Group Computing at CPPM

A.Tsaregorodtsev

▸ The goal of the discussion

  ▸ Evaluate computing needs of the CPPM teams/services and ways for optimisation of group resources usage

  ▸ Propose a cloud based solution for group computing resources

  ▸ Get feedback from the groups on the proposed solution

Computing resources for the CPPM teams

Local resources

- Individual workstations, laptops
  - Final analysis, working with documents
- Team servers
  - Analysis, visualization
  - Data storage
  - *To be discussed today*
- Cloud CPPM
  - *To be discussed today*
- DEC cluster
  - *Not to be discussed today*

External resources

- AMU/Mesocentre
- CC/IN2P3
- Computing grids
  - EGI
    - HTC, Cloud and HPC resources
  - WLCG
- *Not to be discussed today*

- 24 active group servers
  - ~700 CPU cores in total
  - 12 Storage servers (total storage 600TBs)

- 7 groups have their own servers
  - There are plans to increase the capacity
    - By ~5-7 servers
    - By ~200TBs storage
  - 2 groups (Renoir, imXgam) have GPU resources
    - 10 GPU cards in total
  - Almost all the groups are planning to use or increase usage of the GPU resources

- 3 groups do not have dedicated servers but count to have some local computing resources in the (near) future
  - ~3 servers
  - ~30 TBs storage

- About 90 users are using or would use CPPM computing resources (group servers)
  - Plus a variable number of students, 10-20 persons

- Around 40 users access the group servers simultaneously
 - Important number to plan the capacity for the interactive work

▶ **Mostly interactive work**
  ▶ Various interactive analysis e.g.
    ▶ Short application runs
    ▶ ROOT sessions
    ▶ Jupyter notebooks
    ▶ KATIA session (CAO)

▶ **Batch type computations**
  ▶ e.g. MC simulations
  ▶ Done very occasionally
  ▶ No batch system is used
    ▶ Although would be a desirable functionality

- Large part of the time the servers are idle
  - This is normal for the interactive work
    - No use in off hours, on weekends, etc
    - But also a limited use in the working hours

Current display period year   **TOP**

Other possibilities : hour - day - week - month - year -

	Load average	Idle CPU
HR_marlhcb4 2020-10-12		
HR_marlhcb5 2020-10-12		
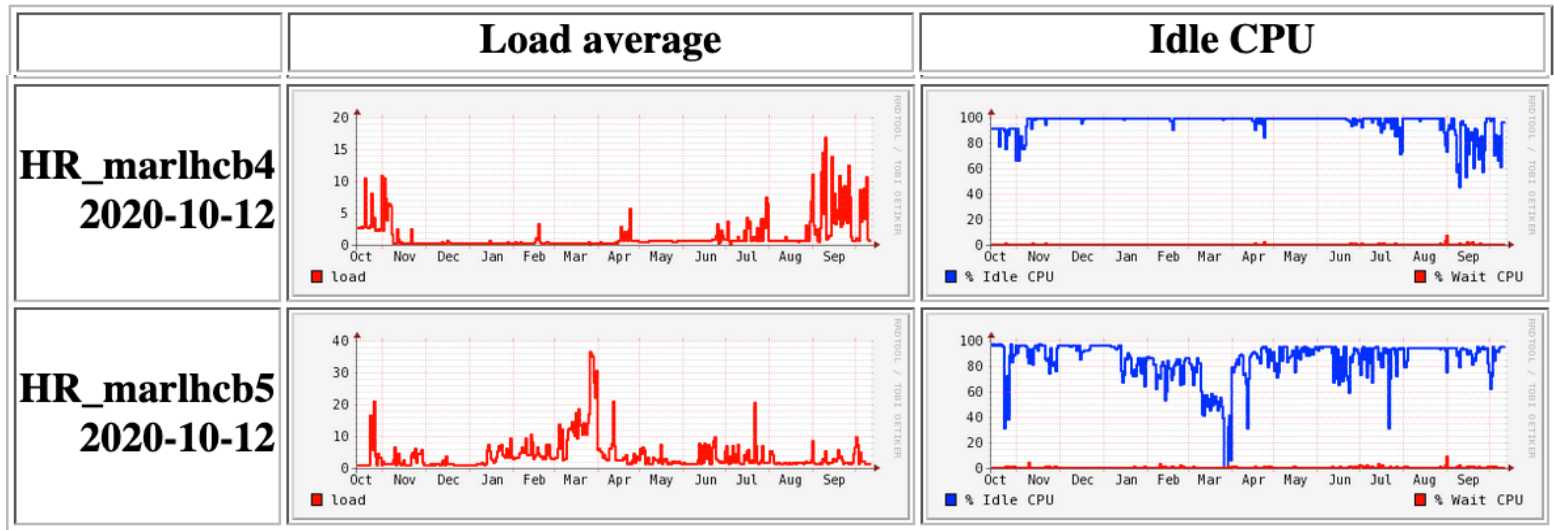
▸ Sharing resources helps optimisation
  ▸ Resources of idle « group servers » could be reused by other active groups

▸ This requires a flexible dynamic mechanism of allocation of computing resources with a suitable configuration and capacity
  ▸ The cloud technology offers such a mechanism

▸ Cloud technology is a trend and not by chance. It will help to address many more issues in the future and not just the group server problem

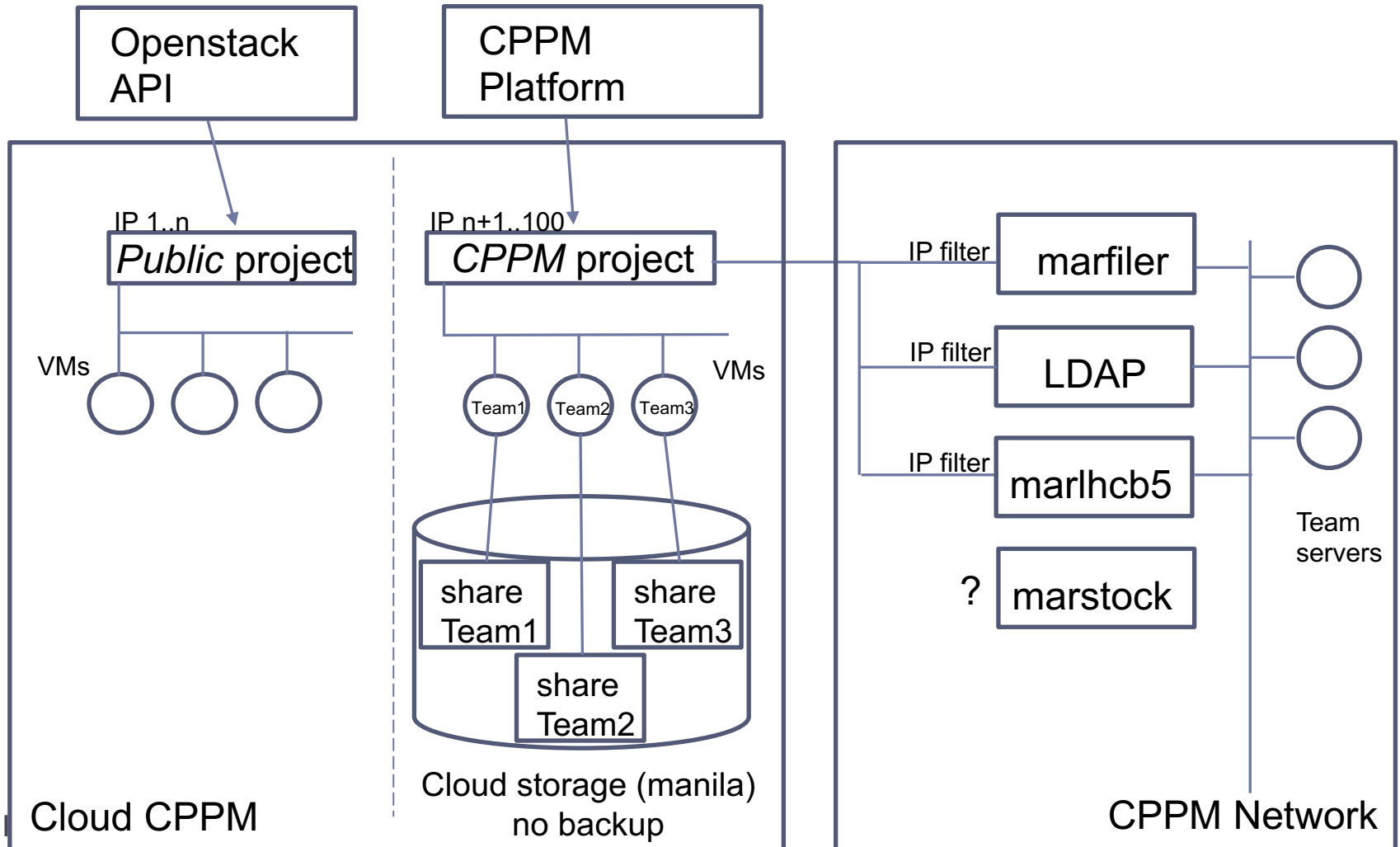# "N'est ce pas un peu tôt de faire cette réunion ?" (M.Barbero)

- The CPPM Cloud prototype is now in a quite mature state
  - Let's consider how this facility can help increasing the computing capacity offered to the CPPM users

- Before proceeding with further development, we would like to synchronize our views with the groups
  - The proposed solution is open to your critics and suggestions

CPPM Computing Platform

CPPM computing platform

Openstack API

CPPM Platform

IP 1..n

Public project

IP n+1..100

CPPM project

VMs

VMs

Team1 Team2 Team3

IP filter — marfiler

IP filter — LDAP

IP filter — marlhcb5

? marstock

Team servers

share Team1

share Team2

share Team3

Cloud CPPM

Cloud storage (manila) no backup

CPPM Network

- The CPPM cloud is split in two projects
  - *Public*: accessible for all the users registered in the Openstack service
    - CPPM and possibly external users
    - The project open to the France-Cloud infrastructure
  - *CPPM*: accessible for all the CPPM users
    - Passing by the CPPM Platform interface

- The projects have non-overlapping VM IP address ranges
  - This allows to control access to the CPPM internal network services

▸ Users of VMs created in the Public project have several advantages and limitations

  ▸ « Free » choice of image and flavour

    ▸ Out of those that are enabled in the Openstack service

  ▸ *root* privileges on the VM

  ▸ No access to the CPPM network

    ▸ No print service

    ▸ No NFS mount of file servers (marfiler, etc)

  ▸ No backed up home directory

  ▸ Possibility of remote FS mount by SSHFS

    ▸ Limited I/O performance

- VMs are created by the CPPM Platform service
    - VMs belong to the service
    - Users do not have *root* access
        - More secure usage

- Limited access to the CPPM Network
    - Marfiler
    - No print service
        - May be eventually but is it important ?

- VMs usage is similar to the use of group servers
  - VMs are created with the image defined in the group profile
    - e.g. CentOS7 official image
    - Users can choose the VM flavour/size:
      - CPU cores, RAM, GPU
  - VMs are contextualized according to a team profile
    - Setting up the team software environment
  - Users log in to the VMs using their CPPM account
    - Using the CPPM LDAP service
      - Or equivalent contextualization
  - User home directories are kept in marfiler
    - Same as in the CPPM network
    - Backup

Team VMs contextualization

- Common VMs contextualization part
  - Common software (to be defined)
    - CVMFS, Singularity, Conda, etc
  - VM monitoring service
    - Monitoring activities on VM
    - Possibly swicthing off the VM if no activity

- Team contextualization
  - Defined in the team profile
  - Mount of team storage shares
    - And possibly file servers from the CPPM internal network
  - Installing team specific software

- VMs allocated in a planified way for long time periods
    - Full analog of the group servers
    - Stable IPs, easy to remember DNS/host names
    - Stopped as necessary by the group administrators to free resources

- VMs allocated dynamically on the user request
    - Allocated if spare capacity is available
        - Opportunistic resource
    - No stable IP or host name
    - Stopped automatically if no activity detected
        - « No activity » is to be defined
            - e.g. nobody is logged in, zero CPU consumption, etc

Group storage shares

- Shares are permanent storage allocations per group
    - From cloud storage resources
- Shares are NFS mounted in the VMs
    - Fast data access from VMs
- Shares are not backed up
    - Equivalent of the group file servers
- Shares can be populated from the group file servers or from external storages
    - I/O rates for data copying are to be tested
- Currently there are 2 shares allocated
    - LHCb (30TBs)
    - DarkSide (20TBs)

CPPM Plarform service

- VMs instantiation in the CPPM project is performed through the Platform service
- For static VMs deployment it can be just a manual operation by the system administrators
  - e.g. LHCb group servers deployed in the CPPM cloud
- For dynamic on-demand VMs deployment the Platform service should provide
  - Dashboard for users
    - User authentication/login: CNRS certificate or local LDAP
    - Choice of the group profile, flavour
    - VM launching and getting IP address
    - Web and command line interface
  - Dashboard for administrators
    - Monitor activities
    - Accounting reports
    - VM manipulation tools
- The advanced features of the Platform service are to be designed and developed
  - Based on VMDIRAC project

- Tests the platform with group servers
    - Deploy at least one server per group
        - With the help of the group computing experts
    - Define the group server profiles
        - Software environment
        - Mounted shares and file servers
    - Measure and optimize the performance of the group server
        - e.g. data access rates
        - optimal flavours
        - special attention to the interactive server performance
    - Improve the performance of the cloud infrastructure
        - E.g. upgrade the network performance
    - Iron out all the features

- Automate creation of group servers with a tested stable profile
- Gradually move to a more intensive use of dynamically allocated servers
- The current group servers stay as they are at the group disposal to ensure continuity

Questions to be answered

- Sharing resources requires clear usage policies
  - Group quotas, priorities, planning
  - Mechanisms to ensure the defined policies

- Production Service operation procedures
  - Trained service administrators available « 24/7 »
    - At least 1 FTE equivalent manpower (to be defined)
  - Reactive user support
    - Sharing usage experience among groups should help

- Providing resources for the CPPM cloud
  - Purchases by the laboratory ?
  - Group contributions ?

Conclusions

▸ The use of local computing resources by CPPM groups can be considerably optimized

▸ The proposed cloud based solution allows to combine static resources allocation similar to the group servers today and dynamic allocation for more flexible usage

▸ If agreed with the groups, resources will be allocated in the cloud to start testing and to define group profiles