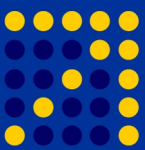


The Information System

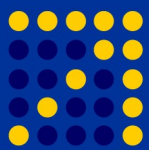
LA ROCCA, Giuseppe
INFN Catania
giuseppe.larocca@ct.infn.it





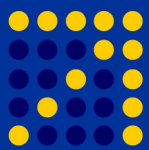
How to discover resources ?

- Once an user is logged into an User Interface (s)he is ready to take advantage of the Grid Power for his/her own application.
- But what are the available resources to accomplish his/her tasks?
- The answer to this question comes through the interactions with the **Information System (IS)**.
- The Information System (IS) provides information about the LCG-2 Grid resources and their status.



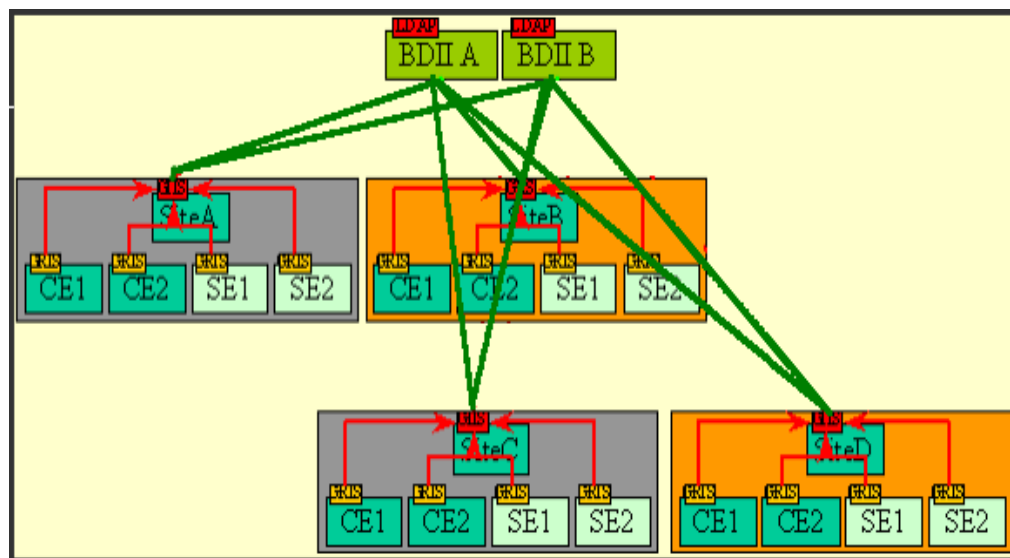
How to discover resources (cont)

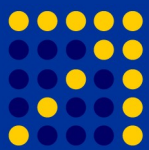
- The data published in the IS conforms to the GLUE (Grid Laboratory for a Uniform Environment) Schema. The **GLUE Schema** aims to define a common conceptual data model to be used for Grid resources
- In LCG-2, the *BDII* (*Berkeley DB Information Index*), based on an updated version of the **Monitoring and Discovery Service (MDS)**, was adopted as main provider of the Information Service
- In gLite, *R-GMA* (*Relational Grid Monitoring Architecture*) is adopted as IS



Monitoring and Discovery Service

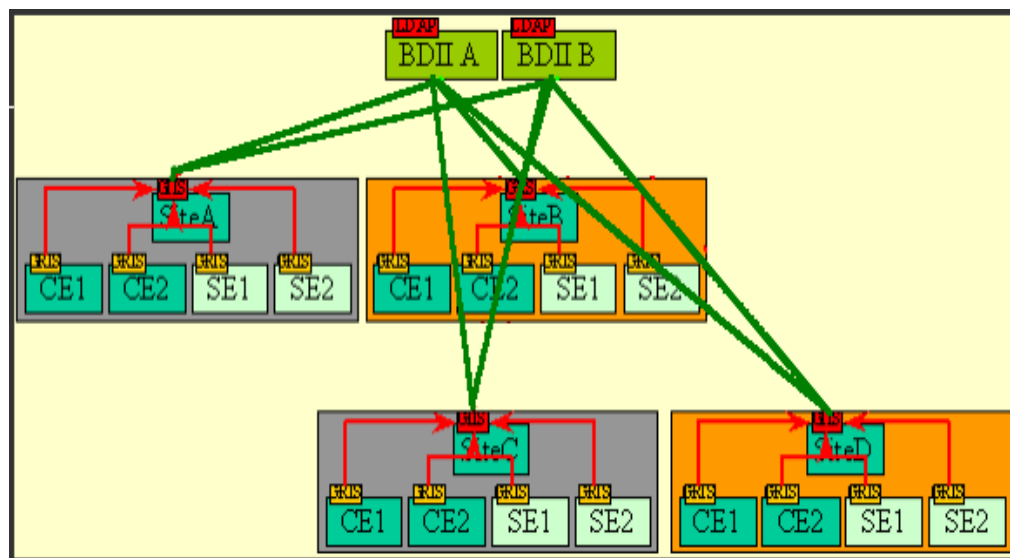
- Computing and storage resources at a site implement an entity called **Information Provider**, which generates the relevant information of the resource (e.g.: the used space in a SE).
- This information is published by the **Grid Resource Information Servers**, or GRISes.

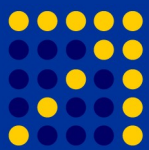




Monitoring and Discovery Service

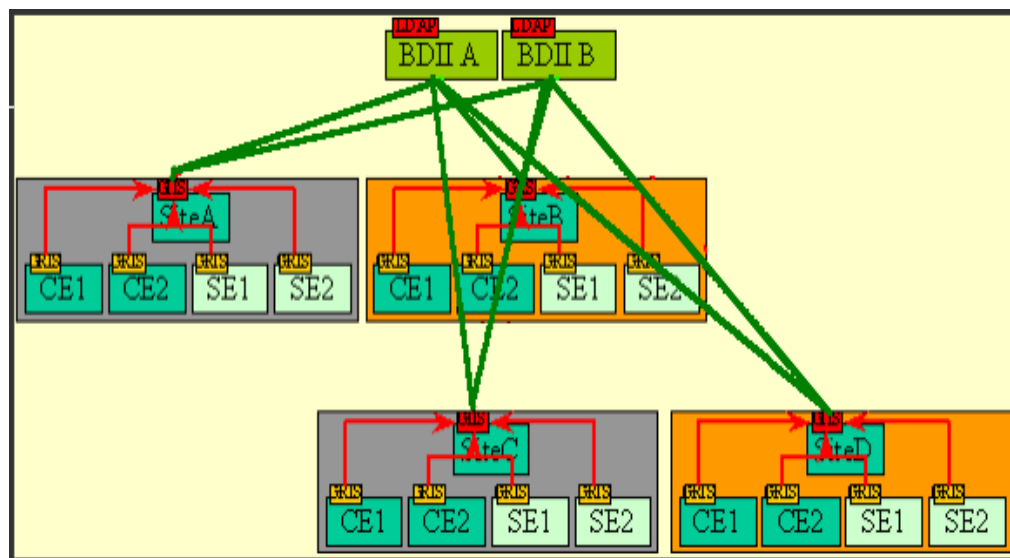
- In each site an element called the Site **Grid Index Information Server** (GIIS) collects all the information of the different GRISes and publishes it.
- This BDII queries the GIISes and acts as a cache, storing information about the Grid status in its database.

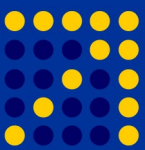




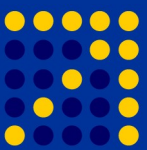
Monitoring and Discovery Service

- Querying the BDII a user or a service has all the available information about the status of the grid resources.
- Moreover in order to get more up-to-date information it is possible to querying directly the GIISes or GRISes.





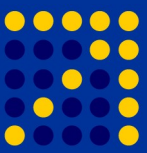
- The local GRISes runs on Computing Elements and Storage Elements and reports information on the characteristics and status of the services.
 - They give both static and dynamic information.
- In order to interrogate the GRIS on a specific Grid Element
 - the *hostname* (- **h**) of the Grid Element and the *TCP port* where the GRIS run must be specified (- **p**).
 - Port is always 2135.
 - - **x** option indicates that simple authentication should be used;
 - - **b** option is used to specify the initial entry from which starts the search in the LDAP tree.



E.g.: `$ ldapsearch -x -h <hostname>`
`-p 2135`
`-b "mds-vo-name=local, o=grid"`

or

`$ ldapsearch -x -H <LDAP_URI>`
`-b "mds-vo-name=local, o=grid"`



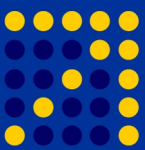
Interrogating the GRIS on a CE

- The command used to interrogate the GRIS located on host `lxn1181.cern.ch` is:

```
$ ldapsearch -x  
-h lxn1181.cern.ch  
-p 2135  
-b "mds-vo-name=local, o=grid"
```

or:

```
$ ldapsearch -x  
-H ldap://lxn1181.cern.ch:2135  
-b "mds-vo-name=local, o=grid"
```



version: 2

#

filter: (objectclass=*)

requesting: ALL

#

lx1181.cern.ch/siteinfo, local, grid

dn: in=lx1181.cern.ch/siteinfo,Mds-Vo-name=local,o=grid

objectClass: SiteInfo

objectClass: DataGridTop

objectClass: DynamicObject

siteName: CERN-LCG2

sysAdminContact: hep-project-grid-cern-testbed-managers@cern.ch

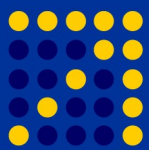
userSupportContact: hep-project-grid-cern-testbed-managers@cern.ch

siteSecurityContact: hep-project-grid-cern-testbed-managers@cern.ch

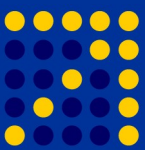
dataGridVersion: LCG-2_0_0beta

installationDate: 20040106120000Z

[..]



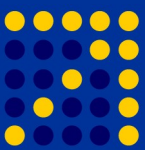
- At each site, a Site GIIIS collects information about all resources coming from all the GRISes.
- Usually a site GIIIS runs on a Computing Element.
- In order to interrogate the Site GIIIS
 - the *hostname* (- **h**) of the Grid Element and the *TCP port* where the GIIIS run must be specified (- **p**).
 - Port is always 2135.
 - - **x** option indicates that simple authentication should be used;
 - - **b** option is used to specify the initial entry from which starts the search in the LDAP tree.
 - *A different base name must be used !*



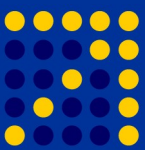
Interrogating the Site GLS

- The command used to interrogate the Site GLS located on lcgce02.ifae.es is:

```
$ ldapsearch -x  
-H ldap://lcgce02.ifae.es:2170  
-b "mds-vo-name=piclcg2,o=grid"
```



version: 2
#
filter: (objectclass=*)
requesting: ALL
#
https://edt003.cnaf.infn.it:7772, infn-cnaf, grid
dn: GlueServiceURI=https://edt003.cnaf.infn.it:7772,Mds-Vo-name=infn-cnaf,o=gr
id
objectClass: GlueService
objectClass: GlueSchemaVersion
GlueServiceURI: https://edt003.cnaf.infn.it:7772
GlueServiceAccessPointURL: https://edt003.cnaf.infn.it:7772
GlueServiceType: ResourceBroker
GlueServicePrimaryOwnerName: LCG
GlueServicePrimaryOwnerContact: mailto:sitemanager@cnaf.infn.it
GlueServiceHostingOrganization: INFN-CNAF
GlueServiceMajorVersion: 1
GlueServiceMinorVersion: 00
[...]



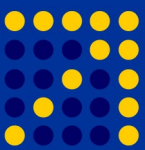
How to query the IS?

- In order to query directly the IS elements two high level tools are provided.

lcg-infosites

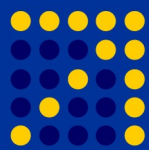
lcg-info

These tools should be enough for most common user needs and will usually avoid the necessary of raw LDAP queries.



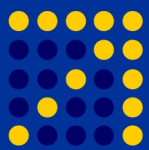
- The **lcg-infosites** command can be used as an easy way to retrieve information on Grid resources for the most use cases.

USAGE: lcg-infosites --vo <vo name> options
-v <verbose level>
--is <BDII to query>



lcg-infosites options and other

ce	The information related to number of CPUs, running jobs, waiting jobs and names of the CEs are provided. All these data group all VOs together. With "-v 1" only the names of the queues will be printed while with "-v 2" The RAM Memory together with the operating system and its version and the processor included in each CE are printed.
se	The names of the SEs supported by the user's VO together with the kind of Storage System, the used and available space will be printed. With "-v 1" only the names of the SEs will be printed.
closeSE	The names of the CEs where the user's VO is allowed to run together with their corresponding closest SEs are provided.
lfc	Name of the lfc Catalog for the user's VO.
tag	The names of the tags relative to the software installed in site is printed together with the corresponding CE.
all	It groups together the information provided by ce, se, lfc and rmc.
is	If not specified the BDII defined in default by the variable LCG GFAL INFOSYS will be queried. However the user may want to query any other BDII without redefining this environment variable. This is possible specifying this argument followed by the name of the BDII which the user wants to query. All options admit this argument.



Obtaining information about CE

BioinfoGRID

```
$ lcg-infosites --vo gilda ce
```

```
*****
```

These are the related data for gilda: (in terms of queues and CPUs)

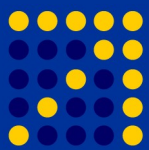
```
*****
```

#CPU	Free	Total Jobs	Running	Waiting	Computing	Element
4	3	0	0	0	0	cn01.be.itu.edu.tr:2119/jobmanager-lcglsf-long
4	3	0	0	0	0	cn01.be.itu.edu.tr:2119/jobmanager-lcglsf-short
34	33	0	0	0	0	grid010.ct.infn.it:2119/jobmanager-lcgpbs-long
16	16	0	0	0	0	grid011f.cnaf.infn.it:2119/jobmanager-lcgpbs-long
1	1	0	0	0	0	grid006.cecalc.ula.ve:2119/jobmanager-lcgpbs-log
2	1	1	0	1	1	gildace.oact.inaf.it:2119/jobmanager-lcgpbs-short

```
$ lcg-infosites --vo gilda ce -v 2
```

RAM	Memory	Operating System	System Version	Processor	CE Name
1024		SLC	3	P4	ced-ce0.datagrid.cnr.it
4096		SLC	3	Xeon	cn01.be.itu.edu.tr
1024		SLC	3	PIII	cna02.cna.unicamp.br
917		SLC	3	PIII	gilda-ce-01.pd.infn.it
1024		SLC	3	Athlon	gildace.oact.inaf.it
1024		SLC	3	Xeon	grid-ce.bio.dist.unige.it





Obtaining information about SE

BioinfoGRID

```
$ lcg-infosites --vo gilda se
```

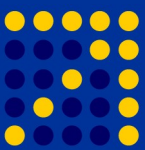
```
*****
```

These are the related data for gilda: (in terms of SE)

```
*****
```

Avail Space(Kb)	Used Space(Kb)	Type	SEs
143547680	2472756	disk	cn02.be.itu.edu.tr
168727984	118549624	disk	grid009.ct.infn.it
13908644	2819288	disk	grid003.cecalc.ula.ve
108741124	2442872	disk	gildase.oact.inaf.it
28211488	2948292	disk	testbed005.cnaf.infn.it
349001680	33028	disk	gilda-se-01.pd.infn.it
31724384	2819596	disk	cna03.cna.unicamp.br
387834656	629136	disk	grid-se.bio.dist.unige.it





Listing the close Storage Elements

```
$ lcg-infosites --vo gilda closeSE
```

Name of the CE: cn01.be.itu.edu.tr:2119/jobmanager-lcglsf-long

Name of the close SE: cn02.be.itu.edu.tr

Name of the CE: cn01.be.itu.edu.tr:2119/jobmanager-lcglsf-short

Name of the close SE: cn02.be.itu.edu.tr

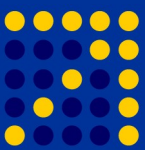
Name of the CE: grid010.ct.infn.it:2119/jobmanager-lcgpbs-long

Name of the close SE: grid009.ct.infn.it

Name of the CE: grid011f.cnaf.infn.it:2119/jobmanager-lcgpbs-long

Name of the close SE: testbed005.cnaf.infn.it





Listing tags of installed software

```
$ lcg-infosites --vo gilda tag
```

Information for gilda relative to their software tags included in each CE

Name of the TAG: VO-gilda-GEANT

Name of the TAG: VO-gilda-GKS05

Name of the CE:cn01.be.itu.edu.tr

Name of the TAG: VO-gilda-slc3_ia32_gcc323

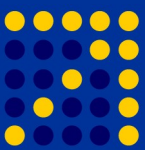
Name of the TAG: VO-gilda-CMKIN_5_1_1

Name of the TAG: VO-gilda-GEANT

Name of the TAG: VO-gilda-GKS05

Name of the CE:grid010.ct.infn.it





- This command can be used to list either CEs or the SEs that satisfy a given set of conditions, and to print the values of a given set of attributes.
- The information is taken from the BDII specified by the **LCG_GFAL_INFOSYS** environment variable.

- The query syntax is like this:

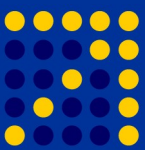
attr1 op1 valueN, ...
attrN opN valueN

After the upgrading
of the new
GLUE SCHEMA it's
not possible
use the operators
'>' and '<'

where *attrN* is an attribute name

op is =, >= or <=, and the cuts are ANDed.

The cuts are comma-separated and spaces are not allowed.

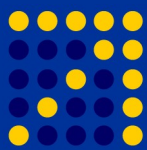


lcg-info --list-ce [--bdii bdii] [--vo vo] [--sed]
[--query query] [--attrs list]

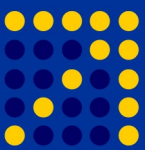
lcg-info --list-se [--bdii bdii] [--vo vo] [--sed]
[--query query] [--attrs list]

lcg-info --list-attrs

lcg-info --help



--list-attrs	Prints a list of the attributes that can be queried.
--list-ce	Lists the CEs which satisfy a query, or all the CEs if no query is given.
--list-se	Lists the SEs which satisfy a query, or all the SEs if no query is given.
--query	Restricts the output to the CEs (SEs) which satisfy the given query.
--bdii	Allows to specify a BDII in the form <code>:. If not given, the value of the environmental variable LCG_GFAL_INFOSYS is used. If that is not defined, the command returns an error.</code>
--sed	Print the output in a "sed-friendly" format.
--attrs	Specifies the attributes whose values should be printed.
--vo	Restricts the output to CEs or SEs where the given VO is authorized. Mandatory when VO-dependent attributes are queried upon.



\$ lcg-info --list-attrs

Attribute name

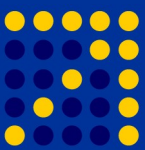
MaxTime
CEStatus
TotalJobs
CEVOs
TotalCPUs
FreeCPUs
CE
WaitingJobs
RunningJobs
CloseCE
CloseSE
SEVOs
UsedSpace
AvailableSpace
Type
SE
Protocol
ArchType
Processor
OS
Cluster
Tag
Memory

Glue object class

GlueCE
GlueCE
GlueCE
GlueCE
GlueCE
GlueCE
GlueCE
GlueCE
GlueCE
GlueCESEBindGroup
GlueCESEBindGroup
GlueSA
GlueSA
GlueSA
GlueSE
GlueSE
GlueSEAccessProtocol
GlueSL
GlueSubCluster
GlueSubCluster
GlueSubCluster
GlueSubCluster
GlueSubCluster

Glue attribute name

GlueCEPolicyMaxWallClockTime
GlueCEStateStatus
GlueCEStateTotalJobs
GlueCEAccessControlBaseRule
GlueCEInfoTotalCPUs
GlueCEStateFreeCPUs
GlueCEUniqueID
GlueCEStateWaitingJobs
GlueCEStateRunningJobs
GlueCESEBindGroupCEUniqueID
GlueCESEBindGroupSEUniqueID
GlueSAAccessControlBaseRule
GlueSAStateUsedSpace
GlueSAStateAvailableSpace
GlueSEType
GlueSEUniqueID
GlueSEAccessProtocolType
GlueSLArchitectureType
GlueHostProcessorModel
GlueHostOperatingSystemName
GlueSubClusterUniqueID
GlueHostApplicationSoftwareRunTimeEnvironment
GlueHostMainMemoryRAMSize

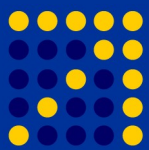


- List all the CE(s) in the BDII satisfying given conditions

```
$ lcg-info --vo gilda --list-ce  
--query 'TotalCPUs=10,OS=SL*'  
--attrs 'RunningJobs,FreeCPUs'
```

- CE: dgt01.ui.savba.sk:2119/jobmanager-lcgpbs-long
 - RunningJobs 0
 - FreeCPUs 10
- CE: dgt01.ui.savba.sk:2119/jobmanager-lcgpbs-short
 - RunningJobs 0
 - FreeCPUs 10
- CE: dgt01.ui.savba.sk:2119/jobmanager-lcgpbs-infinite
 - RunningJobs 1
 - FreeCPUs 10
- CE: gilda-ce-01.pd.infn.it:2119/jobmanager-lcgpbs-long
 - RunningJobs 0
 - FreeCPUs 10





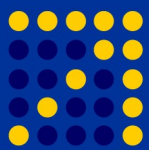
Print all the tags published by a specific query

```
$ lcg-info --vo gilda --list-ce  
--query 'CE=*grid-ce.bio.dist.unige.it*'  
--attrs 'Tag'
```

LCG-2_1_1
LCG-2_2_0
LCG-2_3_0
LCG-2_3_1
LCG-2_4_0
R-GMA
AFS
CMS-1.1.0
ATLAS-6.0.4
GATE-1.0.0-3
LHCb-1.1.1
IDL-5.4
CMSIM-125
ALICE-4.01.00
ALIEN-1.32.14
POVRAY-3.5
DEMTOOLS-1.0

CSOUND-4.13
MPICH
VIRGO-1.0
CMS-OSCAR-2.4.5
LHCb_dbase_common-v3r1
GEANT4-6
VLC-0.7.2
EGEODE-1.0
RASTER3D
SCILAB-2.6
G95-3.5.0
MAGIC-6.19
CODESA3D-1.0
VO-gilda-slc3_ia32_gcc323
VO-gilda-CMKIN_5_1_1
VO-gilda-GEANT
VO-gilda-GKS05





List the CEs with a particular SW

```
$ lcg-info --vo gilda  
      --list-ce  
      --query 'Tag=*MPICH*'  
      --attrs 'CE'
```

- CE: cn01.be.itu.edu.tr:2119/jobmanager-lcglsf-long
 - CE cn01.be.itu.edu.tr:2119/jobmanager-lcglsf-long
- CE: cn01.be.itu.edu.tr:2119/jobmanager-lcglsf-short
 - CE cn01.be.itu.edu.tr:2119/jobmanager-lcglsf-short
- CE: grid010.ct.infn.it:2119/jobmanager-lcgpbs-long
 - CE grid010.ct.infn.it:2119/jobmanager-lcgpbs-long
- CE: grid011f.cnaf.infn.it:2119/jobmanager-lcgpbs-long
 - CE grid011f.cnaf.infn.it:2119/jobmanager-lcgpbs-long
- CE: ced-ce0.datagrid.cnr.it:2119/jobmanager-lcgpbs-long
 - CE ced-ce0.datagrid.cnr.it:2119/jobmanager-lcgpbs-long



