



FENIX

RESEARCH INFRASTRUCTURE

Introduction to the ICEI resources at BSC

10th Fenix Research Infrastructure Webinar

Javier Bartolomé (16th April 2021)



The ICEI project has received funding from the European Union's Horizon 2020 research and innovation programme under the grant agreement No 800858.

Agenda

- Summary of BSC ICEI Resources
- Nord3 cluster (SCC and VM service)
 - Hardware Overview
 - Access and Use
- Interactive Computing Cluster (IAC)
 - Hardware Overview
 - Access and User
- Storage infrastructures (ACD and ARD)
 - Active Storage infrastructure
 - Archive Storage infrastructure
- Access to Fenix Resources
- Questions and Answers

Summary of BSC ICEI

			Quarterly allocation		
Component	Service type	ICEI resources	Total	HBP (25%)	PRACE (15%)
Nord3 cluster	SCC / VM	84 nodes	84 servers	21 servers	12 servers
Interactive Computing Cluster	IAC	3 nodes	5832 node-hrs	1486 node-hrs	874 node-hrs
HPC Storage	ACD	70 TB	70 TB	17.5 TB	10.5 TB
Archive Agora Storage	ARD	10 PB	10 PB	2.5 PB	1.5 PB

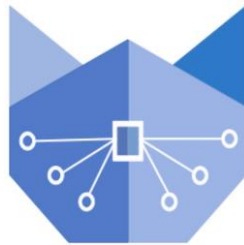
Nord3 Cluster: Hardware description

- 9 iDataPlex compute racks. Each one composed of:
 - 84 IBM dx360 M4 compute nodes
 - 4 Mellanox 36-port Managed FDR10 IB Switches
- All IBM dx360 M4 node contain:
 - 2x E5-2670 SandyBridge-EP 2.6GHz cache 20MB 8-core
 - 500GB 7200 rpm SATA II local HDD
- 3 types of nodes:
 - 500 Default nodes: 32 GB/node
 - 128 Medium memory nodes: 64 GB/node
 - 128 High memory nodes: 128 GB/node
- Interconnection Networks
 - Infiniband Mellanox FDR10
 - 1 and 10 Gigabit Ethernet
- Operating System:
 - Linux - SuSE Distribution 11 SP3 (being upgraded to RHEL8)



Nord3 Cluster: Software description

- Nord cluster is deployed using xCAT software
- Compute nodes could be moved from the operating system HPC image to operating system cloud image
 - Depending on the use-cases and resources requested at each moment
 - Manual procedure a re-deploy of OS image needed
- HPC operating systems is currently based in SLES11 SP3 and LSF as batch scheduling system
 - Being upgraded to RHEL 8 and Slurm
- Cloud operating system is based in RHEL 8
 - Openstack version Train as the software to provide all VM services.



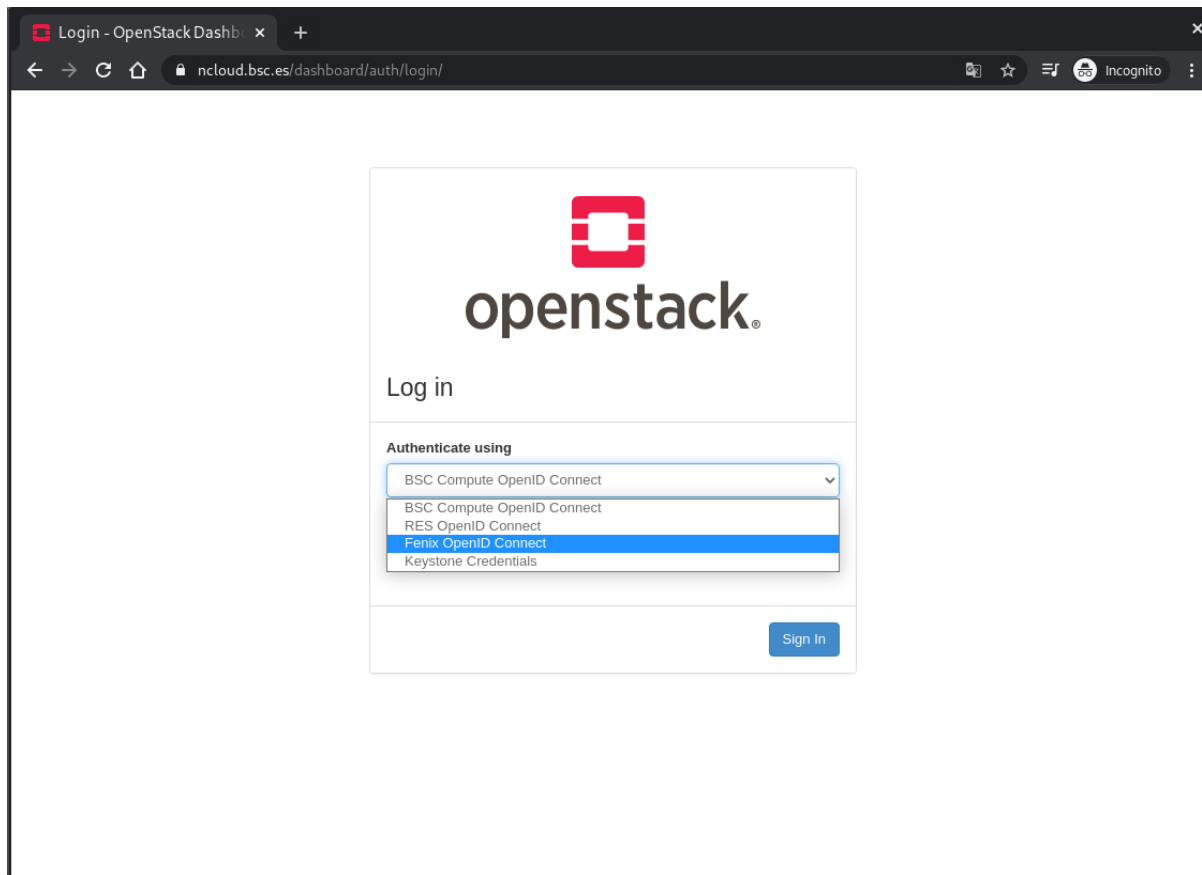
Nord3 Cluster: SCC service

- Access to SCC via SSH
 - nord[1-3].bsc.es
- User manual provided to all users
 - <https://www.bsc.es/support/Nord3-ug.pdf>
- Direct email for HPC support:
 - support@bsc.es



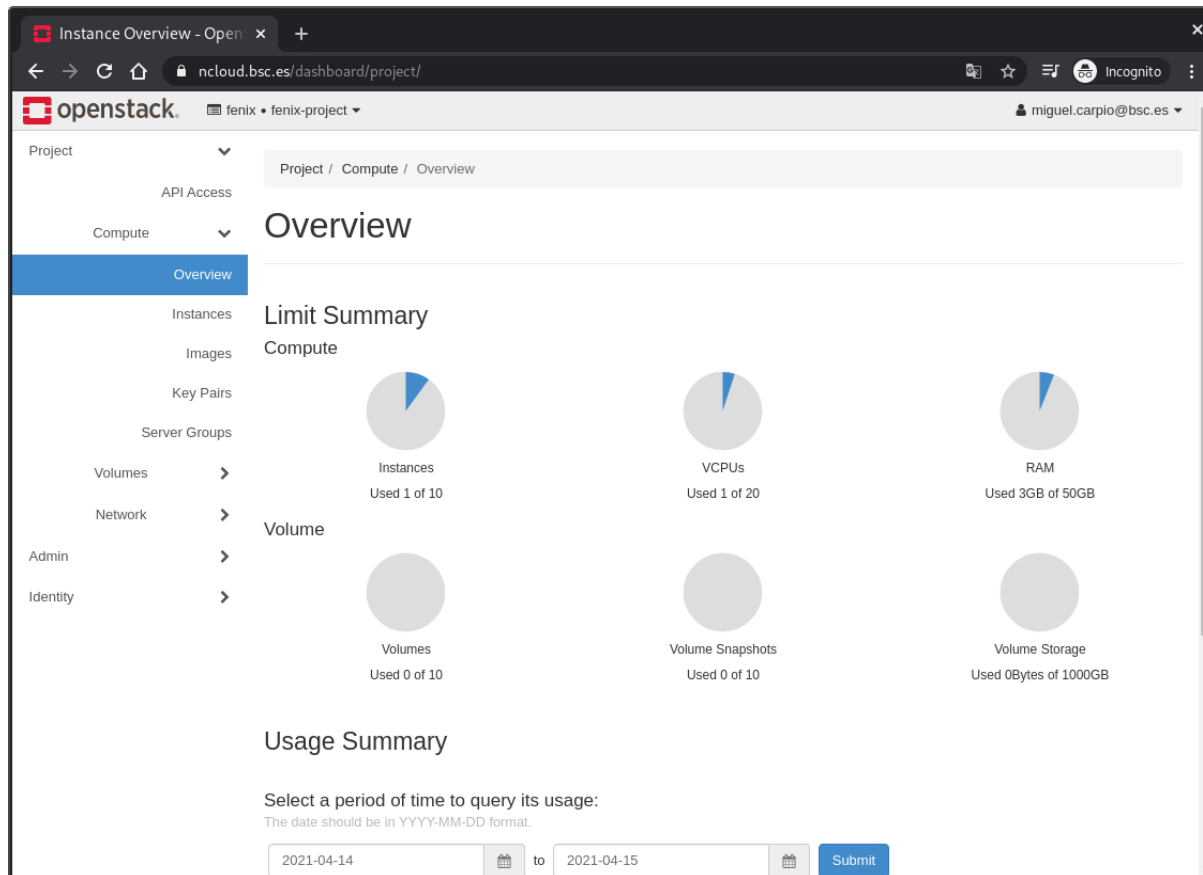
Nord3 Cluster: VM Service

- Dashboard available at: <https://ncloud.bsc.es/>
 - Web Access via Fenix AAI, RES and/or a HPC account



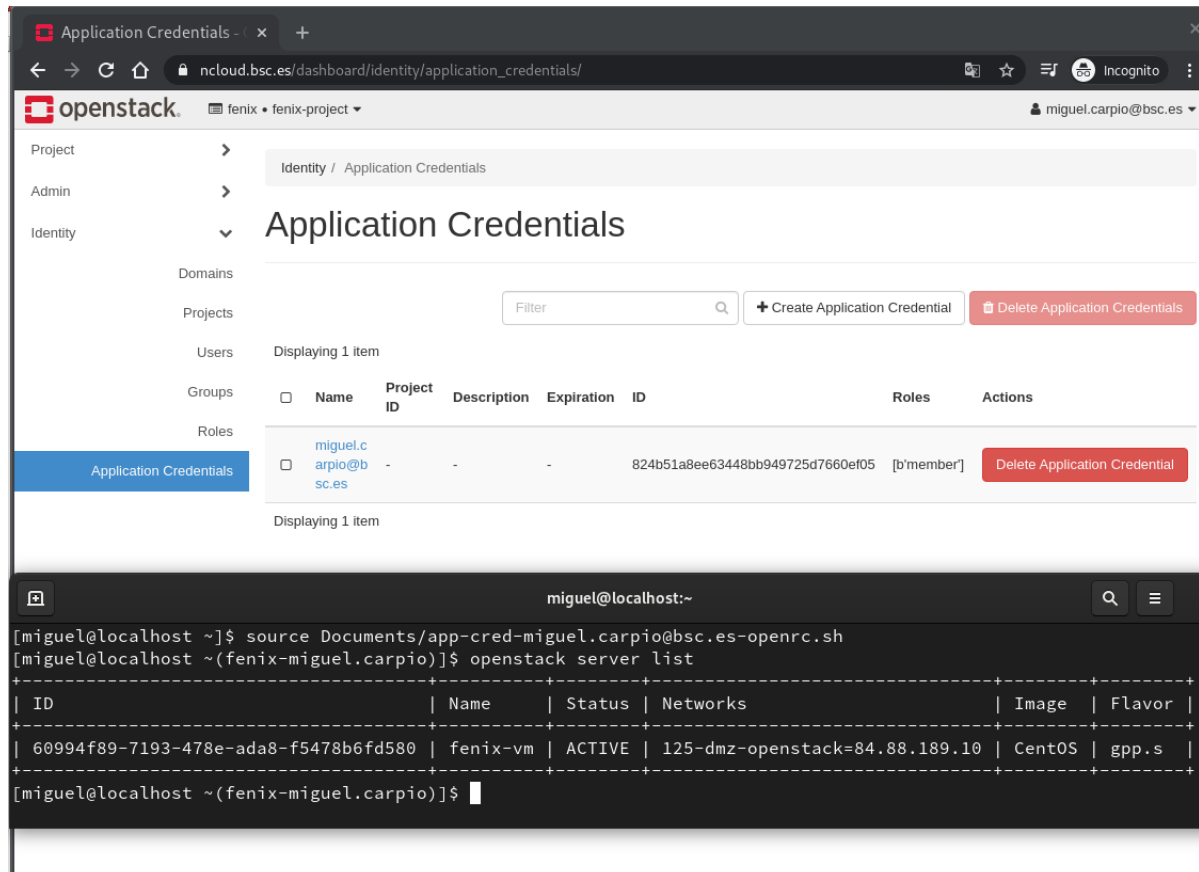
Nord3 Cluster: VM Service

- Project Overview



Nord3 Cluster: VM Service

- CLI Access via Application Credentials



The screenshot displays the OpenStack Identity Application Credentials management interface in a web browser. The browser address bar shows the URL `ncloud.bsc.es/dashboard/identity/application_credentials/`. The page title is "Application Credentials" under the "Identity" section. A sidebar on the left lists navigation options: Project, Admin, Identity, Domains, Projects, Users, Groups, Roles, and Application Credentials (which is highlighted). The main content area shows a table with one application credential. The table has columns for Name, Project ID, Description, Expiration, ID, Roles, and Actions. The single entry is for "miguel.carpio@bsc.es" with Project ID "-", Description "-", Expiration "-", ID "824b51a8ee63448bb949725d7660ef05", and Role "[b'member']". A "Delete Application Credential" button is visible next to the entry. Below the table, it says "Displaying 1 item".

Below the web interface, a terminal window is shown with the prompt `miguel@localhost:~`. The terminal shows the following commands and output:

```
[miguel@localhost ~]$ source Documents/app-cred-miguel.carpio@bsc.es-openrc.sh
[miguel@localhost ~(fenix-miguel.carpio)]$ openstack server list
```

ID	Name	Status	Networks	Image	Flavor
60994f89-7193-478e-ada8-f5478b6fd580	fenix-vm	ACTIVE	125-dmz-openstack=84.88.189.10	CentOS	gpp.s

```
[miguel@localhost ~(fenix-miguel.carpio)]$
```

Nord3 Cluster: VM Service

- OpenStack release: Train
- BSC Cloud - Flavors

Flavour	vCPU	RAM	Disk
gpp.s	1	3 GB	20 GB
gpp.m	2	8 GB	20 GB
gpp.l	4	16 GB	20 GB
gpp.xl	16	64 GB	20 GB

Nord3 Cluster: VM Service

■ Create a VM

```
[miguel@localhost ~(fenix-miguel.carpio)]$ openstack image list
+-----+-----+
| ID | Name | Status |
+-----+-----+
| 3022d568-6e4c-4982-bc2b-d39fcfe430a6 | CentOS | active |
| 75a6ab7c-c083-4960-939c-c0108070b0bd | Ubuntu | active |
+-----+-----+

[miguel@localhost ~(fenix-miguel.carpio)]$ openstack flavor list
+-----+-----+-----+-----+-----+-----+
| ID | Name | RAM | Disk | Ephemeral | VCPUs | Is Public |
+-----+-----+-----+-----+-----+-----+
| 199d413e-2ad4-4855-b23a-7108a95f0224 | gpp.xl | 65536 | 20 | 0 | 16 | True |
| 4d956fbb-f927-4d88-b683-dcc177858a8a | gpp.m | 8192 | 20 | 0 | 2 | True |
| 5c95b21e-cc24-4197-a9f9-ca77e8085085 | gpp.l | 16384 | 20 | 0 | 4 | True |
| b33e63f6-3353-42c0-b11e-d2fdale05d8f | gpp.s | 3072 | 20 | 0 | 1 | True |
+-----+-----+-----+-----+-----+-----+

[miguel@localhost ~(fenix-miguel.carpio)]$ openstack network list
+-----+-----+-----+
| ID | Name | Subnets |
+-----+-----+-----+
| 5ff2eb90-3c56-4522-80e8-8654f55ab091 | 125-dmz-openstack | 0d988bb9d-307d-45c7-baf9-d6f8921593c1 |
+-----+-----+-----+

[miguel@localhost ~(fenix-miguel.carpio)]$ openstack server create --image CentOS --flavor gpp.s --security-group default --network 125-dmz-openstack fenix-vm
+-----+-----+
| Field | Value |
+-----+-----+
| OS-DCF:diskConfig | MANUAL |
| OS-EXT-AZ:availability_zone | |
| OS-EXT-STS:power_state | NOSTATE |
| OS-EXT-STS:task_state | scheduling |
| OS-EXT-STS:vm_state | building |
| OS-SRV-USG:launched_at | None |
| OS-SRV-USG:terminated_at | None |
| accessIPv4 | |
| accessIPv6 | |
| addresses | |
| adminPass | EHGaCc9dm5Tf |
| config_drive | |
| created | 2021-04-15T10:11:31Z |
| flavor | gpp.s (b33e63f6-3353-42c0-b11e-d2fdale05d8f) |
| hostId | |
| id | 60994f89-7193-478e-ada8-f5478b6fd580 |
| image | CentOS (3022d568-6e4c-4982-bc2b-d39fcfe430a6) |
| key_name | None |
| name | fenix-vm |
| progress | 0 |
| project_id | f8e896eab392448193a551217f330bc0 |
| properties | |
| security_groups | name='809b9e5c-95f9-4068-b35e-e705e8075c87' |
| status | BUILD |
| updated | 2021-04-15T10:11:32Z |
| user_id | d27d2e11b4e5468eace75ad4e5c1eb67 |
| volumes_attached | |
+-----+-----+

[miguel@localhost ~(fenix-miguel.carpio)]$ openstack server list
+-----+-----+-----+-----+-----+-----+
| ID | Name | Status | Networks | Image | Flavor |
+-----+-----+-----+-----+-----+-----+
| 60994f89-7193-478e-ada8-f5478b6fd580 | fenix-vm | ACTIVE | 125-dmz-openstack=84.88.189.10 | CentOS | gpp.s |
+-----+-----+-----+-----+-----+-----+
```

Interactive Computing Cluster (IAC): Hardware

- 2 compute nodes, made of:
 - System IBM Power System AC922
 - Processors 2x IBM POWER9 16-core 2.6GHz
 - Memory 16x 64 GB DDR4 2666 MHz DDR4 RDIMM
 - Accelerator 2x NVIDIA Tesla V100 SXM2 16GB
 - Network 1x Ethernet 25Gb Adapter with two ports
 - Disk 2x 960 GB SSD Disks
- 1 compute node, made of:
 - System IBM Power system AC922
 - 2x 20-core 2.4 GHz (3.0 GHz Turbo) POWER9 Processor
 - 512GB memory in 16x 32GB DDR4 DIMMs
 - 2x 960 GB 2.5in SATA/SSD Disk Drive



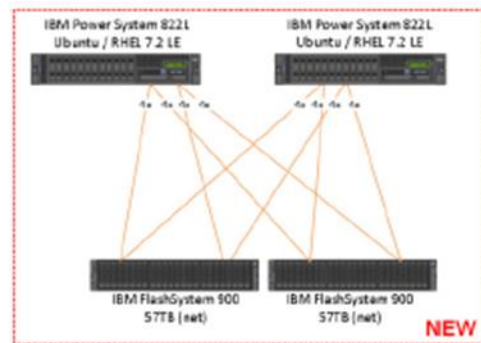
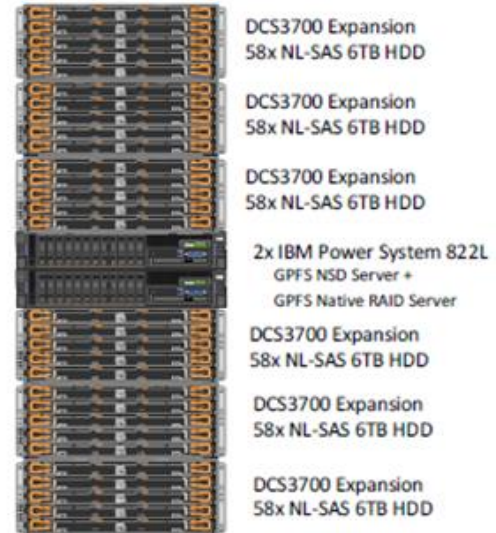
Interactive Computing Cluster (IAC): Software

- Red Hat Enterprise Linux 8.2
- Kernel 4.18.0-193.37.1.el8_2.ppc64le
- xCAT 2.16.1
- IBM Spectrum MPI 10.4.0.3
- IBM Spectrum Scale 5.0.5-4
- NVIDIA CUDA 11 Toolkit 11-0-11.0.3-1
- NVIDIA Driver 450.80.02
- Slurm as Batch scheduling system
 - Possible reservation of nodes

- Status: Being configured, expected to be available in May 2021

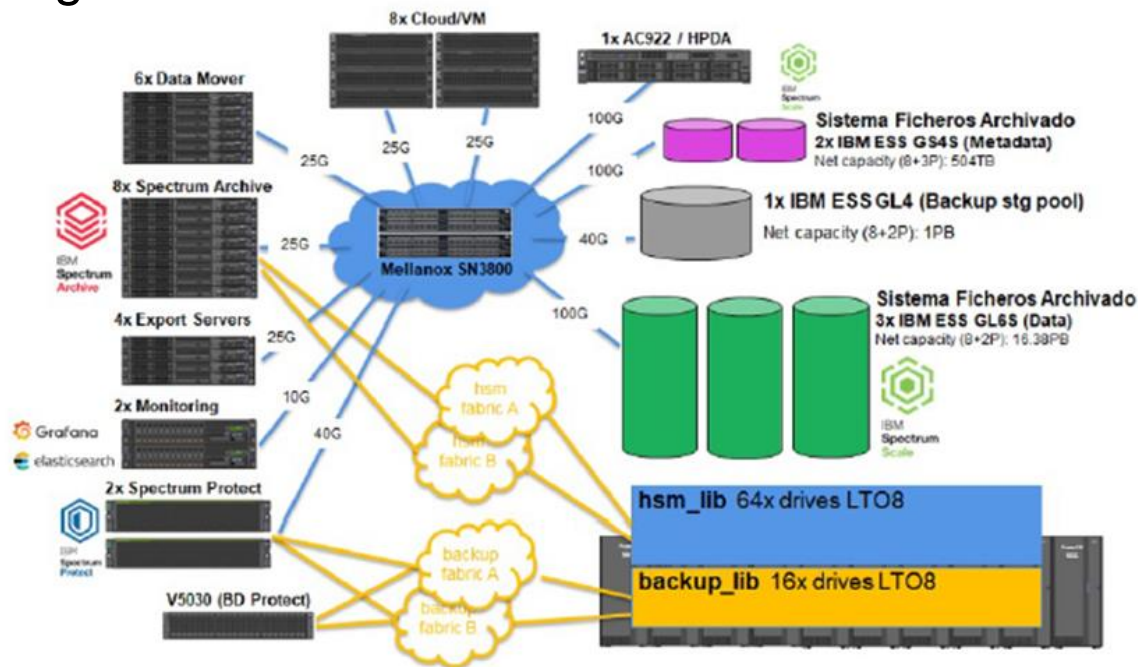
Storage infrastructures (ACD)

- 14 PB Accessible from IAC and SCC resources
- 7 x IBM ESS GL6
 - 2 IBM Power System 822L
 - 256 GB RAM
 - 2 x Intel OPA ports
 - 2 x 40 Gbps Eth ports
 - 6 DCS3700 JBOD expansions
 - 348 NL-SAS 8TB disks
 - 2 SSD 400GB disks
- 2 x Metadata blocks
 - 2 IBM Power Systems
 - 2 x Intel OPA ports
 - 2 x 40Gbps Eth ports
 - IBM Flash System 900
 - 12 x 5.7TB Micro latency Modules
- Spectrum Scale as Parallel filesystem



Storage infrastructures (ARD)

- 3 IBM Elastic Storage Server GL6S providing data storage for the first level of storage
- 2 IBM Elastic Storage Server GS4S providing metadata storage for the first level of storage
- IBM Tape library TS4500 with a total of 80 LTO8 drives and almost 100PB storage



Storage infrastructures (ARD)

- Set of servers to provide the different data services:
 - 8 Cloud Servers to host virtual machines
 - 8 Archive Servers that will implement the movement from the different tiers (HDD and tape) inside Agora storage infrastructure
 - 6 Data Mover Servers that will hold all data transfer operations between different storages (ACD,ARD,Swift)
 - 4 Export servers, that will provide accessibility to data through other protocols such as, Swift, NFS or SMB
 - 2 Monitoring Servers, that will take care of monitoring and alerting for all components Agora is composed of
 - 2 Backup servers, that will take care of backup tasks for different services

Access to Fenix Resources

- All details on access to Fenix Resources:
 - <https://fenix-ri.eu/access>
- Neuroscientists can obtain access through HBP via the EBRAINS/HBP Call
 - <https://wiki.ebrains.eu/bin/view/Collabs/fenix-icei/>
- Also resources can be requested vi PRACE-ICEI calls
 - <https://fenix-ri.eu/news/prace-icei-calls-proposals-call-5>