Transparent extension of INFN-T1 with heterogeneous computing architectures

Stefano Dal Pra[1], Daniele Spiga[2], Tommaso Boccali[3], Lorenzo Rinaldi[4]

[1]INFN CNAF (Bologna),[2]INFN Sezione di Perugia,[3]INFN Sezione di Pisa,[4]INFN Sezione di Bologna

- The INFN-CNAF Tier-1 is engaged for years in a continuous effort to integrate its computing centre with more tipologies of computing resources.
- In particular, the challenge of providing opportunistic access to nonstandard CPU architectures, such as PowerPC or hardware accelerators (GPUs) has been actively exploited.
- We have worked to a solution to transparently integrate access to ppc64 CPUs as also GPUs.
- Further improvements and how this will meet requirements and future plans for the new tecnopolo centre, where the CNAF Tier-1 will be hosted soon.

Integrating opportunistic pledged resources with opportunistic ones can be done almost seamlessly, however a few important details must be considered:

- **Usage Policy** (QoS no more homogeneous through the pool)
- **HW** (The node might have GPUs onboard)
- **Network** (different bandwiths, ACLs and routing)
- **Storage** (No "pledged storage" can be expected to be available)
- **Runtime environment** (less manageability respect to pledged nodes)
- **Fairshare** (opportunistic resources should not impact the balance on the pledged ones)

An example submit file for M100 resources

Upon submission to a HTCondor-CE It matches the JobRouter entry for M100. This assumes that a custom classad "WantRoute" specified the desired resource type.

```
[sdalpra@ui-htc ~]$ cat ce_m100_p9.sub
universe = vanilla
use_scitokens = true
+Owner = undefined
+WantRoute = "cms_m100"
executable = p308/htcp308.p9
output = htcp308_$(ClusterId).$(ProcId).out
error = htcp308_$(ClusterId).$(ProcId).err
log = htcp308_$(ClusterId).$(ProcId).log
arguments = "0 0 1 2001"
queue 1
```

```
JOB_ROUTER_ROUTE_cms_m100 @=jrt
REQUIREMENTS WantRoute =?= "cms_m100") && (AuthTokenIssuer =?=
"https://cms-auth.web.cern.ch/" && AuthTokenSubject == "78f275d5-bb1a-4b2d-9956-f82316a8482e"
 UNIVERSE VANILLA
 SET Requirements (TARGET.Arch =?= "ppc64le")
@jrt

JOB_ROUTER_ROUTE_atlas_m100 @=jrt
REQUIREMENTS WantRoute =?= "atlas_unibo") && (AuthTokenIssuer =?=
"https://atlas-auth.web.cern.ch/" && AuthTokenSubject == "ccff569b-bda5-45ac-91bb-44c8e198a385"
 UNIVERSE VANILLA
 SET Requirements (TARGET.Arch =?= "ppc64le")
@jrt
```

This has been tested to transparently extend the INFN-T1 Grid computing centre with
- Power9 based machines and V100 GPUs from the Marconi 100 HPC cluster managed by CINECA
- Standard resources from the HPC farm of Open Physics Hub project managed by UniBO

Extension model on M100
1. Owned script submit Slurm Jobs
2. The Slurm job launches a singularity container
3. The container starts a HTCondor STARTD
4. The STARTD has token credentials to join the HTC pool at CNAF
5. Jobs requiring M100 resources are properly routed and queued for execution in the M100 node
6. for execution in the M100 node

Easy to replicate for other external resources

Minimum Requirements:
- Outbound connectivity toward a CNAF host (CCB)
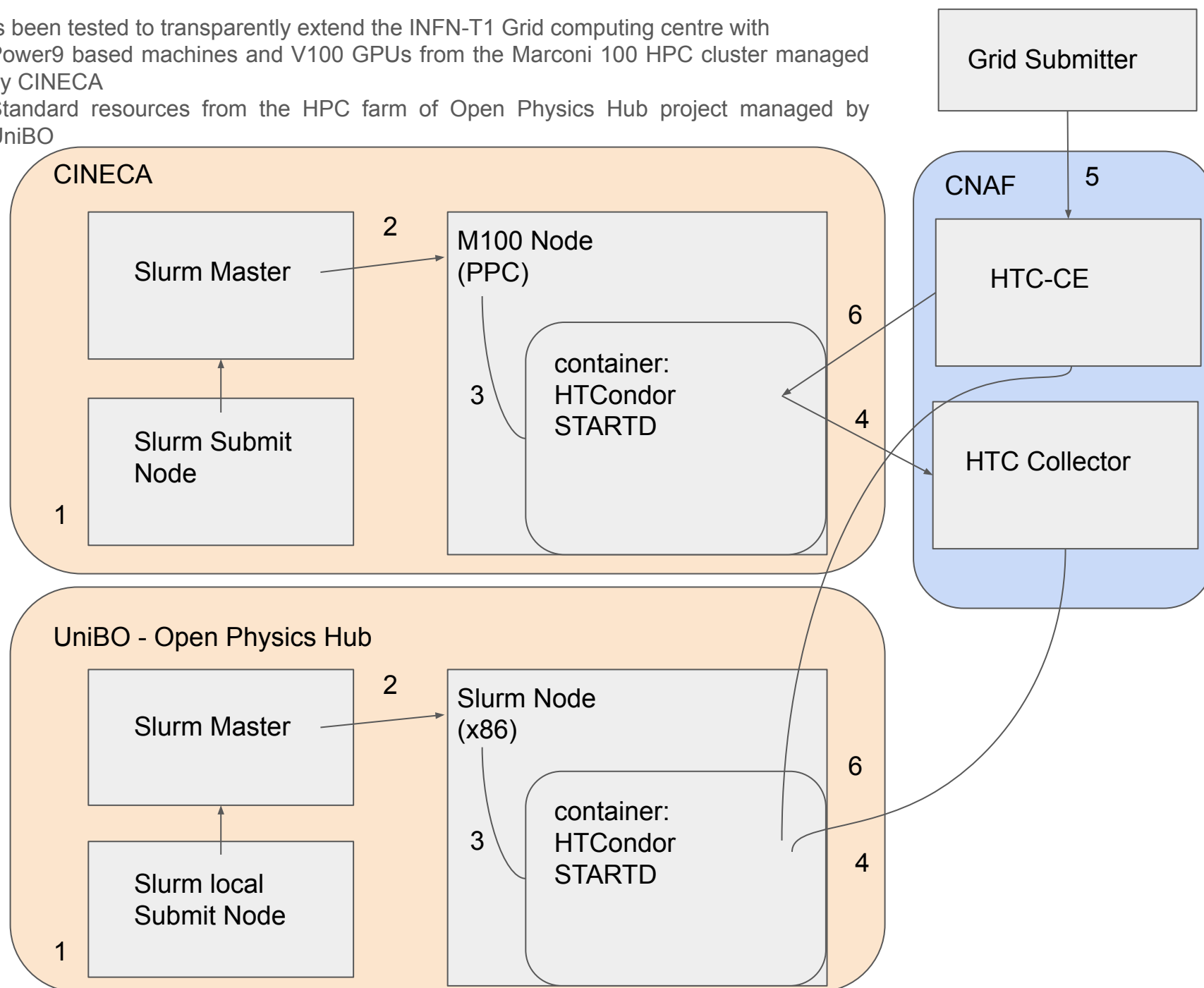
- **The POWER CPU architecture has been now made available to INFN-T1 user communities having suitable payloads.**
- **V100 GPUs have been also made available.**
- **This model can be used to transparently extend the INFN-T1 pool to external computing resources of different typologies (x86, ppc, arm, GPUs; all of these have been tested).**