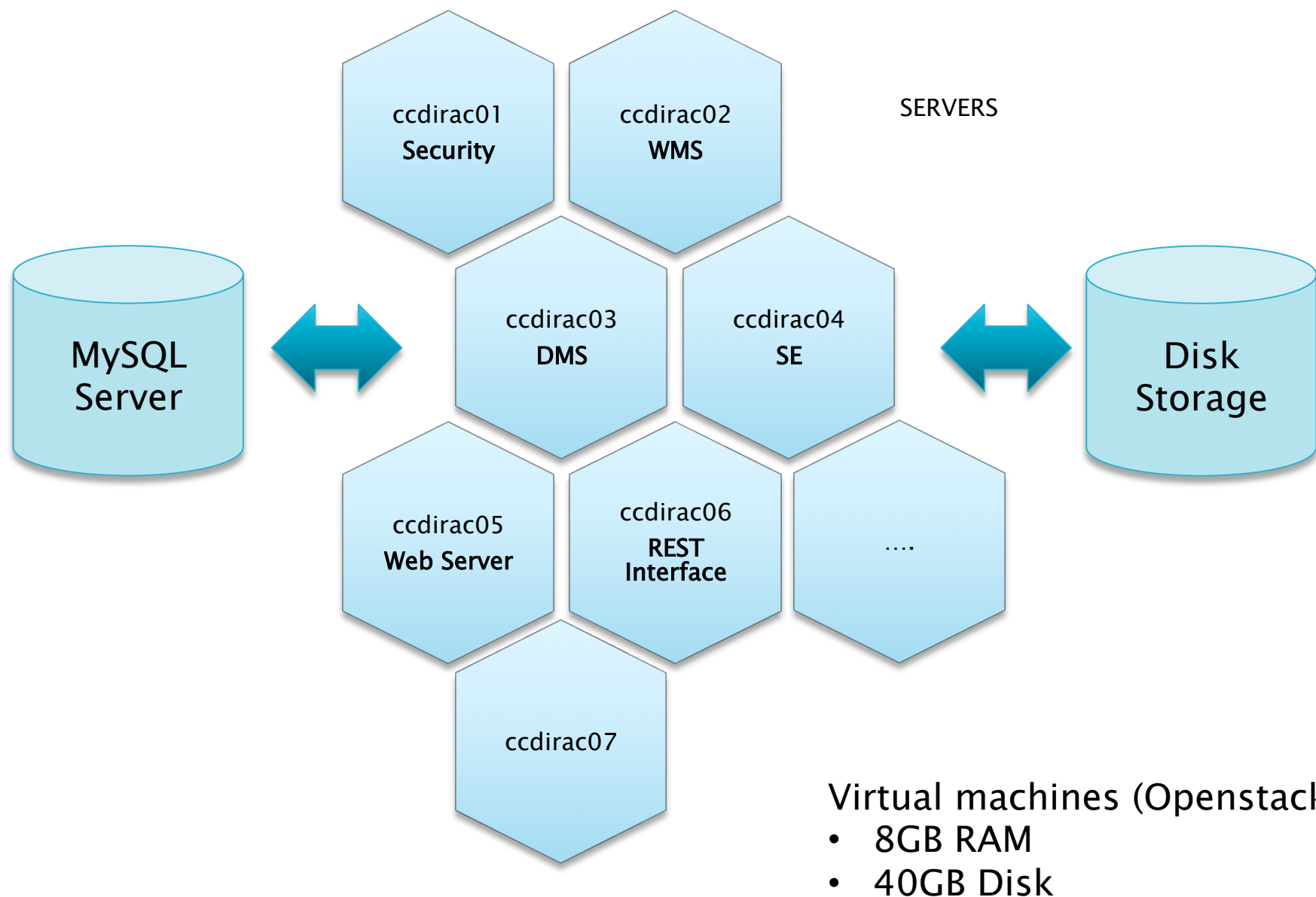


Multi-VO installation based on the FG-DIRAC example

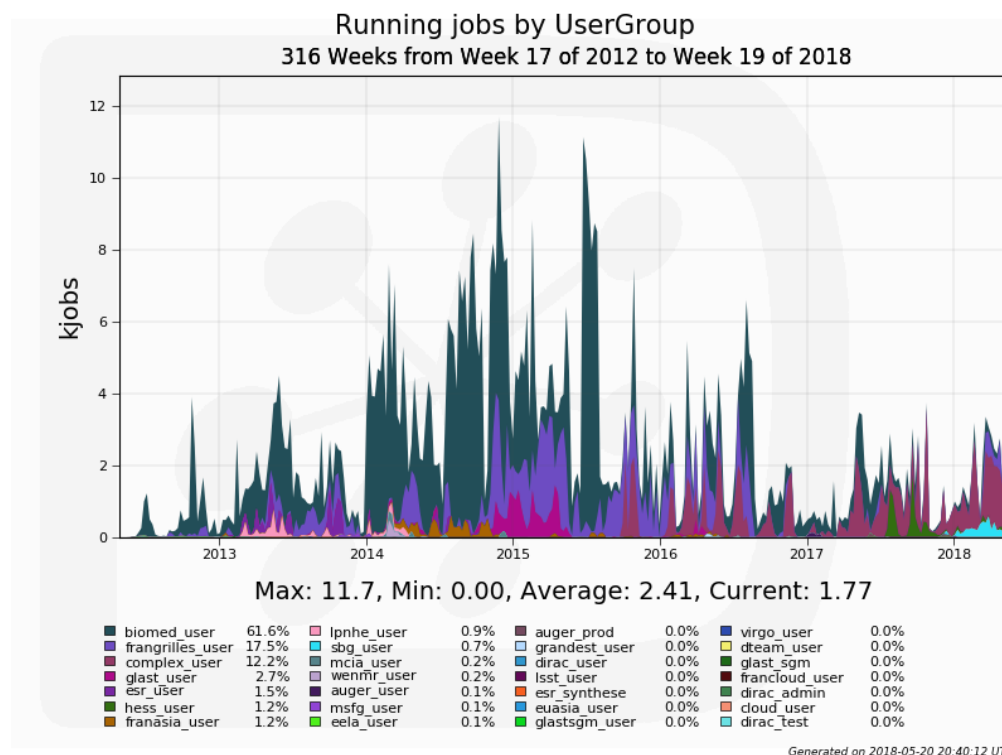
Vanessa HAMAR

- ▶ DIRAC Multi-VO installation – why?
- ▶ FG-DIRAC instance
- ▶ Configuration changes (and a wish list 😊)
 - Registry
 - Resources
 - Operations
- ▶ Conclusions

- ▶ Originally DIRAC was developed as a single community system to manage distributed computing resources
- ▶ Small communities could not afford installation and management of a fully functional DIRAC service
 - No expertise
 - Too complicated
- ▶ France-Grilles was the first grid infrastructure project to offer DIRAC services to its users
 - Starting from 2012



- ▶ 22 VOs configured
- ▶ > 25 user groups
- ▶ > 74.000.000 user jobs since May-2012
- ▶ 2 robot certificats



▶ Definition of a VO

- VO admin
- VOMS information
 - The information on VOMS servers must be kept up to date

▶ Definition of DIRAC groups belonging to the VO

- Minimal set of groups vo_user, vo_pilot, vo_admin
- More groups with special properties can be defined
- For each group should be defined
 - Properties defining rights of the group members
 - VOMS roles mapping

VO administration: Registry

The screenshot displays the Configuration Manager interface for Dirac-Prod. The main pane shows a tree view of the configuration structure. The 'Registry' folder is expanded, revealing several configuration items:

- DefaultGroup**: A list of users including wenmr_user, lsst_user, lphhe_user, mcia_user, superb_user, msfg_user, glast_user, gilda_user, frangrilles_user, franasia_user, euasia_user, esr_user, eela_user, dirac_user, biomed_user, auger_user, astro_user, and grande.
- DefaultProxyLifeTime**: Set to 40000.
- Users**, **Groups**, **Hosts**, and **VO** folders are also visible in the tree.
- Under the **VO** folder, the **biomed** folder is expanded, showing:
 - SubmitPools**: biomedPool
 - VOAdmin**: vhamar
 - VOAdminGroup**: dirac_admin
 - VOMSName**: biomed
 - DefaultGroup**: biomed_user
- Under the **VOMSServers** folder, the **cclcgvomsl01.in2p3.fr** folder is expanded, showing:
 - DN**: /O=GRID-FR/C=FR/O=CNRS/OU=CC-IN2P3/CN=cclcgvomsl01.in2p3.fr
 - Port**: 15000
 - CA**: /C=FR/O=CNRS/CN=GRID2-FR
- Other folders visible include **astro.vo.eu-egee.org**, **auger**, **esr**, and **euasia**.

The interface includes a top toolbar with 'View as Text', 'Download', and 'Reload' buttons, and a right sidebar with 'Show diff.', 'Show history', and 'Manage' options.

- ▶ Adding users
 - By hand
 - Using `dirac-admin-add-user` command
 - Editing CS with the web editor
 - Running VOMS2CSAgent

- ▶ Managing users requires sometimes manual intervention to edit user data
 - Using general CS editor
 - With many hundreds of registered users is quite impractical

- ▶ Wish list:
 - Need for a functional web based Registry editor
 - The existing editor should be considerably improved

▶ VOMS2CSAgent

- Automatically synchronizes VOMS and Registry information
 - Adds users
 - Modifies user info
 - Optionally creates user home directories in the File Catalog
 - Optionally deletes users

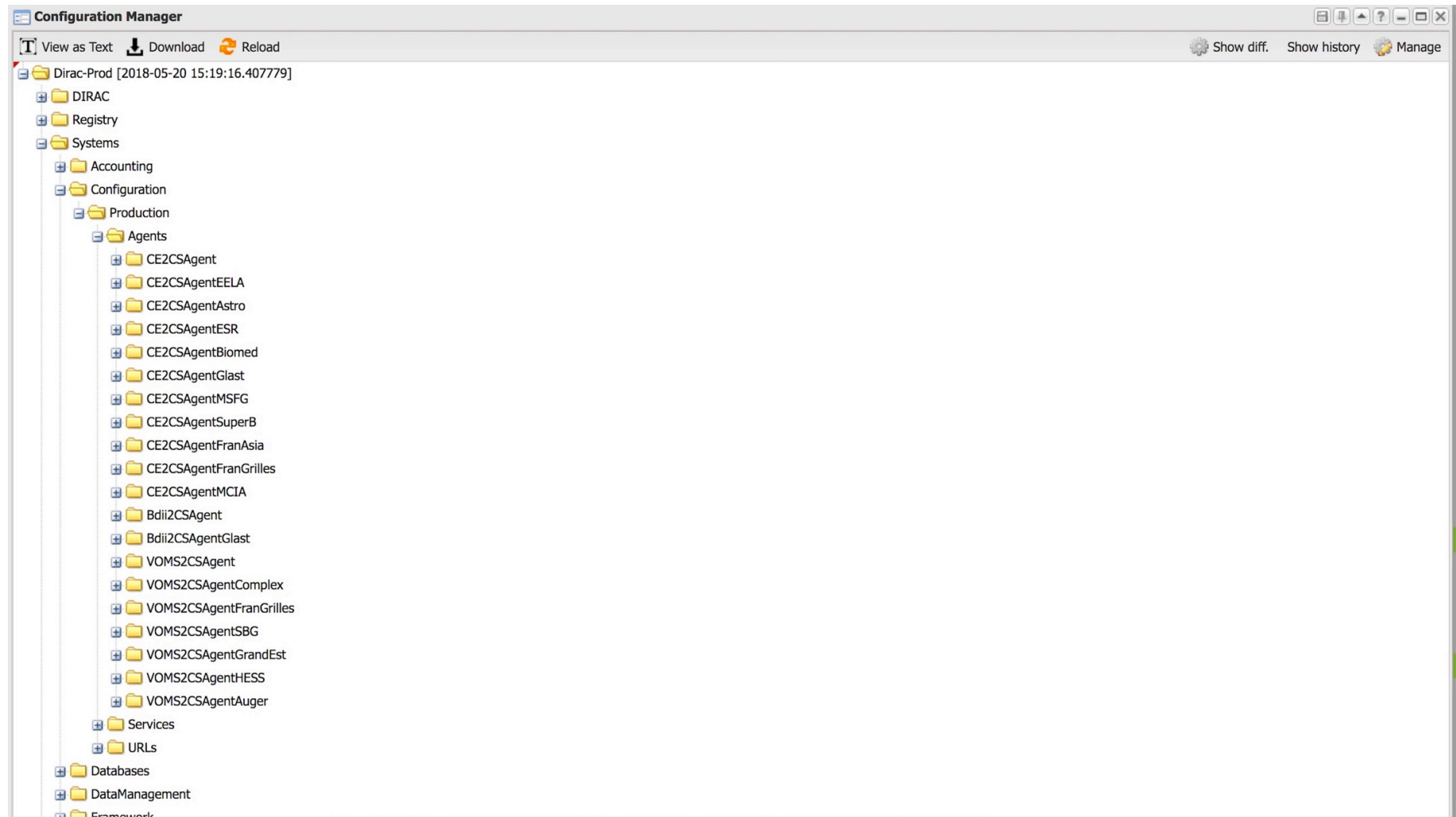
▶ For many VOs the majority of users in the VOMS database are in Suspended status

- Before v6r20 the status information was not available for VOMS2CSAgent and for many VOs automatic synchronization was switched off.

▶ Wish list

- Users can get Suspended status in VOMS after registration in the DIRAC Registry
 - Need to introduce Suspended user status in DIRAC by VO
 - Not allowing most of activities (jobs, uploading data)
 - Keeping existing data

VO Administration: Configuration



► Operational parameters

- Each VO must have a dedicated section under the /Operations CS category
 - For example, /Operations/biomed/Dirac-Production/
- Pilot subsection must be present to define properties of VO specific pilot jobs (see below)
- Services subsection to choose VO specific services
 - For example, auger VO chooses to use AugerFileCatalog service running in FZU, Prague, instead of a general purpose File Catalog common for other VOs
- Cloud subsection for VM VO specific details, e.g. contextualization information
- Shifter section to supply user identities for specific activities
 - Is it obsoleted ?
- Other sections are not defined in FG-DIRAC installation
 - Possibility to define custom sections with information used in VO specific applications

VO administration: Operations

The screenshot shows the Configuration Manager interface with a tree view of operations. The tree is organized as follows:

- Operations
 - Dirac-Production
 - JobDescription
 - RSSConfiguration
 - astro.vo.eu-egee.org
 - auger
 - biomed
 - Dirac-Production
 - Pilot
 - GenericPilotGroup = biomed_pilot
 - GenericPilotDN = /O=GRID-FR/C=FR/O=CNRS/OU=CREATIS/CN=Sorina Camarasu
 - Shifter
 - SAMManager
 - ProductionManager
 - Cloud
 - user_data_commands_base_url = https://ccdirac05.in2p3.fr/defaults/bootstrap
 - user_data_commands = vm-bootstrap, vm-bootstrap-functions, vm-pilot, vm-monitor-agent, pilotCommands.py, pilotTools.py, dirac-install.py, power.sh, parse-jobagent-log, dirac-pilot.py, save-payload-logs
 - GenericCloudGroup = biomed_pilot
 - GenericCloudUser = atsareg
 - biomed
 - Project = DIRAC
 - Version = v6r15
 - vo.complex-systems.eu
 - esr
 - Dirac-Production
 - Pilot
 - Shifter
 - Services
 - Catalogs
 - FileCatalog
 - AccessType = Read-Write
 - Status = Active

▶ Wish list:

- There is a need to clearly define a role of a VO administrator as a very specific activity
 - Need special rights
 - VOMS
 - Usually VO-Admin Role
 - Needed to extract all the user information
 - DIRAC CS
 - Rights to edit specific parts of the CS related to the VO
- Now there is only a possibility to define a right to edit the CS as a whole
 - Access control per section should be introduced

- ▶ Keeping VO Resources description up to date is one of the main VO administrator tasks

- ▶ Tools
 - Manually editing the CS Resources section
 - Need for a specific Resources editor Web application
 - Checking and Adding new resources with `dirac-admin-add-resources` command
 - Synchronization with the BDII index
 - Allows to add new sites, CEs and SEs interactively
 - Running BDII2CS agent per VO
 - Updating existing resources data
 - Sending reports by mail about newly available resources which should be added by hand

VO administration: Resources

The image displays two screenshots of the Configuration Manager interface, showing the configuration for LCG.CPPM.fr and Dirac-Prod.

Left Screenshot (LCG.CPPM.fr):

- Name = IN2P3-CPPM
- Coordinates = 5.441:43.231
- Mall = gridadmin@cppm.in2p3.fr
- CE = marcream01.in2p3.fr, marcream02.in2p3.fr
- SE = CPPM-disk, DIRAC-USER
- CEs
 - marcream01.in2p3.fr
 - wnTmpDir = /tmp
 - architecture = x86_64
 - OS = ScientificSL_Carbon_6.4
 - SI00 = 2720
 - Pilot = True
 - CETType = CREAM
 - SubmissionMode = Direct
 - PilotRunDirectory = /tmp/wrk
- Queues
 - cream-pbs-formation
 - cream-pbs-biomed
 - maxCPUTime = 2880
 - SI00 = 2720
 - VO = biomed
 - MaxWaitingJobs = 92
 - MaxTotalJobs = 922
 - OutputURL = gsiftp://localhost
 - cream-pbs-esr
 - cream-pbs-gisela
 - cream-pbs-france
 - cream-pbs-dteam
 - cream-pbs-asia
 - cream-pbs-cta
 - MaxRAM = 4096

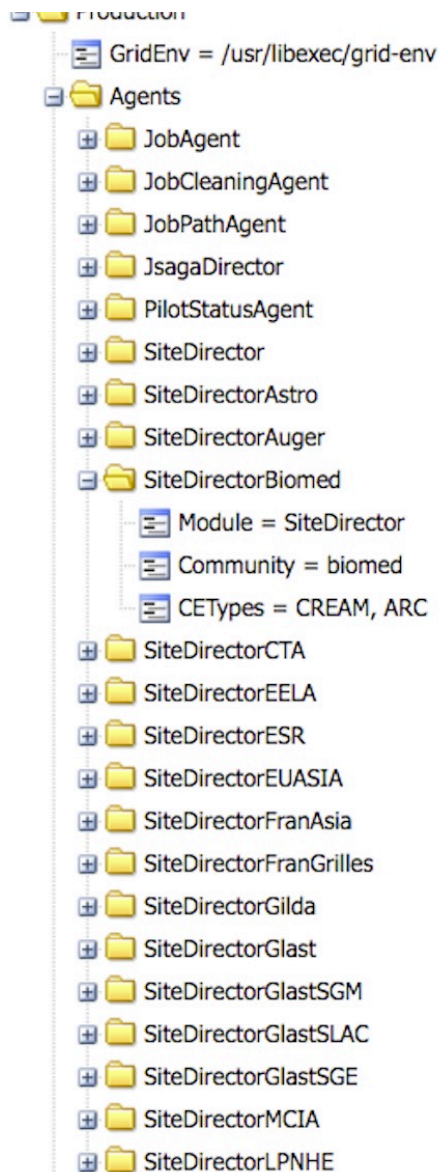
Right Screenshot (Dirac-Prod [2018-05-20 15:19:16.407779]):

- DIRAC
- Registry
- Systems
- Website
- Resources
 - JSAGA
 - FileCatalogs
 - Sites
 - StorageElements
 - DefaultProtocols = file, root, rfi, gsiftp
 - ProductionSandboxSE
 - DIRAC-USER
 - BackendType = DISET
 - ReadAccess = Active
 - WriteAccess = Active
 - SEType = D1
 - VO = esr, biomed, astro.vo.eu-egee.org, auger, euasia, gilda, glast.org, vo.cta.in2p3.fr, vo.formation.idgrilles.fr, vo.france-asia.org, vo.france-grilles.fr, vo.msfg.fr, vo.mcia.fr, vo.lpnhe.in2p3.fr, virgo, enmr.eu, france-fedcloud, vo.grand-est.fr
 - AccessProtocol.1
 - AccessProtocol.2
 - AccessProtocol.3
 - AccessProtocol.4
 - MCIA-Irods
 - CPPM-disk
 - CESNET-disk
 - FZU-proddisk
 - FZU-disk
 - AACHEN-disk
 - SIGNET-disk
 - FERMI-disk
 - IBCP-disk
 - IN2P3-disk
 - IN2P3-disk

► Wish list:

- Many resources in the DIRAC CS are obsoleted
 - Need a tool, e.g. enhancing the BDII2CSAgent, for detecting non-existent resources
- From the user perspective we need a tool to list available resources as defined in DIRAC
 - Analog of the lcg-infosites command
- The existing `dirac-dms-show-se-status` command is not enough
- There is no command to list VO specific CEs/Queues with their properties and status

- ▶ Each VO has a dedicated SiteDirector
 - Configured as an independent WMS agent for a given VO
- ▶ SiteDirectors send pilots according to VO specific configuration parameters:
 - /Operations/<VO>/<Setup>/Pilot CS section
 - Pilot identity
 - DIRAC extensions and versions to make available on WNs
 - Pilot commands to use (if not standard ones)



- ▶ Sites offer resources to a certain set of VOs
 - Having special queues per VO
 - Having common queues for multiple VOs

- ▶ Wish list:
 - Per VO SiteDirectors run completely independently but respect common limits for the number of pilots in a given queue
 - This can create problems with multiple-VO queues
 - Some VOs can be blocked because of activities in other VOs
 - A more complex scheduling algorithm is needed in this case
 - For example, allow a VO to still submit a small number of pilots even if the queue is full with “not-the-same-VO” pilots.

► Wish list:

- Some VO's have subgroups which can require special pilot properties
 - Example: using special VOMS roles in order to have separate accounting in the given grid infrastructure
- In this case, separate SiteDirector per such activity can be setup
 - This can quickly become too complicated to manage, might need enhancing SiteDirectors to manage multiple such case by one agent

- ▶ Multi-VOs instances allows to small VOs to use grid easily
- ▶ Thanks for make life easier to administrators, vo-admins and DIRAC users