Convergence between HPC and Big Data: Simulating the File System

Adrien Faure[1,2], Michael Mercier[1,2], Olivier Richard[1]

UNIVERSITÉ Grenoble Alpes [1] Inria INVENTEURS DU MONDE NUMÉRIQUE [1]

Bull atos technologies [2]

Janvier 2019

Motivation

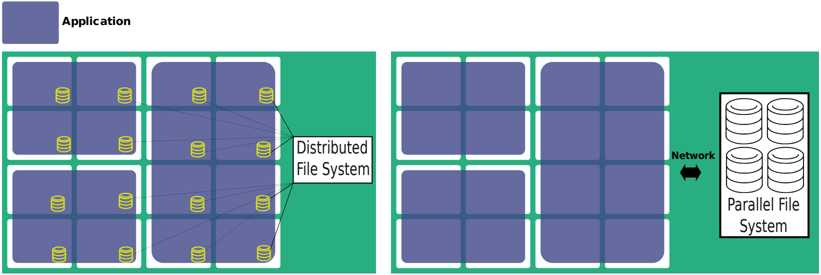
How well a workload composed of Big data applications can run on a HPC platform ?

- Big data applications respond to the need to analyse a large volume of data.
- In a HPC production environment applications share the file system, this can lead to IO contentions.
- **The file system becomes the bottleneck.**
- What file system should we use, a distributed file system or a parallel file system?
- We use our simulator, Batsim to answer the question.

HPC applications versus Big Data applications

Applications	**Big Data**	**HPC**
Infrastructure	Heterogeneous	Dedicated for HPC
Computations	Move computations where the data is located.	Move the data to the computations.
I/O pattern	Read large quantity of data (that does not fit in one node)	Periodic (due to checkpointing) Bursty (sometime short IO bursts occur)
File System	Use a distributed filesystem	Uses a parallel file system (PFS) (shared between all applications)

Parallel File System(PFS) vs Distributed File System (DFS)



1) HPC Cluster with DFS

2) HPC Cluster with PFS

1) Distributed file system uses the local disks of the compute nodes.
2) Parallel file system is shared between all applications.

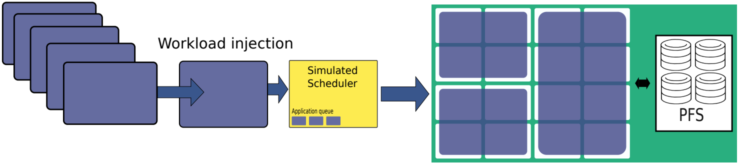
Batsim: Infrastructure simulator

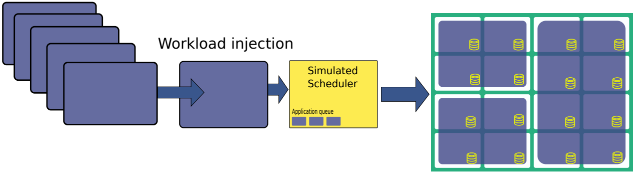
Infrastructure simulator for scheduling

- It simulates the scheduler of the platform.
- It simulates a workload of applications and their IO consumption.
- It simulates the platform, with a PFS or a DFS.

Experimentation

Two platform set up.





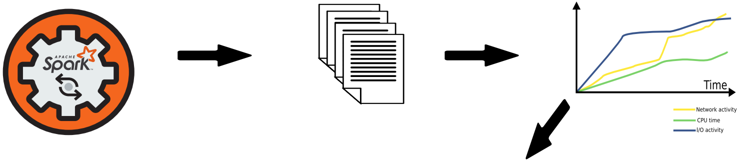
Application model from spark traces

We use the logs to extract models to incorporate in our simulator.

Spark benchmarks executions

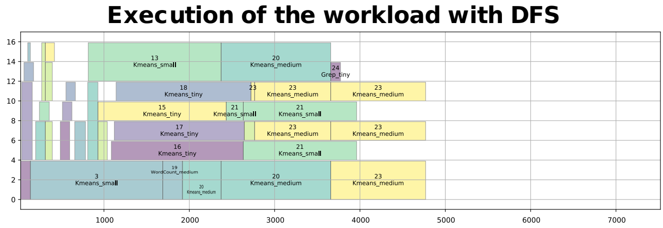
Collects logs of the applications

Extracts metrics from the logs

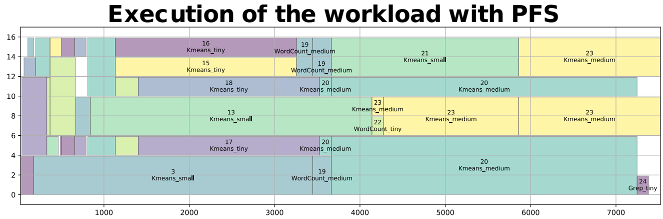


Generates application models compatible with Batsim

Gantt chart of the results

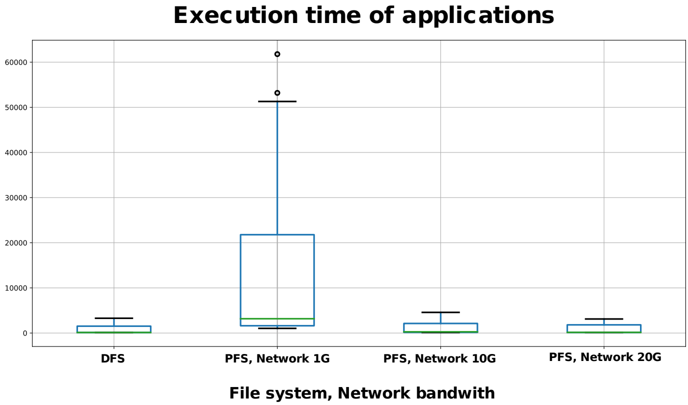


Execution of the workload with DFS

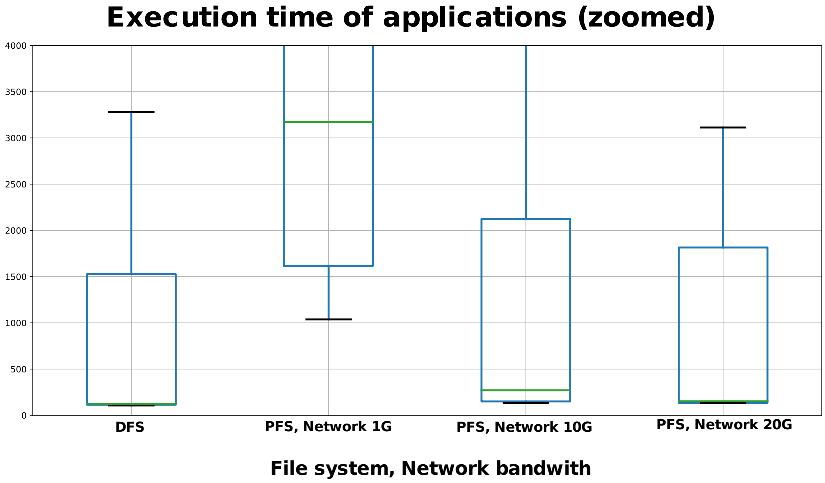


Execution of the workload with PFS

Execution times of applications



Execution times of applications



Conclusions and Future Work

In conclusion:

- We simulate interferences between applications with a PFS versus a DFS.
- The big data applications are slower with a PFS than with a DFS, which supports the results of Yildiz et al.

In future work:

- Create application models for HPC applications and workloads.
- Impact of big data workload on HPC platform with batsim.