



# Grid/Rucio (Introduction + Hands-on Session)

ATLAS-D Physics Meeting Heidelberg 2016

Gen Kawamura

II.Physikalisches Institut, Universität Göttingen

# Overview

- Your Grid environments

- Grid UI with Docker in your Linux Box
- Lxplus at CERN
- Setup CVMFS

- Introduction to Grid computing

- Concepts
- Certificate Authorities and VOMS
- ATLAS Grid computing & WLCG Resources
- Grid job and data
- ATLAS Resources
- Grid user interface (CLI) and CVMFS
- BigPanda Monitoring

- Rucio (ATLAS data management system)

- Basic concept
- Setup Rucio
- Searching for data (CLI)
- Listing metadata (CLI)
- Download data (CLI)
- Moving data (WebUI)

- Links ad references

# Your Grid environments



# Grid UI with Docker in your Linux Box



# Lxplus at CERN



# Introduction to Grid computing



# Concepts - 1

- Uses Public Key Infrastructure (PKI)
- Pluggable computing power and storage like Power Grid
  - Easy to use like electricity
  - Public and open interfaces by common technologies such as Web, Web services etc
  - WWW allows us to access unlimited information, Grid allows us to access unlimited computing power and storage on this planet



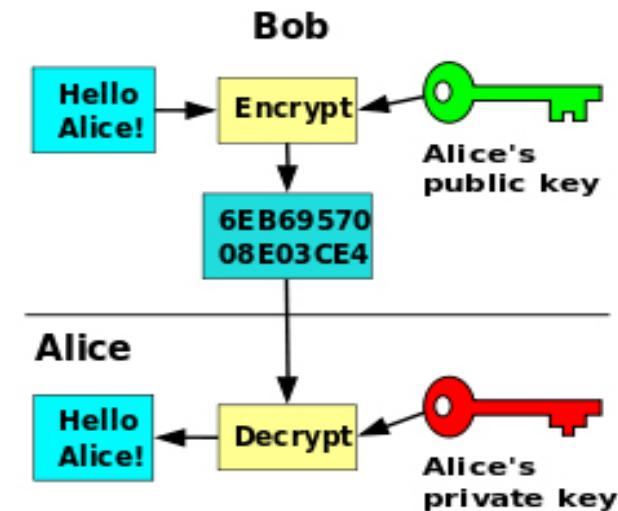
# Concepts - 2

- Some technical terms you may often need (however not in Physics)
  - User certificate, proxy certificate, CA, Virtual Organization (VO), VOMS, authentication, authorization, Computing Element, Storage Element, Worker nodes, Batch system, Workload Management System, data management system, Job, logical file name, data replica, information provider, site



# Concepts - 3

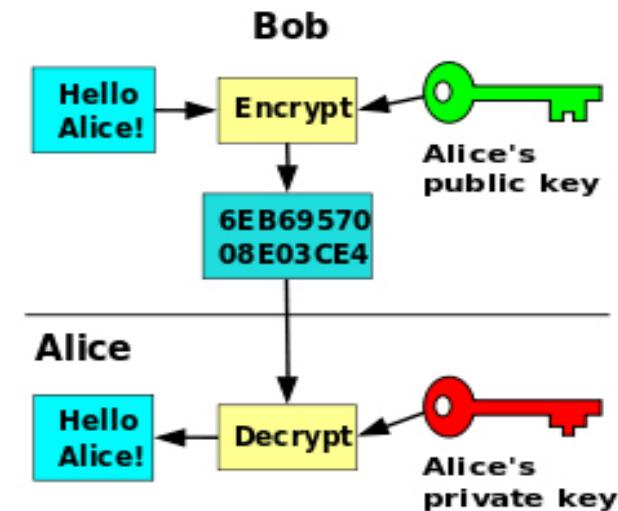
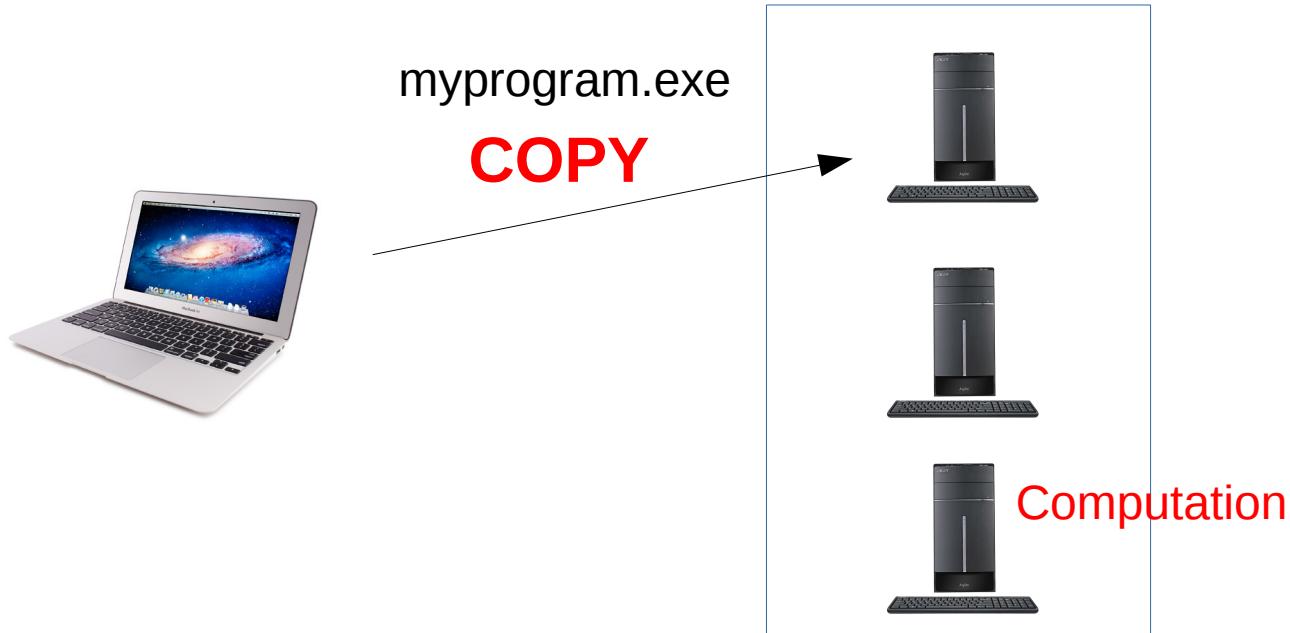
- How it works (without Grid)
  - If PKI by general openssl



# Concepts - 3

- How it works (without Grid)
  - If PKI by general openssl

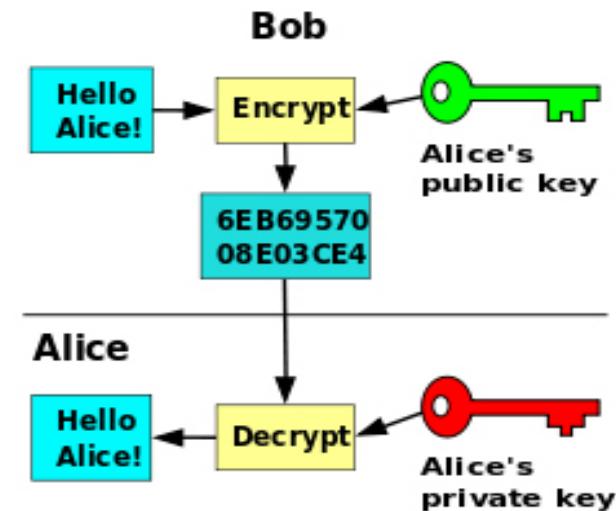
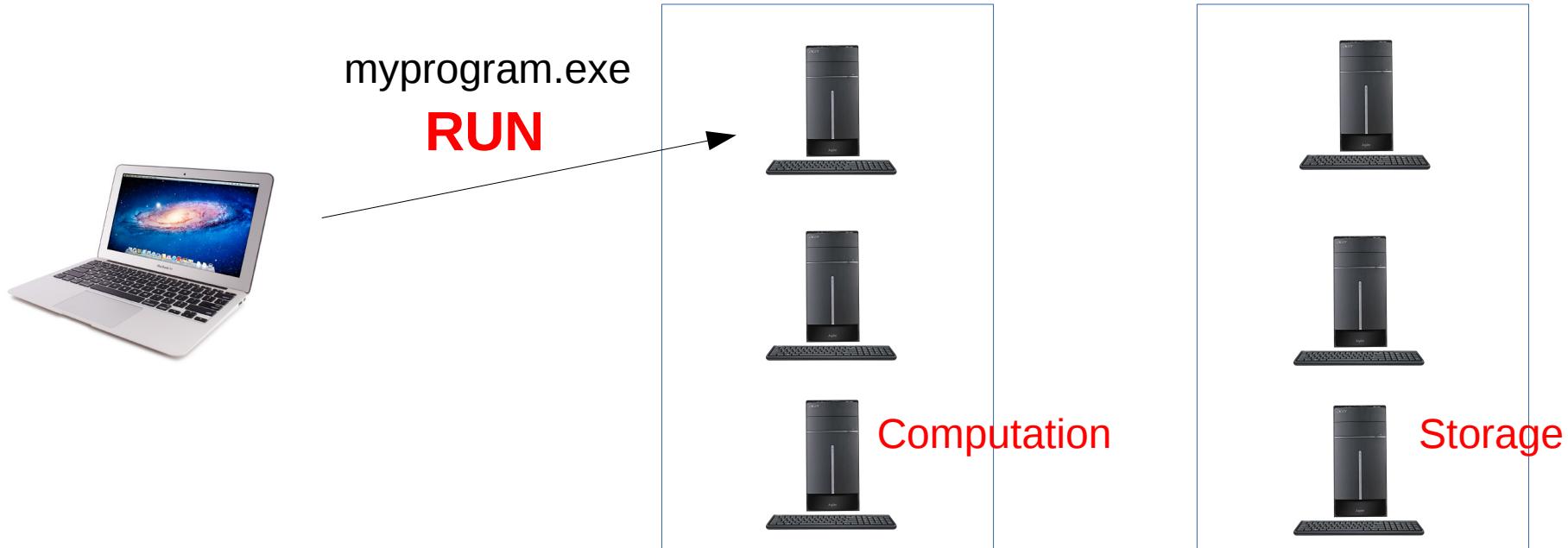
```
$ scp myprogram.exe computer01:~/
```



# Concepts - 3

- How it works (without Grid)
  - If PKI by general openssl

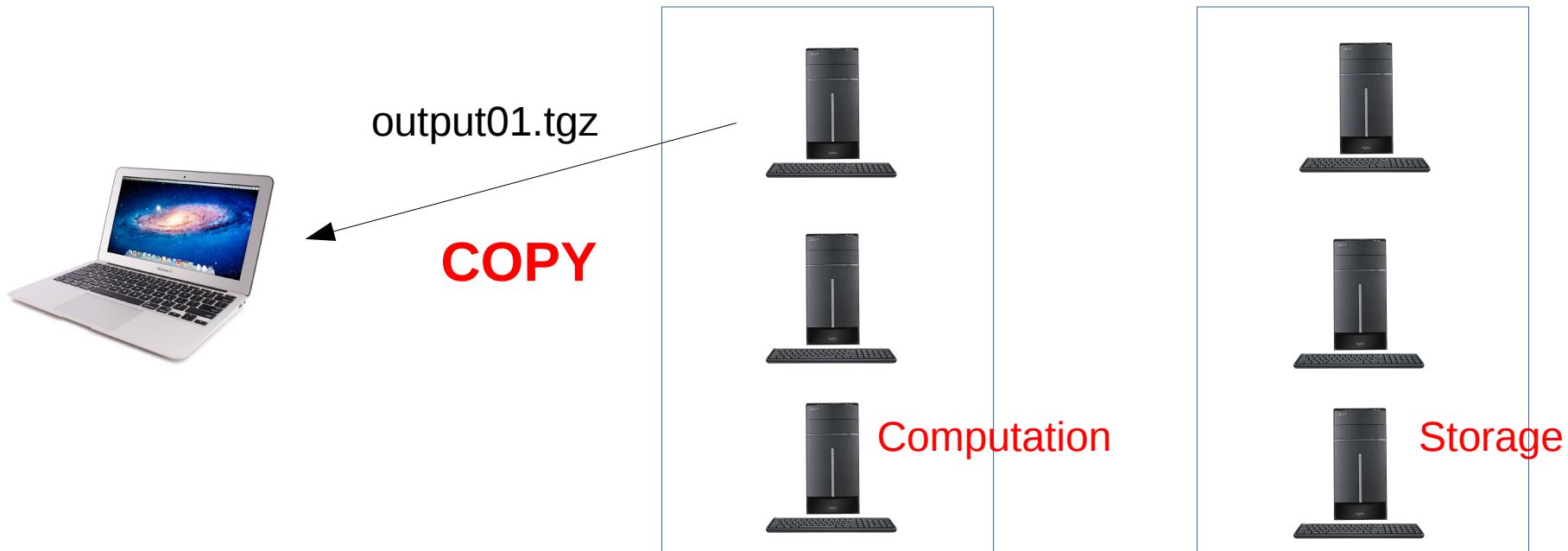
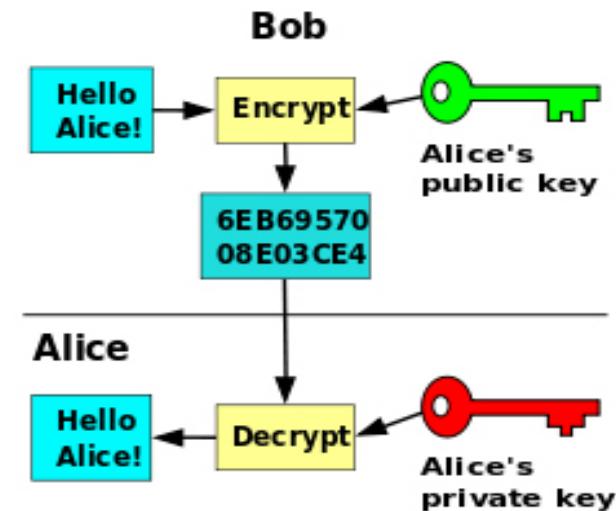
```
$ scp myprogram.exe computer01:~/  
$ ssh computer01 “~/myprogram.exe run 01”
```



# Concepts - 3

- How it works (without Grid)
  - If PKI by general openssl

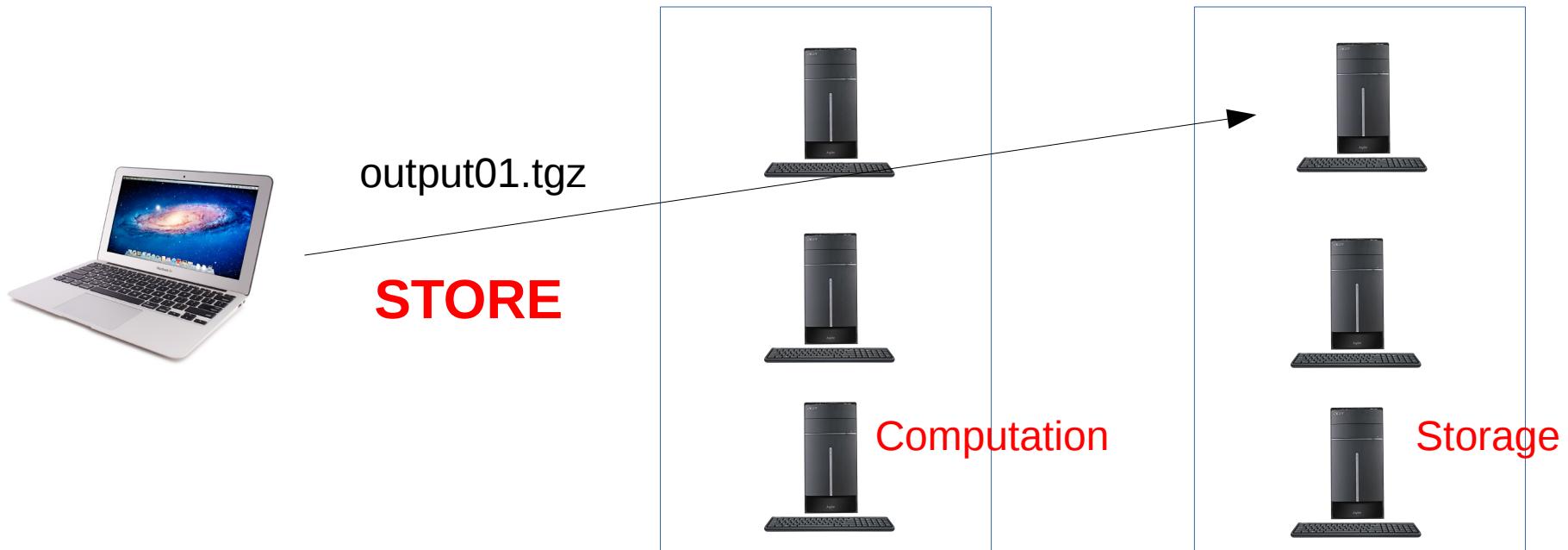
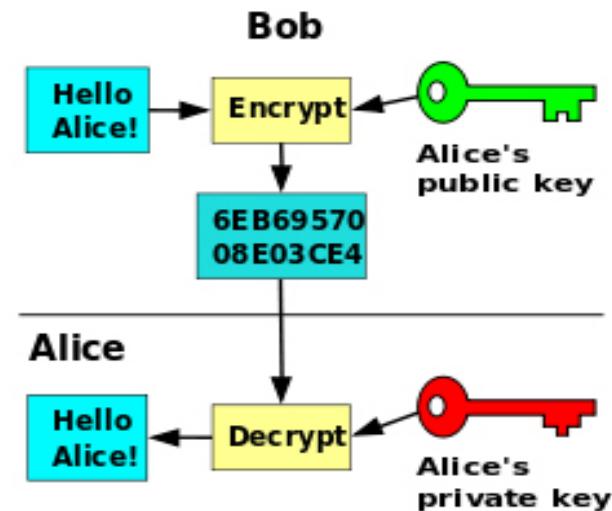
```
$ scp myprogram.exe computer01:~/  
$ ssh computer01 “~/myprogram.exe run 01”  
$ scp computer01:output01.tgz ./output01.tgz
```



# Concepts - 3

- How it works (without Grid)
  - If PKI by general openssl

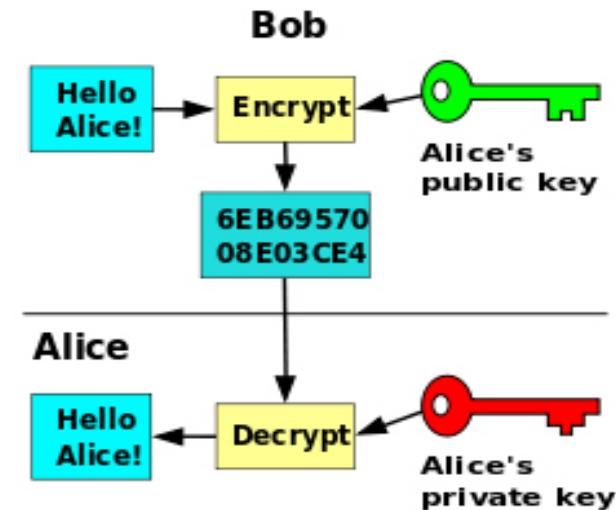
```
$ scp myprogram.exe computer01:~/  
$ ssh computer01 “~/myprogram.exe run 01”  
$ scp computer01:output01.tgz ./output01.tgz  
$ scp output01.tgz storage01:output01.tgz
```



# Concepts - 3

- How it works (without Grid)
  - If PKI by general openssl

```
$ scp myprogram.exe computer01:~/  
$ ssh computer01 “~/myprogram.exe run 01”  
$ scp computer01:output01.tgz ./output01.tgz  
$ scp output01.tgz storage01:output01.tgz
```



It works, but too complicated!



STORE



Computation



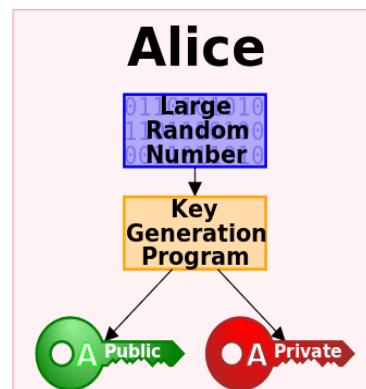
Storage

# Concepts - 3

- How it works (without Grid)
  - You can be inspired!
    - Security policy based on PKI is fine
    - Recognition of a user in affiliation, organization, experiment
    - Control of Job allocation
    - Control of storage space allocation
    - Control among geographically distributed computers
    - Interfacing local computing infrastructure
    - Providing tree-like information of computing clusters
      - E.g. like Domain Name Server
    - End user tools to make our life easy

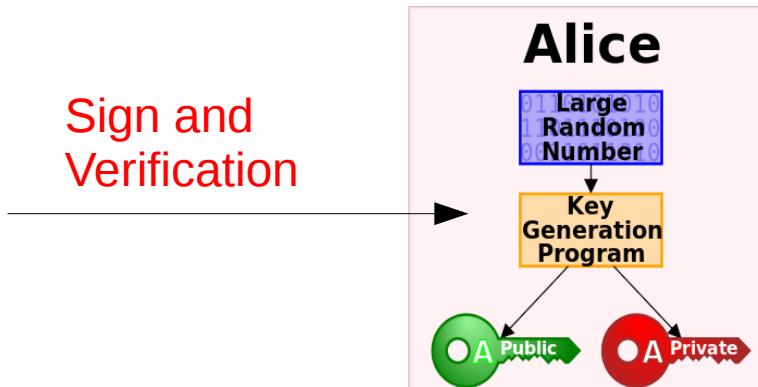
# Concepts - 4

- User certificate, proxy certificate, CA, Virtual Organization (VO), VOMS, authentication, authorization
  - Certificates and verification mechanism
    - User key generation (very briefly)
      - 2 distinct primary numbers → p and q
      - Compute a large number  $N = pq$
      - The p and q are kept in private key (SK), N is in public key (PK)



# Concepts - 4

- User certificate, proxy certificate, CA, Virtual Organization (VO), VOMS, authentication, authorization
  - Certificates and verification mechanism
    - Signing CA identity on the user keys (very briefly)
      - CA has PK and SK as well
      - Sign (S) returns a tag (T) from user private key and a string X
      - Verification (V) returns 1/0 from user public key, X and T



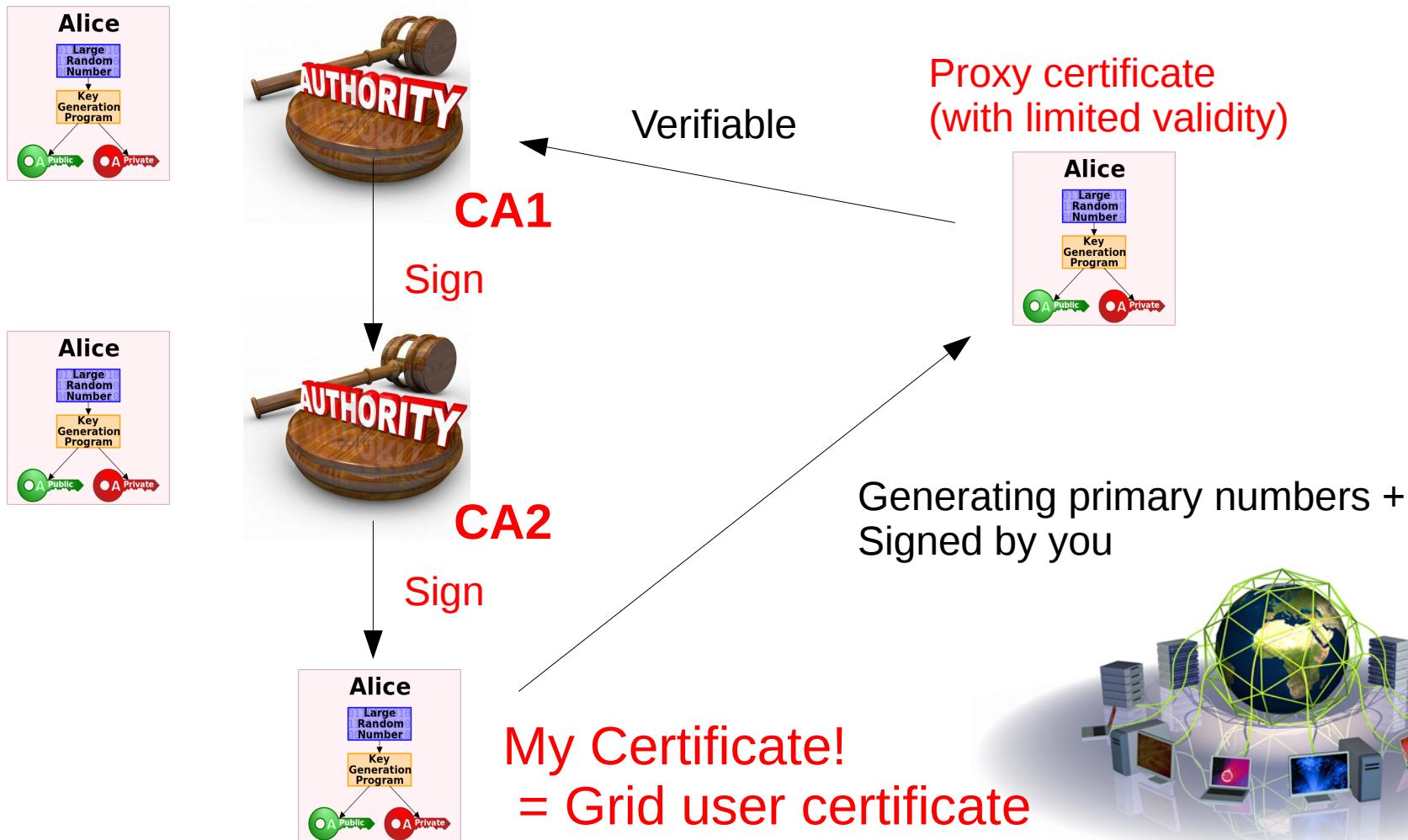
# Concepts - 5

- Then, a process of hierarchical signatures



# Concepts - 5

- Generating a new certificate = proxy certificate



# Concepts - 6

- Virtual Organization (VO), VOMS
  - VOMS extends a certificate with a VO

## Certificate Authority



VOMS server

ATLAS VOMS  
CMS VOMS  
LHCb VOMS

VOs



# Concepts - 7

- A similar analogy
  - Authenticated and authorized for your tasks

Certificate Authority



↓  
Sign



User certificate

↓  
Sign



Proxy certificate  
= a limited copy of your certificate

Controlled



Authentication  
= who you are?



Authorization  
= can enter a new land (Grid)



# Certificate Authorities and VOMS

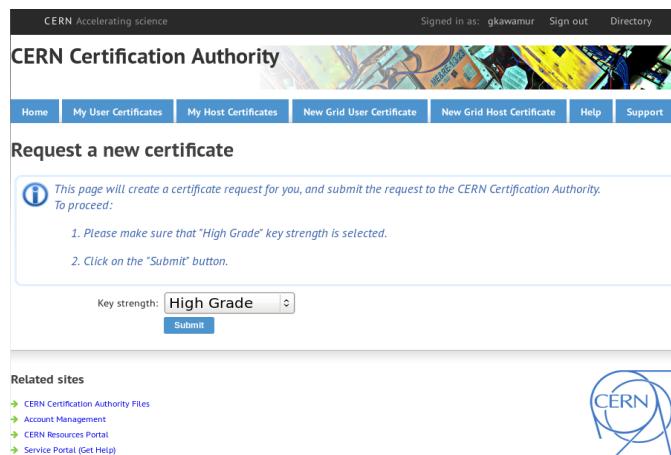
- Germany (FZK)

- <https://gridka-ca.kit.edu/>



- CERN

- <https://ca.cern.ch/ca/user/Request.aspx?template=EE2User>



# Certificate Authorities and VOMS

- VOMS top page
  - <https://voms2.cern.ch:8443/>
- VOMS ATLAS (request your ATLAS VO)
  - <https://voms2.cern.ch:8443/voms/atlas>
- VOMS ATLAS users in Germany
  - <https://voms2.cern.ch:8443/voms/atlas/services/VOMSCompatibility?method=getGridmapUsers&container=/atlas/de>



# Hands-on exercise 1

- Grid Basis

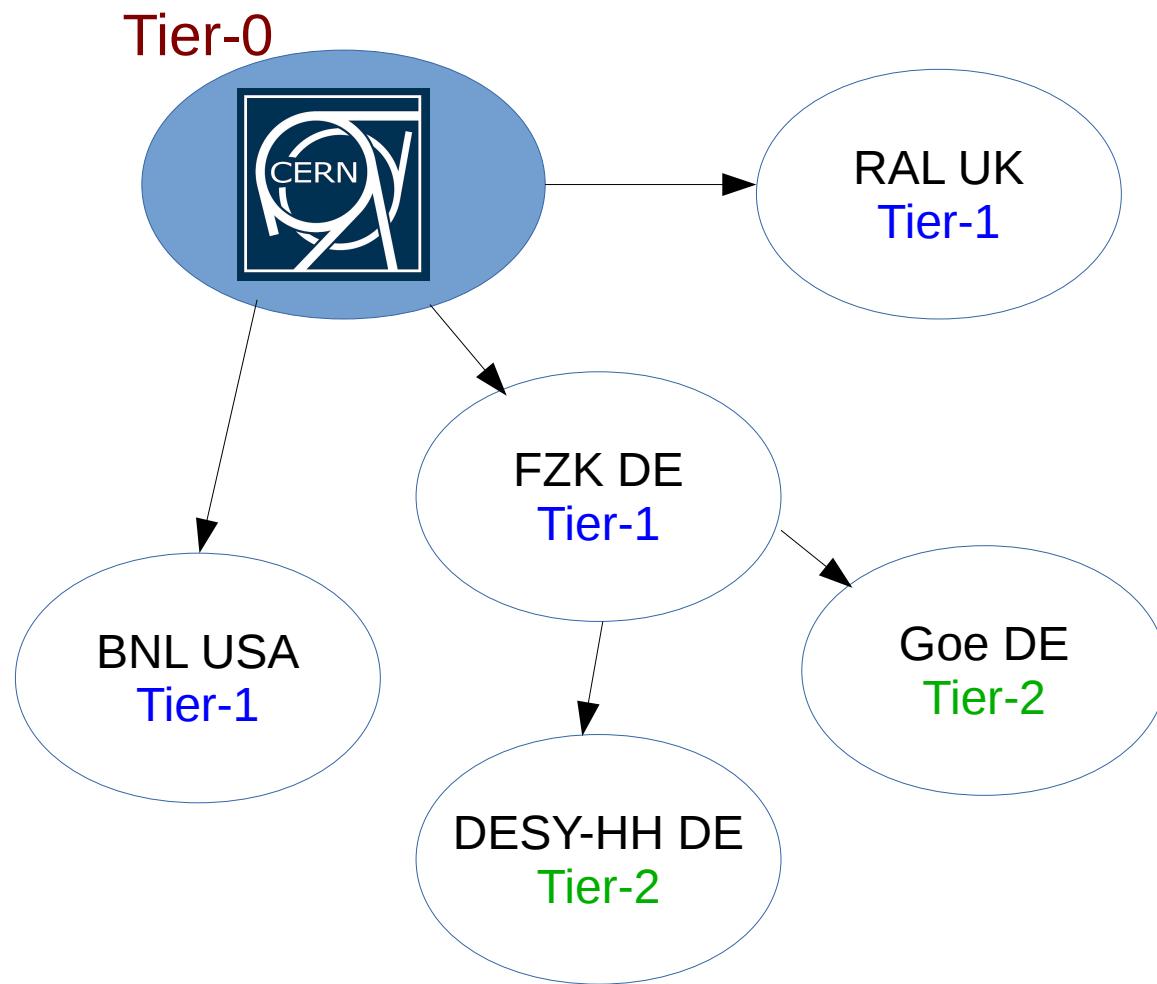
# Hands-on exercise 2

- Grid Basis

# ATLAS Grid computing and WLCG resources - 1

- LHC multi-tier structure

- WLCG = Worldwide LHC Computing Grid



**Tier-0:** Raw data, Data store (in tape), Pre-processing, Reconstruction

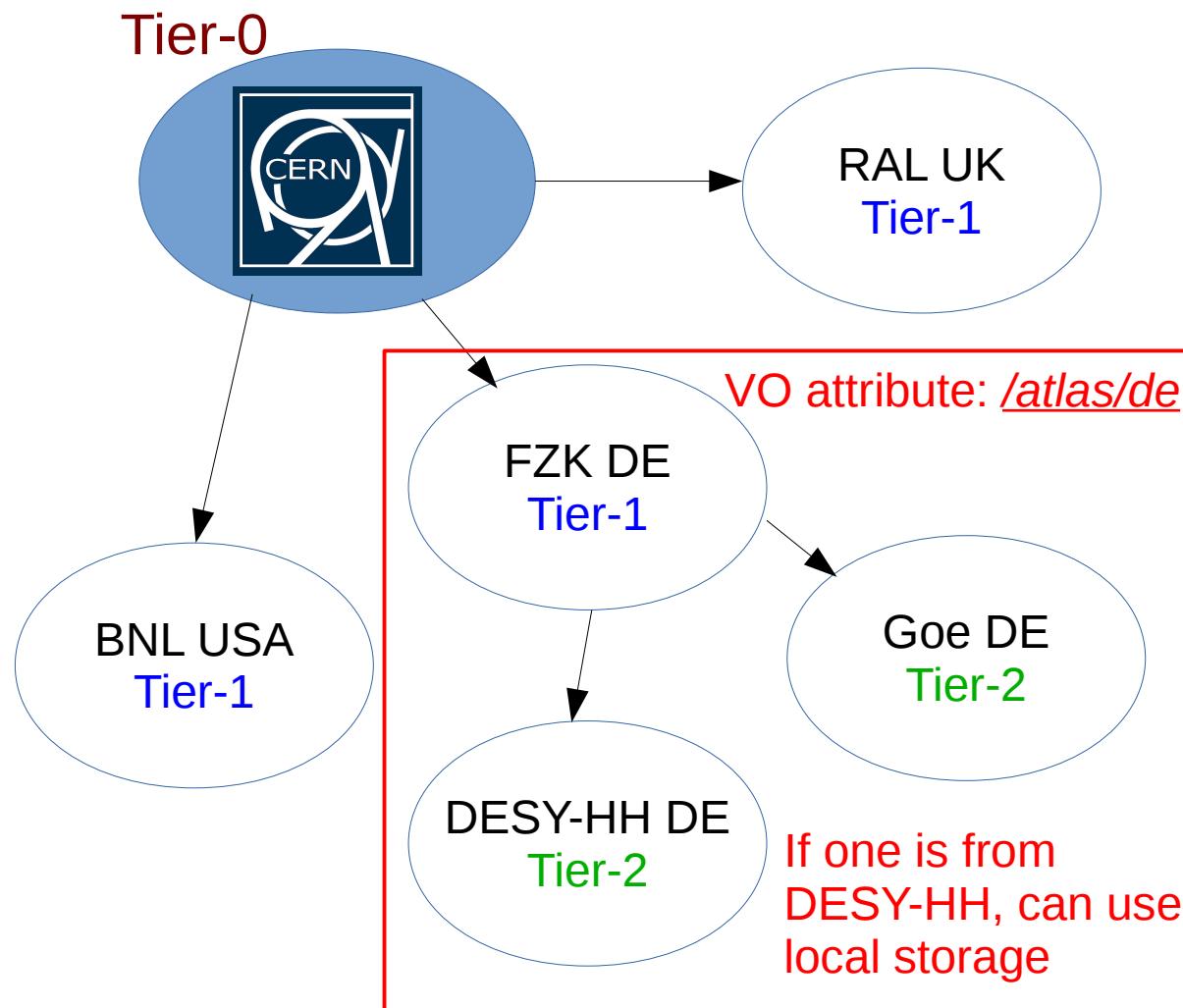
**Tier-1:** National LCG-Centre, faster network connections and larger storage spaces (e.g. Tape), MC production, user analysis, etc.

**Tier-2:** University or Facility level computing sites. MC production, user analysis, etc.

# ATLAS Grid computing and WLCG resources - 1

- LHC multi-tier structure

- WLCG = Worldwide LHC Computing Grid



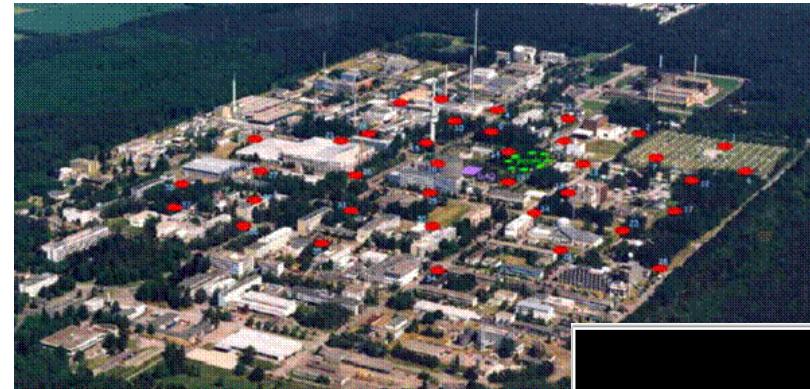
**Tier-0:** Raw data, Data store (in tape), Pre-processing, Reconstruction

**Tier-1:** National LCG-Centre, faster network connections and larger storage spaces (e.g. Tape), MC production, user analysis, etc.

**Tier-2:** University or Facility level computing sites. MC production, user analysis, etc.

# ATLAS Grid computing and WLCG resources - 2

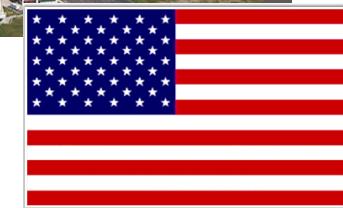
- FZK Tier-1



Forschungszentrum Karlsruhe  
in der Helmholtz-Gemeinschaft

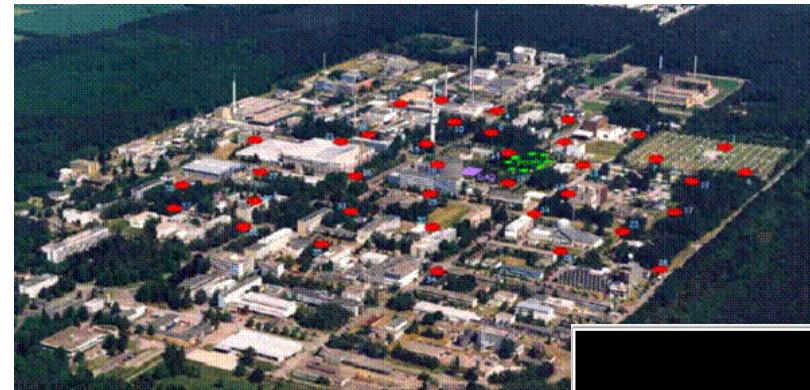


- BNL Tier-1

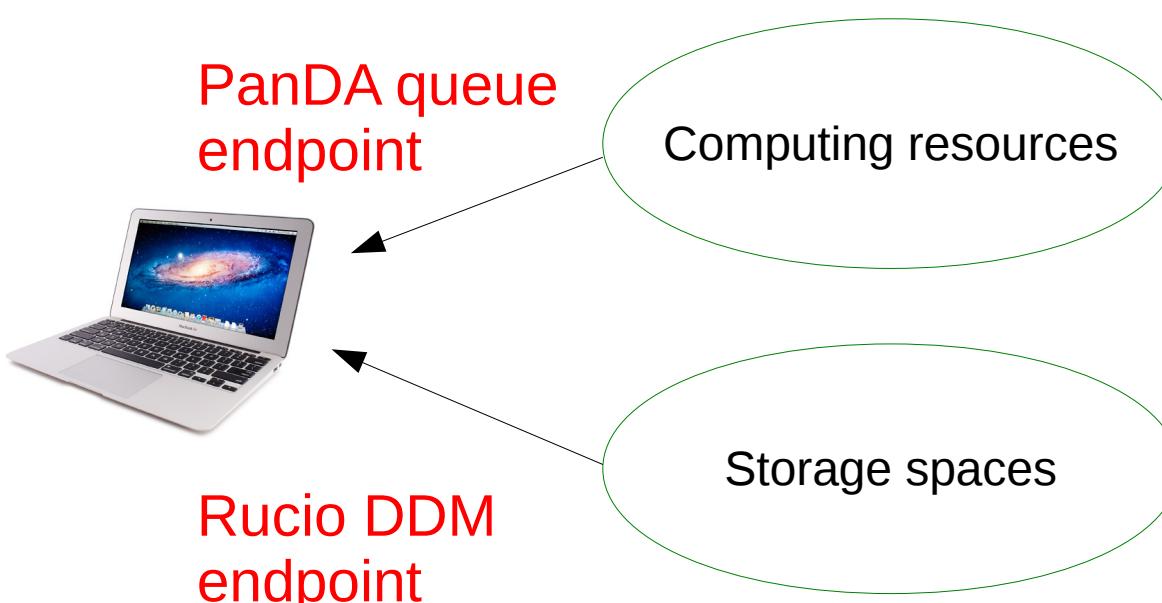


# ATLAS Grid computing and WLCG resources - 2

- FZK Tier-1



Forschungszentrum Karlsruhe  
in der Helmholtz-Gemeinschaft



FZK Computing centre



DDM = Distributed Data Management

# Grid job and data - 1

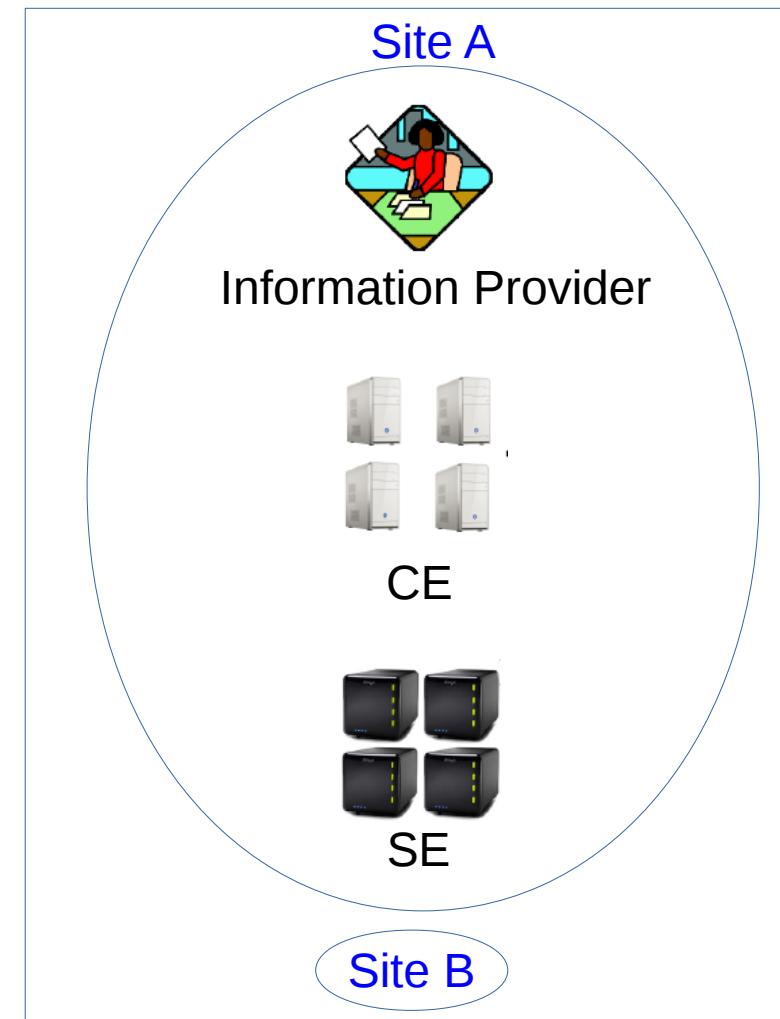
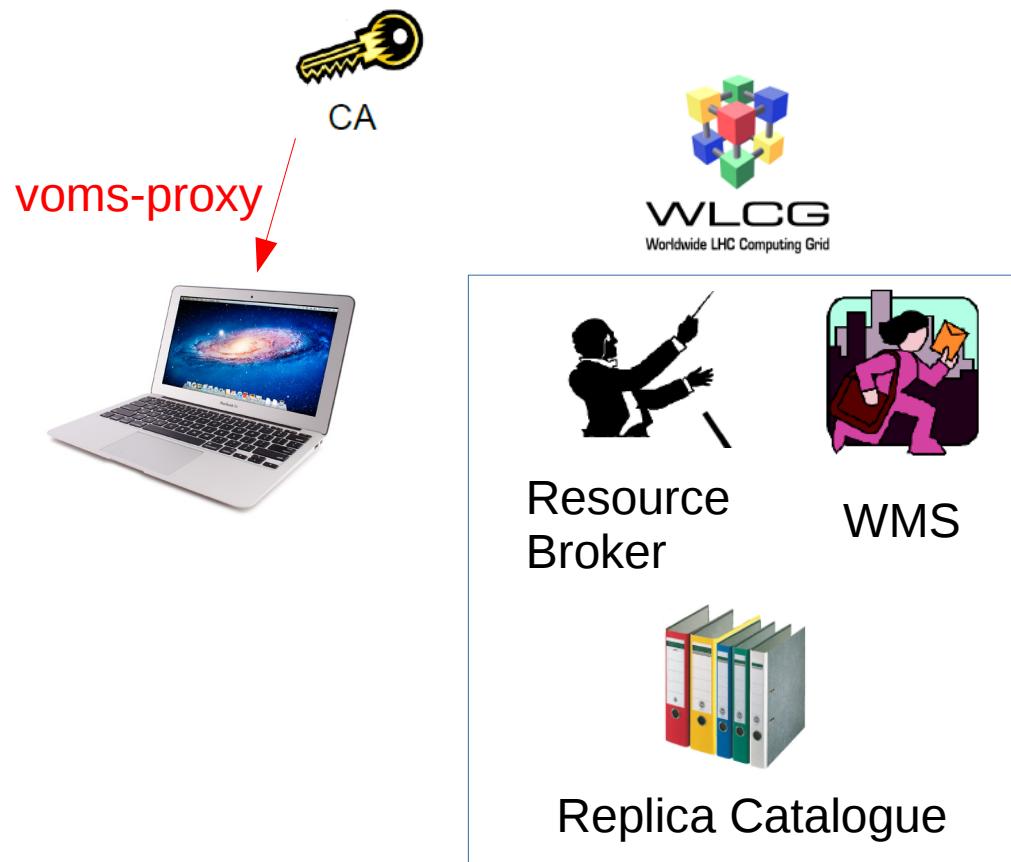
- Computing Element, Storage Element, Worker nodes, Batch system, Workload Management System, data management system, Job, data replica, information provider, site
  - **Job** = Computing task of which user run program in Grid system
  - **Site** = Minimum boundary at which provides computing power and storage space
  - Components of a Grid site
    - UI = User Interface
    - CE = Computing Element
    - SE = Storage Element
    - WN = Worker Node
  - Distribution service
    - DDM = Distributed Data Management System
    - WMS = Workload Management System

# Grid job and data - 2



**WLCG**  
Worldwide LHC Computing Grid

- How do they work?

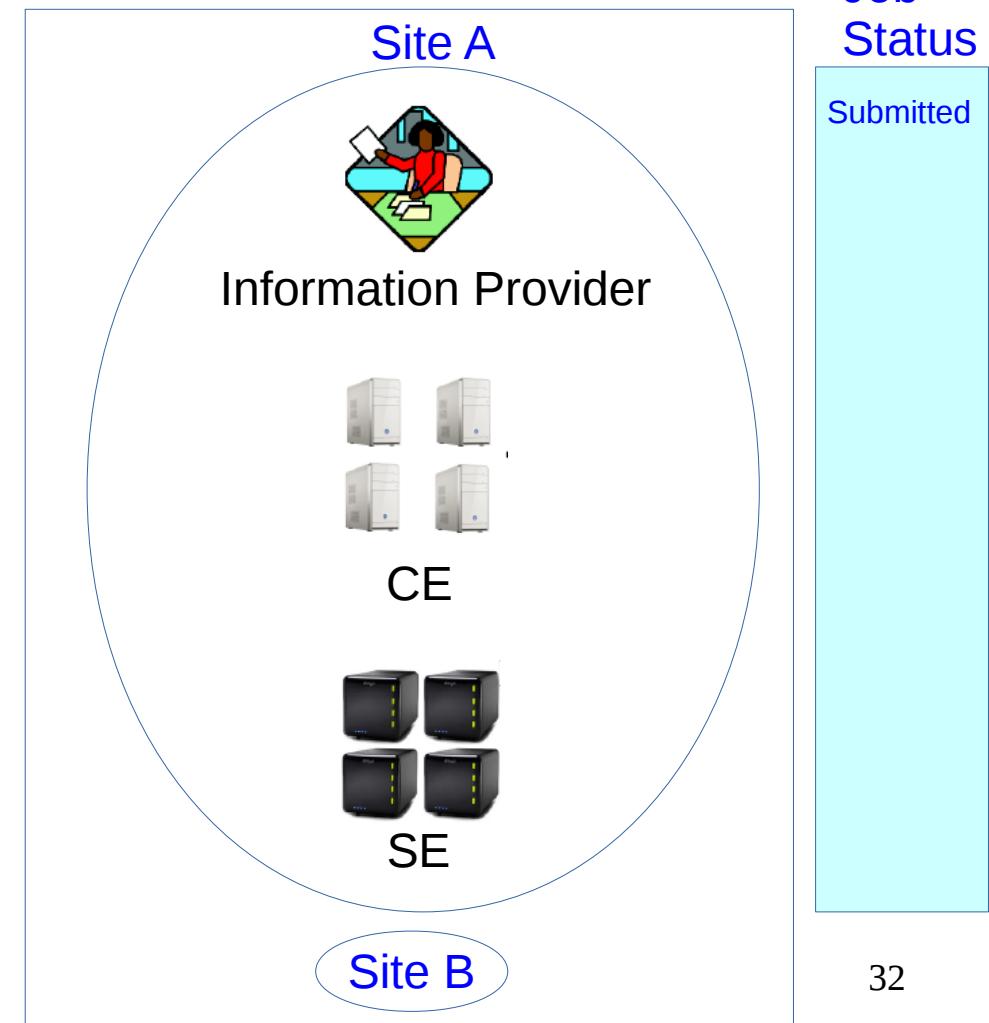
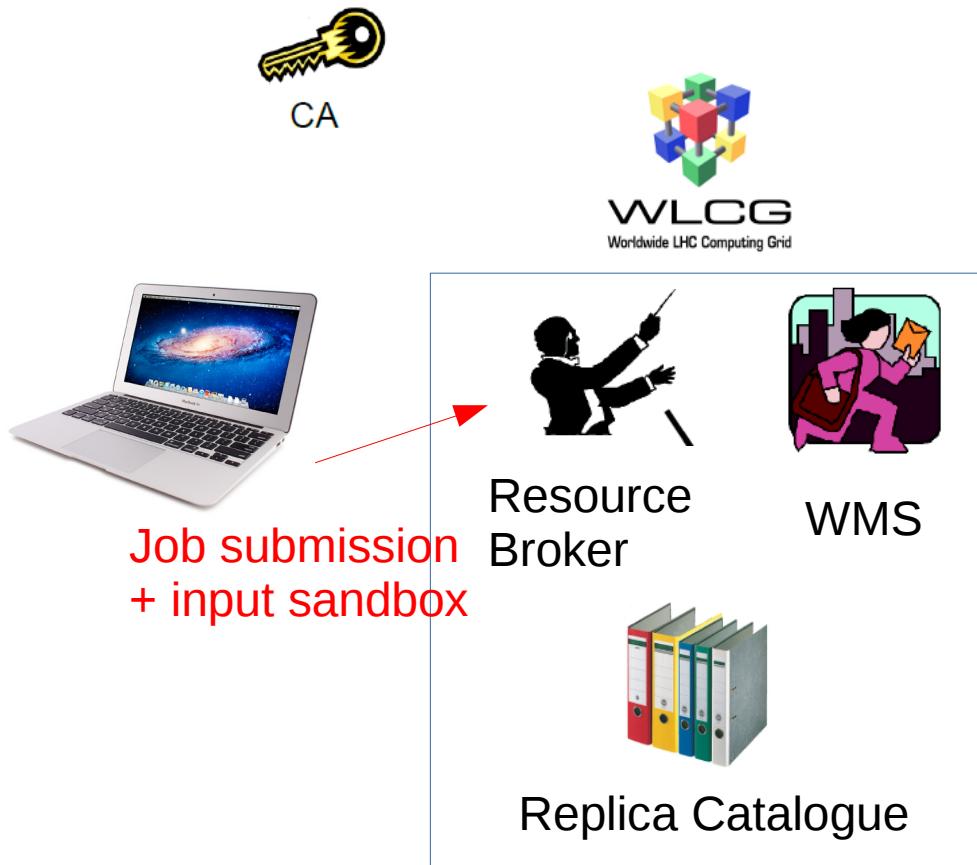


# Grid job and data - 2



**WLCG**  
Worldwide LHC Computing Grid

- How do they work?

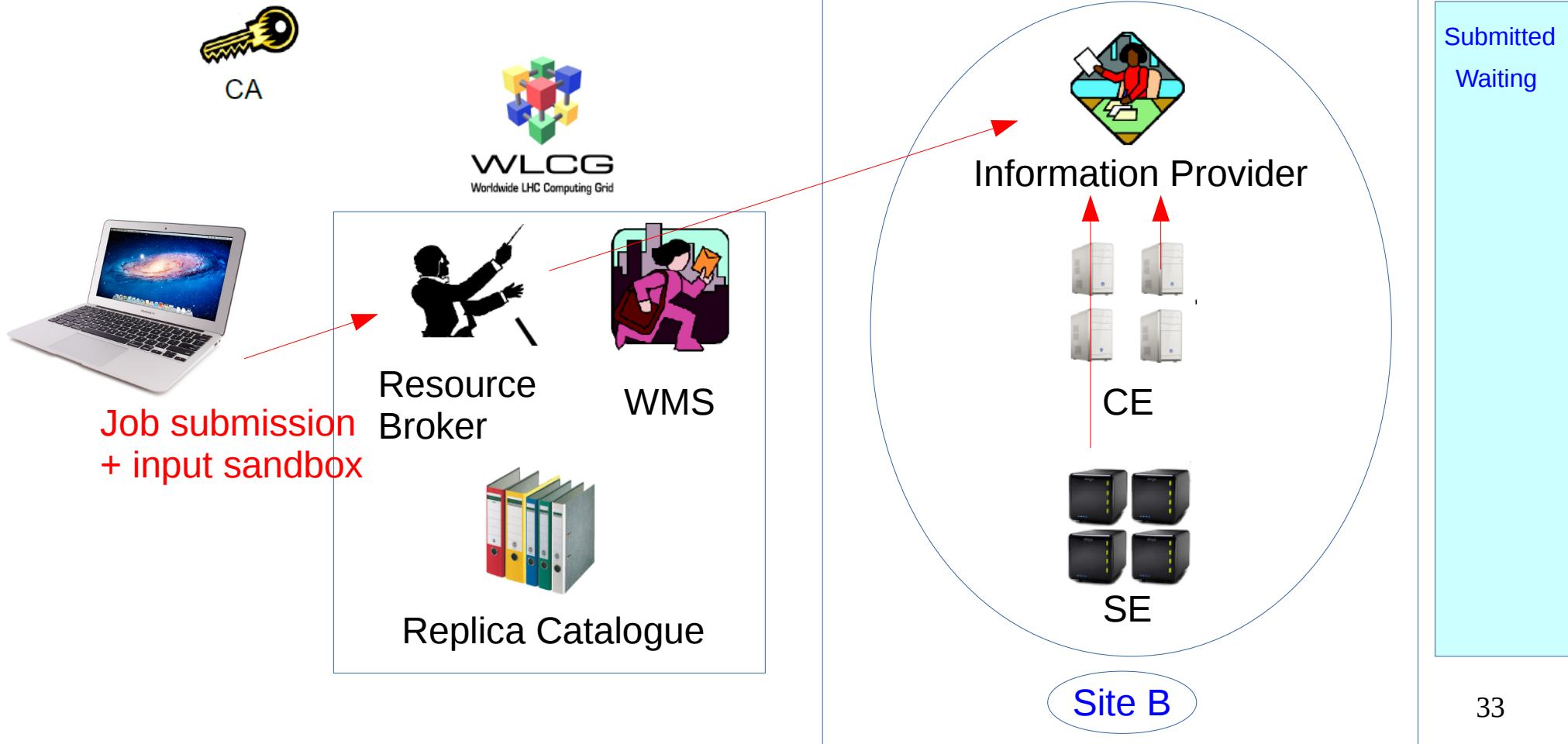


# Grid job and data - 2



**WLCG**  
Worldwide LHC Computing Grid

- How do they work?

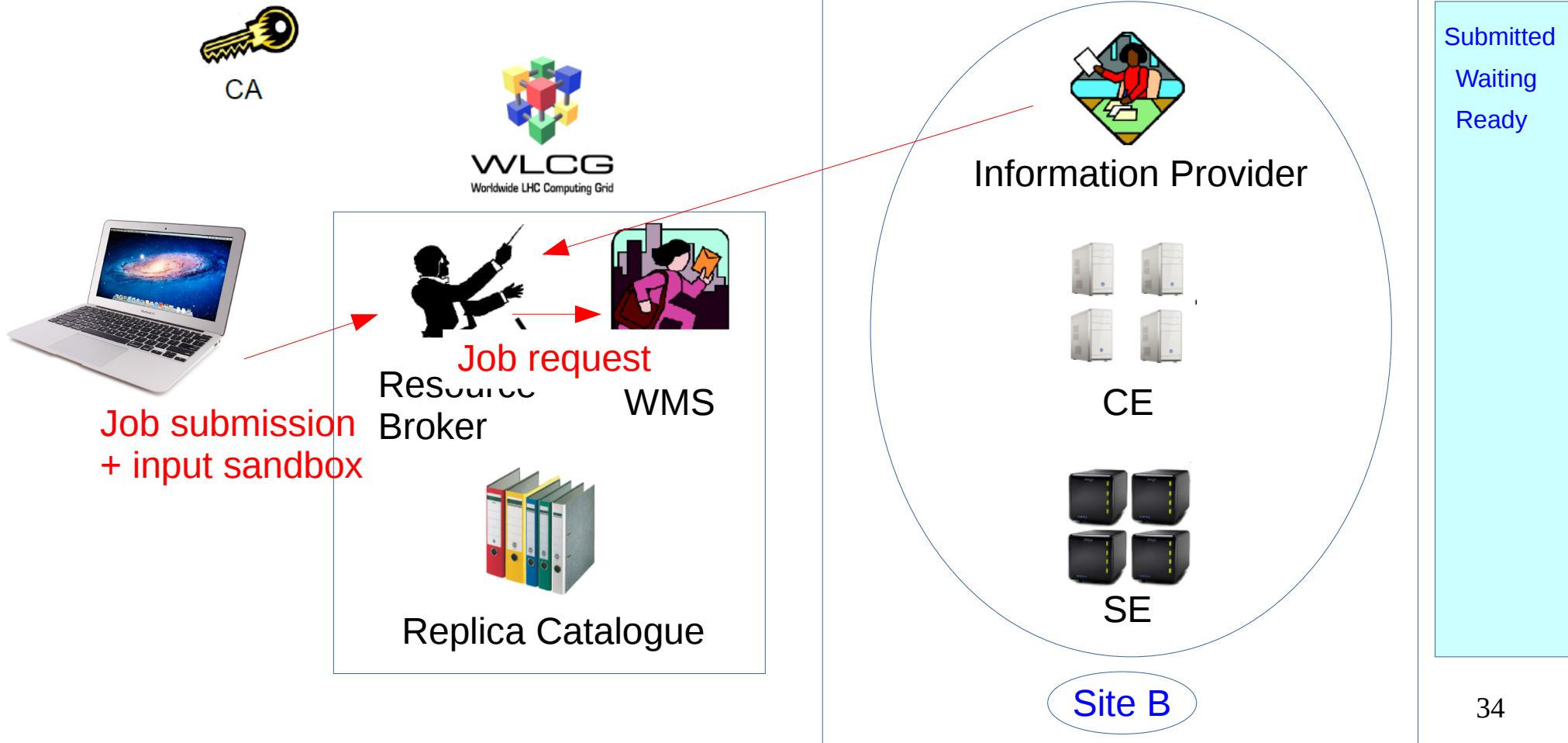


# Grid job and data - 2



**WLCG**  
Worldwide LHC Computing Grid

- How do they work?

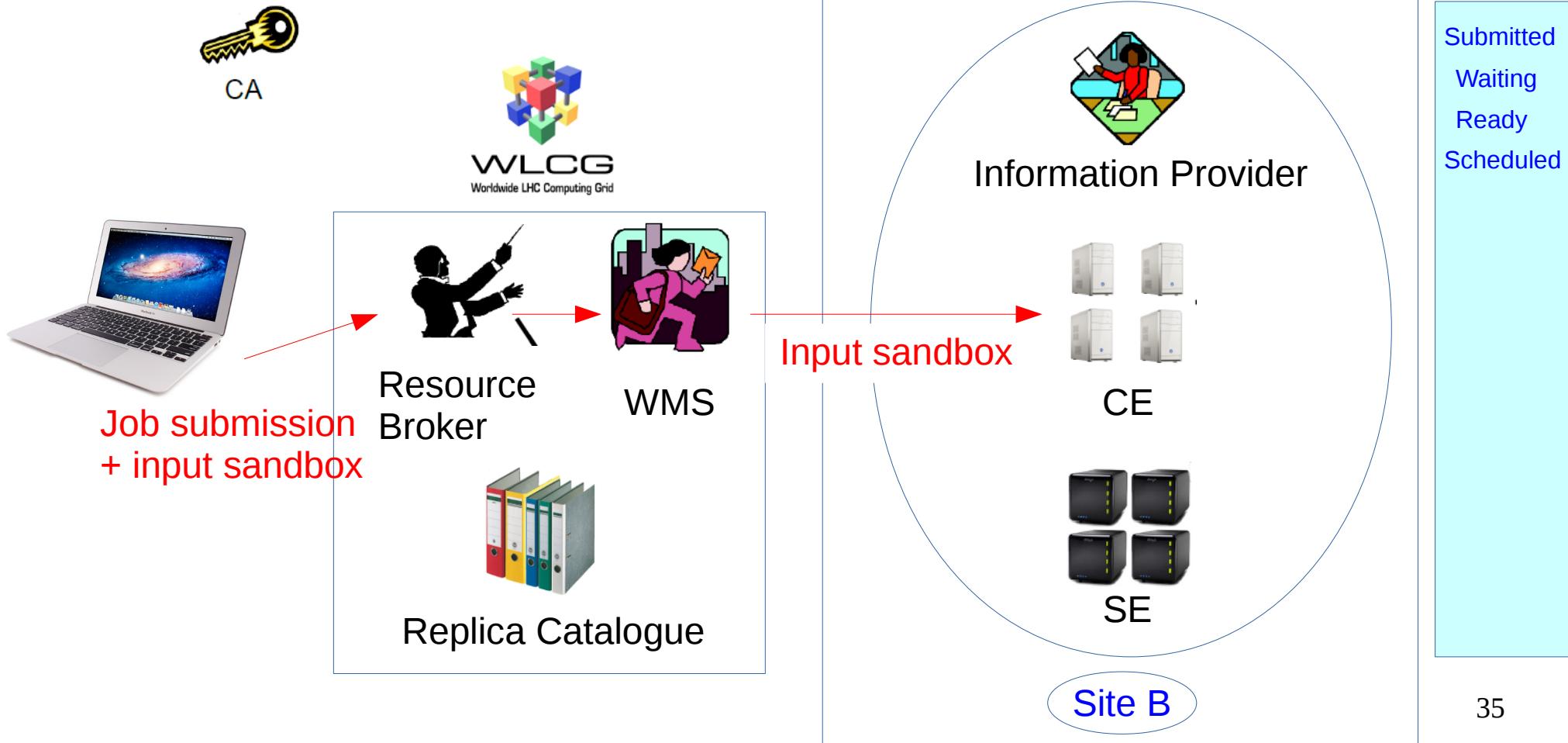


# Grid job and data - 2



**WLCG**  
Worldwide LHC Computing Grid

- How do they work?



# Grid job and data - 2

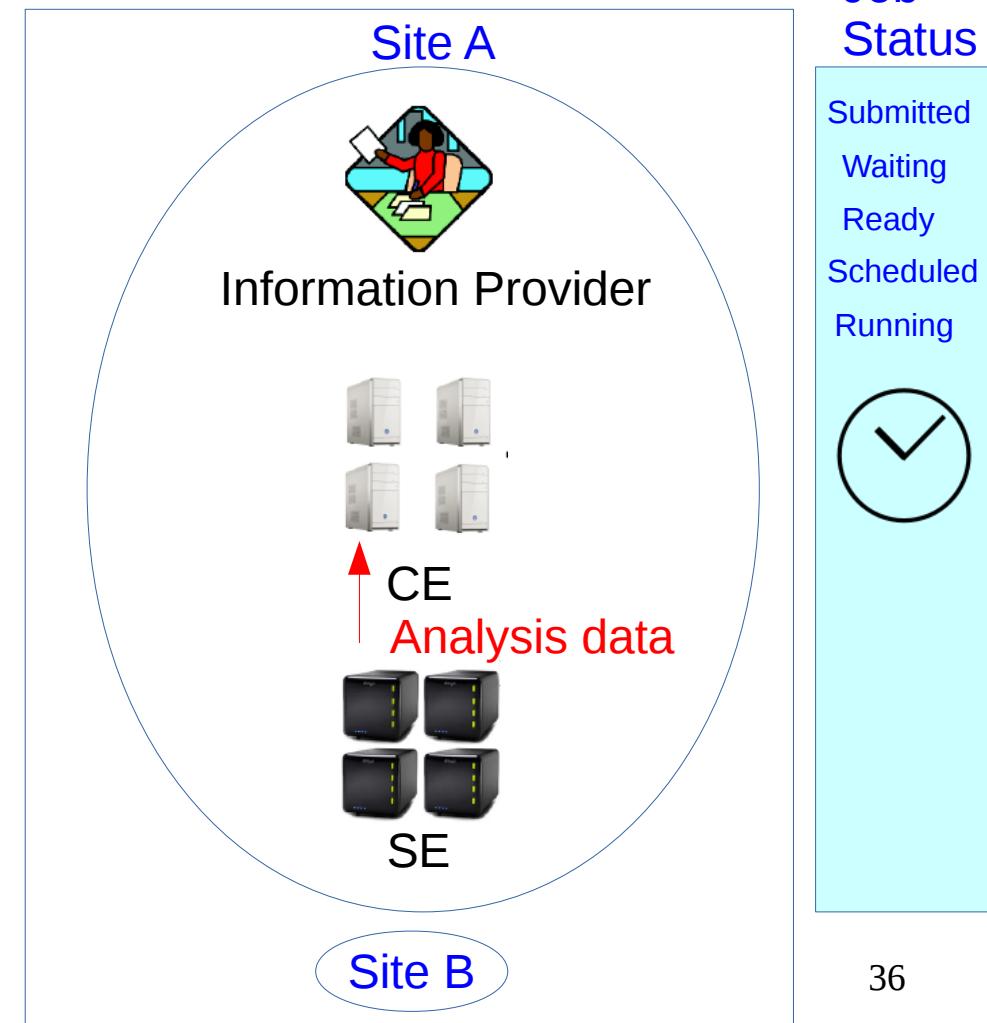
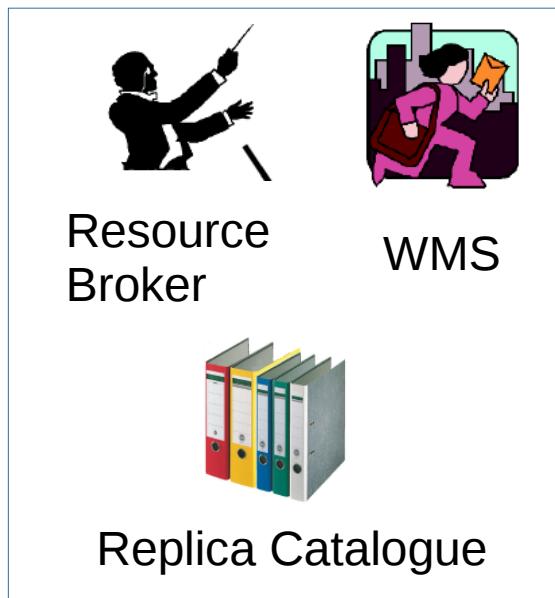


**WLCG**  
Worldwide LHC Computing Grid

- How do they work?



CA



# Grid job and data - 2



**WLCG**  
Worldwide LHC Computing Grid

- How do they work?



CA



**WLCG**  
Worldwide LHC Computing Grid

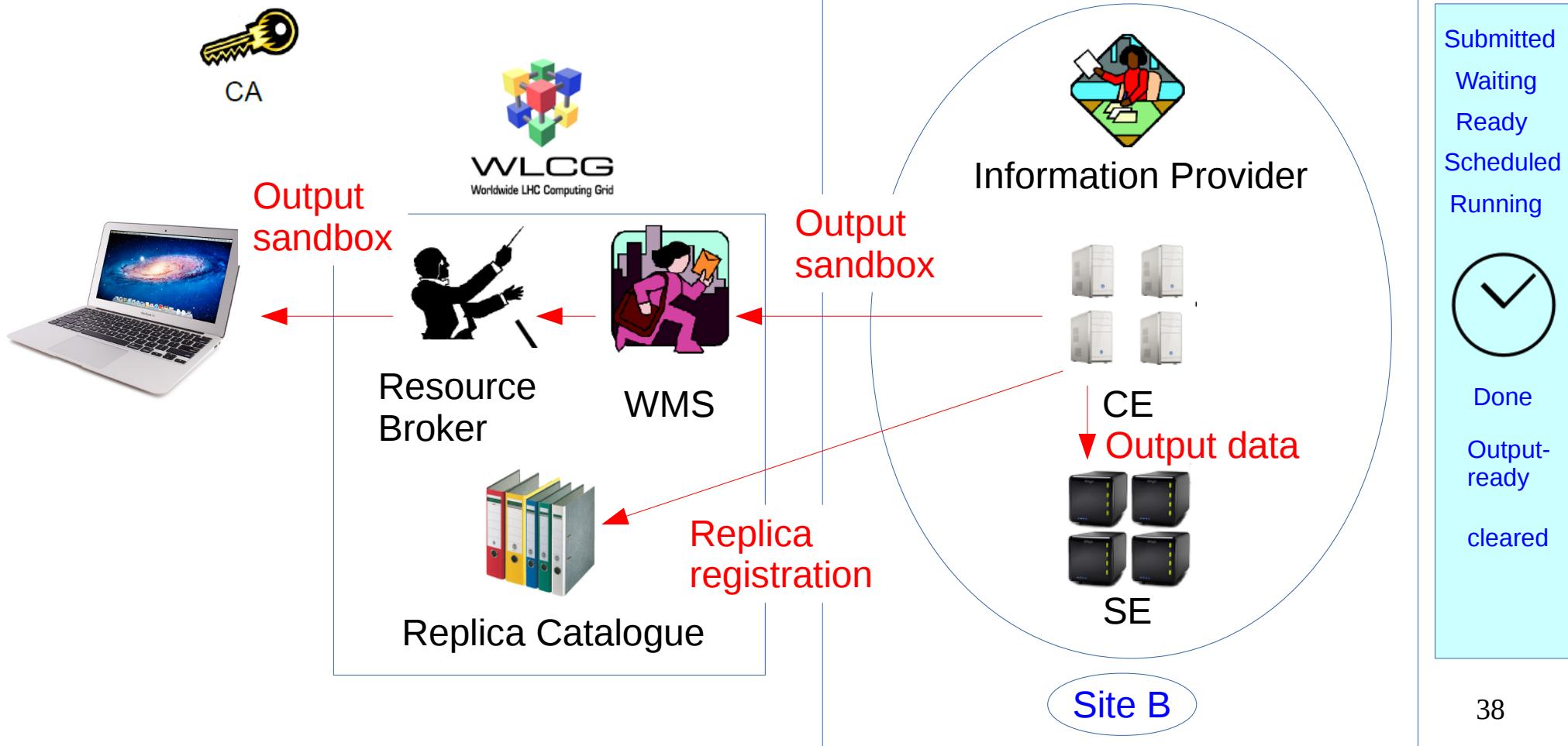


# Grid job and data - 2



**WLCG**  
Worldwide LHC Computing Grid

- How do they work?

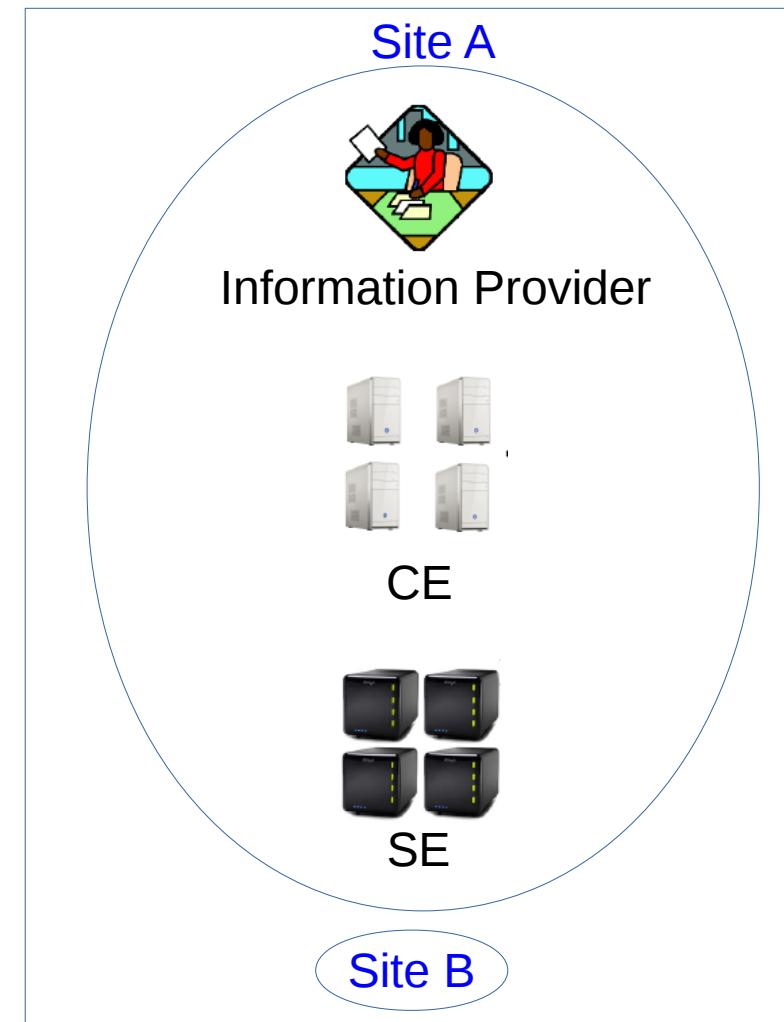
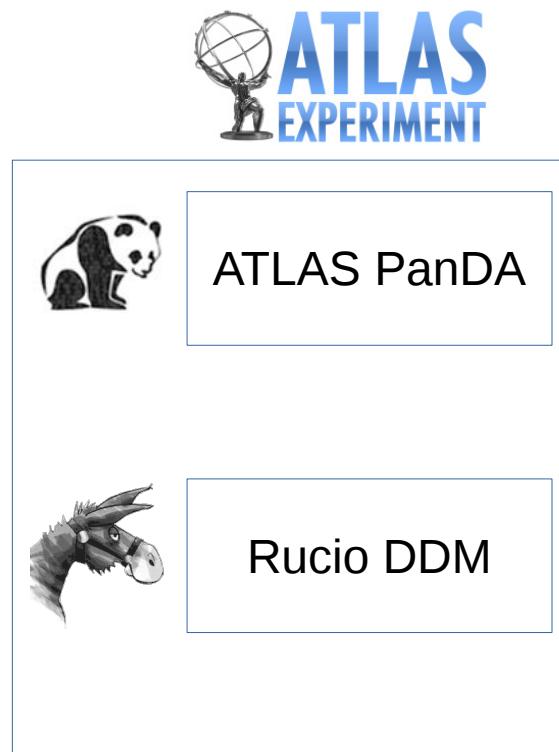


# Grid job and data - 3



**WLCG**  
Worldwide LHC Computing Grid

- How do they work?

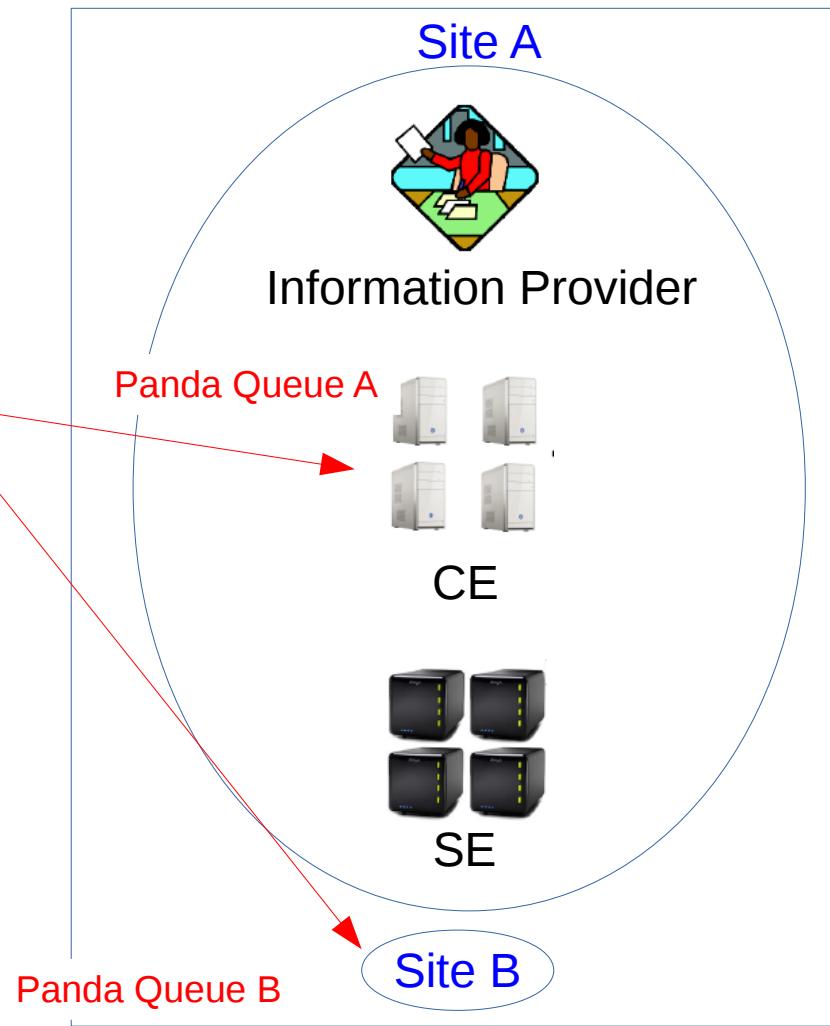
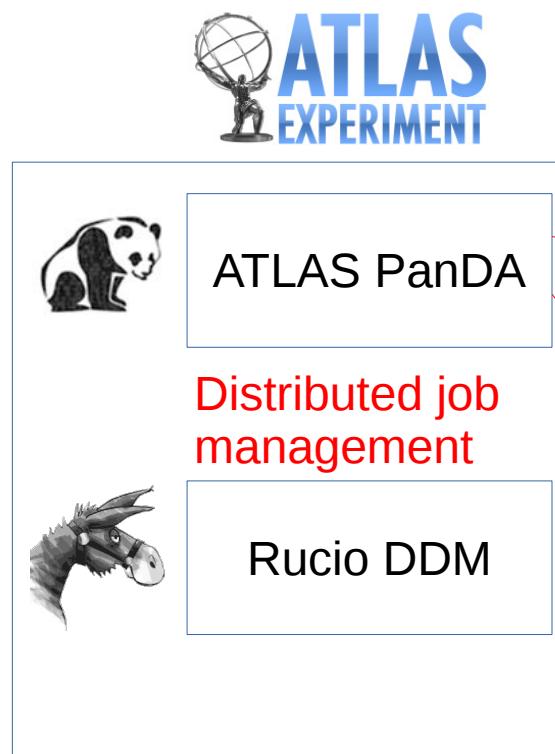


# Grid job and data - 3



**WLCG**  
Worldwide LHC Computing Grid

- How do they work?

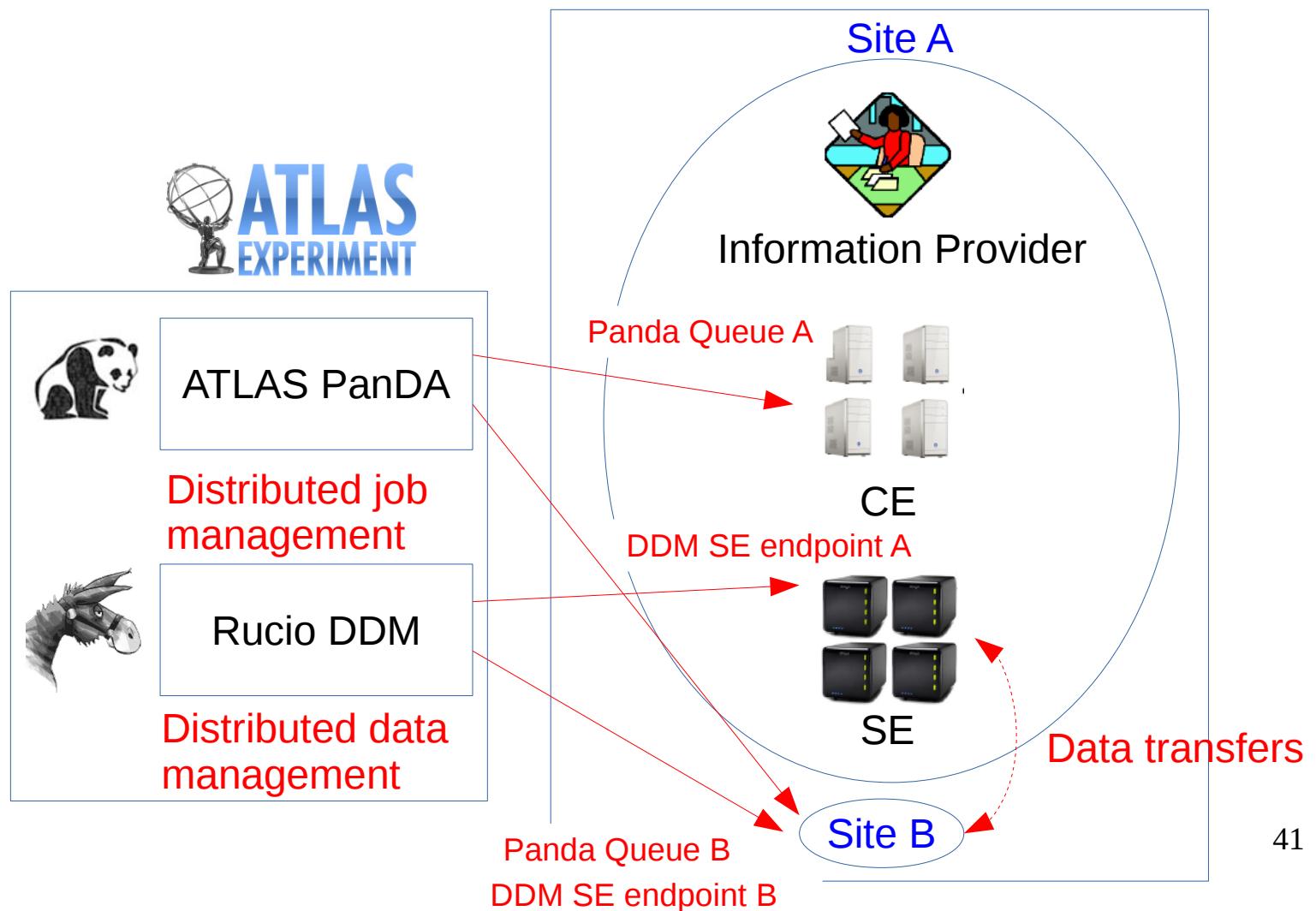


# Grid job and data - 3



**WLCG**  
Worldwide LHC Computing Grid

- How do they work?



# ATLAS Resources - 1

- AGIS (ATLAS Grid Information System)
  - <http://atlas-agis.cern.ch/agis/>

ATLAS Grid Information System			
RC Site	ATLASSite	DDMEndpoint	PANDA Queue
Service	Central Services	DDM Groups	Docs
<ul style="list-style-type: none"><li>▪ Define RC site</li><li>▪ Define Experiment site</li><li>▪ Define DDM endpoint</li><li>▪ <b>Define OS RSE endpoint (new implementation)</b></li><li>▪ Define PANDA site</li><li>▪ Define PANDA queue</li><li>▪ RC pledges</li><li>▪ Find DDM endpoints links</li><li>▪ Find TransferMatrix links</li></ul>	<ul style="list-style-type: none"><li>▪ <b>Define OS service</b></li><li>▪ Define LFC service</li><li>▪ Define SE service</li><li>▪ Define CE service</li><li>▪ Define Redirector service</li><li>▪ Define PerfSonar service</li><li>▪ Define Frontier service</li><li>▪ Define Squid service</li><li>▪ Define Central service</li><li>▪ <b>SE protocols (DDM/Panda activities)</b></li></ul>	<ul style="list-style-type: none"><li>▪ Cron list</li><li>▪ ADMINs list</li><li>▪ Changes log</li><li>▪ <b>Request ADMIN privileges</b></li></ul>	<ul style="list-style-type: none"><li>▪ Main TWiki</li><li>▪ TWiki WEBUI instructions</li><li>▪ API Docs</li></ul>
DOWNTIMES	TOACACHE EXPORT	COMPARISON & VALIDATION TOOLS	
<ul style="list-style-type: none"><li>▪ Downtime calendar</li><li>▪ DDM Blacklisting data</li><li>▪ PANDA Blacklisting data</li></ul>	<ul style="list-style-type: none"><li>▪ <b>dynamic ToACache (changes are immediately propagated):</b> <a href="http://atlas-agis-api.cern.ch/request/toacache/TiersOfATLASCache.py">http://atlas-agis-api.cern.ch/request/toacache/TiersOfATLASCache.py</a></li><li>▪ <b>static ToACache:</b> <a href="http://atlas-agis-api.cern.ch/ToACache/TiersOfATLASCache.py">http://atlas-agis-api.cern.ch/ToACache/TiersOfATLASCache.py</a></li><li>▪ <b>previous caches:</b> <a href="http://atlas-agis-api.cern.ch/ToACache/cache/">http://atlas-agis-api.cern.ch/ToACache/cache/</a><ul style="list-style-type: none"><li>▪ ToACache with Extra data</li></ul></li><li>▪ View/Modify ToACache ExtraData (RSE integration)</li></ul>	<ul style="list-style-type: none"><li>▪ Consistency checker</li><li>▪ ToAComparator</li><li>▪ AGIS-BDII CE comparison</li><li>▪ AGIS-Schedconf-PF mon CE comparison</li><li>▪ AGIS-DIMGOCDB sites+services comparison</li><li>▪ AGIS-PANDA PandaResource+SWReleases comparison</li><li>▪ AGIS-Schedconfig (topology) comparison</li><li>▪ AGIS-Schedconfig JSON comparison</li><li>▪ AGIS-GSR services comparison</li></ul>	

# ATLAS Resources - 2

- PanDA queue end points

ATLAS Grid Information System

RC Site	ATLASSite	DDMEndpoint	PANDA Queue	Service	Central Services	DDM Groups	PandaQueue combined resources						Docs	TWiki	OLD	JSON						
Show 200 entries				First	Previous	1	Next	Last														
<a href="#">give me url of this page</a> <a href="#">hold shift click column for Multi-column ordering</a>															Status	Manual	HC	Switcher	Panda Integration	CLOUD	TIER	
atlas			FZK						ACTIVE													
VO name	▲	ATLAS Site	▲	PanDA Site	▲	Template object	▲	PanDA Resource	▲	PanDA Queue	▲	state	▲	(current) status	▲	type	▲	capability	▲	CLOUD	▲	TIER
atlas		FZK-LCG2		FZK-LCG2		FZK-LCG2_VIRTUAL		ANALY_FZK		Clone ANALY_FZK		ACTIVE		online		analysis		score		DE		T1
atlas		FZK-LCG2		FZK-LCG2		FZK-LCG2_VIRTUAL		ANALY_FZK_HI		Clone ANALY_FZK_HI		ACTIVE		online		analysis		score		DE		T1
atlas		FZK-LCG2		FZK-LCG2		FZK-LCG2_VIRTUAL		ANALY_FZK_SHORT		Clone ANALY_FZK_SHORT		ACTIVE		online		analysis		score		DE		T1
atlas		FZK-LCG2		FZK-LCG2		FZK-LCG2_VIRTUAL		FZK-LCG2		Clone FZK-LCG2-all-prod-CEs		ACTIVE		online		production		score		DE		T1
atlas		FZK-LCG2		FZK-LCG2		FZK-LCG2_VIRTUAL		FZK-LCG2_HIMEM		Clone FZK-LCG2_HIMEM		ACTIVE		online		production		himem		DE		T1
atlas		FZK-LCG2		FZK-LCG2		FZK-LCG2_VIRTUAL		FZK-LCG2_MCORE		Clone FZK-LCG2_MCORE		ACTIVE		online		production		mcose		DE		T1
atlas		FZK-LCG2		FZK-LCG2		FZK-LCG2_VIRTUAL		FZK-LCG2_MCORE_HI		Clone FZK-LCG2_MCORE_HI		ACTIVE		online		production		mcose		DE		T1
atlas		FZK-LCG2		FZK-LCG2		FZK-LCG2_VIRTUAL		FZK-LCG2_MCORE_LO		Clone FZK-LCG2_MCORE_LO		ACTIVE		online		production		mcose		DE		T1

Active PanDA (job) queues  
ANALY\_ .... = Analysis queue

# ATLAS Resources - 3

- DDM end points

**ATLAS Grid Information System**

RC Site ATLASTSite **DDMEndpoint** PANDA Queue Service Central Services DDM Groups **DDM Endpoints** Docs TWiki OLD JSON

Show 200 entries **FZK** First Previous 1 Next Last

give me url of this page hold shift + click column for Multi-column ordering

DDM Endpoint	State	DDM Site	ATLAS Site	ATLAS TIER	CLOUD	type	Full Endpoint	FTS Master	FTS Test
FZK-LCG2_DATADISK	ACTIVE	FZK-LCG2	FZK-LCG2	T1	DE	DATADISK	token:ATLASDATADISK:srm://atlassrm-fzk.gridka.de:8443/srm/managerv?SFN=/pnfs/gridka.de/atlas/disk-only/atlasdatadisk/	CERNFTS3_REST   https://fts3.cern.ch:8446	CERNFTS3PILOT_REST   https://fts3-pilot.cern.ch:8446
FZK-LCG2_DATATAPE	ACTIVE	FZK-LCG2	FZK-LCG2	T1	DE	DATATAPE	token:ATLASDATATAPE:srm://atlassrm-fzk.gridka.de:8443/srm/managerv?SFN=/pnfs/gridka.de/atlas/atlasdatatape/	CERNFTS3_REST   https://fts3.cern.ch:8446	CERNFTS3PILOT_REST   https://fts3-pilot.cern.ch:8446
FZK-LCG2_GROUPTAPE_PERF-EGAMMA	ACTIVE	FZK-LCG2	FZK-LCG2	T1	DE	GROUPTAPE	token:ATLASMCTAPE:srm://atlassrm-fzk.gridka.de:8443/srm/managerv?SFN=/pnfs/gridka.de/atlas/atlasgrouptape/perf-egamma/	CERNFTS3_REST   https://fts3.cern.ch:8446	CERNFTS3PILOT_REST   https://fts3-pilot.cern.ch:8446
FZK-LCG2_LOCALGROUPDISK	ACTIVE	FZK-LCG2	FZK-LCG2	T1	DE	LOCALGROUPDISK	token:ATLASLOCALGROUPDISK:srm://dgridsrm-fzk.gridka.de:8443/srm/managerv?SFN=/pnfs/gridka.de/datas/atlaslocalgroupdisk/	CERNFTS3_REST   https://fts3.cern.ch:8446	CERNFTS3PILOT_REST   https://fts3-pilot.cern.ch:8446
FZK-LCG2_LOCALGROUPTAPE	ACTIVE	FZK-LCG2	FZK-LCG2	T1	DE	LOCALGROUPTAPE	token:ATLASLOCALGROUPTAPE:srm://dgridsrm-fzk.gridka.de:8443/srm/managerv?SFN=/pnfs/gridka.de/datas/atlaslocalgrouptape/	CERNFTS3_REST   https://fts3.cern.ch:8446	CERNFTS3PILOT_REST   https://fts3-pilot.cern.ch:8446
FZK-LCG2_MCTAPE	ACTIVE	FZK-LCG2	FZK-LCG2	T1	DE	MCTAPE	token:ATLASMCTAPE:srm://atlassrm-fzk.gridka.de:8443/srm/managerv?SFN=/pnfs/gridka.de/atlas/atlasmctape/	CERNFTS3_REST   https://fts3.cern.ch:8446	CERNFTS3PILOT_REST   https://fts3-pilot.cern.ch:8446
FZK-LCG2_PERF-EGAMMA	ACTIVE	FZK-LCG2	FZK-LCG2	T1	DE	GROUPDISK	token:ATLASDATADISK:srm://atlassrm-fzk.gridka.de:8443/srm/managerv?SFN=/pnfs/gridka.de/atlas/disk-only/atlasgroupdisk/perf-egamma/	CERNFTS3_REST   https://fts3.cern.ch:8446	CERNFTS3PILOT_REST   https://fts3-pilot.cern.ch:8446
FZK-LCG2_PERF-IDTRACKING	ACTIVE	FZK-LCG2	FZK-LCG2	T1	DE	GROUPDISK	token:ATLASDATADISK:srm://atlassrm-fzk.gridka.de:8443/srm/managerv?SFN=/pnfs/gridka.de/atlas/disk-only/atlasgroupdisk/perf-idtracking/	CERNFTS3_REST   https://fts3.cern.ch:8446	CERNFTS3PILOT_REST   https://fts3-pilot.cern.ch:8446
FZK-LCG2_PERF-TAU	ACTIVE	FZK-LCG2	FZK-LCG2	T1	DE	GROUPDISK	token:ATLASDATADISK:srm://atlassrm-fzk.gridka.de:8443/srm/managerv?SFN=/pnfs/gridka.de/atlas/disk-only/atlasgroupdisk/perf-tau/	CERNFTS3_REST   https://fts3.cern.ch:8446	CERNFTS3PILOT_REST   https://fts3-pilot.cern.ch:8446
FZK-LCG2_PPSSCRATCHDISK	ACTIVE	FZK-LCG2	FZK-LCG2	T1	DE	SPECIAL	token:ATLASPPSSCRATCHDISK:srm://ppssrm-kit.gridka.de:8443/srm/managerv?SFN=/pnfs/gridka.de/atlas/atlasppsscratchdisk/	CERNFTS3_REST   https://fts3.cern.ch:8446	
FZK-LCG2_SCRATCHDISK	ACTIVE	FZK-LCG2	FZK-LCG2	T1	DE	SCRATCHDISK	token:ATLASSCRATCHDISK:srm://atlassrm-fzk.gridka.de:8443/srm/managerv?SFN=/pnfs/gridka.de/atlas/disk-only/atlasscratchdisk/	CERNFTS3_REST   https://fts3.cern.ch:8446	CERNFTS3PILOT_REST   https://fts3-pilot.cern.ch:8446

Showing 1-11 of 11 entries

**Active DDM storage end points**

# ATLAS Resources - 4

- **SCRATCHDISK** (Tier1 + Tier2s in Germany)
  - FZK-LCG2\_SCRATCHDISK
  - DESY-HH\_SCRATCHDISK
  - DESY-ZN\_SCRATCHDISK
  - LRZ-LMU\_SCRATCHDISK
  - WUPPERTALPROD\_SCRATCHDISK
  - UNI-FREIBURG\_SCRATCHDISK
  - GOEGRID\_SCRATCHDISK
- **LOCALGROUPDISK** (e.g. DESY-HH and UniGoettingen)
  - DESY-HH\_LOCALGROUPDISK
  - GOEGRID\_LOCALGROUPDISK
  - ..... \_LOCALGROUPDISK

# ATLAS Resources - 4

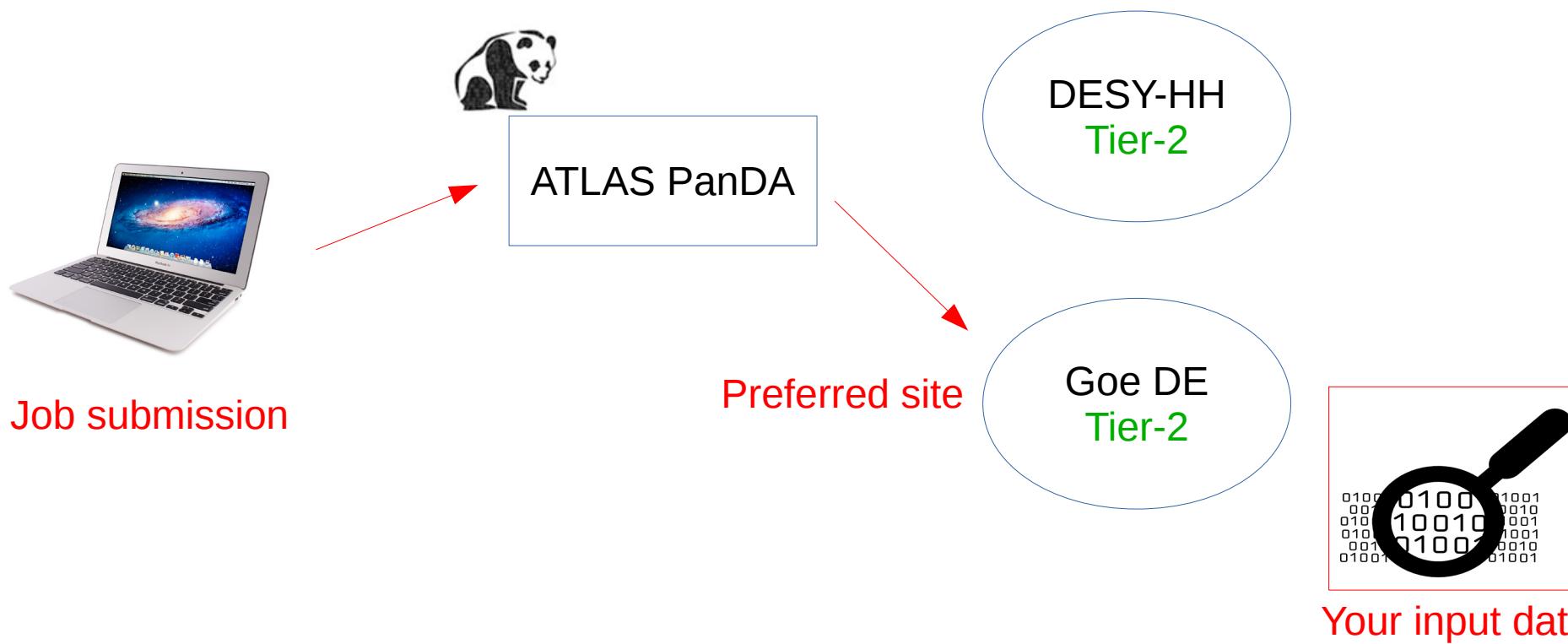
- **SCRATCHDISK** (Tier1 + Tier2s in Germany)
  - FZK-LCG2\_SCRATCHDISK
  - DESY-HH\_SCRATCHDISK
  - DESY-ZN\_SCRATCHDISK
  - LRZ-LMU\_SCRATCHDISK
  - WUPPERTALPROD\_SCRATCHDISK
  - UNI-FREIBURG\_SCRATCHDISK
  - GOEGRID\_SCRATCHDISK

Storages for temporary data of PanDA jobs. Would be automatically ***REMOVED!***
- **LOCALGROUPDISK** (e.g. DESY-HH and UniGoettingen)
  - DESY-HH\_LOCALGROUPDISK
  - GOEGRID\_LOCALGROUPDISK
  - ..... \_LOCALGROUPDISK

Permanently ***KEPT.*** Generally speaking, in total a few hundred TB in each site

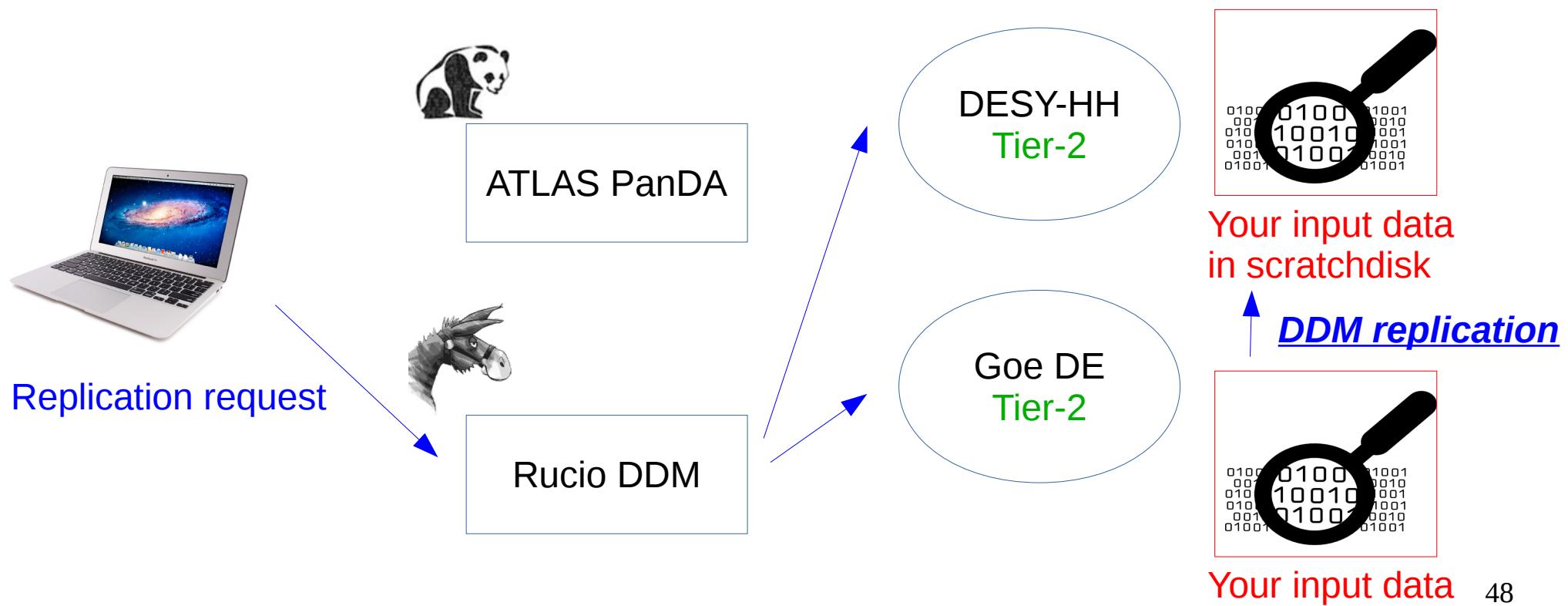
# ATLAS Resources - 5

- A tip: User job allocation policy among sites
  - Rule: Grid jobs (should) go to their data locations



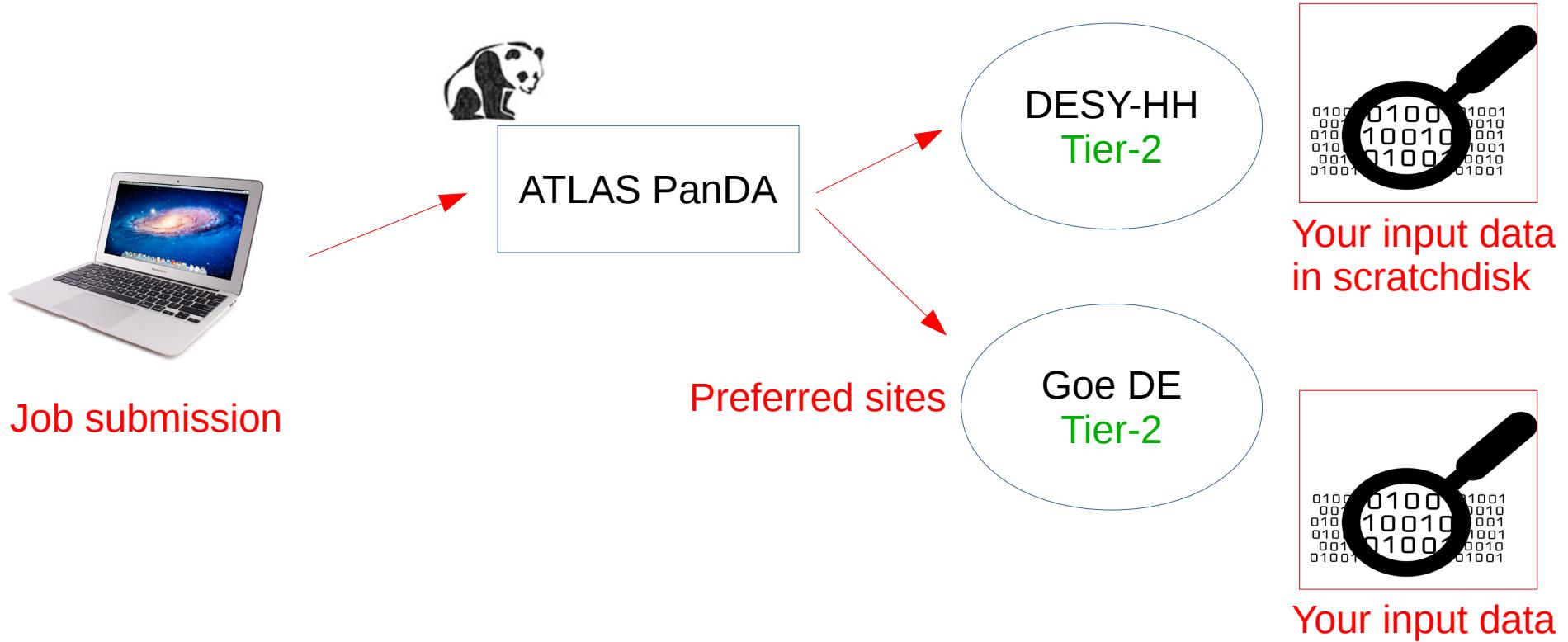
# ATLAS Resources - 5

- A tip: User job allocation policy among sites
  - Rule: Grid jobs (should) go to their data locations



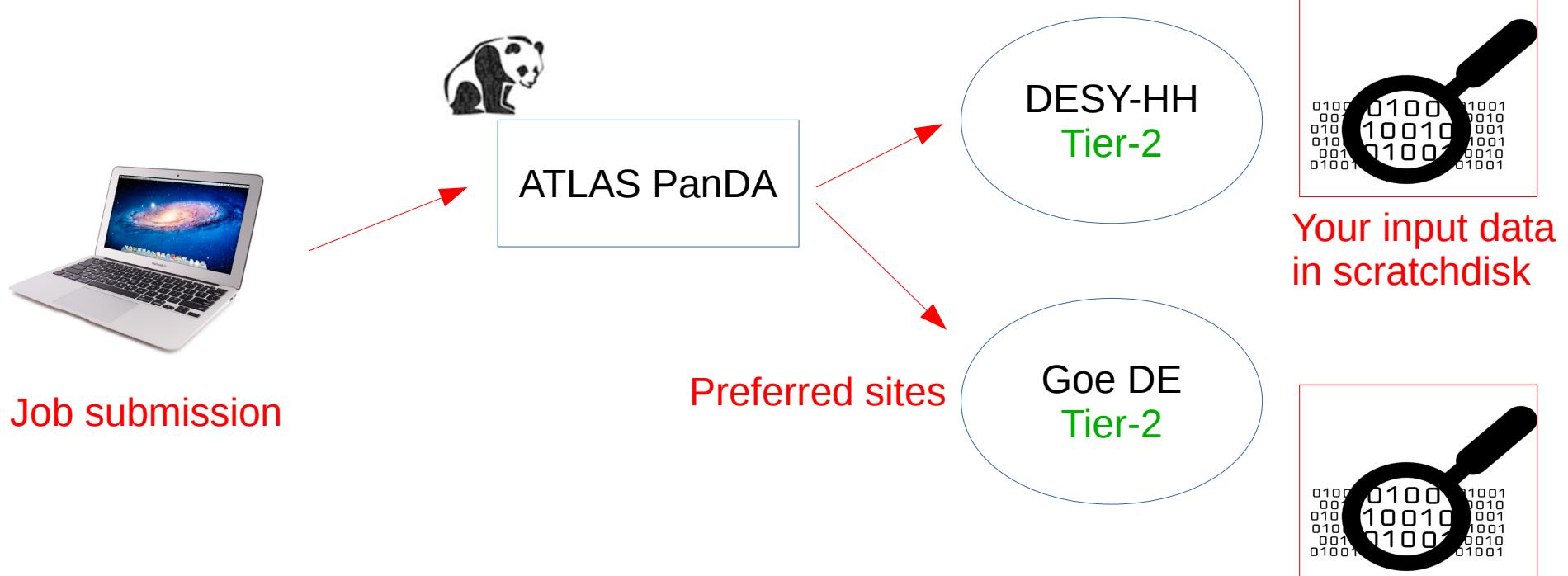
# ATLAS Resources - 5

- A tip: User job allocation policy among sites
  - Rule: Grid jobs (should) go to their data locations



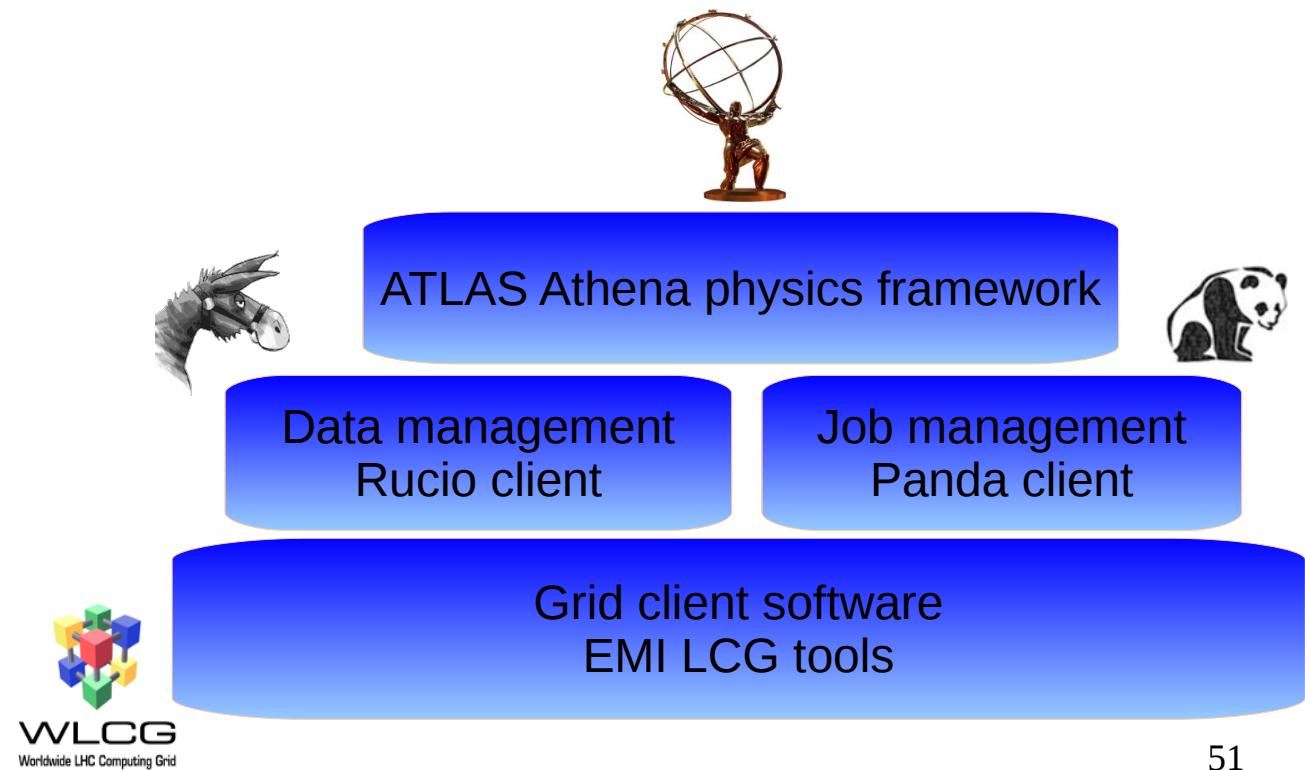
# ATLAS Resources - 5

- A tip: User job allocation policy among sites
  - Rule: Grid jobs (should) go to their data locations



**Local access is better than remote access**  
(or to avoid an error in PanDA, “local data does not exist”)

# Grid User Interface and CVMFS



# Hands-on exercise 3

- Grid UI

# Hands-on exercise 4

- Grid UI

# Hands-on exercise 5

- First PanDA job submission

# Rucio (ATLAS data management system)



# Basic concepts

- A replacement of data management system for Run-2 in ATLAS computing system
  - Old one was “DQ2”
    - DQ2 will be completely removed in Dec., 2016

# Setup Rucio

# Searching for data (CLI)

# Listing metadata (CLI)

# Download data (CLI)

# Moving data (WebUI)

# Links and references

- RucioUI
  - <https://rucio-ui.cern.ch/>
- Rucio Documentation
  - <http://rucio.cern.ch/index.html>
- Software twiki tutorial
  - <https://twiki.cern.ch/twiki/bin/viewauth/AtlasComputing/SoftwareTutorialGettingDatasets>
- Athena Docker setup
  - <https://twiki.cern.ch/twiki/bin/viewauth/AtlasComputing/AthenaMacDockerSetup>
- Docker container for CVMFS
  - <https://github.com/sbinet/docker-containers/tree/master/cvmfs-atlas>
- Binet, Sébastien, and Ben Couturier. "*docker & HEP: Containerization of applications for development, distribution and preservation.*" Journal of Physics: Conference Series. Vol. 664. No. 2. IOP Publishing, 2015.
  - <http://iopscience.iop.org/article/10.1088/1742-6596/664/2/022007/meta>
- B
  -
- C
  -
- D
  -