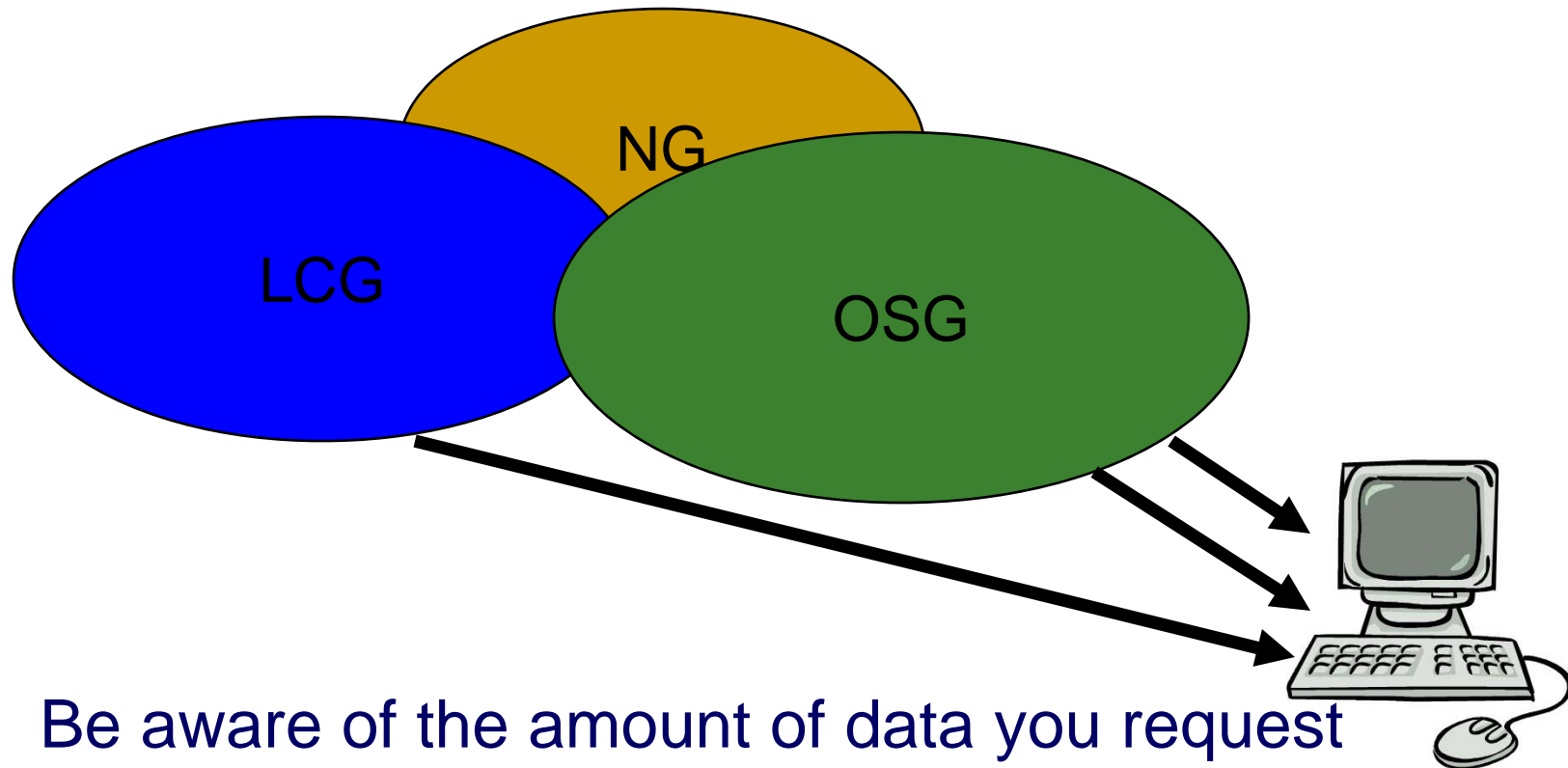

Extended OSG client for WLCG



US ATLAS Tier2/Tier3 Workshop
May 28, 2008
University of Michigan, Ann Arbor, MI

Marco Mambelli
University of Chicago

Moving (lots of) data



- Be aware of the amount of data you request
- 20TB overnight? Shipped?
- 1TB takes one day at 100Mb/s sustained rate

Client tools

- Only for machines where you need a local installation of client software (Tier2, Tier3 or PC)
- WLCG-Client - grid client
 - Authentication (Grid proxy, VOMS proxy)
 - Data movement (globus-url-copy, srmcp, lcg-cp, lcg-ls, ...)
 - Job submission (to OSG and Pathena)
- DQ2 enduser tools
 - Data movement (dq2_get, dq2_ls, ...)
- DQ2 client
 - Central and Site services interaction
 - Data movement (dq2-get, dq2-ls, ...)

WLCG-Client

■ Installation

- `pacman -get http://www.mwt2.org/caches/:wlcg-client`

■ Use

- `source /installation_dir/setup.sh`

■ Documentation

- <http://www.usatlas.bnl.gov/twiki/bin/view/Admins/WlcgClient>

■ Content

- VDT-Client (Globus and much more)
 - For more info: <http://vdt.cs.wisc.edu/releases/1.10.1/>
- OSG vomses information
- LCG-Utills (`gfal_*` and `lcg-*` commands)
- LFC-min (python API to interact with LFC)
 - <http://twiki.mwt2.org/bin/view/DataServices/PackageLFCmin>

Enduser tools: dq2_*

■ Install enduser tools

- Python code
- Copy files from AFS (or CVS)
 - `/afs/cern.ch/atlas/offline/external/GRID/ddm/endusers/`
 - `/afs/usatlas.bnl.gov/Grid/Don-Quijote/dq2_user_client/`
 - `http://atlas-sw.cern.ch/cgi-bin/viewcvs-atlas.cgi/offline/DataManagement/DQ2_0_2/endusers/?pathrev=HEAD`

■ Use

- Requires Grid client (wlcg-client)
- `source setup.sh`
- `voms-proxy-init` required

DQ2 Client: dq2-*

- New DQ2 Client (M. Lassnig, M. Branco at CERN) is provided also as pacman package
 - <https://twiki.cern.ch/twiki/bin/view/Atlas/DistributedDataManagement>
 - <http://ddm-build.cern.ch/ddm/build/testing/doc/guides/dq2-clientapi-cli/html/user/>
 - `pacman -trust-all-caches -allow tar-overwrite -get`
<http://atlas.web.cern.ch/Atlas/GROUPS/DATABASE/project/ddm/releases/pacman/cache:DQ2Clients>
- The package is not fully working with OSG but there is a repackaged version, integrated with WLCG-Client
 - `pacman -get http://www.mwt2.org/caches/:ATLAS-Client`
- Commands are similar (`dq2_get` -> `dq2-get`) but not identical (**there is some difference in behavior**)!

Final (personal) notes

- Tools for data retrieval work OK
 - Generally you can find and retrieve datasets
 - A single command may issue long and expensive operations (staging from tape, transfer many big files)
 - May cause overload on your local system
- Use carefully the tools to modify datasets
 - E.g. tools to write and register new datasets, remove existing datasets, ...
 - Many confusing options (dq2-rm, dq2-erase, ...)
 - Not really designed for casual users, may create inconsistencies
- **dq2-ls prints the size of the dataset!**
- **= BE AWARE OF YOUR DATASET SIZE**