egee

Open Science Grid

LCG

VO BOX Summary

Conclusions from

Joint OSG and EGEE Operations Workshop - 3

Abingdon, 27 - 29 September 2005

Editors S. Traylen, M. Schulz and J. Templon

www.eu-egee.org

Information Society

Enabling Grids for E-sciencE

- Sites want LCG to succeed and hence agree to deploy VO boxes where / when they are needed!!!

- Running a grid site is not easy; giving free reign to unknown software could make it much worse. Hence Sites are placing some constraints on what can be deployed

- **installation, configuration and maintenance of privileged and untrusted software;**

- **inbound and outbound connectivity maybe through ports that are normally closed (due to site policies);**

- **potentially impossibility to trace real users due to the use of service credentials;**

- **use of standard X.509 digital certificates (i.e. no proxy) to run services, or requests to directly use the host certificate rather than service certificates. Raises concerns about responsibilities, authorization leaks and stolen / misused credentials.**

VO Box Deployment Must Follow guidelines in "**VO Box Security Recommendations and Questionnaire**" from Joint Security Policy Group

- Root cause analysis of operational failures is already tough; adding another active, dynamic, unknown and uncontrolled element can make life much harder if not sufficiently isolated

- Firm agreements are made on what is expected from the two parties operating a VO box, as the site maintains the hardware and the OS, the VO the rest … where *exactly* do we draw the line?

- Operational Holy Services:
  - **Workload will be accepted ONLY through the standard gatekeeper**
    - **P.s. heads up gLite**
  - **Data for storage will be accepted ONLY through the standard SRM interface**
  - **If the VOBox requires the UI s/w to be installed on the base system, this will be the LCG UI.**

VOBox deployment must follow guidelines in "**LCG VOBox Operations Recommendations and Questionnaire**"

Enabling Grids for E-sciencE

- **VO specific monitoring have to communicate**
  - What?
  - How often?
  - How?
  - Stored where?
  - Accessible by whom?
  - Is it monitoring a quantity that is already monitored?
- **Should not be taken as a given that all monitoring will be allowed**
- **Privacy**
  - Current approaches break national rules (unfortunately different ones in each country)
  - Need to find some way (interaction sites/VOs) to let sites control monitoring info via policy specification …

Enabling Grids for E-sciencE

- Capacity inflation
    - **Request VO box to have 3 GB space per job slot == terabyte for medium site!!**
- Resource Drain
    - **Sites already have ~ five-box 'tax' to join grid; VO boxes increase this**
    - **Proportionally larger for smaller sites … sites may choose to drop 'low-priority' VOs**
    - **Answer: don't ask for more than you need, and try to make it possible to run your stuff on shared box if at all possible**
- Is the need to have "own" box due to a
    - **Resource problem (really need whole machine) or**
    - **Organizational problem (so a VM would be OK)?**
- Service Duplication
    - **Not accepted**
    - **Extensions are OK if needed.**

Enabling Grids for E-sciencE

- **We will deploy the VO boxes as long as the VOs do their best to be good citizens (mostly a question of education, not character :)**

- **The VO box should be re-discussed by Baseline Services group (is now significantly expanded from original idea in report)**

- **Also needs to be discussed in forum where site representation is on equal representation with that of experiments (was NOT the case in BS group).**

Enabling Grids for E-sciencE

- **Need for VO-specific services will likely not go away**

- **How to deal with this in the future?**
    - OSG edge services (VM image management)
    - Container technology (java applets in site-managed tomcat)
    - gsissh login that we now have
    - "Local" appearance of services via proxy mechansim?

- **Keep pressure on generic middleware**
    - Some things (monitoring) CAN be generic
    - A strong EGEE generic middleware is a great benefit to HEP, not only for our work but for being able to donate it to other communities … don't underestimate how valuable this is