



Enabling Grids for E-scienceE



Where we are today and where we're headed

Stephen Childs
Trinity College Dublin & Grid-Ireland

6th quattor workshop
27th September 2008

www.eu-egee.org



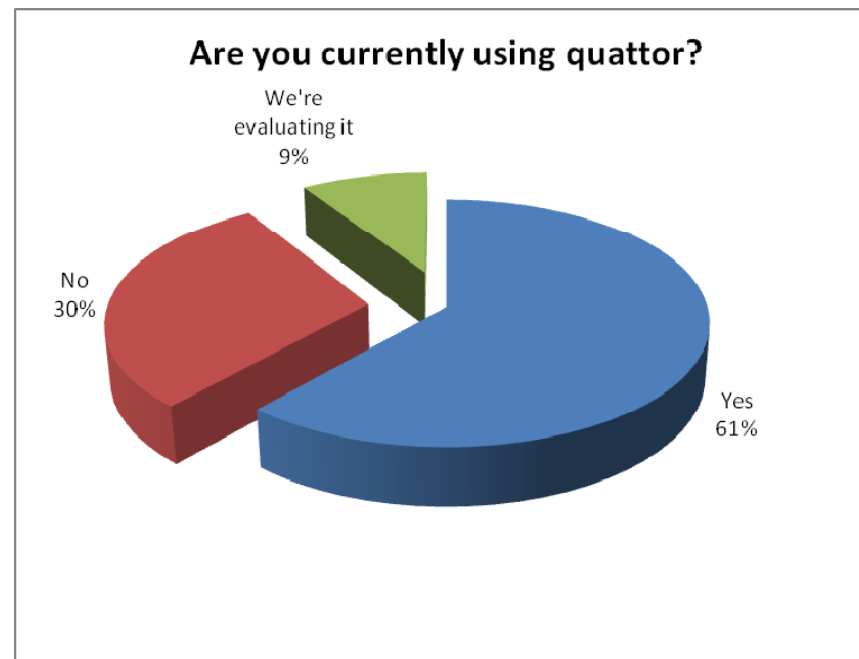
Information Society
and Media



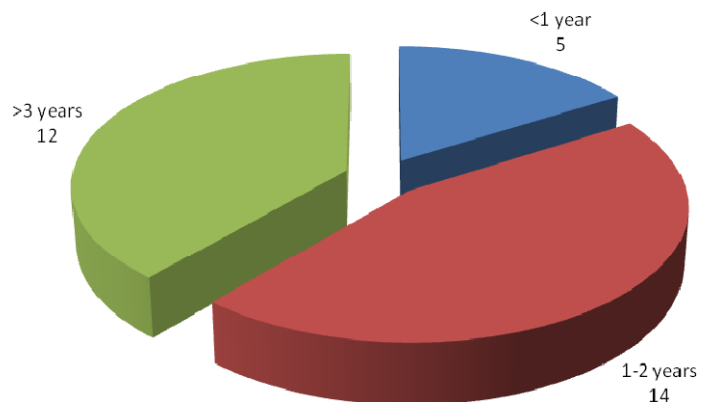
- **Development started in European Data Grid project (2001-2004)**
 - Intended as comprehensive fabric management solution for science grids
 - But NOT adopted as generic grid deployment tool for EGEE
- **quattor community historically centred on CERN**
 - Largest deployment (still?) at CERN (~8000 today)
 - Development hosted at CERN since beginning
 - Core services, service management components, web site, mailing lists
- **Significant development based at LAL since 2005**
 - Subversion configuration database
 - Grid (gLite) configuration templates (Quattor Working Group – QWG)
 - pan compiler
- **Now have sites from outside science grid community involved**

- **Based on survey circulated on quattor and grid mailing lists**
 - Conducted 6th-21st October 2008
 - 48 responses (44 “real” responses after merge/removal of duplicates)

- **Results**
 - 27 using quattor (some represent multiple sites)
 - 14 countries: FR, UK, BE, DE, ES, IT, RU, CA, PT, CH, IE, GR, CH, JP, IN

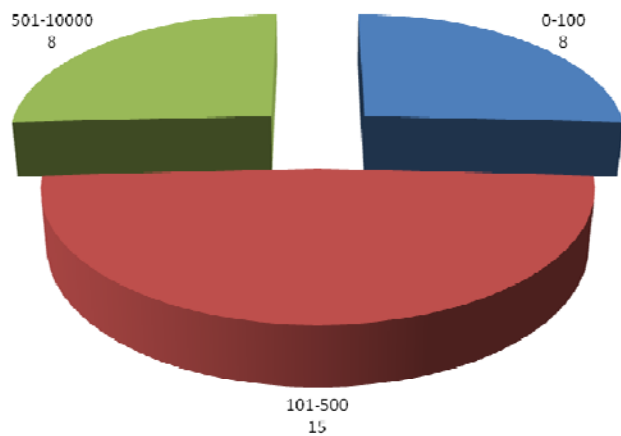


How long have you been using quattor for?

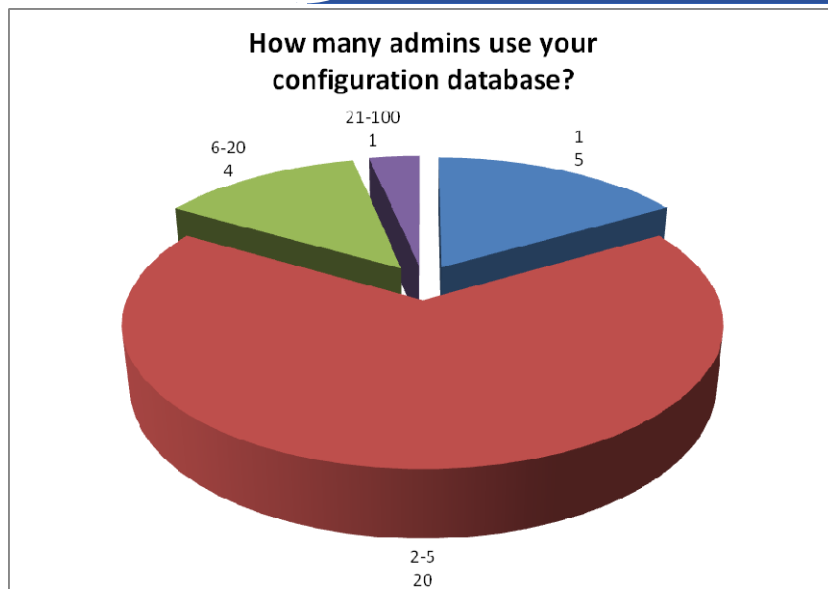


- **Experience with quattor**
 - 61% two years or less
 - 39% more than three years

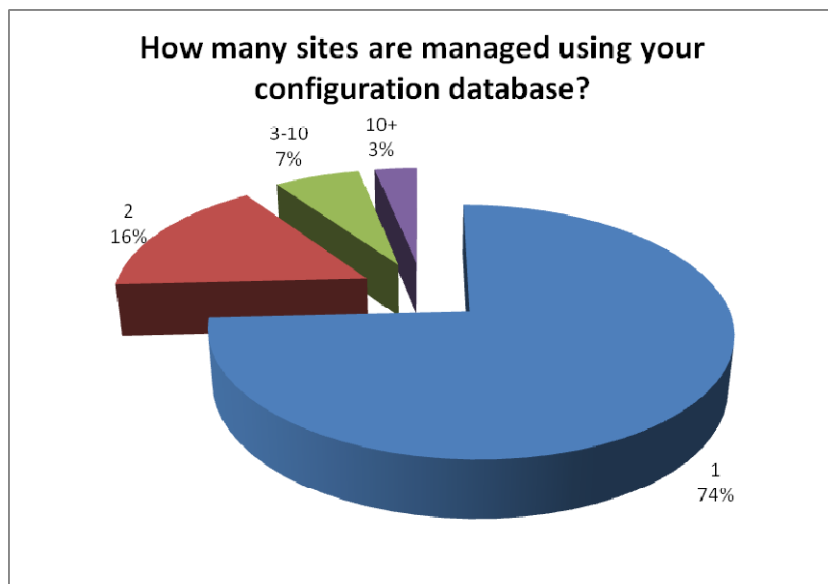
How many nodes do you manage using quattor?



- **Size of installations**
 - ~25% small, ~50% medium, ~25% large

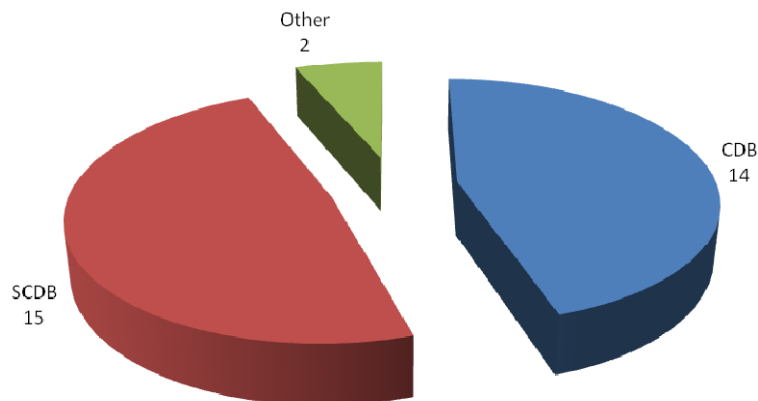


- **Admins per DB**
 - 83% <5 admins
 - 13% have 6-20
 - Only one greater than 20



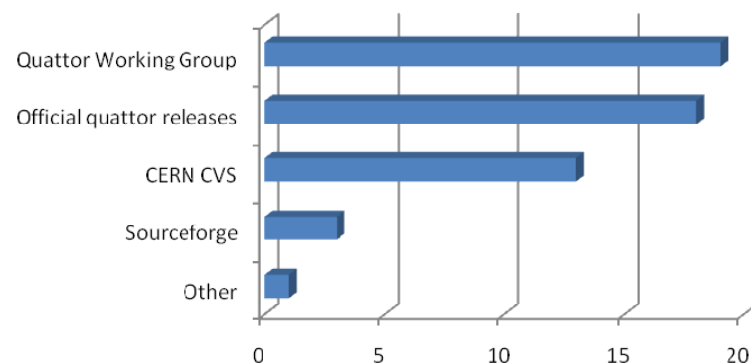
- **Number of sites per DB**
 - 74% single-site
 - 23% co-manage 2-10 sites
 - Only one co-manages >10 sites

Which configuration database do you use?

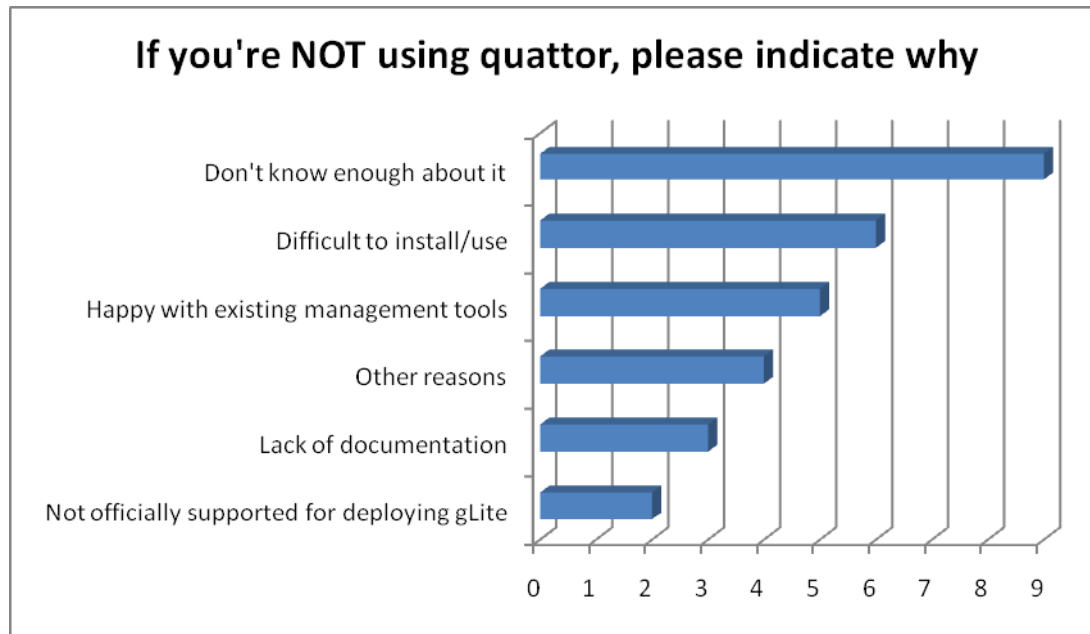


- **CDB is alive and kicking (roughly 50-50 with SCDB)**

Where do you get new quattor software from?



- **A mix of sources**
 - 4 only use CERN CVS + quattor releases
 - 5 use 4-5 different sources
 - 9 depend totally on QWG



- **Lack of knowledge about quattor**
- **Difficult to use (unclear if this is based on experience or perception)**
- **Existing solutions that work well**

Apparently one user is very unhappy

- A site that uses quattor responded to the question “*Which tools would you like to see developed by the quattor community?*” as follows:

“I would like quattor developers and lcg-qwg to get some brains into their empty heads.”

- 1. quattor is being used seriously by sites all over the world**
 - Justification when applying for funding for further development
- 2. Complexity and lack of information put people off**
 - Need coherent documentation, especially quick-start style
- 3. Perceived as difficult to use**
 - Need to consolidate tools, make sure they work, develop server VM?
- 4. Both large and small sites use quattor**
 - Evidence that it scales and isn't too complex for smaller sites
- 5. Only one site currently requires many concurrent users**
 - This was previously stated to be a reason to use CDB
- 6. CDB is widely used outside CERN and CNAF**
 - Are the sites that use it willing to maintain and develop it?
 - Community still needs different DB implementations?
- 7. While multi-site deployments work, most are single-site**
- 8. The community is looking for tools**
 - In some cases, functionality exists but could be packaged better

- **Strengths**

- Open and supportive community (applies across the board: core, QWG)
- Improving integration with related systems (Nagios, Lemon, Xen, OpenVZ)
- Large suite of configuration components (254 at Q4 2008)
- Well-defined language model for describing configuration
- Good support for configuring gLite grid middleware

- **Weaknesses**

- Lack of dedicated effort for release management, “tidying up”
- No automated build and test
- Lack of documentation (especially quick-start style)
- Difficult to contribute (access to code repositories, ...)
- Legacy architecture makes writing components difficult
- Hard to bootstrap: many external dependencies (Apache, SVN, ant, DHCP, PXE, ...)
- Level of knowledge still low among prospective “customers”
- Not officially supported for gLite grid configuration

Where to focus efforts?

- **Transition from CERN-centred, grid-focussed community to generic open source project**
 - In progress, see later talk on migration to SourceForge
 - May need to discuss software architecture as well?
- **Concentrated effort on documentation, packaging**
 - Need integration-focussed workshop to do this work
 - Quattor virtual appliance for site bootstrap
- **Dissemination**
 - Paper in USENIX LISA November 2008 ✓
 - Poster at EGEE September 2008 ✓
 - Supercomputing, HEPIX, Linux/open-source/sysadmin events ?
- **Seek funding for dedicated development effort?**
 - EU, national “commercialisation” funds, ..., ?

- **Finalise structures for opening up community collaboration**
 - Ensure access rights are easy to manage (OpenID or similar)
- **Coordinate tool development**
 - Centralise and catalogue (and package?) existing tools
 - Distribute responsibility for developing new tools (and allocate effort)
- **Decide on sustainable build and release procedures**
 - Not dependent on one person, either for expertise or access reasons
 - Must be automated