

INFN Dynamic
Extension to the
Azure Cloud

Antonio
Falabella(1) -
Stefano Dal Pra(1)
- Vincenzo
Ciaschini(1) -
Tommaso
Boccali(2) - Daniele
Spiga(3)

INFN - CNAF

Farm Extensions

Farm Extension to
MS Azure

Dynfarm

Summary

INFN Dynamic Extension to the Azure Cloud

Antonio Falabella(1) - Stefano Dal Pra(1) - Vincenzo Ciaschini(1) -
Tommaso Boccali(2) - Daniele Spiga(3)

INFN - CNAF (Bologna)(1), INFN Pisa(2), INFN Perugia(3)
ISGC - Academia Sinica in Taipei, Taiwan from 16-23 March 2018

March 23, 2018

INFN Dynamic
Extension to the
Azure Cloud

Antonio

Falabella(1) -

Stefano Dal Pra(1)

- Vincenzo

Ciaschini(1) -

Tommaso

Boccali(2) - Daniele

Spiga(3)

INFN - CNAF

Farm Extensions

Farm Extension to
MS Azure

Dynfarm

Summary

1 INFN - CNAF

2 Farm Extensions

3 Farm Extension to MS Azure

4 Dynfarm

5 Summary

INFN Dynamic
Extension to the
Azure Cloud

Antonio
Falabella(1) -
Stefano Dal Pra(1)
- Vincenzo
Ciaschini(1) -
Tommaso
Boccali(2) - Daniele
Spiga(3)

INFN - CNAF

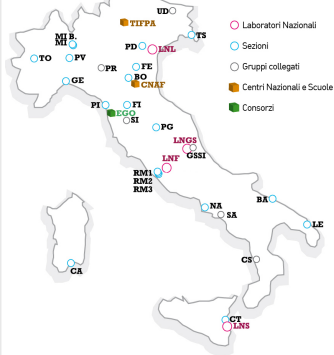
Farm Extensions

Farm Extension to
MS Azure

Dynfarm

Summary

- INFN is a public research agency which mission is the study of the fundamental constituents of matter and the laws ruling their interaction
- 4 Nationals laboratories
- 20 Units/Divisions (Sezioni) hosted at University Departments
- 2 National Centers
- 1 National Computing Center (CNAF - Bologna)
- 11 Associated groups
- > 1700 staff
- > 300 fixed term
- ~ 100 post docs
- ~ 500 fellowships
- ~ 1000 associated



INFN Dynamic
Extension to the
Azure Cloud

Antonio
Falabella(1) -
Stefano Dal Pra(1)
- Vincenzo
Ciaschini(1) -
Tommaso
Boccali(2) - Daniele
Spiga(3)

INFN - CNAF

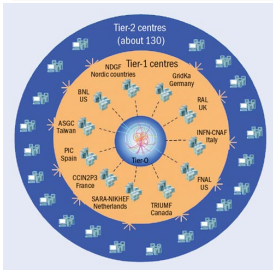
Farm Extensions

Farm Extension to
MS Azure

Dynfarm

Summary

- Supports four major LHC experiments
 - and more than 30 other collaborations
- The CNAF Tier1 is part the WLCG



INFN Dynamic
Extension to the
Azure Cloud

Antonio
Falabella(1) -
Stefano Dal Pra(1)
- Vincenzo
Ciaschini(1) -
Tommaso
Boccali(2) - Daniele
Spiga(3)

INFN - CNAF

Farm Extensions

Farm Extension to
MS Azure

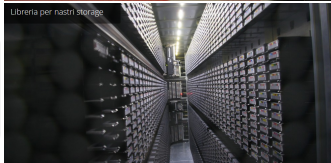
Dynfarm

Summary

- >25000 CPU Cores
- ~ 30PB Disk storage
- ~ 70PB Tape storage
- Fiber Channel + InfiniBand
SAN / TAN

- Plus a small HPC farm
InfiniBand interconnected with
accelerators (GPUs and PHIs)
30TFlops

- Total load is currently 1.4 MW



INFN Dynamic
Extension to the
Azure Cloud

Antonio
Falabella(1) -
Stefano Dal Pra(1)
- Vincenzo
Ciaschini(1) -
Tommaso
Boccali(2) - Daniele
Spiga(3)

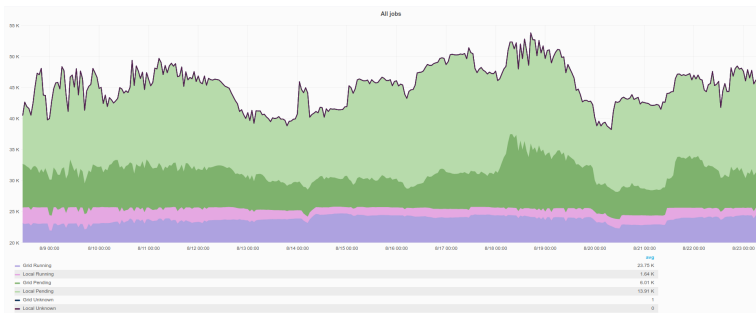
INFN - CNAF

Farm Extensions

Farm Extension to
MS Azure

Dynfarm

Summary



- Peak ~ 26K Grid + Local Job Running
- Avg ~ 15K Grid + Local Job Pending



INFN Dynamic
Extension to the
Azure Cloud

Antonio
Falabella(1) -
Stefano Dal Pra(1)
- Vincenzo
Ciaschini(1) -
Tommaso
Boccali(2) - Daniele
Spiga(3)

INFN - CNAF

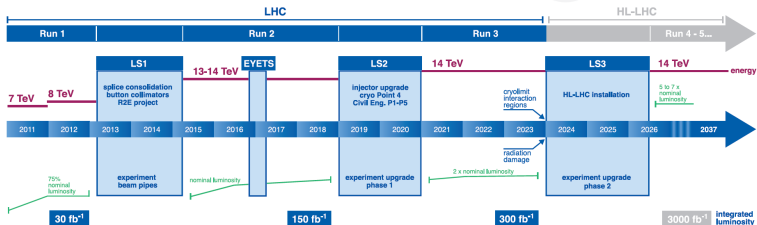
Farm Extensions

Farm Extension to
MS Azure

Dynfarm

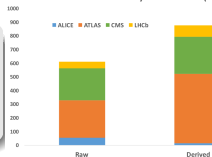
Summary

LHC / HL-LHC Plan

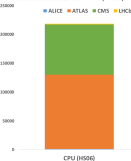


- An internal review shows that CNAF needs to be prepared to increases of more than a factor 50 in the next 10 years

Data estimates for 1st year of HL-LHC (PB)



CPU Needs for 1st Year of HL-LHC (HIS06)



INFN Dynamic
Extension to the
Azure Cloud

Antonio
Falabella(1) -
Stefano Dal Pra(1)
- Vincenzo
Ciaschini(1) -
Tommaso
Boccali(2) - Daniele
Spiga(3)

INFN - CNAF

Farm Extensions

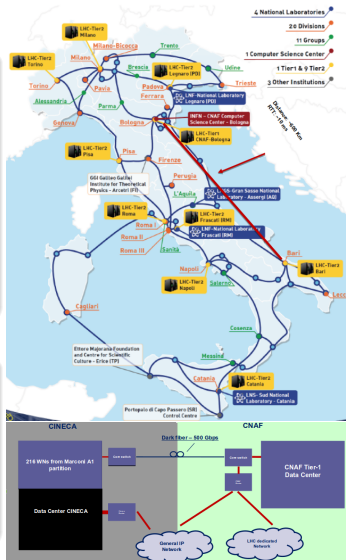
Farm Extension to
MS Azure

Dynfarm

Summary

- ~ 13% of CPU resources pledged to WLCG experiments are located in Bari-RECAS data center
- Transparent access for WLCG experiments
- Similar to CERN/Wigner extension
- 20 Gbps Level 2 VPN
- Disk cache provided via GPFS-AFM

- In 2018 ~ 170 kHS06 will be provided by CINECA
- Setup on going
- 500 Gbps VPN ready
- No disk cache, direct access to CNAF storage



INFN Dynamic
Extension to the
Azure Cloud

Antonio
Falabella(1) -
Stefano Dal Pra(1)
- Vincenzo
Ciaschini(1) -
Tommaso
Boccali(2) - Daniele
Spiga(3)

INFN - CNAF

Farm Extensions

Farm Extension to
MS Azure

Dynfarm

Summary



- Another solution to resource scarcity is the utilization of external clouds resources

- In the last few years at CNAF we are getting experience with the extension of the farm to external Clouds and HPC providers



aruba
CLOUD



CINECA



HNLIX
NUBULA
THE SCIENCE CLOUD



Microsoft
Azure

INFN Dynamic
Extension to the
Azure Cloud

Antonio
Falabella(1) -
Stefano Dal Pra(1)
- Vincenzo
Ciaschini(1) -
Tommaso
Boccali(2) - Daniele
Spiga(3)

INFN - CNAF

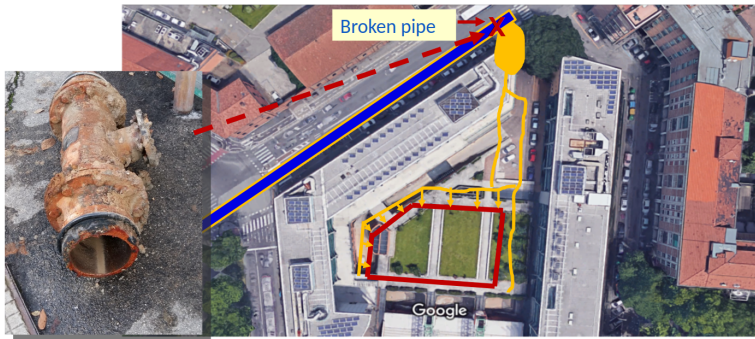
Farm Extensions

Farm Extension to
MS Azure

Dynfarm

Summary

- The flood happened on November 9 early in the morning
- Breaking of one of the main water pipelines in Bologna
- Also the road near CNAF seriously damaged



INFN Dynamic
Extension to the
Azure Cloud

Antonio
Falabella(1) -
Stefano Dal Pra(1)
- Vincenzo
Ciaschini(1) -
Tommaso
Boccali(2) - Daniele
Spiga(3)

INFN - CNAF

Farm Extensions

Farm Extension to
MS Azure

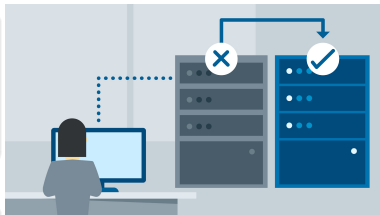
Dynfarm

Summary

- Computing farm: ~ 34 kHS06 are now lost (~ 14% of the total 2018 capacity)
- Disk storage in recovery
- 1 Tape drive damaged

- Tapes : 136 tapes damaged

- The extension to Bari-Recas and CINECA helped us to be online faster



INFN Dynamic
Extension to the
Azure Cloud

Antonio
Falabella(1) -
Stefano Dal Pra(1)
- Vincenzo
Ciaschini(1) -
Tommaso
Boccali(2) - Daniele
Spiga(3)

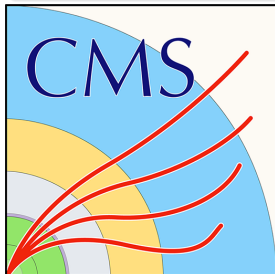
INFN - CNAF

Farm Extensions

Farm Extension to
MS Azure

Dynfarm

Summary



In 2017 we receive a 20000
USD grant to use with MS
Azure Cloud Infrastructure



The test has been done for CMS which is one of biggest users of the CNAF tier1 with 66KHS06 allocated CPU resources, 5.8PB of disk and 21PB of tape

- DODAS - Dynamic on-Demand Analysis Service
- Among the different CMS activities (Simulation, Reconstruction, User analysis) we chose to submit only MC mutli core jobs to Azure in order to simplify the exercise

INFN Dynamic
Extension to the
Azure Cloud

Antonio

Falabella(1) -

Stefano Dal Pra(1)

- Vincenzo

Ciaschini(1) -

Tommaso

Boccali(2) - Daniele

Spiga(3)

INFN - CNAF

Farm Extensions

Farm Extension to
MS Azure

Dynfarm

Summary

DODAS - Dynamic on-Demand Analysis Service

- Developed in the context of INDIGO-DataCloud
- Automatic generation of container based clusters to deploy an HTCondor cluster

Architecture

- TOSCA templates for the data center description
- Ansible used for automatic configuration
- Cluster management : Mesos and Marathon
- INDIGO Identity Access Management (IAM) for AuthN/AuthZ

INFN Dynamic
Extension to the
Azure Cloud

Antonio
Falabella(1) -
Stefano Dal Pra(1)
- Vincenzo
Ciaschini(1) -
Tommaso
Boccali(2) - Daniele
Spiga(3)

INFN - CNAF

Farm Extensions

Farm Extension to
MS Azure

Dynfarm

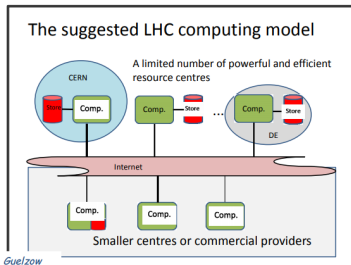
Summary

DODAS - Use cases

- Opportunistic resources integration: Targeting both private and public cloud providers
- **Extension of already existing facility** : Both to absorb peaks of usage and to generate mission specific facilities
- User-friendly computing infrastructure generation and management : of computing cluster to analyse research data

Look to the future:

- **DODAS design compliant with future paradigms**



See Daniele Spiga's Tuesday morning talk slides for details

INFN Dynamic
Extension to the
Azure Cloud

Antonio
Falabella(1) -
Stefano Dal Pra(1)
- Vincenzo
Ciaschini(1) -
Tommaso
Boccali(2) - Daniele
Spiga(3)

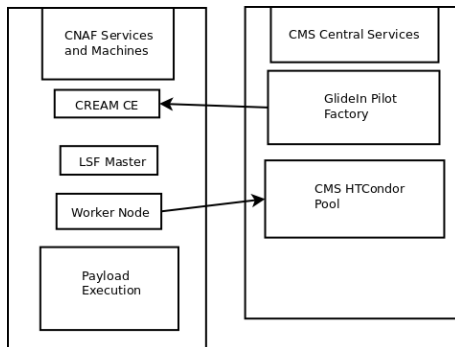
INFN - CNAF

Farm Extensions

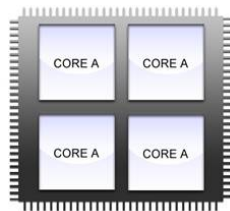
Farm Extension to
MS Azure

Dynfarm

Summary



- GlideInWMS submits jobs to CREAM CE
- CMS jobs are all multi-core



- The actual processing task is dispatched by the HTCondor Pool

INFN Dynamic
Extension to the
Azure Cloud

Antonio
Falabella(1) -
Stefano Dal Pra(1)
- Vincenzo
Ciaschini(1) -
Tommaso
Boccali(2) - Daniele
Spiga(3)

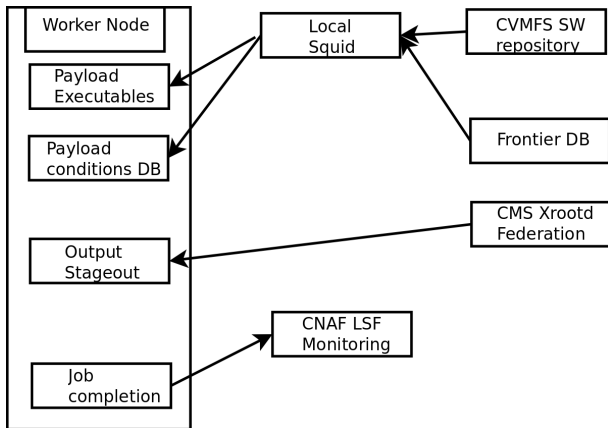
INFN - CNAF

Farm Extensions

Farm Extension to
MS Azure

Dynfarm

Summary



- Conditions and software via CVMFS through a Squid Proxy
- stage output files and logs to a SRM enabled storage supporting CMS

INFN Dynamic
Extension to the
Azure Cloud

Antonio
Falabella(1) -
Stefano Dal Pra(1)
- Vincenzo
Ciaschini(1) -
Tommaso
Boccali(2) - Daniele
Spiga(3)

INFN - CNAF

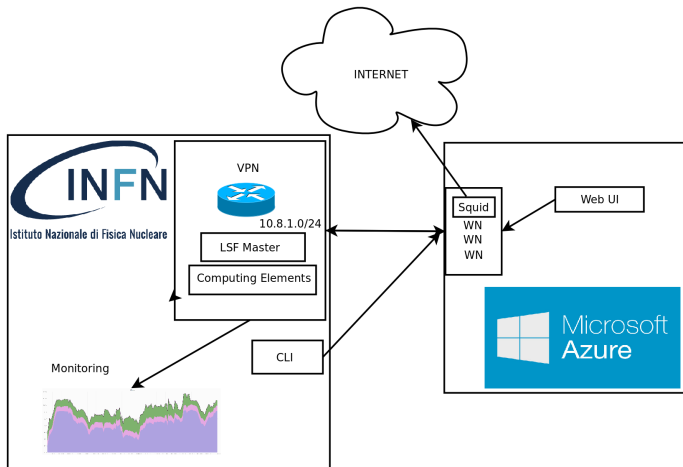
Farm Extensions

Farm Extension to
MS Azure

Dynfarm

Summary

- Extension to Azure has been obtained using in-house developed application : *Dynfarm*
- Python software that allows to create a VPN (OpenVPN) for a group of selected machines



INFN Dynamic
Extension to the
Azure Cloud

Antonio
Falabella(1) -
Stefano Dal Pra(1)
- Vincenzo
Ciaschini(1) -
Tommaso
Boccali(2) - Daniele
Spiga(3)

INFN - CNAF

Farm Extensions

Farm Extension to
MS Azure

Dynfarm

Summary

- The architecture consists of two servers:
dynfarm-server and OpenVPN server
- The machine contacts the *dynfarm-server*
to receive the VPN configuration
- Only the traffic to selected nodes are
routed to the VPN



```
/etc/remotebatch/azure.cnf
{"address": "52.178.162.194",
 "class": "azure",
 "mask": "24",
 "azure": {"hosts": ["10.8.1.1 ce03-lcg",
 "131.154.192.106 vpnserver",
 "131.154.194.24 lsf-1",
 "131.154.193.196 ui-tier1",
 "131.154.192.6 ce01-lcg",
 "131.154.193.10 ce04-lcg",
 "131.154.195.213 ce05-lcg",
 "131.154.193.23 ce06-lcg",
 "131.154.193.9 ce08-lcg",
 "131.154.193.3 ce07-lcg",
 "131.154.193.14 argus"],
 "services": ["lsf"],
 "version": "2.2.1"},
 "vpnconfig": "/etc/remotebatch/azure-config"}
```

INFN Dynamic
Extension to the
Azure Cloud

Antonio
Falabella(1) -
Stefano Dal Pra(1)
- Vincenzo
Ciaschini(1) -
Tommaso
Boccali(2) - Daniele
Spiga(3)

INFN - CNAF

Farm Extensions

Farm Extension to
MS Azure

Dynfarm

Summary

- CEs, Argus and the batch master must include the proper route

131.154.192.106

```
route add -net 10.8.1.0 netmask 255.255.255.0 gw 131.154.192.106
```

- The Azure worker nodes are configured by dynfarm to include the corresponding routes:

```
/etc/openvpn/server_azure.conf
push "route 131.154.193.9 255.255.255.255"
push "route 131.154.193.196 255.255.255.255"
push "route 131.154.195.213 255.255.255.255"
push "route 131.154.192.6 255.255.255.255"
push "route 131.154.193.10 255.255.255.255"
push "route 131.154.192.95 255.255.255.255"
push "route 131.154.193.3 255.255.255.255"
push "route 131.154.193.14 255.255.255.255"
push "route 131.154.192.242 255.255.255.255"
push "route 131.154.194.24 255.255.255.255"
push "route 131.154.193.23 255.255.255.255"
```

INFN Dynamic
Extension to the
Azure Cloud

Antonio
Falabella(1) -
Stefano Dal Pra(1)
- Vincenzo
Ciaschini(1) -
Tommaso
Boccali(2) - Daniele
Spiga(3)

INFN - CNAF

Farm Extensions

Farm Extension to
MS Azure

Dynfarm

Summary

- The provisioning has been performed with azure CLI (node.js scripts)
- Worker node image prepared in advance and uploaded
- Template json (A8 v2: 8 cores 16.00 GiB RAM 80 GiB Disk):



```
"properties": {
  "hardwareProfile": {
    "vmSize": "Standard_A8_v2"
  },
  "storageProfile": {
    "osDisk": {
      "name": "osdisk",
      "OSType": "Linux",
      "createOption": "FromImage",
      "image": {
        "uri":
          "[concat('http', '://', 'cnatf1standard',
            '.blob.core.windows.net', '\\
            /system/Microsoft.Compute/Images/disk31may2017/
            disk31may2017-osDisk.vhd')]"
      }
    }
  }
}
```

INFN Dynamic
Extension to the
Azure Cloud

Antonio

Falabella(1) -

Stefano Dal Pra(1)

- Vincenzo

Ciaschini(1) -

Tommaso

Boccali(2) - Daniele

Spiga(3)

INFN - CNAF

Farm Extensions

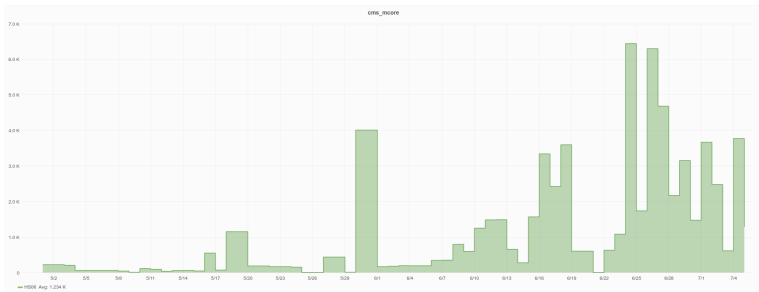
Farm Extension to
MS Azure

Dynfarm

Summary

- Created a total of 34 worker nodes
- Total number of Jobs submitted:
6331

- Consumed CPT: 62960h
- Consumed WCT: 150778h



INFN Dynamic
Extension to the
Azure Cloud

Antonio
Falabella(1) -
Stefano Dal Pra(1)
- Vincenzo
Ciaschini(1) -
Tommaso
Boccali(2) - Daniele
Spiga(3)

INFN - CNAF

Farm Extensions

Farm Extension to
MS Azure

Dynfarm

Summary

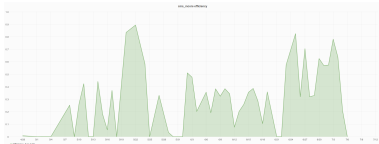
CMS on local nodes

- Average Efficiency: 0.8



CMS on Azure nodes

- Average Efficiency: ~ 0.4
- Peak efficiency: ~ 0.9



Why?

- Mainly due to the provider network performances
- Dynfarm VPN not a bottleneck

INFN Dynamic
Extension to the
Azure Cloud

Antonio

Falabella(1) -

Stefano Dal Pra(1)

- Vincenzo

Ciaschini(1) -

Tommaso

Boccali(2) - Daniele

Spiga(3)

INFN - CNAF

Farm Extensions

Farm Extension to
MS Azure

Dynfarm

Summary

- CNAF Tier1 is the main data center of INFN serving WLCG experiments as long as around 30 other collaborations
- The increasing demand of computing resources led to the investigation of several techniques to dynamically extend the existing farm
- An approach based on the dynfarm software to extend the farm to the MS Azure Cloud
- We created 34 nodes and submitted > 6000 CMS multi-core jobs
- The test was successful although the average efficiency of these jobs was ~ 0.4 mainly due to QoS of network resources of the providers