A Large Ion Collider Experiment

European Organisation for Nuclear Research

CERN

ALICE

AliEn²
@GRID

AliEn status report

Miguel Martinez Pedreira

Infrastructure: MySQL

- New TaskQueue machine (same as catalogue)
  - SSDs, lots of memory…

- Updated MySQL version based on system to 5.6 latest Ubuntu (Costin)
  - Very few changes to [j]AliEn (reserved word 'partition')
  - Changes on some behaviours
    - Timeouts, config items, backup method ignored, readline usage…

- Significant problems early after
  - Change ordering of result sets (see later)
  - Deadlocks
    - Solved with locking (most of it)
    - Randomly happening on small and 'simple' tables
  - Simple queries taking very long
  - MySQL 5.7 in the to-do
  - Alerts added

New || Fixed: Broker

- Related to previous point: fix to **job priorities**
  - As said, new MySQL doesn't behave as before regarding the order of result sets
  - JobAgents where ordered by priority (one per user), and then the oldest queueId was taken (implicitly)
    - What about pushing the execution a part of my own jobs ? (e.g. merging)

- Introduced the use of 'Price' and 'Oldest Queue Id'

⬆ Priority   ⬆ Price   ⬇ oldestQueueId

New || Fixed: Broker

- Broker query simplification
  - Sending flag instead of all packages
  - Extra checks in other places in AliEn

+ New || Fixed

- TLS communications
  - Finally all services TLS-only

- OCDB optimizer fix
  - Performing very slowly after long runtime
  - Last reported mid-October
    - Module pushed to ISA on each `get` request
  - Reference: +100K files published to CVMFS
  - Convention to detect publication failures, and sending mail notifications

- From previous Offline: now latest as default AliEn version
  - New one to test soon, after next point below is OK

- Software updates
  - Latest OpenSSL-HTTPD combination
    - Proxy certificates support in future ?
  - Xrootd 3.3.6+xalienfs 1.0.14r1 built, under test
  - ApMon 2.2.23, IGTF Cas 1.69
  - Effort to move to new buildsystem (together with Dario)

New || Fixed

- APISCONFIG automatically loaded if not defined in Packages
  - Thus not needed anymore in the JDL

- CVMFS package environment parsing fixed
  - Miraculously nothing broke until the patch ☺

- Batch interfaces reviewed
  - SLURM, SGE…

- JobAgent checks if payload killed
  - Asks Authen each ~10mins
    - Same as for sending procinfo
    - Over a very small table (tokens)
  - To be tested gradually on a few sites
    - E.g. when deploying new AliEn
  - Authens are ready

New || Fixed

- Sorting InputData (SE split)
  - Suggested by an analysis expert
  - Has it done any change ?

- CVMFS not in default location
  - Can be configured from LDAP or environment

- Fix: not leaving empty roles after resync

+ Catalogue

- Catalogue load spikes comes from bad usage (though not very frequent anymore)
  - Coming from OCDB object finds from jobs
  - Or from jobs thrasing with too many queries (event loop?)
    - Need to control: could this be added easily and soon to the framework ?
      - Options in APIs to limit results, but not for Hz, common user…

- Very fast growth!
  - O(4-5M) PFNs per day ☺
  - ~45M last 7 days, ~2B entries
    - More activity [+ output pattern changed lately ?]

AliEn development - Miguel Martinez Pedreira
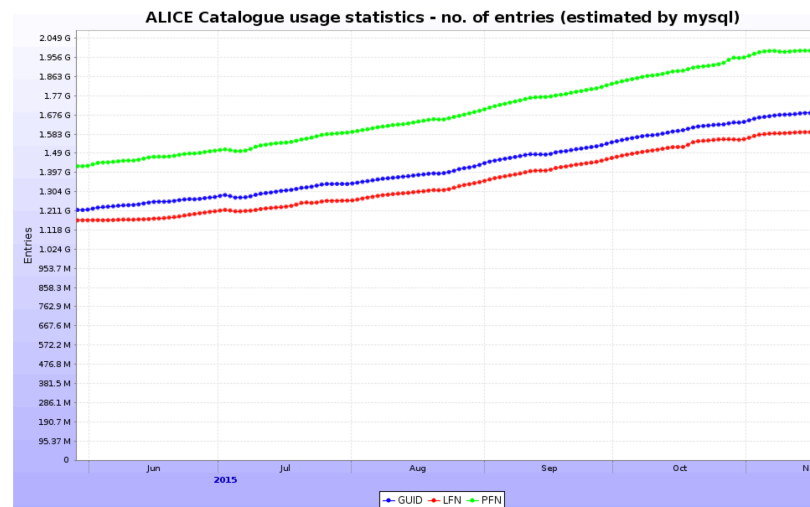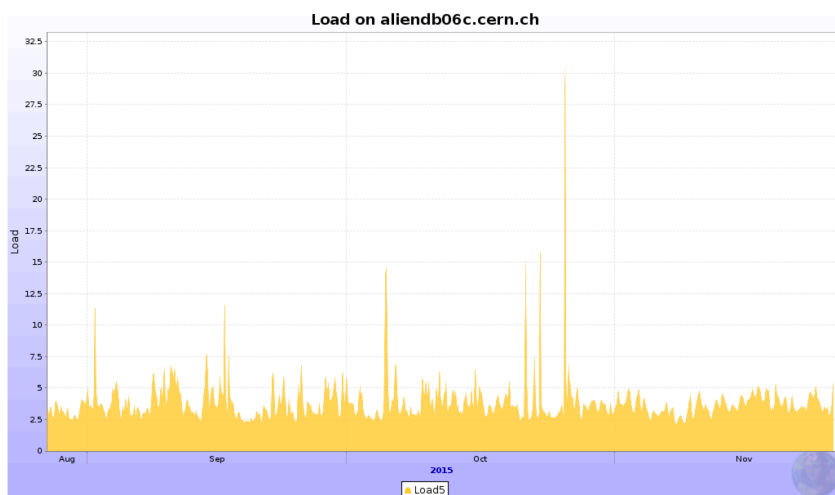
+ Catalogue

- PFN growth implies creating more G#L tables
  - Pseudo-manual operation so far…
  - Because the calculation of the value to insert in GUIDINDEX is not trivial

```perl
$filename = sprintf "%s/%02.2d/%05.5d/%s",$prefix,$self->{GUID}->GetCHash($guid),$self->{GUID}->GetHash($guid),$guid;

  $self->do(
"create function string2binary (my_uuid varchar(36)) returns binary(16) deterministic sql security invoker return unhex(replace(my_uuid, '-', ''))"
  );
  $self->do(
"create function binary2string (my_uuid binary(16)) returns varchar(36) deterministic sql security invoker return insert(insert(insert(insert(hex(my_uuid),9,0,'-'),14,0,'-'),19,0,'-'),24,0,'-')"
  );

  $self->do(
"create function string2date (my_uuid varchar(36)) returns char(16) deterministic sql security invoker return upper(concat(right(left(my_uuid,18),4), right(left(my_uuid,13),4),left(my_uuid,8)))"
  );

  $self->do(
"create function binary2date (my_uuid binary(16))  returns char(16) deterministic sql security invoker
return upper(concat(right(left(hex(my_uuid),16),4), right(left(hex(my_uuid),12),4),left(hex(my_uuid),8)))"
  );
```

  - Need to automate it !
    - Digest all these formulas (or look for a reasonable alternative)

+ Optimizers

- ToDo: improve splitting/inserting forking algorithm
  - Now: the parent splits 1 process and gives 15 jobs to it
    - Loops like this, making more children
  - Good when there are many masterJobs to split, not in the opposite case
    - Parent to calculate how many to split and fork total/N with a maximum of N per child
    - A masterJob can take potentially *long* time to split

- ToDo: tag to optimize number of baskets?

- ToDo: do we want to block a queueId-site when failed N times?

- We need more monitoring for the Optimizers
  - Right now, checked through the parent → not enough!
  - A possibility is to use PIDs of the processes and subprocesses and establish status and thresholds checks
    - Depending the type of optimizer
  - ML offers something that fits ? -> ApMon parameters + notifications from ML
    - Or AliEn itself, cronjob, or similar
    - jAliEn re-implementation

+ Extra

- Xrootd proxy working ok so far in GSI
  - But users *near* GSI are assigned to it as site, and cannot go through the proxy…
    - The few users made a quick hack, but proxy was supposed to be open in principle ?
    - How easy to add parameters in API calls [TAliEn, etc] ?

- Document on 'Decommissioning ALICE native xrootd servers and dealing with data loss' accesible here:
  - http://alien2.cern.ch/index.php?option=com_content&view=article&id=56&Itemid=96
  - Merci to Christophe Diarra!

jAliEn (+Pavlo's and Andrey's talk)

- JobAgent
  - Full chain now working!
    - With CVMFS environment setup
    - Without connection through ROOT
  - But many things yet pending:
    - Control process
      - Memory, disk and CPU checks…
    - Not all output options yet
    - JDL options and environment
    - And more…

- Brokering fully working as well
  - With all the changes seen in previous slides

+ Questions / Comments

- Record of 97K+ running jobs ☺

- We ~~still miss~~ have a *real* test environment
    - Replica of TQ, Catalogue, ADMIN dbs…
    - Using production LDAP (could be cloned too)
    - Running usual central services!
        - Of course, without all the concurrency of the GRID…

- Questions?