

Moving to HTCondor (and fighting covid in the middle)

S. Dal Pra and CNAF colleagues,

Also starring:

T. Boccali (INFN-Pisa)

Alberto Boldrini (Sibylla Biotech srl)

At a glance

- In the beginning (2018)
 - ~ 400 KHS06, 36500 slot, 850 physical WN
 - (5+2) x CREAM-CE / LSF 9.1.3
 - ~ 40 User groups: 24 Grid VOs, ~ 25 local
- A minimal HTCondor testbed (May, 2018)
 - Practice, get some pilots submitted by LHC exp,
 - writing scripts and tools, adapt to work with existing facilities.
 - Plan: start a small production cluster, then move WN there.
- In the end (Jun. 2020)
 - (6+1) x HTC-CE, 1 x CM, 850 WN
 - (2 x Ovirt, 2 x Vmware, 2 x Bare metal)
 - 1 x SN, shared for Remote Submission (FS_REMOTE)

Migration plan

1. The switch should be (at most) transparent for users
 - LHC VOs not a big issue: ready to access local resources through a CE
 - Local users: move from **bsub** to **condor_submit**
2. Minimize impact on site management tools
3. The switch should be reversible (rollback to LSF possible as extreme ratio)
4. Allow coexistence of two distinct prod clusters (allow users to take their time)
5. Cluster management should remain “similar”, at least initially.

Adapting Management

- LSF had two nice features:
 - few text files on a shared fs to configure the whole cluster
 - Hostgroups: hierarchical sets of named hosts, can be combined with simple set operations (union and difference)
- These were adapted to HTCondor:
 - We keep HTCondor configuration files on a shared fs
 - a node in the groups: `wn`, `cpu2019`, `atlas` read from `wn.conf`, `cpu2019.conf`, `atlas.conf` and finally from `<my_hostname>.conf`
- Flexible: temporarily adding a few classAd to an arbitrary set of WNs is a matter of defining the hostgroup `<groupname>` and the classads into `<groupname>.conf`

Using hostgroups

```
[root@farm-ops conf]# cat htc_hostgroup.conf
[...]
cn61006 (cn-610-06-01 cn-610-06-02 ... cn-610-06-12)
rack610 (cn61001 cn61002 cn61003 cn61004 cn61005 cn61006)
cineca (rack608 rack609 rack610)

# cn-610-06-04 : broken disk; wn-200-11-01-01-a : testing
quarantena (cn-610-06-04 wn-200-11-01-01-a)
sl7_nodes (sl7_t1 cineca cpu2017 cpu2019 ~quarantena)

[root@farm-ops conf]# cat cineca.conf
t1_wn_hs06 = 832
StartJobs = ((AcctGroup == "atlas" || AcctGroup == "cms") && RequestCpus == 8)

[root@farm-ops conf]# cat quarantena.conf
START = False
```

Adapting cmdline tools

- **HTC:** `condor_q` and `condor_status`. Very powerful, yet easy to get cumbersome. Most common commands were emulated by using python bindings:

LSF	HTC
<code>bjobs</code>	<code>hjobs.py</code>
<code>bqueues</code>	<code>Hqueues.py</code>
<code>bhosts</code>	<code>Hhosts.py</code>

```
[root@htc-2 ~]# /usr/share/htc/cnaf/bin/hjobs.py | head
JobId RemoteOwner GlobalJobId JobStart Cpus Machine TotalCpus LoadAvg CPUsUsage
13946863.0 somelocaluser sn-01 2021-03-15:03:03:47 1 wn-204-13-05-05-a 40.0 1.0 0.999946352996
2061220.0 pillhcb031 ce05-htc 2021-03-15:07:17:40 1 wn-205-10-20-02-a 64.0 1.0 0.99692285813
3304976.0 alicesgm008 ce03-htc 2021-03-15:00:55:24 1 wn-205-13-13-06-a 40.0 0.69 0.437431741832
3307461.0 pillhcb031 ce03-htc 2021-03-15:09:33:20 1 wn-205-11-39-01-a 40.0 1.05 0.999249707998
2484979.0 alicesgm008 ce01-htc 2021-03-14:19:10:26 1 wn-205-08-09-03-a 64.0 0.98 0.985553282691
2483132.0 alicesgm008 ce01-htc 2021-03-14:10:36:10 1 wn-205-11-13-01-a 40.0 0.97 0.983154470612
2489162.0 atlasprd011 ce01-htc 2021-03-15:10:42:32 8 cn-608-01-07 72.0 1.91 1.02848197081
2210728.0 pillhcb031 ce06-htc 2021-03-15:09:06:33 1 wn-204-13-01-05-a 40.0 1.05 0.99931512371
3344123.0 alicesgm008 ce02-htc 2021-03-15:00:23:23 1 wn-205-08-07-01-a 64.0 0.78 0.791479697315
```

Adapting Accounting

- **LSF:** Using custom accounting system from Oct. 2013. Usage records collected daily from LSF and CREAM-CE and stored on a PostgreSQL database
- **HTC:** We configure PER_JOB_HISTORY_DIR
 - One text file per job, as key/value pairs
 - HS06 of the WN present as custom Job Classad
 - CRON script (**/3 min*) to parse the file and "INSERT INTO" the database. Files are then moved to a backup storage
 - This enables us to "merge" monitoring and accounting
 - Todo: we are going to renew the parser to use
 - `condor_q -jobads history.<ClusterId>.<ProcId> -af <classad list>`

Adapting Accounting

Apel records obtained as a SQL VIEW:

```
acct=> SELECT * FROM apelhtjob WHERE "Processors"=8 LIMIT 1;
-[ RECORD 1 ]-----+-----
Site                | INFN-T1
SubmitHost          | ce01-htc.cr.cnaf.infn.it#994620.0#1598009345
MachineName         | htc-2.cr.cnaf.infn.it
Queue               | atlas
LocalJobId          | 994620
LocalUserId         | atlasprd011
GlobalUserName      | /DC=ch/DC=cern/OU=Organic Units/OU=Users/CN=atlpilo1/CN=614260/CN=Robot: ATLAS Pilot1
FQAN                | /atlas/Role=production/Capability=NULL
VO                  | atlas
VOGroup             | /atlas
VORole              | Role=production
WallDuration        | 5484
CpuDuration         | 38555
Processors          | 8
NodeCount           | 1
StartTime           | 1598012883
EndTime             | 1598018367
InfrastructureDescription | APEL-HTC-HTC
InfrastructureType   | grid
ServiceLevelType    | HEPSPEC
ServiceLevel        | 13.419
```

```
acct=> █
```


Migration start

Having built a set of tools and accounting in place we could start, by growing the small production HTC cluster:

- Start with 16 WNs, have LHC VOs using it
- Add more HTC-CEs
- Add more WNs, Migrate VOs and local groups.
- Leave LSF with a small bunch of WNs for last late users (May to Jun 2020).
- **Note:** during transition, providing correct shares was troublesome.
- **Migrating WNs:** No need to drain the machine: LSF and HTC jobs can coexist in the same WN, by reducing node slots in LSF and setting NUM_CPUS accordingly at the STARTD

Sibylla Covid-19 Project

Fighting covid

- **Sibylla:** INFN spin-off, operating on drugs design
- **Covid-19 Project:** simulating the folding of the ACE2 protein
- **Goal:** find a drug to treat a sick person
- INFN-T1 was requested (March 20, 2020) to provide $\sim 50\%$ of its computing power for a period up to a month
- We were just in the middle of the migration: 50% on LSF and 50% on HTC

The Sybilla use case

1. 32 cores jobs
2. need for AVX2 compatible machines
3. Several days runtime ($\sim 5 \dots \sim 15$)
4. Need to have a single run of jobs, in parallel.
5. Need write access for a shared area (for checkpointing)
6. Need for GCC 9.x
7. User side more used with LSF than HTCondor

Problems

1. need for AVX2 machines on LSF
 - Most of these already moved to HTC
 - 72 slot each --> 2 jobs and 8 unused slots
2. 32 cores and several days of runtime
 - Long draining time before moving back to HTC
3. Need for GCC 9.x
 - Available from /cvmfs/
4. Write access need on a shared area
 - A GPFS volume was set up with rw permission
5. Needed Memory and storage size not an issue
6. Severe underpledge for other experiments
 - LHC VOs promptly agreed to halve their expected workload.

Provisioning actions

1. HTC: Draining

- (START = False) on the AXV2 WNs

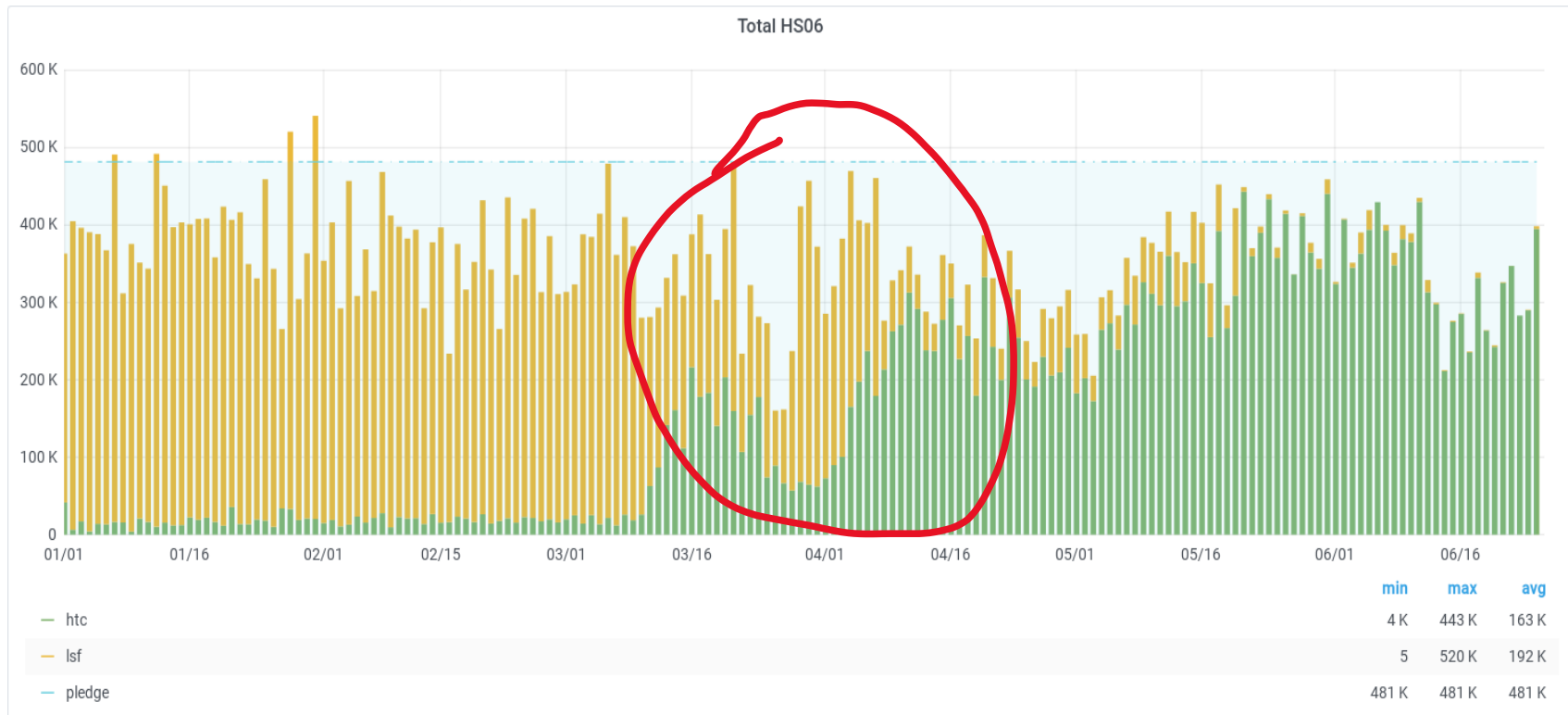
2. LSF:

- Setup **biotech** queue and dedicated hostgroup (**closed_Adm**, initially)
- Enable WNs (**badmin hopen**) whose HTCondor jobs are done (STARTD remains active anyway).

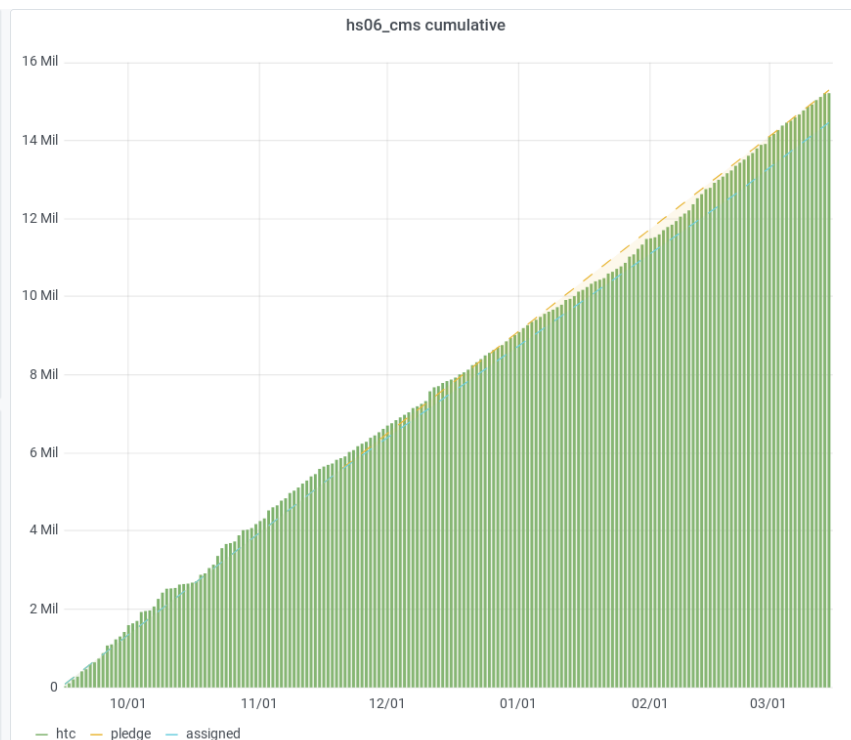
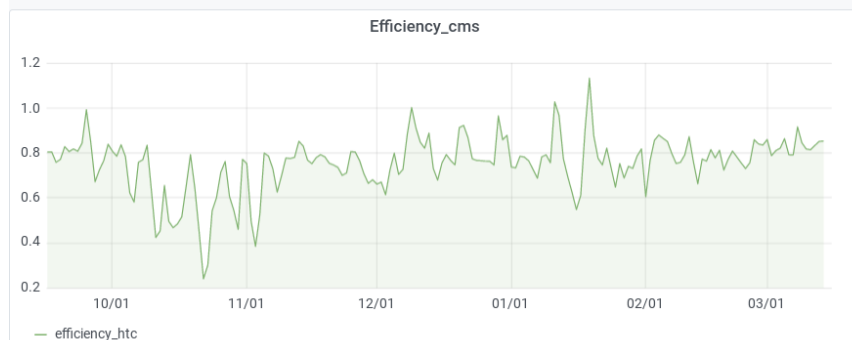
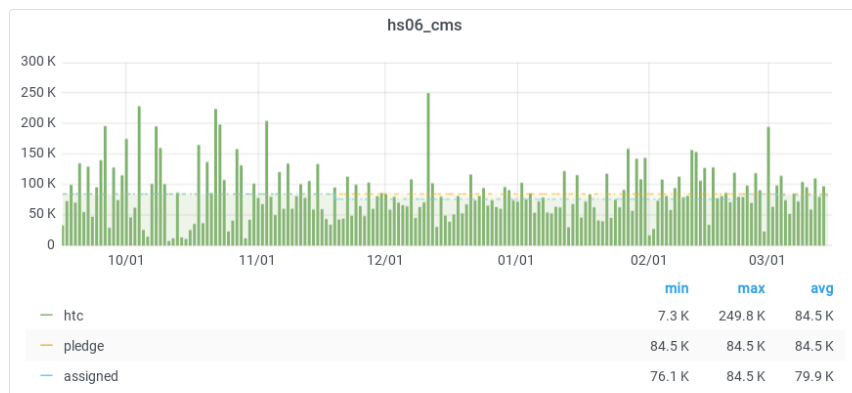
3. Sybilla: Set up and tune their workload on a test machine in the meanwhile

Results

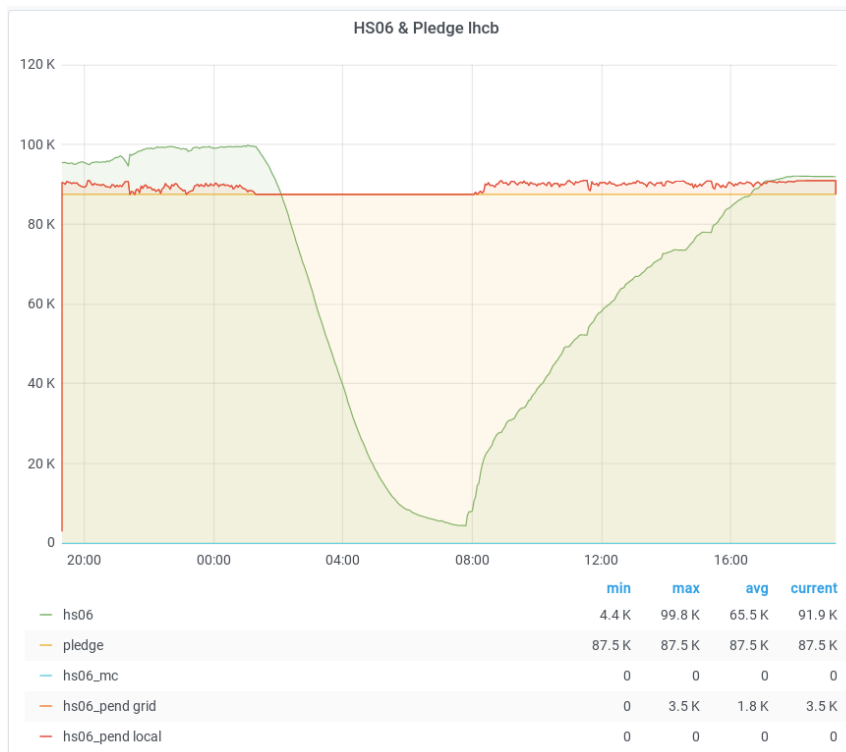
- First Sybilla jobs running by **March 23**, Last jobs ended by **April 7**
- After jobs submission end, WN were closed in LSF (**admin hclose**) and enabled in HTCondor (**NUM_CPUS = 32** initially) to reduce cputime loss due to draining



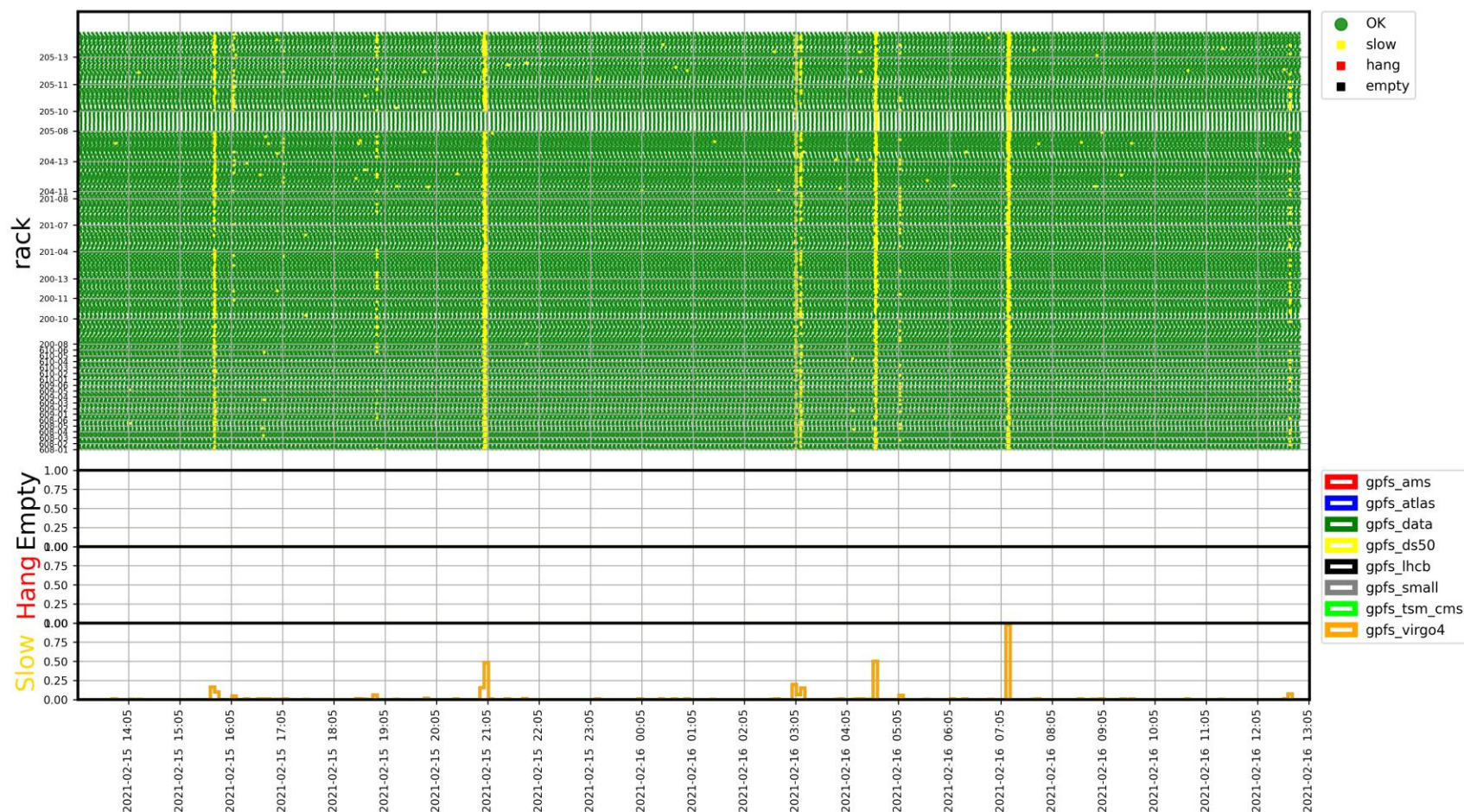
Accounting dashboard



Monitoring / accounting



WNs and GPFS mounts



Experience so far

- HTCondor is an improvement for us
 - Stable and robust, with a rich and powerful set of capabilities
 - New use cases being supported (DAG jobs, and other being investigated)
- Configuring can be difficult
 - Help from developers and community
 - Most of the initial difficulties have been solved
 - Some issues with fairshare with multicore and singlecore mix still being addressed

The End