VOboxes and all that jazz

ALICE Tier-1/Tier-2 Workshop in Seoul

April 17, 2024

M. Litmaath

The VObox in a nutshell

- Site VOboxes run JAliEn CE and MonALISA (ML) services
  - The CE submits jobs to the site resources via HTCondor CE, ARC CE or directly to the site's batch system
  - ML monitors jobs, services, SE and network
- Crontab entries restart the services as needed, with the right environment
- Command: */cvmfs/alice.cern.ch/scripts/vobox/jalien-vobox.sh*
  - By default it prints a help message

VObox credentials (1)

- The JAliEn CE needs an extra credential to obtain job tokens with which job scripts are submitted: a *JAliEn token certificate*
  - Example:

    $ openssl x509 -noout -subject -in **~/.globus/hostcert.pem**
    subject=/C=ch/O=AliEn2/OU=ALICE/CN=voboxalice10.cern.ch
  - See next page

- For jobs submitted to an **HTCondor** CE, the JAliEn CE also needs:
  - A ***WLCG token*** in **~/.globus/wlcg.dat** (see next pages)
  - An X509 proxy managed by the WLCG VObox proxy renewal service
    - To allow **APEL** to keep depending on that for the time being!
    - Our jobs do not need it anymore

- For jobs submitted to an **ARC** CE, the JAliEn CE also needs:
  - Just the X509 proxy managed by the WLCG VObox proxy renewal service

VObox credentials (2)

- For VOboxes submitting directly to the site's batch system, the JAliEn token certificate is renewed manually once per year

- For VOboxes submitting to HTCondor or ARC CEs, there are several central token renewal service instances that renew the token credentials through *gsissh* several times per day:
  - JAliEn token certificate
    - Current lifetime typically 30 days
  - WLCG token
    - Current lifetime 4 days

Details about ARC vs. tokens are documented in this presentation thanks very much to Andrey Zarochentsev!

- WLCG tokens are **not** yet used for ARC CEs
  - **ARC 7** still supports VOMS proxies
  - The JAliEn ARC module code would need major changes not only for the use of tokens, but also to remain usable with **ARC 8**
  - Instead, we intend to let an **HTCondor** installation on the **VObox** deal with ARC CEs, imitating what is already done for HTCondor CEs

WLCG token example

```
$ ls -l .globus/wlcg.dat
-rw------- 1 alicesgm alicesgm 899 Apr 13 19:30 .globus/wlcg.dat
$ decode-token.sh .globus/wlcg.dat
{
  "wlcg.ver": "1.0",
  "sub": "a4f952ab-6e43-059c-c530-80df119a018b",
  "aud": [
    "ce01.some.site:9619",
    "ce02.some.site:9619"
  ],
  "nbf": 1713029430,
  "scope": "compute.create compute.read compute.cancel compute.modify",
  "iss": "https://alice-auth.web.cern.ch/",
  "exp": 1713375030,
  "iat": 1713029430,
  "jti": "f97d2db9-eb09-4418-b3f9-27736f99c560",
  "client_id": "ae76ab27-cc15-4082-a9bc-50ad587a73d6"
}
$
```

From X509 + VOMS to WLCG tokens

- WLCG Authorization WG is coordinating the transition from X509 + VOMS to WLCG tokens during Run 3 (link)
  - Inspired by the use of tokens in ALICE since many years!
  - And by common practice in industry and academia
    - To support federated identities and token standards

WLCG Token Transition Timeline v1.0 published August 22, 2022

- VO management will go from VOMS-Admin to (INDIGO-) **IAM**
  - **Will happen in May! More news in the next weeks…**
  - IAM can provide finely scoped tokens to trusted entities
  - It also has a VOMS endpoint for backward compatibility

- See March GDB update for the latest developments

Batch system aspects

- On most VOboxes, the JAliEn CE submits 8-core batch jobs
  - For some sites or resources, single-core jobs may need to be used
  - At HPC and some grid sites, each batch job can get a ***whole node***
- At any time, each 8-core batch job can run either of these:
  - One 8-core task
  - Any set of 1-, 2- and 4-core tasks that fit within 8 cores
- Legacy workflows for Run 1 & 2 only have single-core tasks
- Each task is run in an *Apptainer* container
  - This has become a requirement as of recent JAliEn versions
- A proposal to go to ***16-core*** standard slots will be discussed at the WLCG / HSF workshop, May 13-17

https://alimonitor.cern.ch/stats?page=proxies

MonALISA Repository for ALICE

MonALISA — MONitoring Agents using a Large Integrated Services Architecture

My jobs | My home dir | LEGO Trains | Hyperloop | Administration Section | Alert XML Feed | JAliEn docs | MonaLisa GUI

A single page to see how the VOboxes are doing

Status of proxies and AliEn and LCG tests

	AliEn Tests										LCG Tests									
	LDAP		CVMFS		JAliEn cert		CE				AliEn proxy		WLCG token		Renewal	Proxy Server		Host proxy		Host cert
Service	Stat	Cores	Stat	Rev	Stat	Time left	Stat	Config	Running	JobAgent	Stat	Time left	Stat	Time left	Stat	Stat	Time left	Stat	Time left	Stat	Time left
1. Altaria		8		18415		59d 22:47		pro	1.8.6-1	1.8.6-1		1d 23:35		3d 22:55			373d 16:57		21:46		331d 1:27
2. Bari_HTC		8		18415		29d 23:59		pro	1.8.6-1	1.8.6-1		1d 23:15		3d 23:58			373d 16:47		23:36		18d 3:16
3. Birmingham		8		18415		29d 22:49		pro	1.8.6-1	1.8.6-1		1d 23:20		3d 23:58			373d 16:47		22:36		208d 17:33
4. Bratislava_ARC		8		18415		29d 22:53		new	1.8.6-1	1.8.6-1		1d 23:58		3d 22:57			162d 15:24		21:47		89d 18:07
5. Capella		0		18415		59d 22:56	CE N...	pro	n/a	pro		1d 23:22		3d 22:49			373d 16:50		21:40		42d 0:20
6. Catania-VF		8		18415		29d 22:49		pro	1.8.6-1	1.8.6-1		1d 23:55		3d 22:55			373d 16:56		21:46		80d 3:26
7. CBPF		8		18415		29d 22:50		pro	1.8.6-1	1.8.6-1		1d 23:27		3d 22:58			373d 17:00		21:49		230d 17:25
8. CCIN2P3_HTC		8		18415		29d 23:59		1.8.6-1	1.8.6-1	1.8.6-1		1d 23:17		3d 22:56			373d 16:58		21:47		329d 3:27
9. CCIN2P3_HTC_2		8		18415		29d 22:54		pro	1.8.6-1	1.8.6-1		1d 23:23		3d 22:50			373d 16:52		21:41		335d 3:21
10. CERN-AURORA		1		18415		59d 22:53		pro	1.8.6-1	1.8.6-1		1d 23:18		3d 22:49			373d 16:51		21:40		183d 1:00
11. CERN-CORONA		8		18415		59d 22:55		pro	1.8.6-1	1.8.6-1		1d 23:22		3d 22:49			373d 16:50		21:39		42d 0:20
12. CERN-MIRAGE		8		18415		59d 22:51		pro	1.8.6-1	1.8.6-1		1d 23:15		3d 22:56			373d 16:58		21:47		19d 23:27
13. CERN-SIRIUS		8		18415		59d 22:58		pro	1.8.6-1	1.8.6-1		1d 23:26		3d 23:00			373d 17:01		21:50		42d 0:30
14. CERN-TRITON		8		18415		59d 23:59		1.8.6-1	1.8.6-1	1.8.6-1		1d 23:11		3d 22:57			373d 16:59		21:48		183d 1:08
15. CERN-ZENITH		8		18415		59d 22:47		pro	1.8.6-1	custom		1d 23:06		3d 22:53			373d 16:54		21:44		365d 1:54
16. Cibinong		8		18415		29d 23:58		pro	1.8.6-1	1.8.6-1		1d 23:16		3d 23:00			71d 18:56		22:50		181d 8:26
17. Clermont_ARC		8		18415		29d 22:56		pro	1.8.6-1	1.8.6-1		1d 23:50		3d 22:58			373d 16:59		21:48		203d 3:28
18. CNAF		8		18415		29d 23:58		1.8.6-1	1.8.6-1	1.8.6-1		1d 23:22		3d 23:59			138d 7:03		21:37		346d 3:17
19. CNAF-DUE		8		18415		29d 22:51		pro	1.8.6-1	1.8.6-1		1d 23:01		3d 23:58			138d 7:02		21:36		178d 3:16
20. Cyfronet_ARC		8		-		29d 22:57		pro	1.8.6-1	1.8.6-1		1d 23:22		3d 22:48			373d 16:49		23:38		373d 19:18
21. DCSC_KU		8		18415		29d 22:59		pro	1.8.6-1	1.8.6-1		1d 23:36		3d 22:53			373d 16:54		21:43		276d 12:00
22. EPN		64		18415		195d 0:59		1.8.6-1	1.8.6-1	1.8.6-1	-	-	-	-	-	-	-	-	-	-	-
23. EPN_MI100		96		18415		143d 15:30		pro	1.8.6-1	1.8.6-1	-	-	-	-	-	-	-	-	-	-	-
24. FZK		8		18415		29d 22:51		1.8.6-1	1.8.6-1	1.8.6-1		1d 23:49		3d 22:55			373d 16:56		21:45		240d 15:45
25. FZK_HTC		8		18415		29d 22:52		pro	1.8.6-1	1.8.6-1		1d 23:48		3d 22:55			373d 16:56		21:45		95d 9:40

JAliEn versions at sites

	LDAP		CVMFS		JAliEn cert		CE			
Service	Stat	Cores	Stat	Rev	Stat	Time left	Stat	Config	Running	JobAgent
7. CCIN2P3_HTC		8		18415		29d 23:59		1.8.6-1	1.8.6-1	1.8.6-1
13. CERN-TRITON		8		18415		59d 23:59		1.8.6-1	1.8.6-1	1.8.6-1
17. CNAF		8		18415		29d 23:58		1.8.6-1	1.8.6-1	1.8.6-1
21. EPN		64		18415		195d 0:44		1.8.6-1	1.8.6-1	1.8.6-1
23. FZK		8		18415		29d 23:49		1.8.6-1	1.8.6-1	1.8.6-1
60. RAL		8		18415		29d 22:49		1.8.6-1	1.8.6-1	1.8.6-1
4. Bratislava_ARC		8		18415		29d 23:51		new	1.8.6-1	1.8.6-1
26. GRIF_IPNO_IJCLAB		8		18415		29d 23:47		new	1.8.6-1	1.8.6-1
51. Oxford		8		18415		29d 23:55		new	1.8.6-1	1.8.6-1
70. Strasbourg_IRES		8		18415		29d 23:50		new	1.8.6-1	1.8.6-1
72. Trieste		8		18415		29d 22:50		new	1.8.6-1	1.8.6-1
77. Vienna		8		18415		29d 23:56		new	1.8.6-1	1.8.6-1
78. Wigner_KFKI_AF		1		18415		29d 22:49		new	1.8.6-1	1.8.6-1
79. Wigner_KFKI_AF_8core		8		18415		29d 23:59		new	1.8.6-1	1.8.6-1
81. Yerevan		1		18415		59d 23:51		new	local	custom
14. CERN-ZENITH		8		18415		59d 22:47		pro	1.8.6-1	custom
29. GSI_8core		8		18415		307d 1:44		pro	1.8.6-1	custom
32. HPCS_Lr		0		18415		104d 0:08		pro	1.8.6-1	custom
50. ORNL		0		18415		355d 11:50		pro	1.8.6-1	custom
71. Torino-HTC		8		18415		29d 23:57		pro	1.8.6-1	custom
1. Altaria		8		18415		59d 23:45		pro	1.8.6-1	1.8.6-1
2. Bari_HTC		8		18415		29d 23:59		pro	1.8.6-1	1.8.6-1
3. Birmingham		8		18415		29d 23:46		pro	1.8.6-1	1.8.6-1
5. Catania-VF		8		18415		29d 23:47		pro	1.8.6-1	1.8.6-1

Fixed: final testing or temporary need

New: early testing

Development site

Custom JA code

Default production

VObox configuration

```
$ cat .alien/config/CE.env
# CE environment bootstrap
export X509_USER_PROXY=$(
    ls -t /var/lib/vobox/alice/proxy_repository/*lcgadmin | sed q
)

$ cat .alien/config/ml.env
# MonaLisa environment bootstrap
export X509_USER_PROXY=$(
    ls -t /var/lib/vobox/alice/proxy_repository/*lcgadmin | sed q
)
export MYPROXY_SERVER=myproxy.cern.ch

$ cat .alien/config/version.properties
#MonaLisa=yyyymmdd-r
#jalien=x.y.z-r
#jobagent.version=x.y.z-r
#custom.jobagent.jar=/cvmfs/alice.cern.ch/java/jar/jalien/...
#...
```

Adjust the X509 proxy and/or other environment variables as needed

ALICE grid team experts will adjust these and other contents as needed

Batch script customization

```
$ cat ALICE/tmp/agent.startup.843443
#!/bin/bash
[...]
export partition=",PDC08_2,multicore_8,"

#
# customization 0 start
#
[...]
#
# customization 0 end
#

#
# customization 1 start
#
[...]
#
# customization 1 end
#

export JALIEN_JOBAGENT_CMD="/cvmfs/[...] alien.site.JobRunner"
eval $JALIEN_JOBAGENT_CMD

#
# customization 2 start
#
[...]
#
# customization 2 end
#
```

Contents of ~/JA-custom-0.sh

Contents of ~/JA-custom-1.sh

Contents of ~/JA-custom-2.sh

- Up to 3 shell scripts can be created in the VObox home directory to customize the JAliEn batch script
  - JA-custom-0.sh
    - Contents get inserted right after the initial environment setup
  - JA-custom-1.sh
    - Contents get inserted right before JALiEn is started
    - Currently there is no code (anymore) between the first two customization points
  - JA-custom-2.sh
    - Contents get inserted right after JALiEn has finished
- In use at some sites to deal with special conditions
- Can also be handy for debugging

Site middleware evolution

- CentOS 7 EOL is June 30 this year
- The most convenient next platforms are RHEL 9 and its derivatives AlmaLinux 9 and Rocky Linux 9
  - Collectively referred to as “EL 9”
  - All are deemed equivalent for our use cases
  - CentOS Stream 9 is **not** recommended because of its slowness in making critical patches available
- Most of the relevant middleware is available for EL 9
  - APEL client & parsers are expected in May
- The EGI UMD-5 repository for EL 9 products is expected in May
  - It will have the UI and WN meta packages
- The WLCG VObox is available from the WLCG repository
- It can also be built as a Docker container, in use e.g. at CERN

Conclusions & outlook

- The VObox is a vital component for ALICE Grid sites
- Thanks to JAliEn and other improvements, it has evolved to be ready for Run 3 and beyond
- The transition from CentOS 7 to EL 9 is a major milestone for **sites** to achieve in the next months!
- The transition from VOMS-Admin to IAM is a major milestone for **WLCG** to achieve in the next months!

Thanks for your kind attention!