### JAliEn status

ALICE-USA meeting LBL, 2024-09-18

costin.grigoras@cern.ch



## Deployment status

Currently at tag 1.9.2 (testing 1.9.3) Fully deployed on all sites 12 whole node queues All others are 8+ CPU cores queues In the process of migrating to EL 9 62% RHEL and Alma 9 29% CentOS and SL 7 9% Alma 8 (EPN nodes) and other variants OS not a problem any more with containers used for everything But only **EL 9** fully supports *cqroupsv2* 





## Worker node execution

#### Configurable slot size (no. of CPU cores)

Automatically adapt to whole node resources where available

Our preferred strategy to accommodate varying computing and memory needs from jobs Free resources advertised to the JobBroker

Available CPU cores, RAM and disk space

Isolated instances

Container execution (container image picked up from CVMFS *f*(software deps)) Per job credentials (X.509 pair issued by our CA), with limited capabilities Constrained resources (Core pinning with *taskset -> cgroupsv2* when possible) **Continuous monitoring of payload resource usage** Accounting and preempting them if running over boundaries

### Architecture matching

New dimensions added to the package definitions

- Platform (i.e. el9-x86\_64, el7-aarch64)
- CVMFS revision
- Dependencies per platform

With automatic import of new packages from CVMFS Allows us to match CPU architecture and run on ARM since Friday also for RISC-V

And to pick the right container image for the job

### Job containers & environment in JAliEn

Images located in CVMFS: /cvmfs/alice.cern.ch/containers/fs/singularity/<arch>/
Build recipe available on Gitlab

PRs possible for package requests

Specific image selected by JAliEn based on required packages for job Attempts to find a common OS build across the packages, e.g EL7 Combined with the system architecture to form a "requirement string"

- Symlinks in CVMFS to the corresponding combination:

compat\_el9-x86\_64 -> el9
el9 -> x86 64/alma9-alice-20240712

compat el9-aarch64 -> alma9-alice-20231212-aarch64

- If a requirement string has no matching symlink, a fallback container is used

GPUs are supported across the multiple layers

Compatibility check likewise done by JAliEn, with flags/mounts added as needed

## ERROR\_E subcategories

#### Added numeric subcodes to ERROR\_E jobs (killreason) Available in command line `ps`

jsh:[alice] [29]	/alice/cern	.ch/user/g/grigoras/	> ps	-f	ERROR	Е -	·l 30	-a	-X
------------------	-------------	----------------------	------	----	-------	-----	-------	----	----

pwg_pp	3152546120	 ALICE::LBL
pwg pp	3152546121	ALICE::
pwg pp	3152546126	ALICE::LBL
qq pwq	3152546131	 ALICE::CER
pwg pp	3152546248	ALICE::CER

ALICE::LBL\_HPCS::LBL\_AFP ALICE::LBL\_HPCS::HPCS ALICE::LBL\_HPCS::LBL\_AFP ALICE::CERN::CERN-SIRIUS ALICE::CERN::CERN-ZENITH

0 -a	-X	
	n0050.alice0	
	n0056.alice0	
	n0050.alice0	
9p08j	p5225.cern.ch	
9p10	p4272.cern.ch	

EE.

143 50 TaskEbyeIterPIDMC LHC20e3a pass3.sh
143 50 TaskEbyeIterPIDMC LHC20e3a pass3.sh
143 50 TaskEbyeIterPIDMC LHC20e3a pass3.sh
143 80 TaskEbyeIterPIDMC LHC20e3a pass3.sh
134 0 TaskEbyeIterPIDMC\_LHC20e3a\_pass3.sh

#### And also as monitoring stream

Idle CPU, Over TTL or Over memory

CERN/ALIEN KILL REASON Nodes Summary/sum/<user> <code>

## Kill reasons (code)

#### 0 - ok

10 - JW JVM out of memory error

20 - Job running longer than the TTL

30 - Killed by the user

40 - Preempted due to memory constraints (cgroups)

50 - Idle for more than 15 minutes

60 - core file (or directory) detected

70 - Over disk space limits

80 - Over the memory limits

90 - Detected kernel oom killing the payload

201 - Failed to create the sandbox directory

202 - Failed to download input files

203 - Cannot set up the packages environment from CVMFS

## CPU brokering

JDL tag: CPUCores (default 1)
Subtracted from the amount of available cores
Keep advertising while available>0
LDAP options (cerequirements CE field)

- cpulsolation
- cpuOversubscription
- memoryPreemption

## Memory brokering

#### JDL tags:

RequiredMemory - minimum amount (default 2 GB \* CPUCores) MemorySize (default 8 GB \* CPUCores; maximum 2x that)

#### RequiredMemory

- Matched against machine resources
  - Subtract from machine's total (whole node)
  - Or from the cgroups v2 hard limit (or BQ policy, or machine RAM)
- When memoryPreemption, sort by the largest offender wrt this value

#### MemorySize

- hard limit, the JobAgent will kill the payload when reached

## Disk brokering

JDL tag: WorkDirectorySize default 10 GB \* CPUCores; max 2x that Matched against available space on the cwd partition Some recent machines don't have 10 GB / core => cannot fill the machine with jobs

## Oversubscription

When enough CPUs idle for enough time and enough memory free and enough disk free

advertise the resources and start a new job

- reniced
- first to be killed when resources are too low

Experimentally enabled on some whole node queues

- hit the limit on the disk size actually

### **Oversubscription** @ LBL\_HPCS

CPU Model	Total memory	CPU cores	Memory per core	
Intel(R) Xeon(R) Gold 6330 CPU @ 2.00GHz	252  GB	112	2.25 GB	To enable
Intel(R) Xeon(R) CPU E5-2680 v4 @ 2.40GHz	126 GB	56	2.25 GB	oversubscription, memory/CPU core
Intel(R) Xeon(R) CPU E5-2680 v4 @ 2.40GHz	$63~\mathrm{GB}$	28	$2.25~\mathrm{GB}$	needs to be > 2 GB
Intel(R) Xeon(R) Gold 6130 CPU @ 2.10GHz	93 GB	64	1.47 GB	

Running additional jobs with complementary resource usage patterns Pick up MC jobs when IO-bound jobs don't use the machine efficiently Extra pressure in CPU is limited in time – oversubscribed jobs are moved to the regular pool as soon as a spot is available

## Extra jobs at LBL\_HPCS

Core count	Node count	Total oversubscribed jobs	Completed job rate	
112	8	563	98.17%	
56	4	30	93.3%	
28	54	20	16.4%	
64	37	0	52	

Table 6.2: Oversubscribed jobs executed on each node type

#### 613 extra jobs, >96% successful 23 jobs killed due to CPU pressure Limiting factor on the 112-core nodes was the disk space

## **TTL optimization**

JDL tag:

TTLOptimizationType="TTLStatistics";

#### Collect history per

- production ID
- CPU model name
- Site name

# Apply weighted history to subsequent jobs



# Ex 1: MC LHC24g3 @ LBL

Production Id / Site / CPU (Number of

Jobs)

aliprod - Production 30705 Speedup 3.1 30705/LBL HPCS/Intel(R) Xeon(R) Gold 6330 CPU @ 2.00GHz (274) Speedup 30705/LBL HPCS/Intel(R) Xeon(R) CPU E5-2698 v3 @ 2.30GHz (312) 2.9 Speedup 30705/LBL\_HPCS/Intel(R) Xeon(R) Gold 6130 CPU @ 2.10GHz (328) 4.6 Speedup 30705/LBL HPCS/Intel(R) Xeon(R) CPU E5-2680 v4 @ 2.40GHz (1207) 5.2 10000 50000 80000 30000 400.00 70000

Time (s)

■ requested TTL ■ mean ■ maximum Time ■ predicted TTL

## Ex 2: MC LHC24g3 @ ORNL



■ requested TTL ■ mean ■ maximum Time ■ predicted TTL

## Ex 3: RAW merging jobs



requested TTL mean maximum Time predicted TTL



## IPv6 support

SiteSonar WN probe

67% of the WNs are ok (3% can't resolve IPv6 and 30% cannot connect - not dual stacked) All components are IPv6 ready

Java, Python, Xrootd 5+ (client and server) Central services see ~70% connections on IPv6 These are job slots

IPv4 still required for legacy binaries (ROOT5 and Xrootd 3.x)57% of the VoBoxes are dual stacked97% of the storage volume is dual stacked & working

6 sites still don't have it (ORNL included), others show various IPv6 errors



### Data management

File crawler for a sampling of current problems <u>Health</u> (exists, can be read, checksum matches) Performance (throughput and stat time) Remote operations to repair content Dark data can be inferred with recursive `ls` Lost data recovery from lists provided by sites 3rd party transfers to re-establish consistency Catalogue cleanup if that's not possible



### Data and ops over http

**EOS** features fsck for reporting and repairing HTTP endpoint for data access Access to fsck reports over the same http port Unprivileged account with access just to this list Available from 5.2+ See Andreea's talk at the EOS workshop



## fsck report and repair

Set scaninterval for space and all filesystems

space config <space-name>(ex:default) space.scaninterval=<sec>

fs config <fs-id> scaninterval=<sec>

Activate collection or repair threads

fsck config toggle-collect [<threads number>]

fsck config toggle-repair [<threads\_number>]
HTTP(s) data and REST API

xrd.protocol XrdHttp: 1094 /usr/lib64/libXrdHttp.so

http.exthandler EosMgmHttp /usr/lib64/libEosMgmHttp.so
eos::mgm::http::redirect-to-https=1

xrd.tls /etc/grid-security/daemon/hostcert.pem /etc/grid-security/daemon/hostkey.pem

xrd.tlsca certdir /etc/grid-security/certificates/

http.gridmap /etc/grid-security/grid-mapfile

EOS MGM ENABLE REST API=1 (in /etc/sysconfig/eos\_env)

Xrootd and HTTP can run on the same port, no need to set up a different firewall for it



## **SciTags**

Experiment and activity accounting of network usage Two methods of reporting

- IPv4: flow marking, UDP firefly sent to R&E collectors
- IPv6: flow label, part of the IPv6 header, no need for extra packets

EOS only supports flow marking, for both protocols All ALICE operations are tagged with values from <u>here</u>

- xrdcp root://....?authz=<token>&**scitag.flow=330**&eos.app=JobWrapper
- 5<<6 + 10 (5==ALICE, 10==Data access)

Can be enabled on recent EOS and Xrootd versions (5.6.7+) Data to be sent to the nearest collector

{eu,us,global}.scitags.org - more to come



## SciTags config in EOS/Xrootd

xrootd.pmark use firefly scitag xrootd.pmark domain any xrootd.pmark debug xrootd.pmark trace xrootd.pmark ffdest 198.128.151.27:10514

For now use the IPv4 address of global.scitags.org Be ready to change the config at a later time eosalice.cern.ch is running it in production



### JWT for data access

</authz>

Industry standard, cross-experiments implementation No need for an extra auth plugin for ALICE
Similar to our current tokens, 1-to-1 mapping of operations
Under development in collaboration with the EOS team

authz>	
<file> <access>read</access> <turl>root://eosalice.cern.ch:1094//2c44fd849358</turl> <lfn>/alice/cern.ch/user/g/grigoras/wn.xml</lfn> <size>755</size> <guid>CF54C1A0-5E90-11E8-BEE3-2C44FD849358</guid> <md5>6f0d829a0f3fc8295f48c204a8053a75</md5> <pfn>/00/44960/cf54c1a0-5e90-11e8-bee3-2c44fd849358</pfn> <se>ALICE::CERN::EOS</se> </file>	<pre>"aud": "https://wlcg.cern.ch/jwt/v1/any", "sub": "aliprod", "nbf": 1711492372, "scope": "storage.write:/eos/dev/alice/test1", "iss": "https://alice-jcentral.cern.ch:8098/", "exp": 1711495972, "iat": 1711492372, "jti": "NDdjNmYyhmMmE1" }</pre>