
  - NL-SCIA-DC (KNMI, SRON)
  - eNMR (Bijvoet Centre, UU)

- Contributed to JSPG for improvements to policy, see

  https://edms.cern.ch/document/972973

- 'gut feeling' requires well-identified credentials for Function1 to Function3 portals

- A service/host cert does *not* fulfill these requirements!

- Robot certs, issued on hardware tokens are
  - Simple and cheap
  - NL gives them out 'for free', supported by VL-e and BiG Grid
  - see http://ca.dutchgrid.nl/etokens for documentations and software
  - Well secured – and protect against abusing the keypair off the portal machine somewhere else
  - Middleware cannot verify 'source of origin' in a reliable way in a system that supports delegation
    (binding to a source address does not survive first delegation)

Robot certificate support needed 'globally' to enable compliant portals …

… do you support them already?