

The EP's Disk-iplinary Resource Management

Managing Storage at the EP

By: Cole Bollig

Software Developer for CHTC

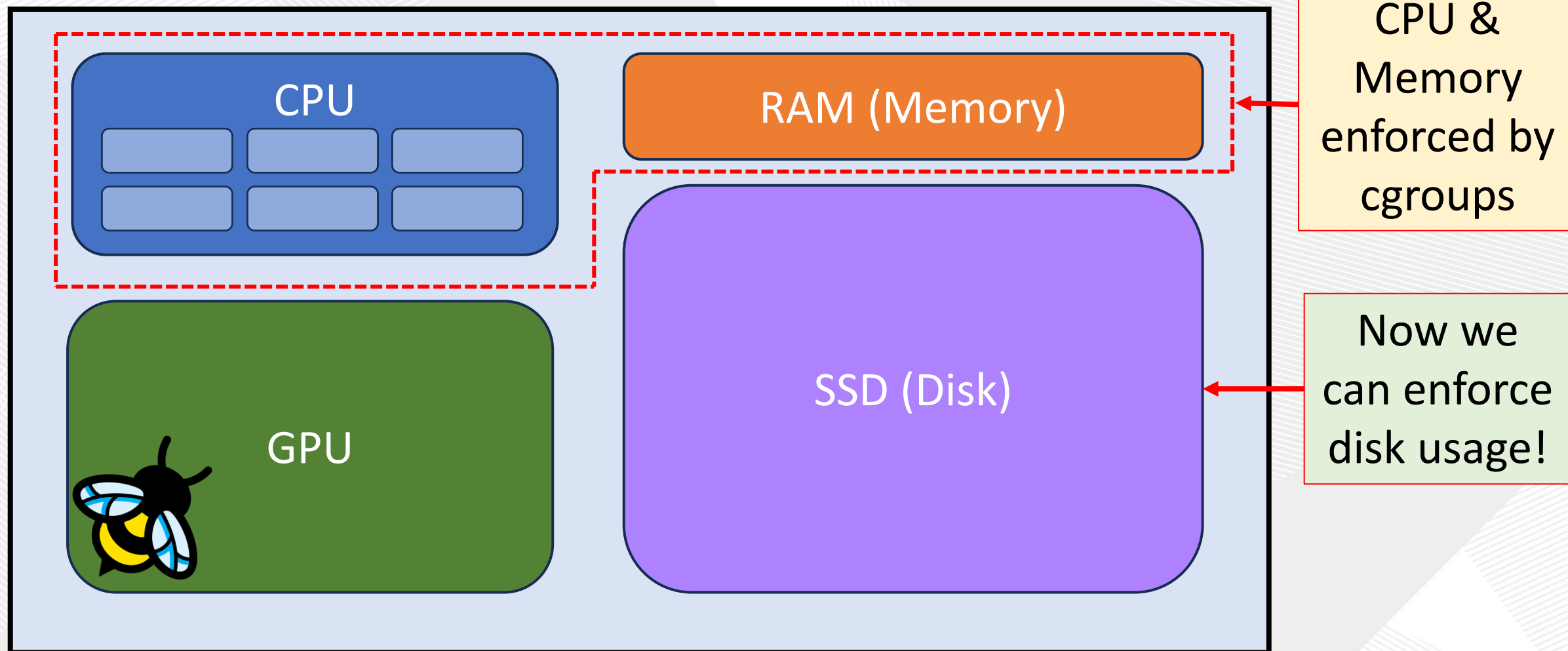
European HTCondor Workshop 2024

CHTC

HTCondor
Software Suite

PATh

Execution Point (EP)

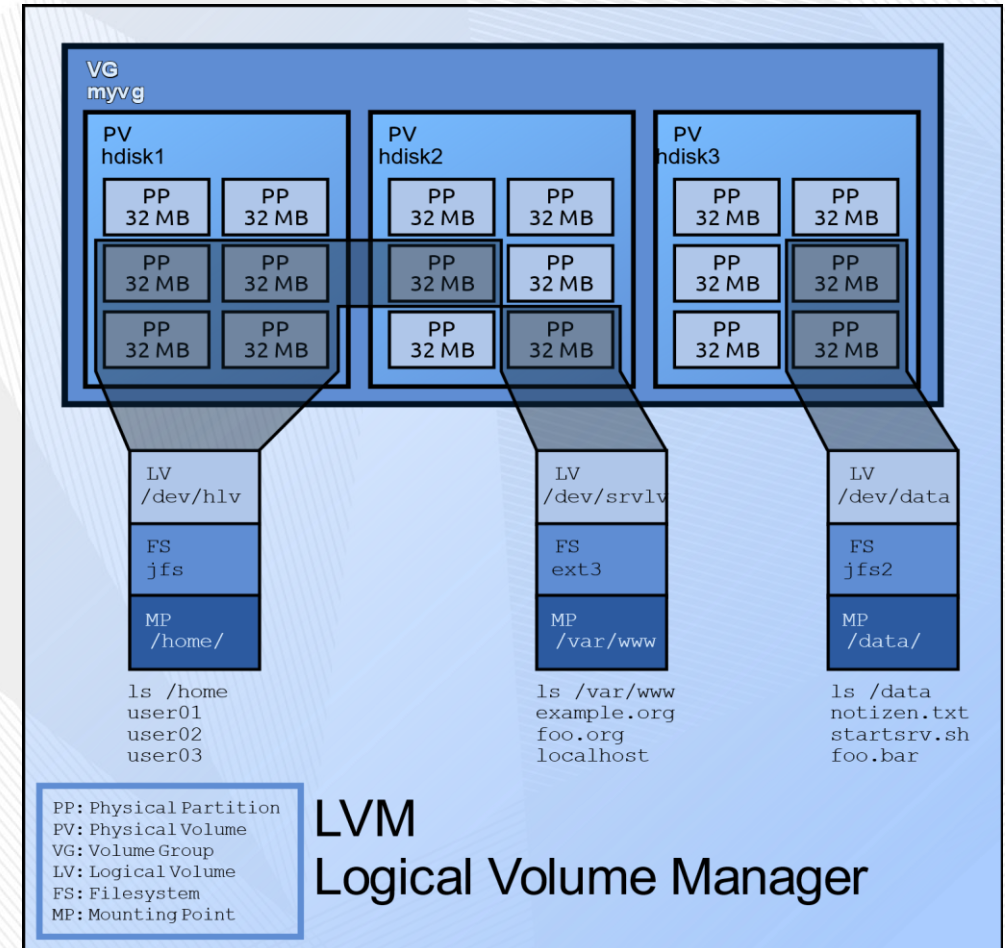


How does the EP enforce disk usage?

- Using the Logical Volume Manager (LVM).
- The EP will create a unique ephemeral Logical Volume (LV) for each job the EP runs.

Requirements

1. Linux OS
2. HTCondor running as root

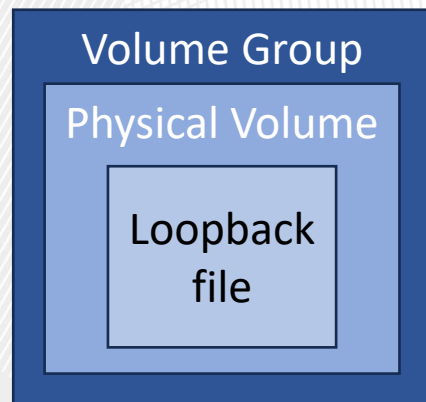


[Wikipedia - Various elements of the LVM](#)

How to Enable Disk Enforcement?

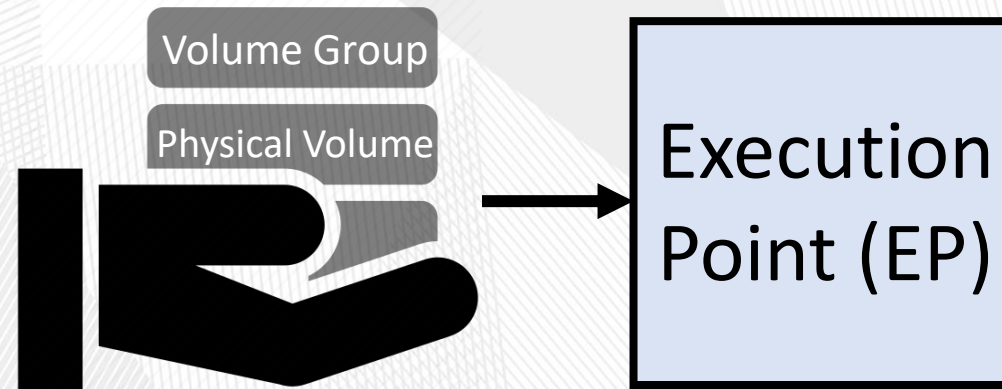
- Add the configuration **STARTD_ENFORCE_DISK_LIMITS =**
- **True** May also want to set the following configuration options:
 - **LVM_BACKING_FILE_SIZE_M B** (Defaults to 10GB)
 - **LVM_BACKING_FILE**

Note: The EP needs a restart to enable or disable LVM enforcement.



Setup LVM for the EP!

- If specified, the EP will use LVM components it is informed about rather than setting everything up itself
- Inform the EP of Volume Group name to use via **LVM_VOLUME_GROUP_NAME**



Benefits of Ephemeral LV's

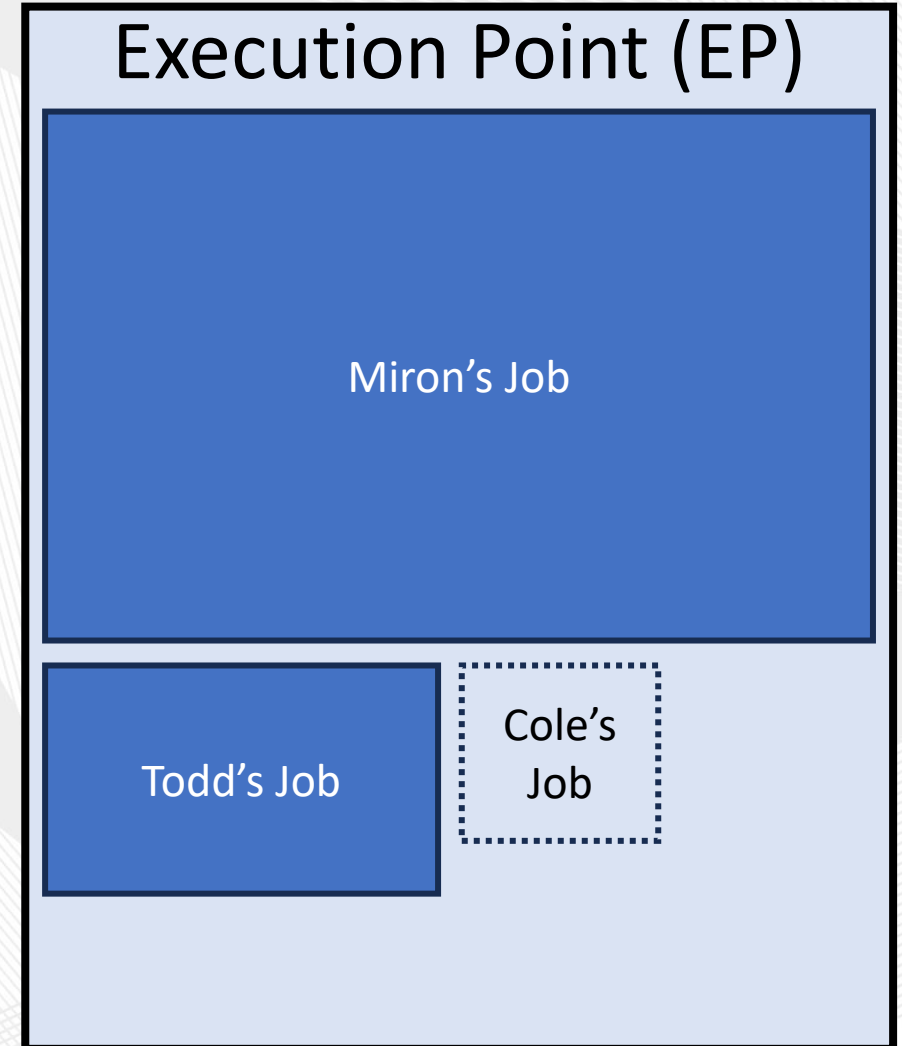
1. Improved Job Isolation
2. Improved Disk Management
3. Improved EP Efficiency
4. Data Encryption



Benefit: Improved Job Isolation

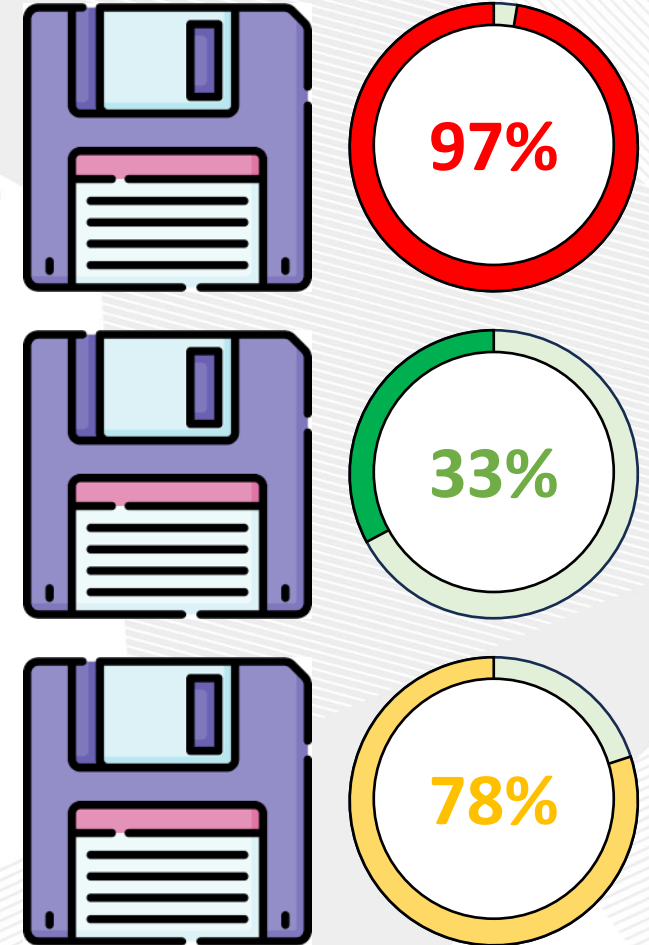
- Each job runs in its own filesystem.
 - The FS lives in the ephemeral LV created for the job.
- Make the FS only visible to the user job with **LVM_HIDE_MOUNT**
 - Note: This option does not play well with Docker Universe Jobs!

[Configuration Knob - LVM_HIDE_MOUNT](#)



Benefit: Improved Disk Management

- Jobs get what they ask for!
 - With the ephemeral LV & filesystem each job has a maximum amount of available disk to use (`request_disk`).
 - Meaning a job can not run away and use more disk.
- With a hard cap comes a greater risk of ENOSPC.
 - The EP will monitor the LV usage.
 - If the job is using more disk space than requested the job is put on hold with a message informing the user that the job needs to request more disk.



Benefit: Improved EP Efficiency

- No more pesky and slow traversal of the job's sandbox!
 - Currently an EP must traverse the entire job sandbox to report the job's disk usage (counting each file manually) and to cleanup after the job.
- With an LV we can do a simple query for disk usage.
- Once the job is gone, we can simply delete the LV in one step.



Benefit: Data Encryption

- The entire LV can be encrypted with **cryptsetup**
 - Administrator can enable LV encryption for all jobs via configuration:
ENCRYPT_EXECUTE_DIRECTORY
 - Users can request encryption with the a submit command: **encrypt_execute_directory**
- Used to have encryption with eCryptFS before it was deprecated



[Submit Command - encrypt_execute_directory](#)

[Configuration Knob - ENCRYPT_EXECUTE_DIRECTORY](#)

Thin Vs Thick Provisioning

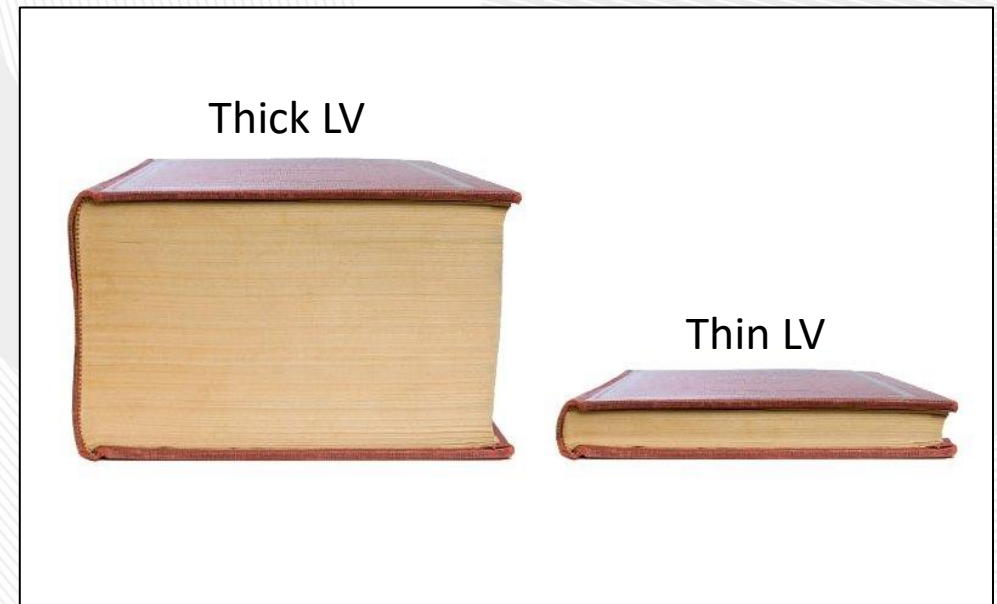
Thin Provisioning

- Disk is provisioned as needed
- Allows the EP to overprovision the LV
 - Better over usage reporting to user
- Requires backing thinpool LV

Thick Provisioning

- Disk is provisioned at LV creation time
- Just requires a Volume Group

LVM_USE_THIN_PROVISIONING



Questions?