

Improvement of a Parallel File System Simulator by Data Striping Implementation

Vivien Michel

Polytech'Grenoble, France

Francieli Zanon Boito

Universidade Federal do Rio Grande do Sul, Brazil

Mario Dantas

Universidade Federal de Santa Catarina, Brazil

Workshop de Processamento Paralelo e Distribuído - 2015

Improvement of a **Parallel File System Simulator** by Data Striping Implementation

- Some people don't have **supercomputers**...
- We could want to make deep modifications to the system
- Money and energy consumption

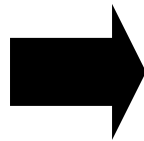
Parallel File System

- **Transparent access:** applications do not need to know where the file is
- Files = **data + metadata** (information about data)
 - Data servers and metadata server(s)
- Examples: Lustre, PVFS

The simulator

- **PFSSim** – PFS simulation over Omnet++
- Interface to describe and simulate
I/O scheduling algorithms
- Did not implement **data striping**

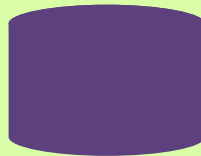
Parallel File Systems



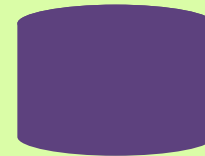
Data striping



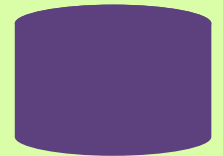
Data Server



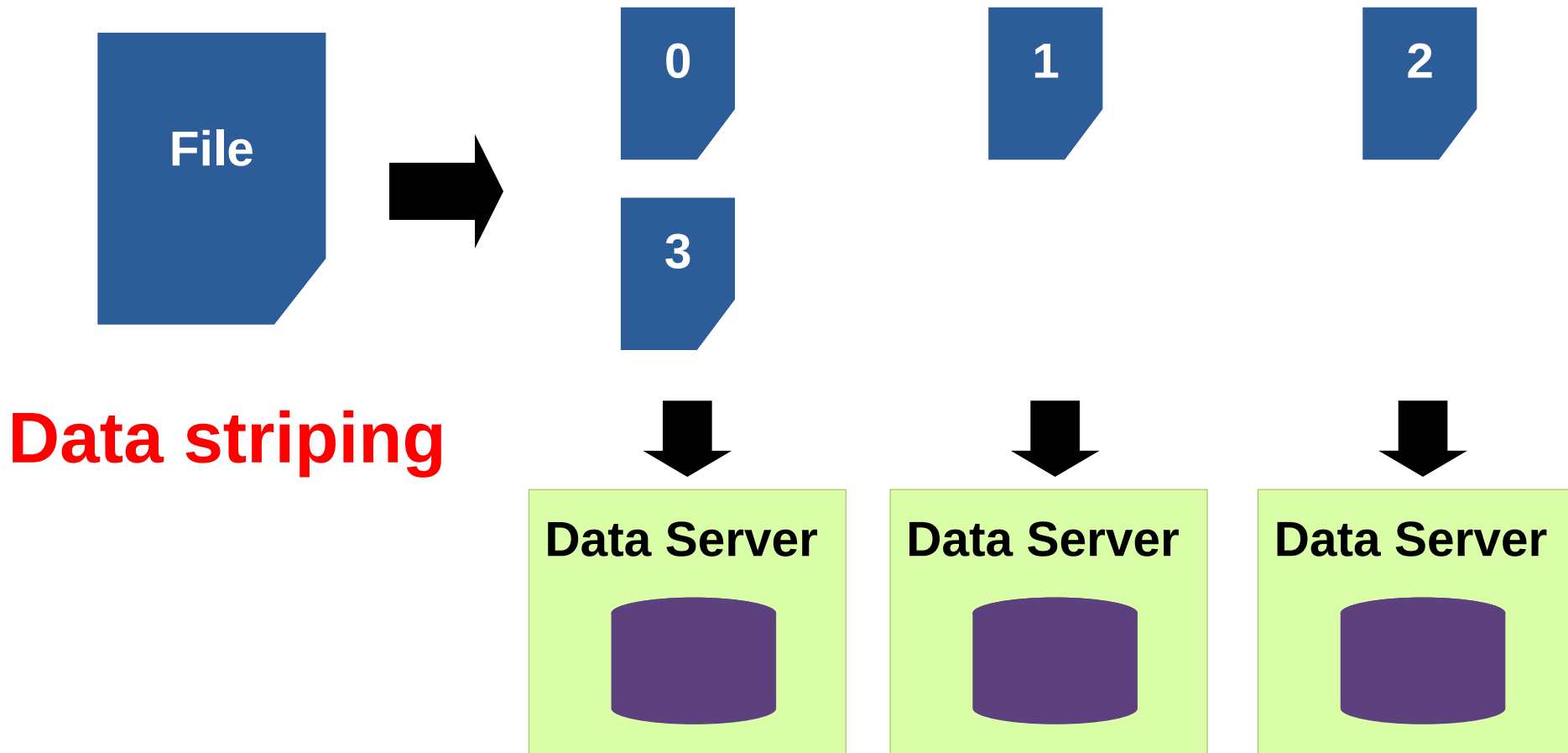
Data Server



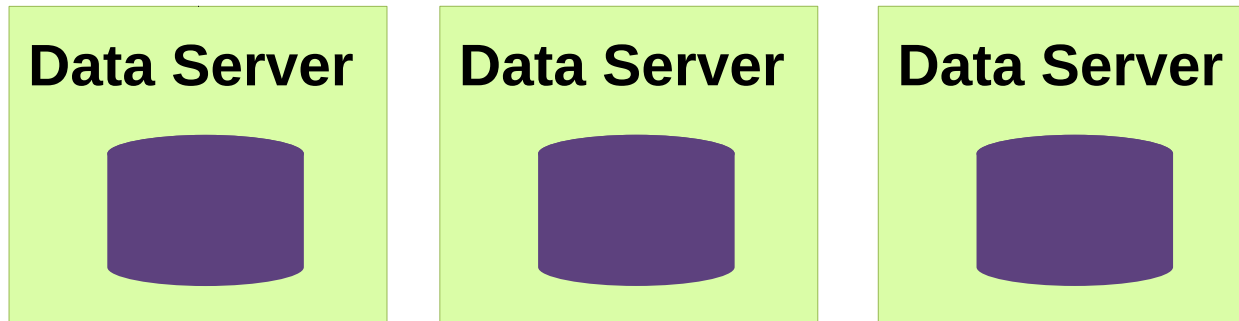
Data Server



Parallel File Systems



Parallel File Systems

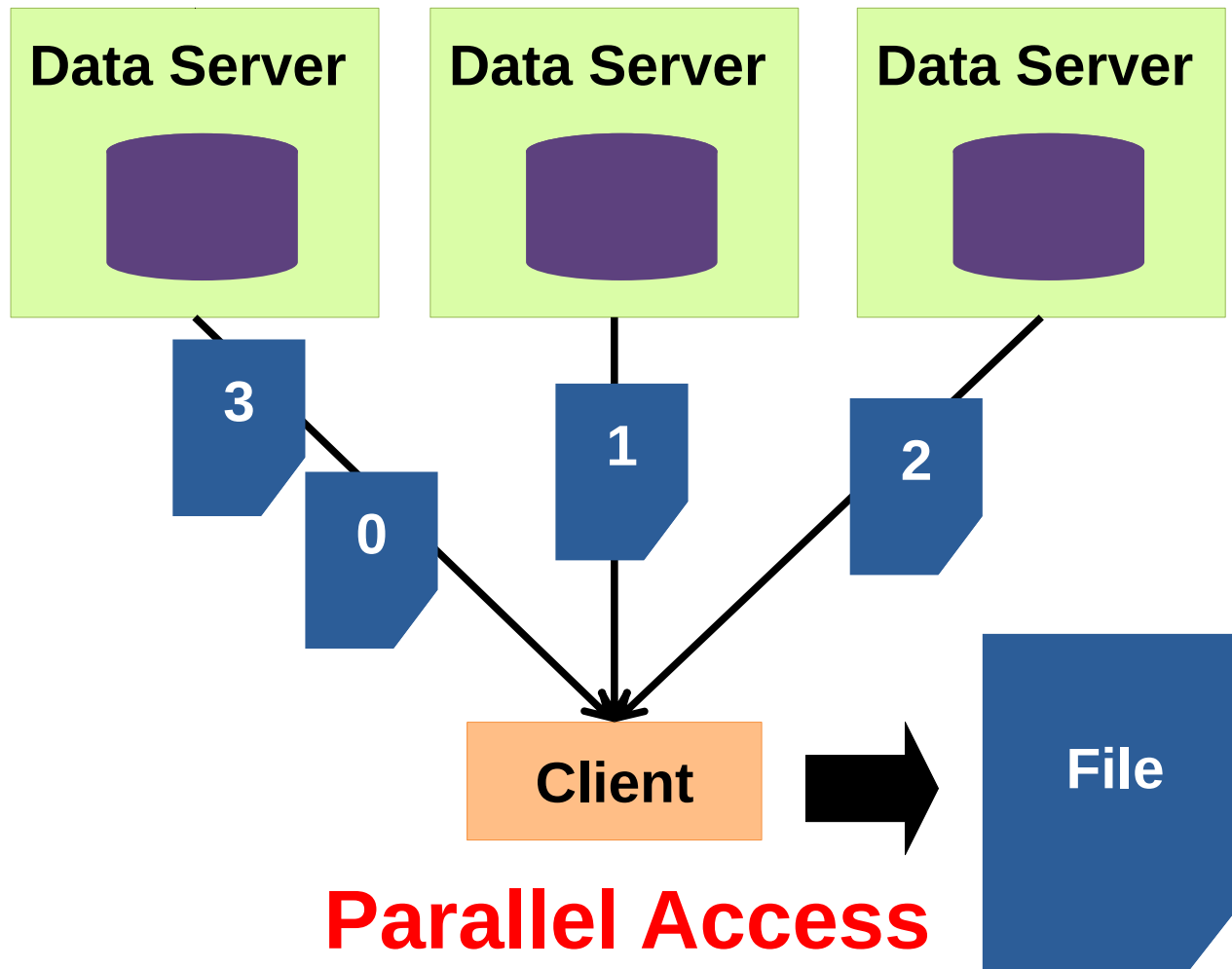


File?

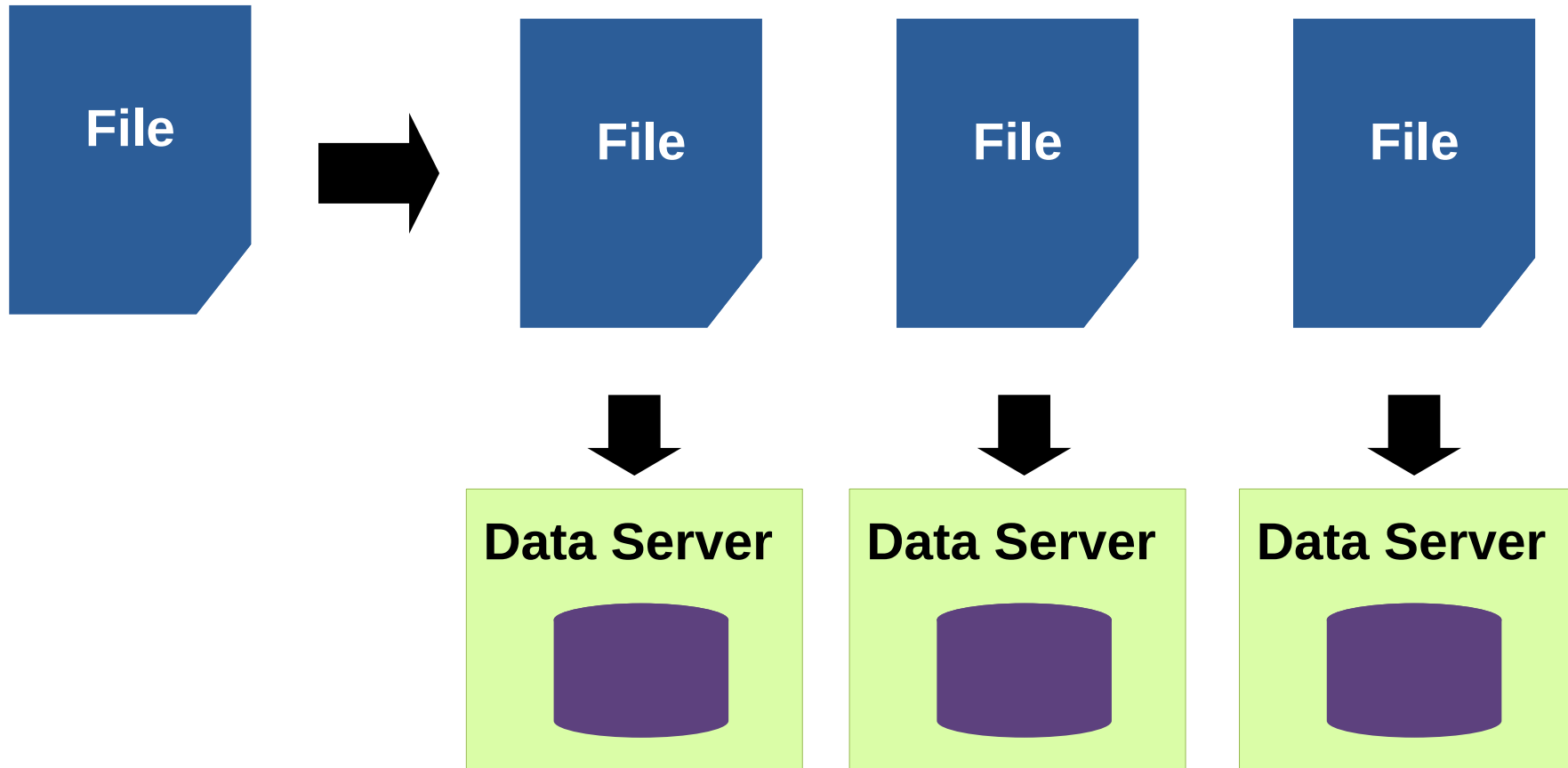
Client

Parallel Access

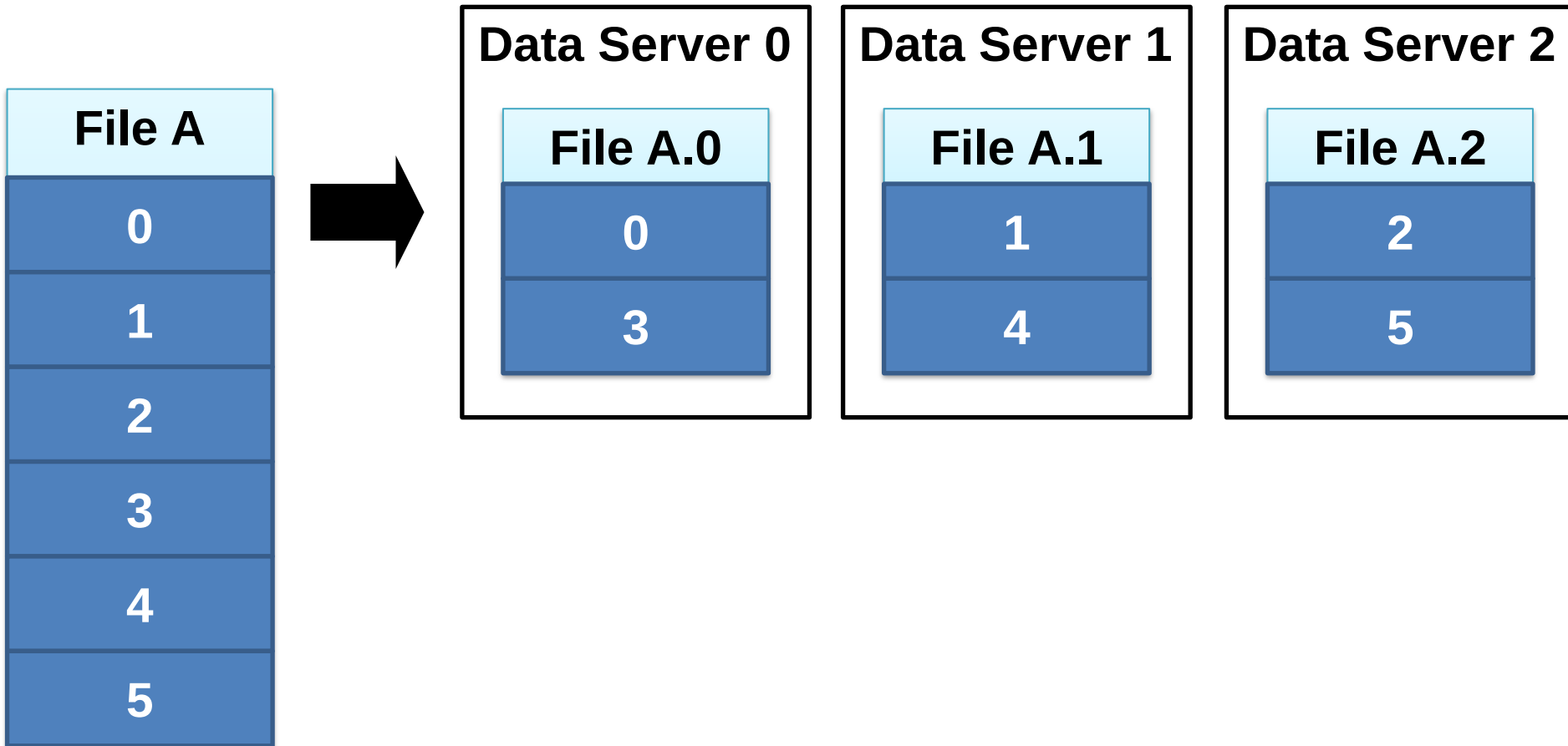
Parallel File Systems



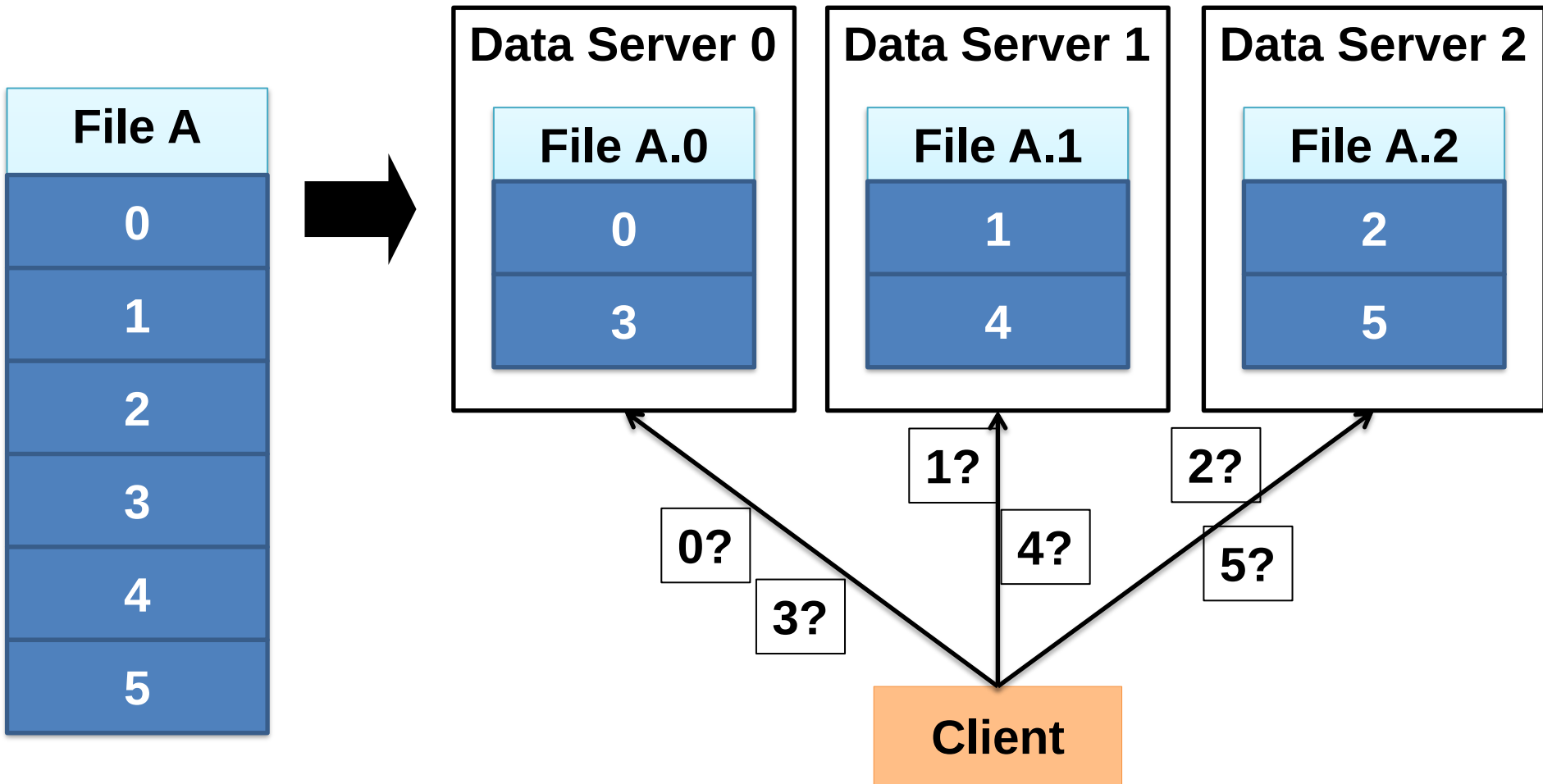
“Data striping” on **PFSSim**



Why is that a problem?



Why is that a problem?



Why is that a problem?

- Access **sequentiality** usually **matters**
- **I/O scheduling** algorithms reorder requests to improve the generated access pattern
 - Results will make **no sense!**

The project

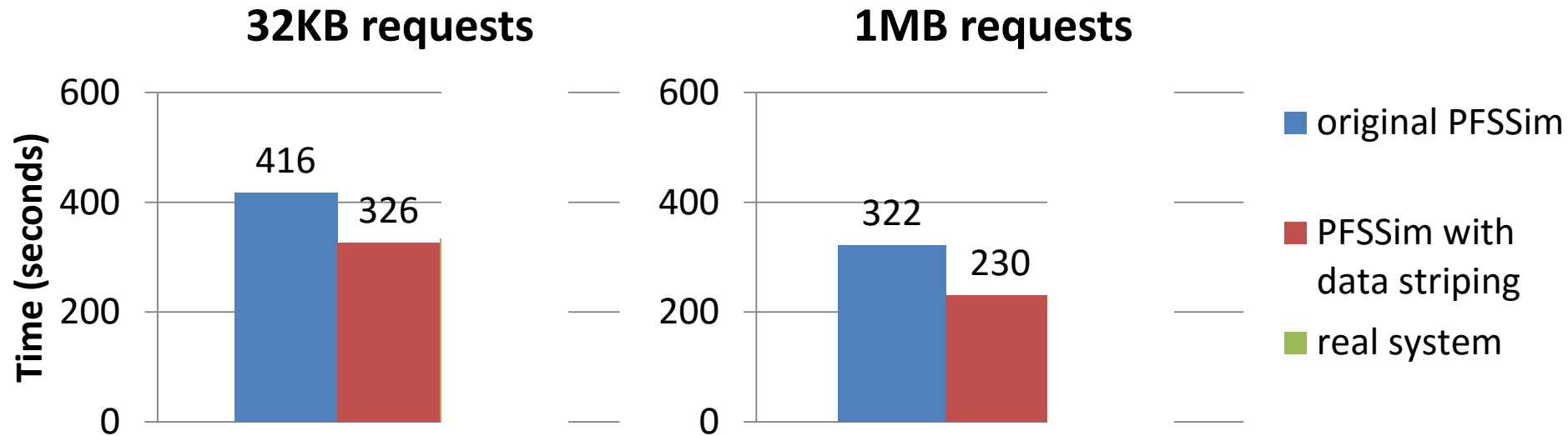
To implement **data
striping on PFSSim**

Did that work?

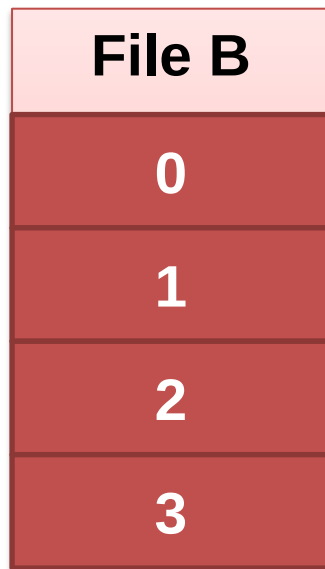
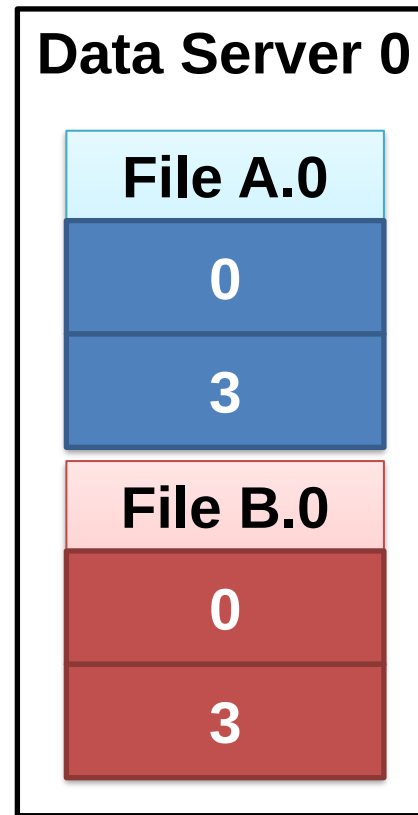
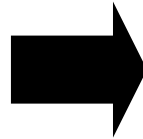
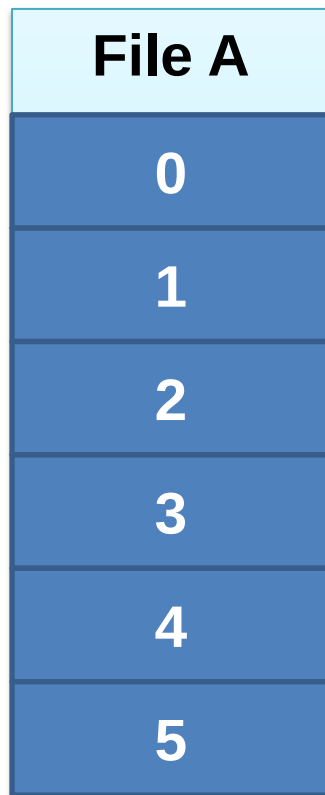
- Compared **both PFSSim versions with a real system** (Graphene cluster from Grid'5000)
 - Described its interconnection to Omnet++
 - Matched SW versions as much as possible
 - Tested with **multiple access patterns**

Results – **file per process** approach

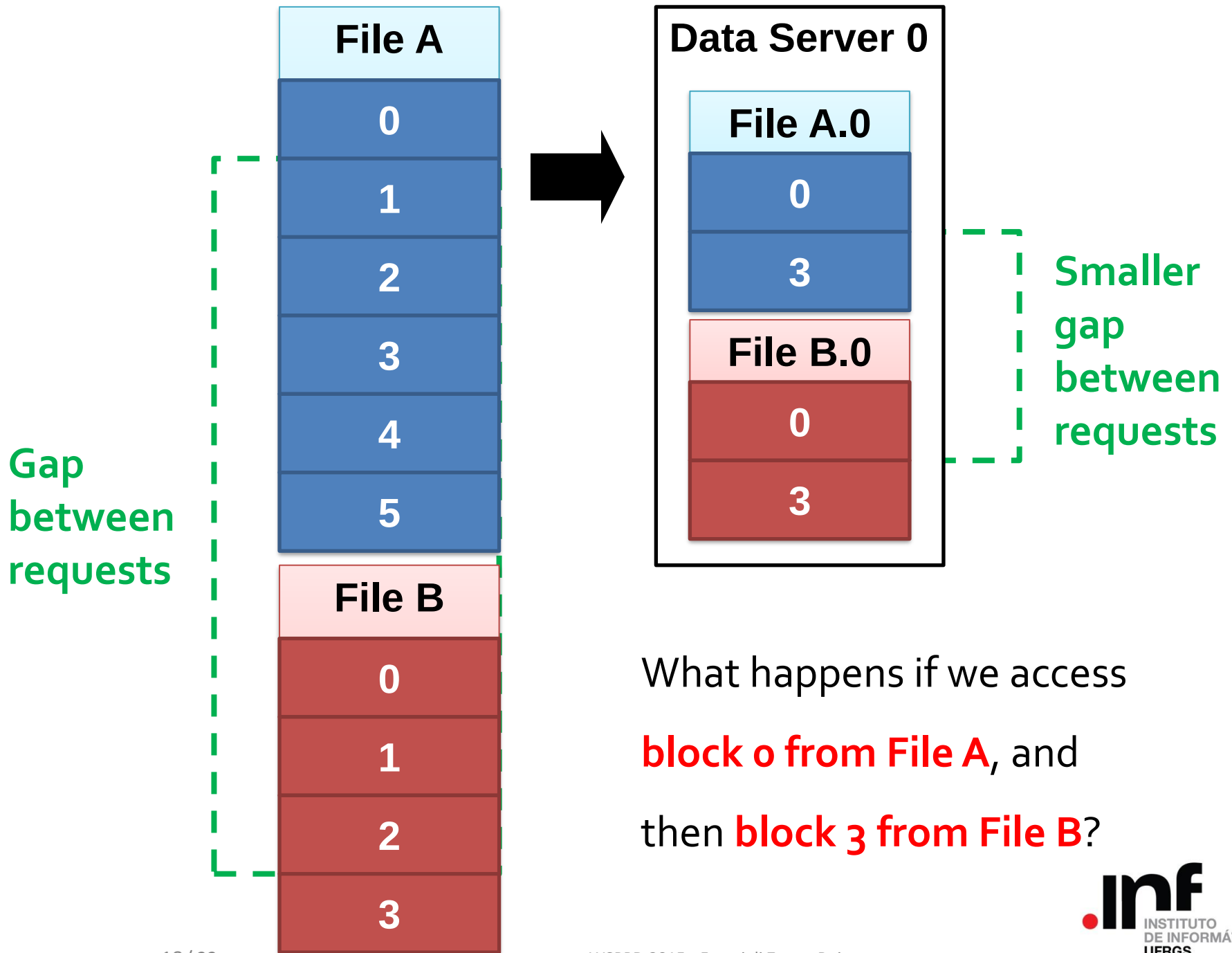
Results – **file per process** approach



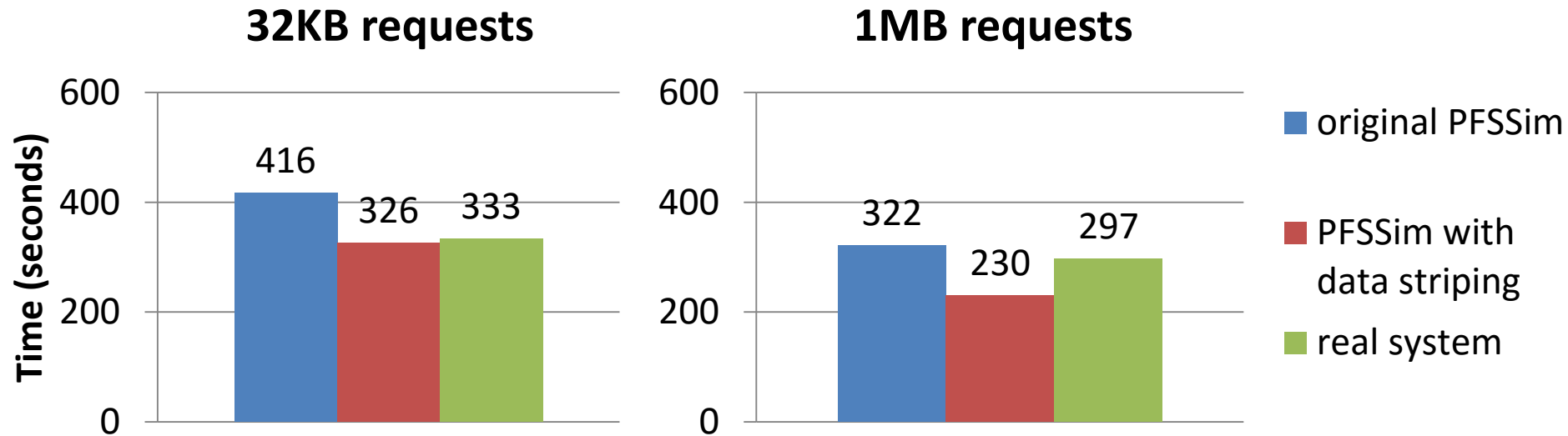
- Simulated **time is always shorter** for the simulator with **data striping**



What happens if we access
block 0 from File A, and
then **block 3 from File B**?



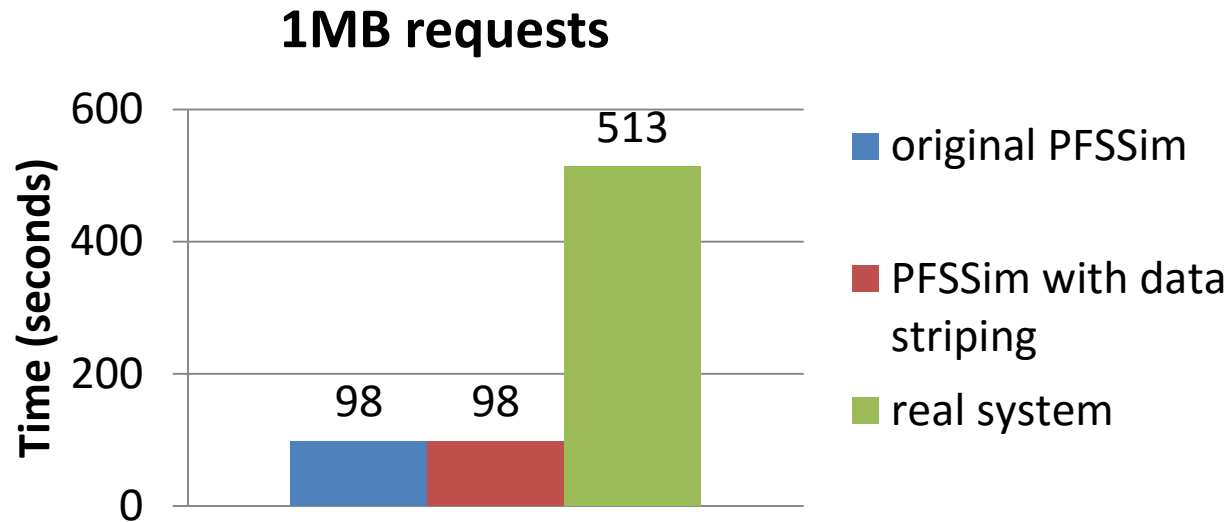
Results – **file per process** approach



- Simulated time is always shorter for the simulator with data striping
- **Good for small requests** - error grows with the **request size**

Results – **shared file** approach

Results – **shared file** approach



- **No difference** between the PFSSim versions
- **Underestimation** of I/O time

Conclusion

- We would like a good **parallel file system simulator**
- PFSSim had some problems
- We have **implemented data striping**
- Good results for small requests and file per process approach
- We still have problems
 - **Disk** module
 - **Cache** module

Improvement of a Parallel File System Simulator by Data Striping Implementation

Vivien Michel

Polytech'Grenoble, France

Francieli Zanon Boito

Universidade Federal do Rio Grande do Sul, Brazil

Mario Dantas

Universidade Federal de Santa Catarina, Brazil

Workshop de Processamento Paralelo e Distribuído - 2015