VS user meeting 2025: Recommendations for our users

Alexandr Mikula

FZU Fyzikální ústav Akademie věd České republiky

Outline:

- UI server resources
- NFS Servers (includes /home/)
- HTCondor batch resource constraints
  - Memory management
  - Disk management
  - CPU and runtime management
- Miscellanea

UI Servers resources

- 3 login servers for access to the farm
  - Ui1 virtual machine (12 Cores, 30GB RAM, 300GB HDD, 2×25Gbit LAN shared with other Vms)
  - Ui2 & Ui3 physical machines (64 Cores, 1TB RAM, 960GB HDD, 25Gbit LAN)
- Maximum resource limits per user:
  - RAM/10 up to 64GB
  - floor(Cores/10), minimum 2 Cores

UI fair usage

- There are no general restrictions on how to use the UIs
- Few points of interest:
  - UIs are not for extensive computing tasks
  - Big (TBs +) data transfers* should be coordinated with us, to achieve better efficiency
  - Please tend to your vs-code servers and similar stuff which may run astray
    - We will end all processes of the offending user indiscriminately

* This request does not include the grid resources

NFS servers

- 8 physical nfs servers with varying capacity and age
  - 2 for general users
  - 6 for the astroparticle experiments
  - Quotas to measure the occupancy per user, no enforcement
  - **!!! NO BACKUP OF THE DATA !!!**
    - We will try to do our best to salvage in case of the accident but without guarantee
    - Use the DU service from CESNET for your extra important data (https://du.cesnet.cz/en/start)

NFS01 (/home)

- 1 virtual machine for /home directory (similar constraints as ui1)
  - Only this enforces the quotas (default: 10GB, 1M files)
- Data backed up daily, except for the 2nd of month
- The demand for responsiveness of the nfs server grew, but there are also more clients using the same resource without upgrade for last 4+ years
- NFS inherently slow for lot of small files

HTCondor: Memory

- Default memory limit (RequestMemory) is 2GB × Requested CPUs (RequestCPUs)
- Memory is managed with CGroups with an upper limit of 1.1 × RequestMemory
  - Your job will be automatically terminated upon reaching this limit, with the reason of termination stated
- No swap is allowed for the jobs

HTCondor: Disk management

- Job disk space managed by LVM on growing part of our infrastructure
- Default is (20GB of disk (RequestDisk) × RequestCPUs) + 2GB fail safe
- Maximum is 1TB
- Caveat: some of older machines report 50% more of disk than available (possible new flag for the machine expression)
- Job disk space fully isolated, jobs do not see other job disk spaces

HTCondor: CPU and runtime

- CPUs are limited to your RequestCPUs (default is 1)
  - Your job can get more if there is a free CPU time on machine (you can try to leverage this spare time)
  - Maximal CPU count is 128
- Default max runtime is 14 days
  - Please try to set sensible run time lower than this

HTCondor: Wrap up

- Set as low resource requests as possible, you job will start faster
  - Resource metrics to look for:
    - DiskUsage
    - CpusUsage
    - ResidentSetSize_RAW
    - HoldReason (not a metric, but useful info, why was your job stopped)
- Documentation:
  - https://www.farm.particle.cz/go/doc

HTCondor: submitting lots of jobs

- Limit jobs requesting single nfs mount!
  - concurrency_limits https://www.farm.particle.cz/go/concurrency
- Use materialize:
  - https://htcondor.readthedocs.io/en/latest/users-manual/submitting-a-job.html#submitting-lots-of-jobs

Miscellanea

- SELinux slowly applied on machines
- LVM on all jobs
- Second GPU server
  - Currently we can’t provide support for GPU workloads due to the HR constraints
- MOSH on UIs available
- We can’t provide support for the user workloads
- Future plans and necessities:
  - Two factor login, not yet decided how extensive
    - Probably only ssh keys on UIs (no passwords)
  - User /cvmfs repositories, is there a demand for it?

Thank you for your attention

FZU Fyzikální ústav Akademie věd České republiky