Brokering to heterogeneous resources on the Grid: LHCb
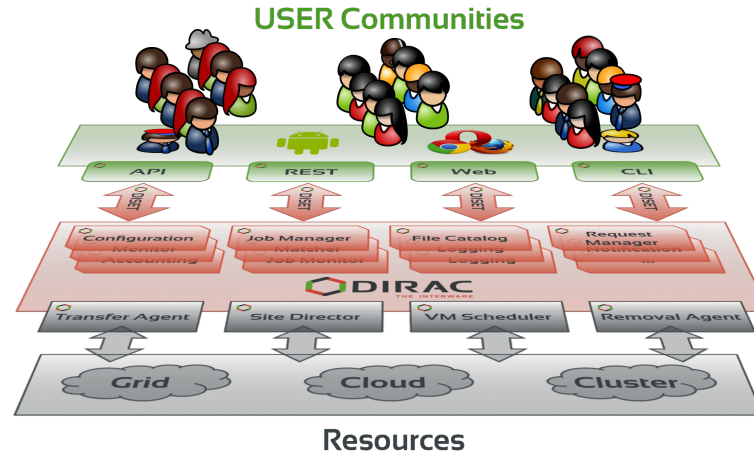
Concezio Bozzi (Federico Stagni)

concezio.bozzi@fe.infn.it

federico.stagni@cern.ch

Everything through DIRAC

- A not-only LHCb software framework for distributed computing
- Other experiments using DIRAC can profit of developments done
  - But there's one bit (see later) that is, for now, LHCb specific

[WMS] basics

Pilots-based WMS, with late binding

- **Users** define and submit **jobs**. Jobs have **requirements**. Job descriptions are stored in DIRAC's Job DB.
- Independently, **Pilots** are started (1) on the **sites' worker nodes (WN)**
- Pilots will try to **match** (2) the worker nodes' capabilities to the jobs requirements.
- Jobs are started on WNs. DIRAC monitors their progress.
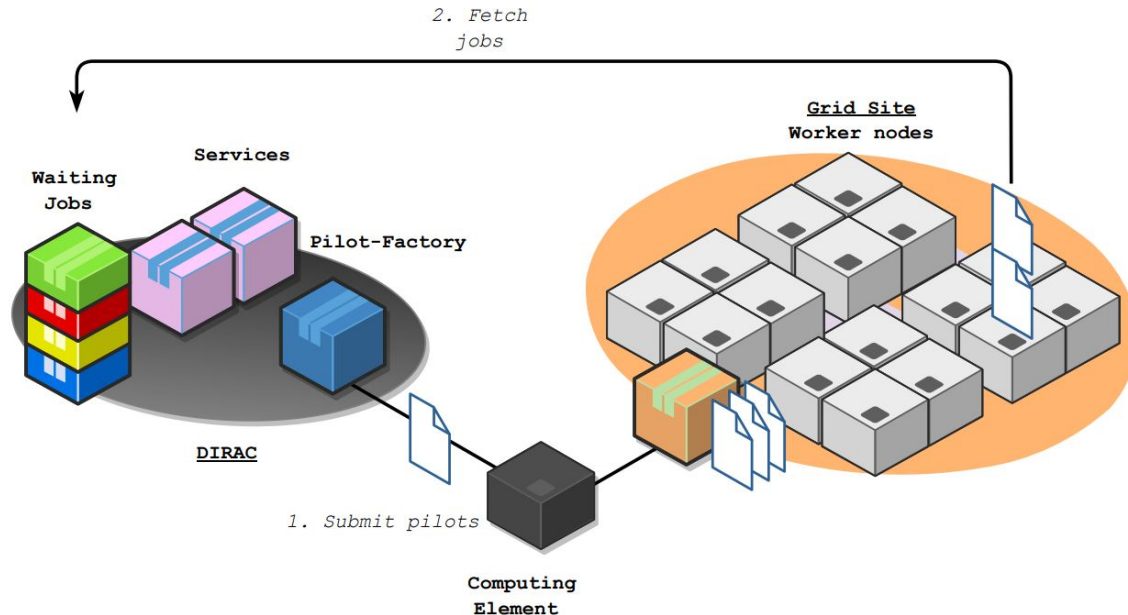
[WMS] Computing resources
where to run the pilots

- **Grids** (nowadays: HTCondor, ARC)
  - through proxies and/or tokens
- **Clusters** behind a BS
  - access through SSH/GSISSH tunnel
    - a really thin layer that we call "SSH CE"
  - and then talks with batch system
- **VMs scheduler**:
  - Based on apache libcloud
  - Contextualization from standard images
    - with, at least, the DIRAC pilot
- **Vacuum**:
  - Vac resources
  - HLT farm (LHCb)
- **HPC sites**
  - it often means at least SSH+Slurm
    - → more in CHEP pres

Not only pilot model
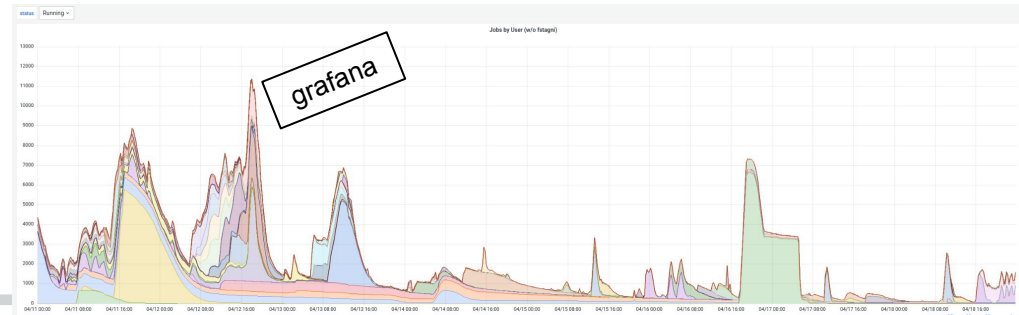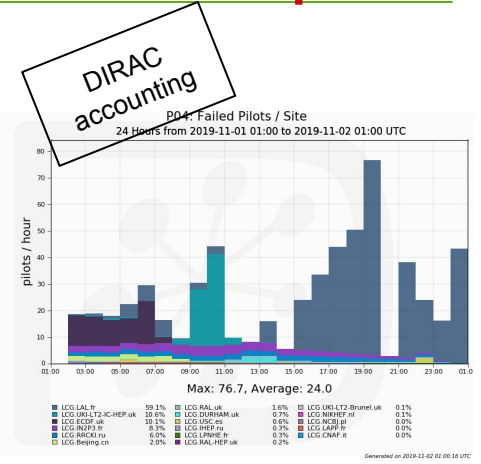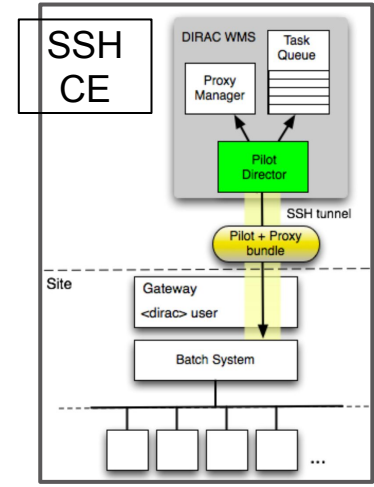
PushJobAgent

- For "closed" HPCs
- Works as a Pilot-Job that would be executed outside of the HPC
- Fetches jobs, manages their input and output data, and solely submit the application (i.e. Gauss, for LHCb) to the HPC.
- Requires a direct access to the LRMS.

A. Boyer, CHEP talk

Matching job slots to job requirements
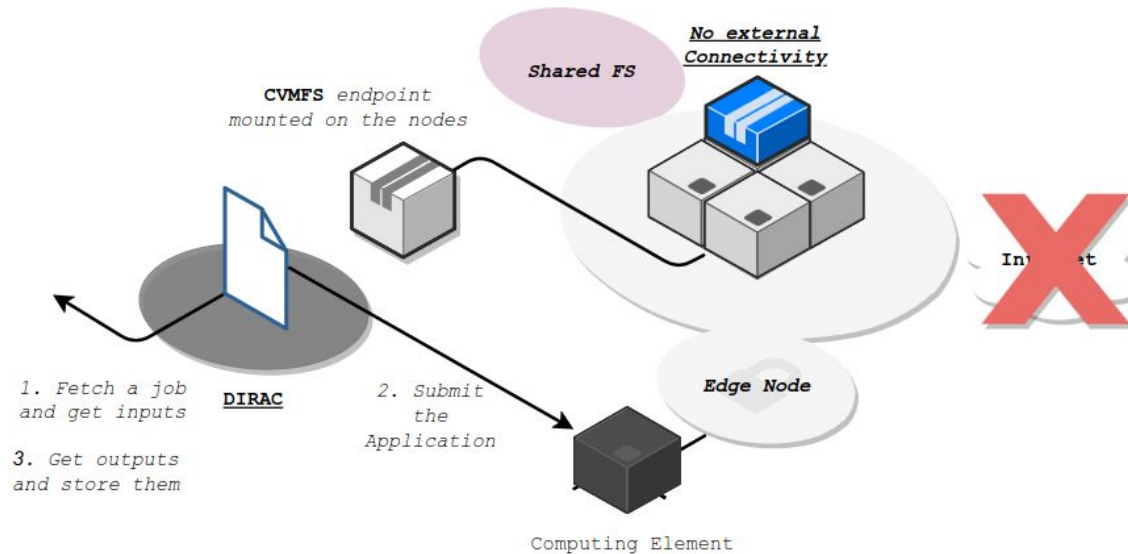
- Jobs have an optional "platforms" requirement (a list of strings) – meaning the job CAN run here, e.g.:
  - `Linux_x86_64_glibc-2.27` [DIRAC generic]
  - `skylake_avx512-centos7` [LHCb "specific" – see later]
  - `volta_cuda-alma9` [a free string I just invented]
- The pilots run a script (`LbPlatformUtils`) to identify the local platform
- The DIRAC job matching process matches platform job slot capabilities with jobs' platforms requirements

LbPlatformUtils

- ## LHCb-developed library (but really generic)
  - identifies the platform, of the node where it ran
  - used for matching of LHCb jobs
    - actually, very little of LHCb in there
  - handles compatibilities
  - finds out the possibility to use appcontainer (which will extend the available platform's list)
  - recognizes the instructions set available

```
LHCb DIRAC platforms examples:

'cannonlake-centos7', 'skylake_avx512-centos7', 'skylake-centos7'...
'skylake_avx512-slc6', 'skylake-slc6', 'broadwell-slc6',...

'cannonlake-slc5', 'skylake_avx512-slc5'...
```

https://gitlab.cern.ch/lhcb-core/LbPlatformUtils

The DIRAC approach: "Tags"

- <u>On top</u> of the "JDL platforms requirement" DIRAC provides a **<u>generic</u>** mechanism for matching computing capabilities with resource providers
- Can be used, on one side, by the users to "mark (tag)" their jobs with requirements, and on the other to "mark (tag)" Sites or CEs or Queues.
- Use cases include:
 - MP jobs
  - Queues with "special" platforms

On the server side, it's a matter of what's written in the DIRAC Configuration:

- A Site or CE or Queue that has **Tag=X** is *capable* of running jobs that *require* **Tag=X**.
- A Site or CE or Queue that has **RequiredTag=X** will *match only* jobs that *require* **Tag=X**.

- Gauss simulation application successfully built on ARM
  - Simulation takes 90% of CPU work on the grid
- Physics validation needed
  - LHCb performance & regression test suite (LHCbPR)
- Continuous integration / nightly builds a prerequisite
  - server available from CERN/IT
- LbPlatformUtils need to be properly extended

9

- Allen application for HLT1 on GPUs in the online event building farm
  - Allen CPU version used to emulate HLT1 on simulation productions running on the grid
    - Slower than GPU version, but compute work still negligible wrt to G4 simulation
- Both LbPlatformUtils and DIRAC tags can be used to access GPUs on the grid
  - ML training another possible use case

10

Few bullet points
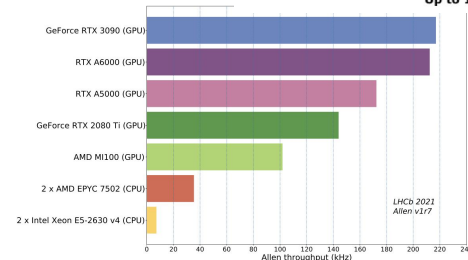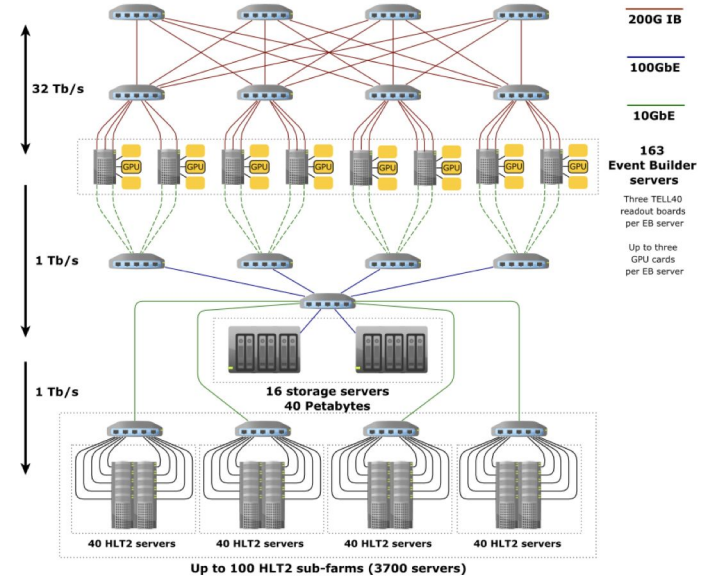
- All "non-x86_64" computing resources are, at the moment, opportunistic
    - and, in general, they are not many…
- (LHCb)DIRAC (client, so the Pilot) can be installed on `x86_64, ppc64le, aarch64`
- `LbPlatformUtils` will be expanded for supporting all LHCb "platforms of interest"
- LHCb does not have, at the moment, interest in Power9

Questions/comments?