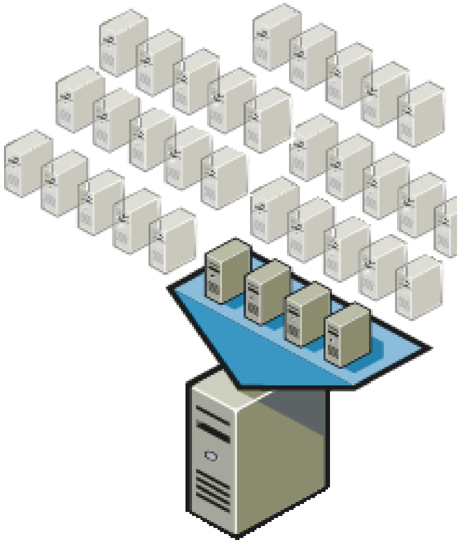
Planning live-migrations to prepare servers for maintenance

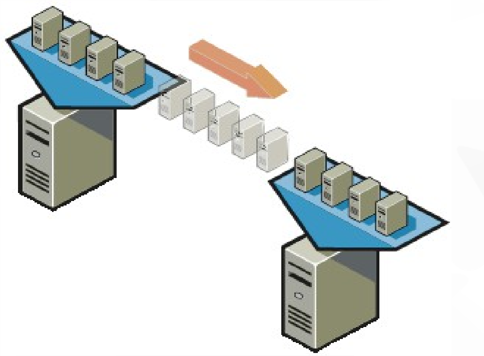
Vincent Kherbache, *Fabien Hermenier,*
Eric Madelaine

Servers maintenance tasks in virtualized DCs

- Hosting capabilities
  - Performance isolation
  - Consolidation techniques
    Increasing number of VMs per server



- Server maintenance
  - Need to operate on idle or offline servers
  - **Live-Migrations** to prepare for the operation
  - Impact all running VMs
  - Server preparation is a critical task !

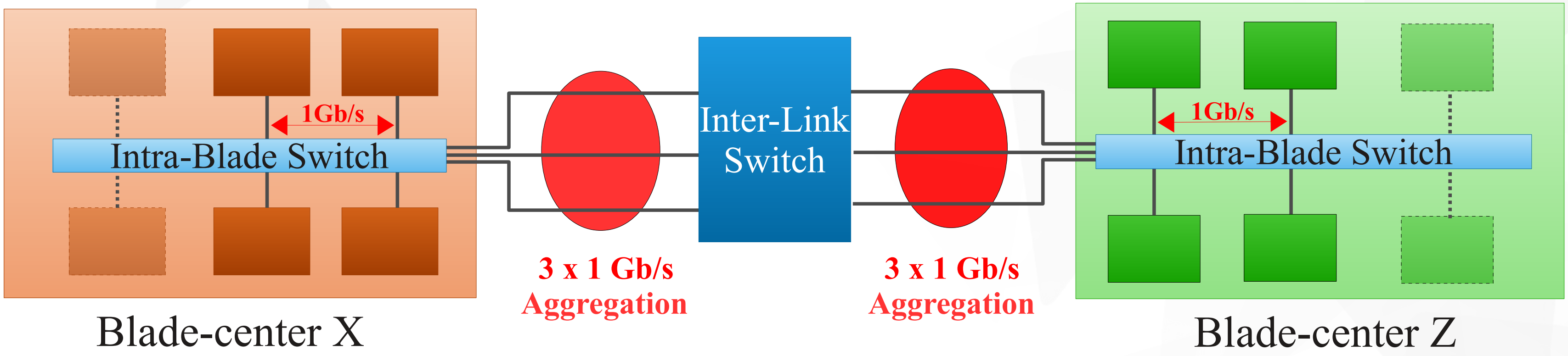


Problematic: How to prepare servers maintenance efficiently

- A problem with many facets
  - Completion times
  - Migration duration
  - Energy usage
  - Technical, environmental, human aspects
- Our contribution: analysis of realistic migrations plans
  - Exhibit common pitfall
  - Deduce levers to improve their quality/efficiency
  - Propose improvements

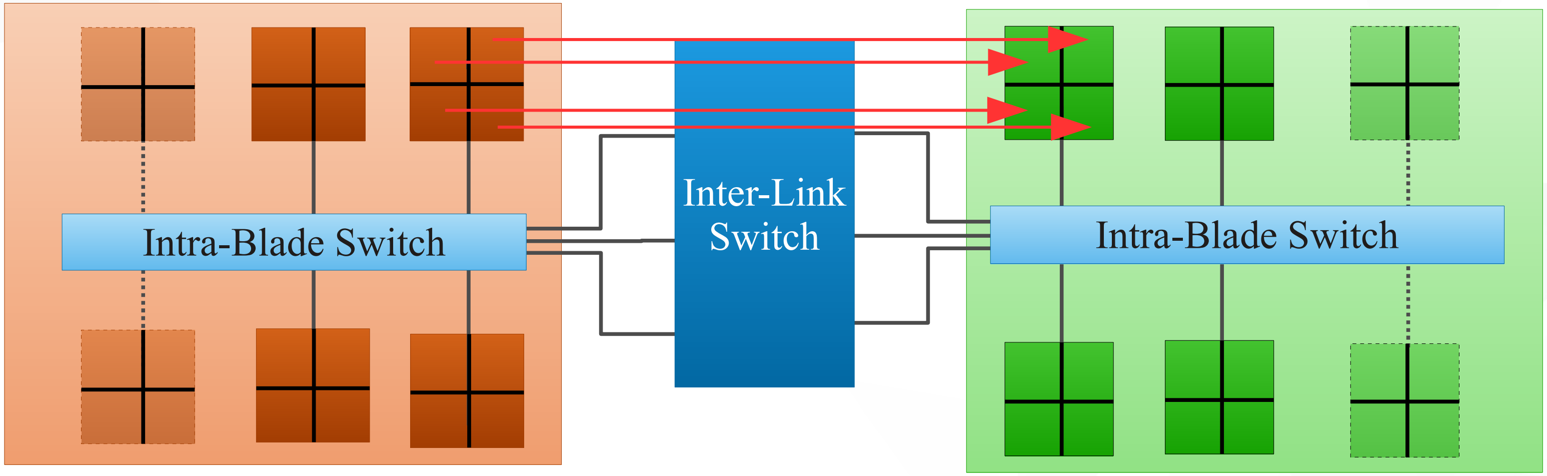
Experimenting servers preparation

- Experimental testbed
  - 3 blade-centers: Bullx B500
  - 15 servers per blade-center: - 2 CPU quad-core @ 2.27GHz,
    - 24 GB ram
- Network
  - 3 x 1 GB/s Inter-link between blade-centers



Experimenting servers preparation

- Scenario: Blade-center maintenance
  - 4 VMs per server
  - 60 VMs to relocate to spare servers

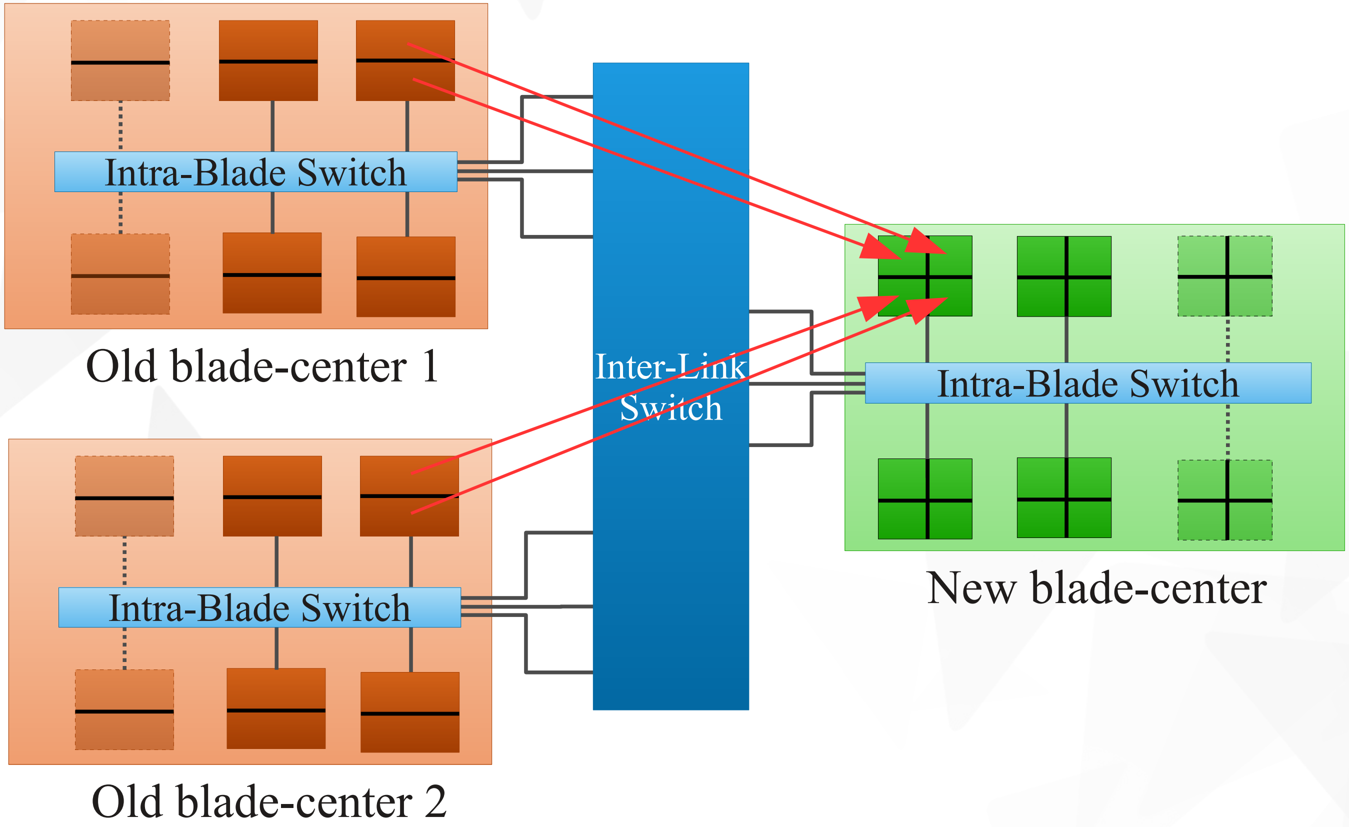


Prod blade-center

Spare blade-center

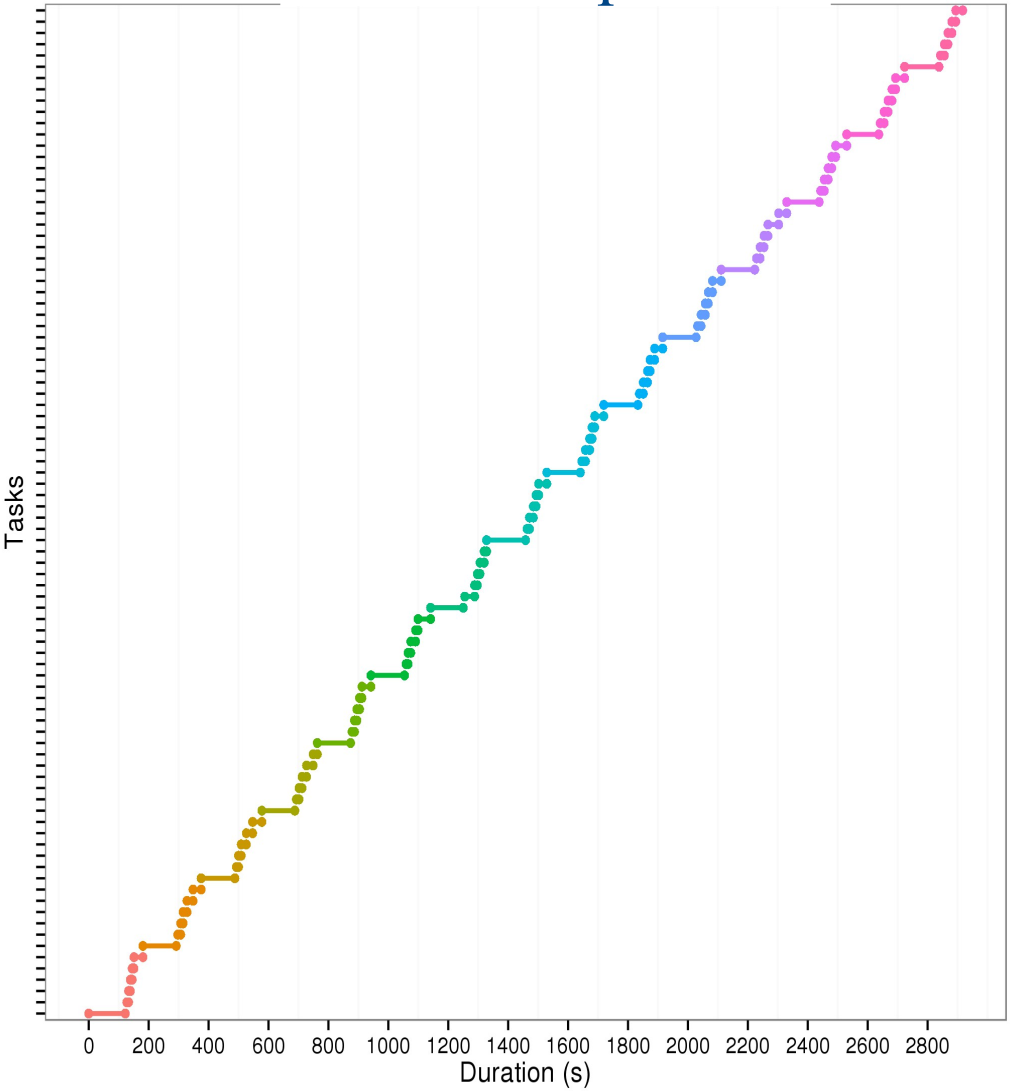
Experimenting servers preparation

- Scenario: Server upgrading
  - Blade-centers replacement
  - From 2 to 4 VMs per server

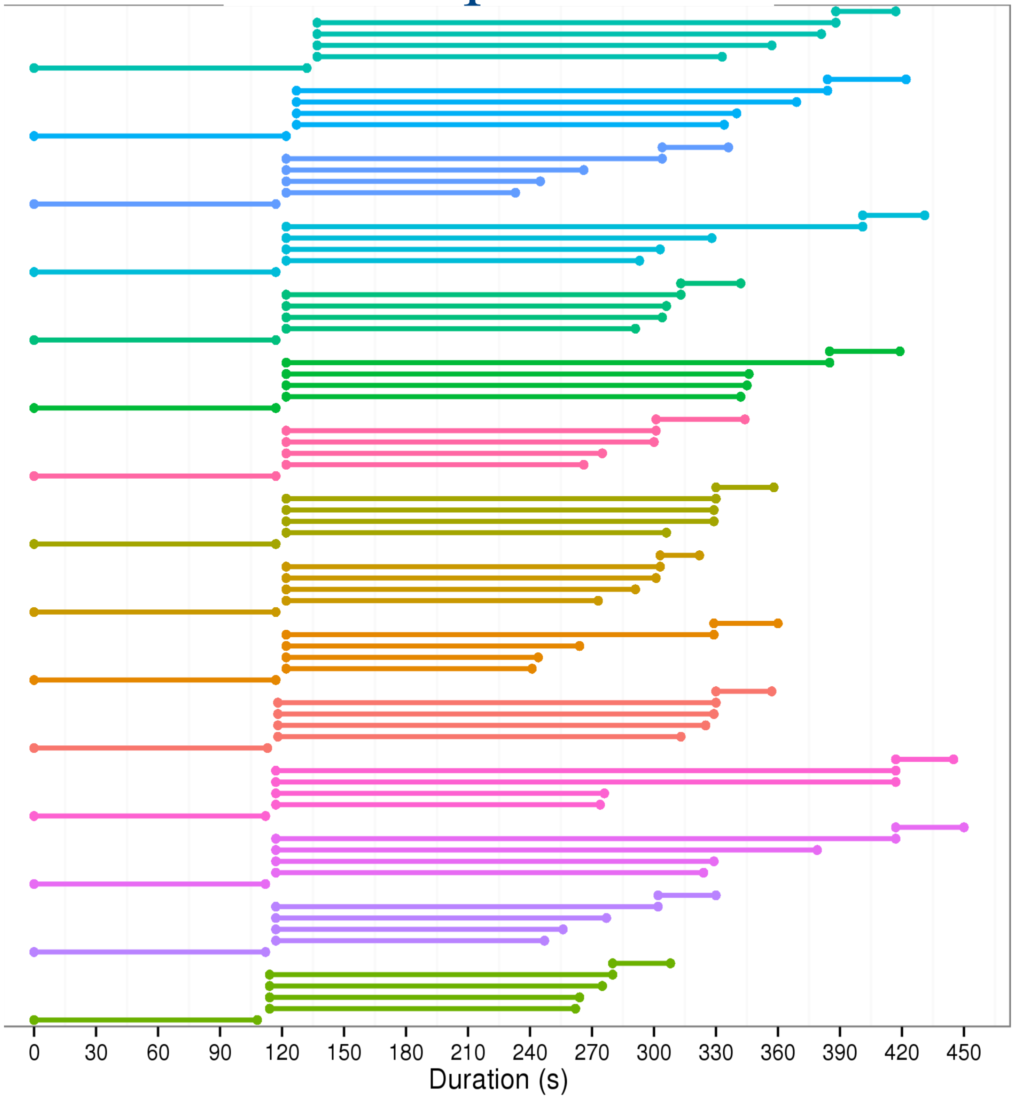


Blade-center maintenance

Full sequential

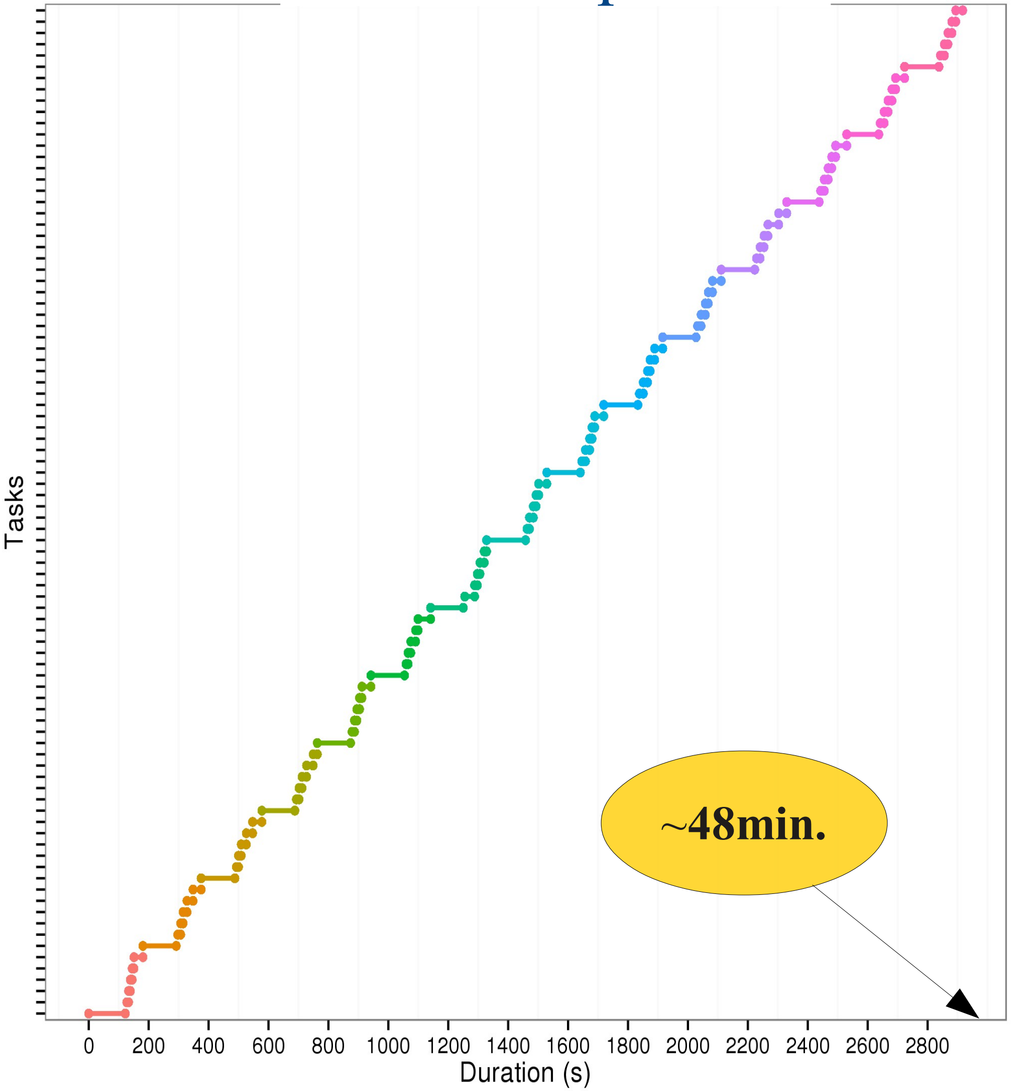


Full parallel



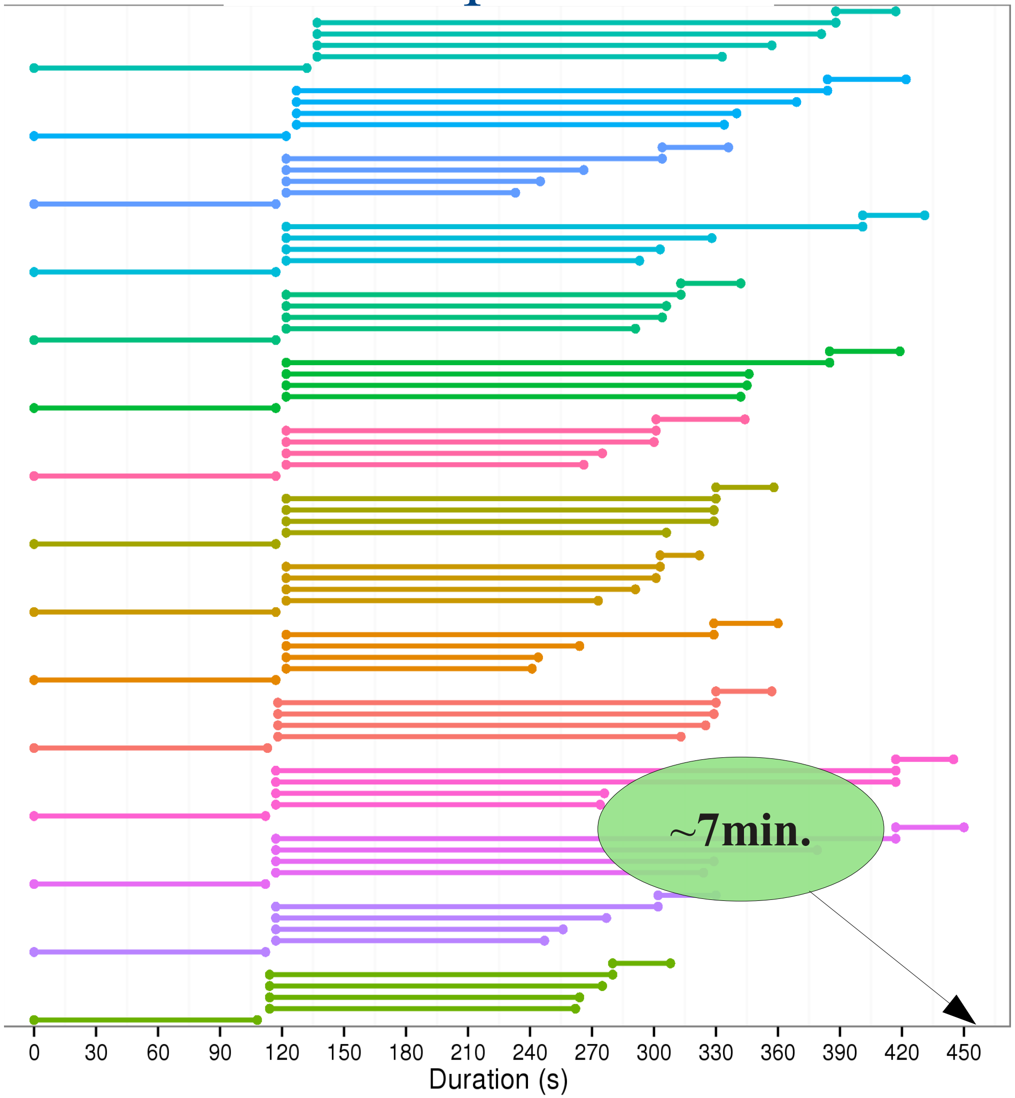
Blade-center maintenance

Full sequential



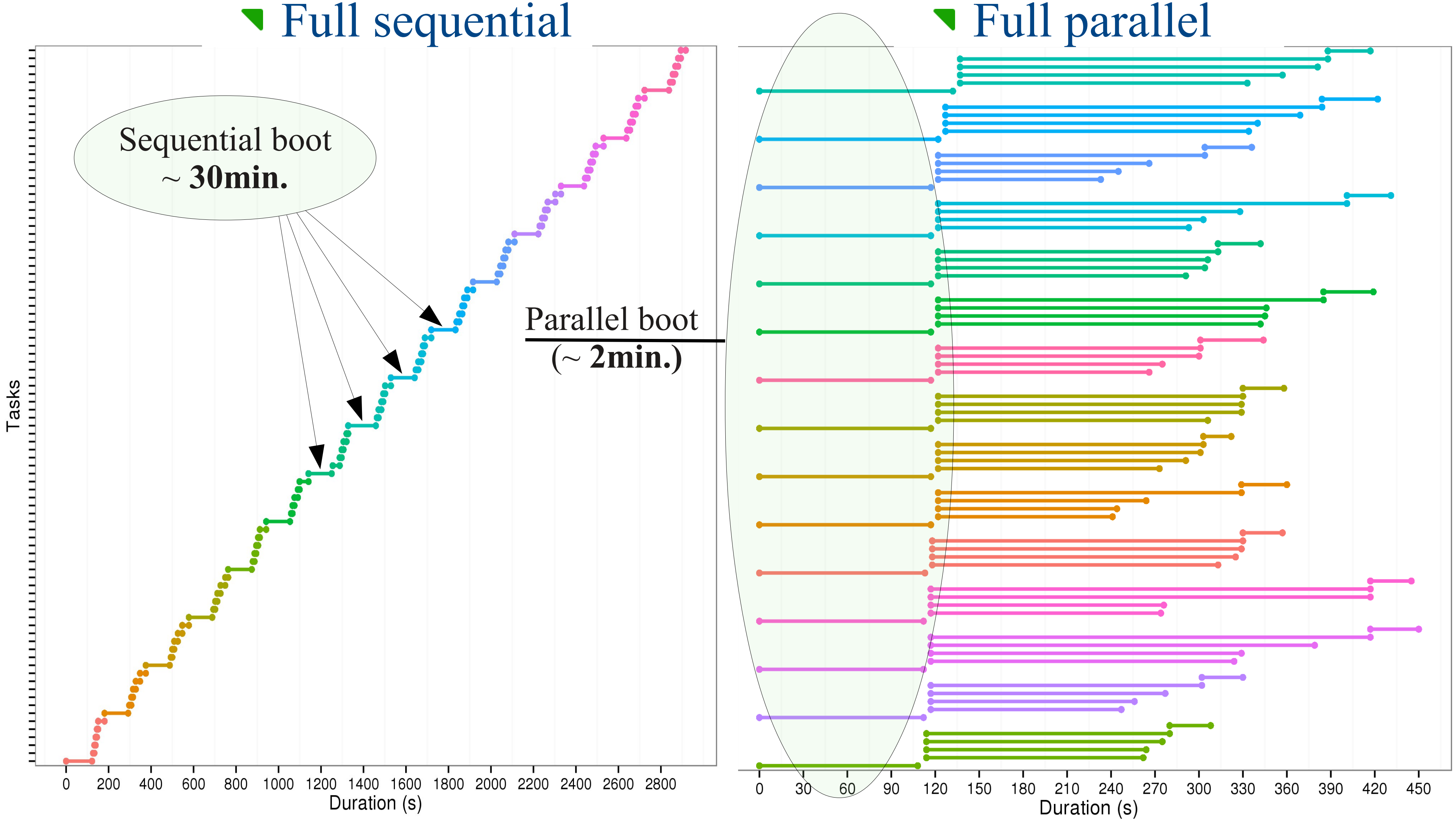
~48min.

Full parallel



~7min.

Blade-center maintenance

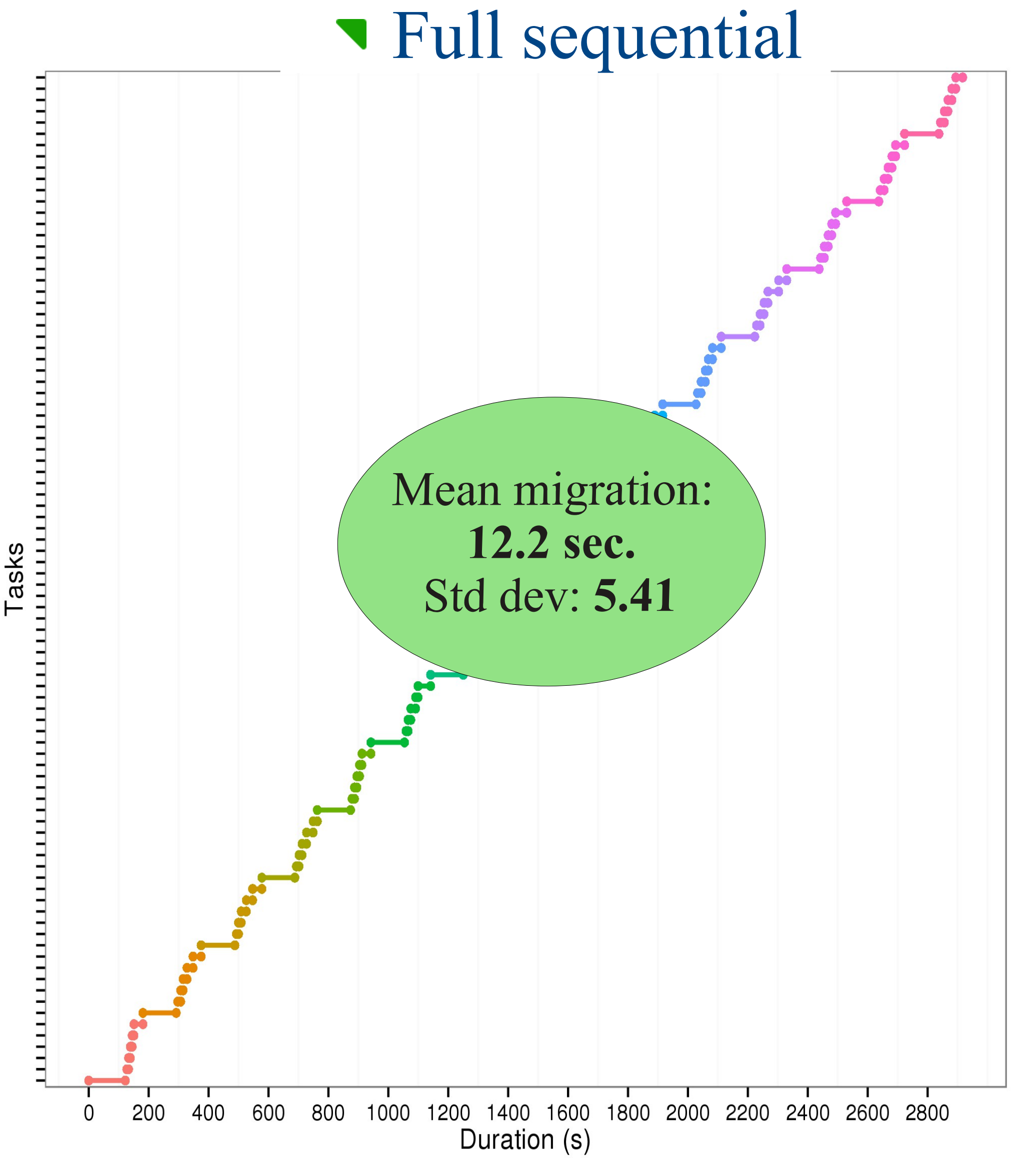


- Full sequential
- Full parallel

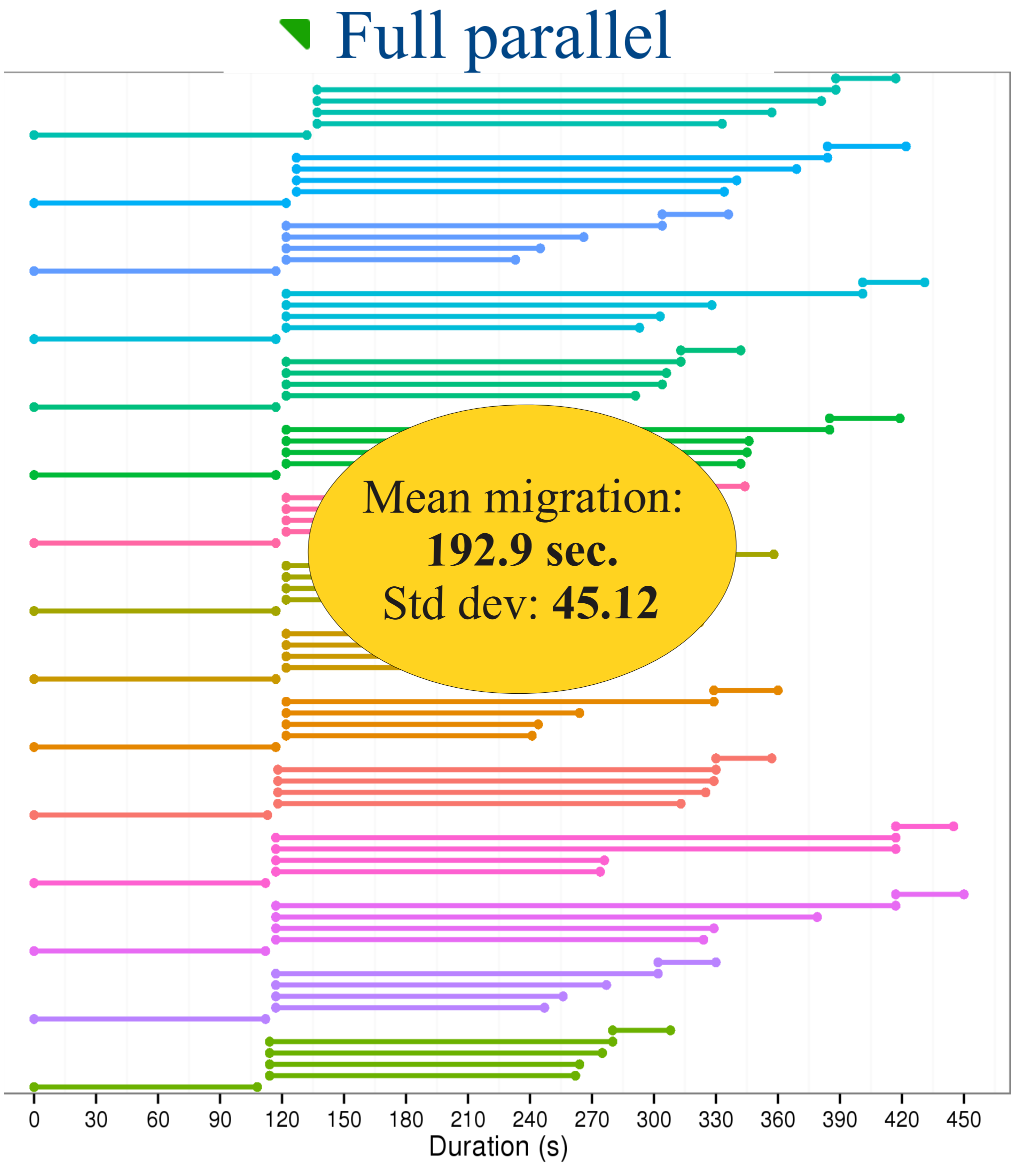
Sequential boot
~ **30min.**

Parallel boot
(~ **2min.)**

Blade-center maintenance



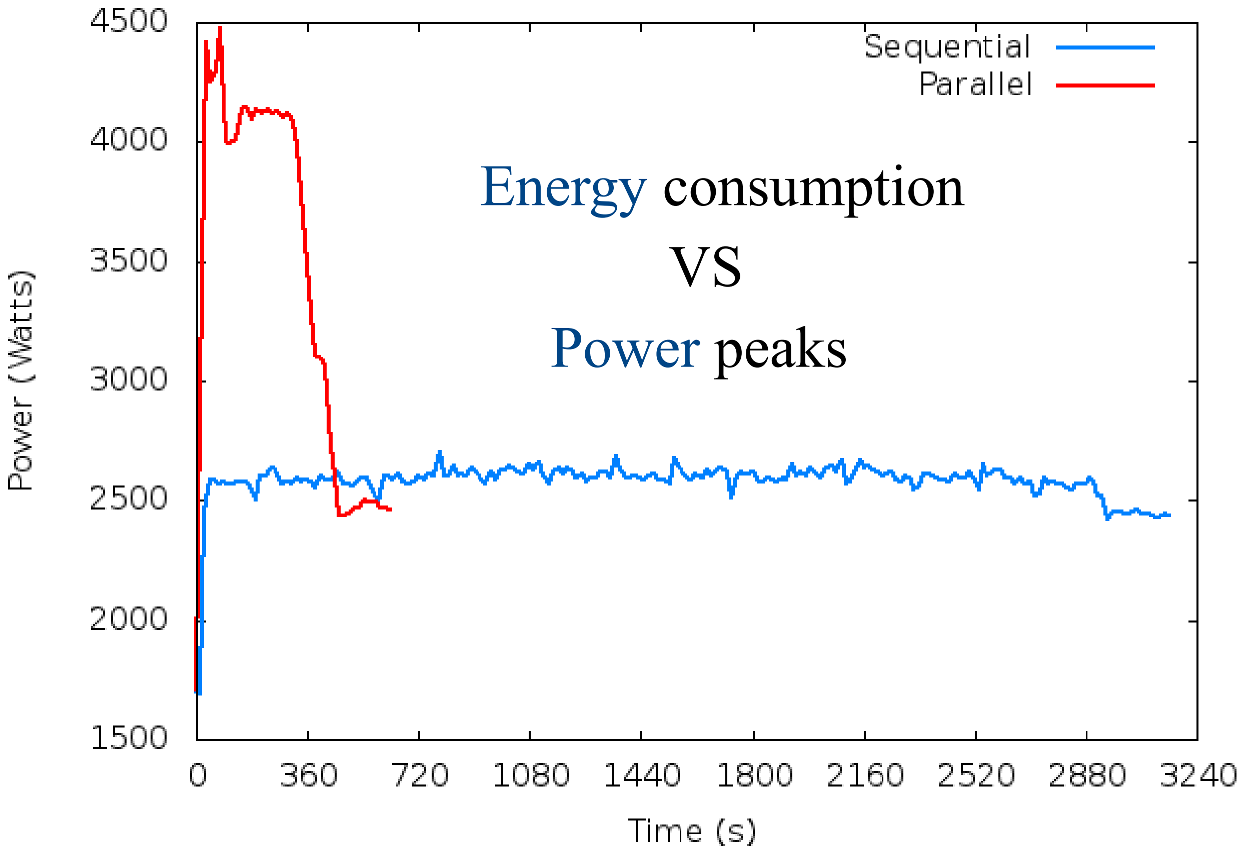
Mean migration: **12.2 sec.** Std dev: **5.41**



Mean migration: **192.9 sec.** Std dev: **45.12**

Energy consumption

- **Blade-center maintenance**



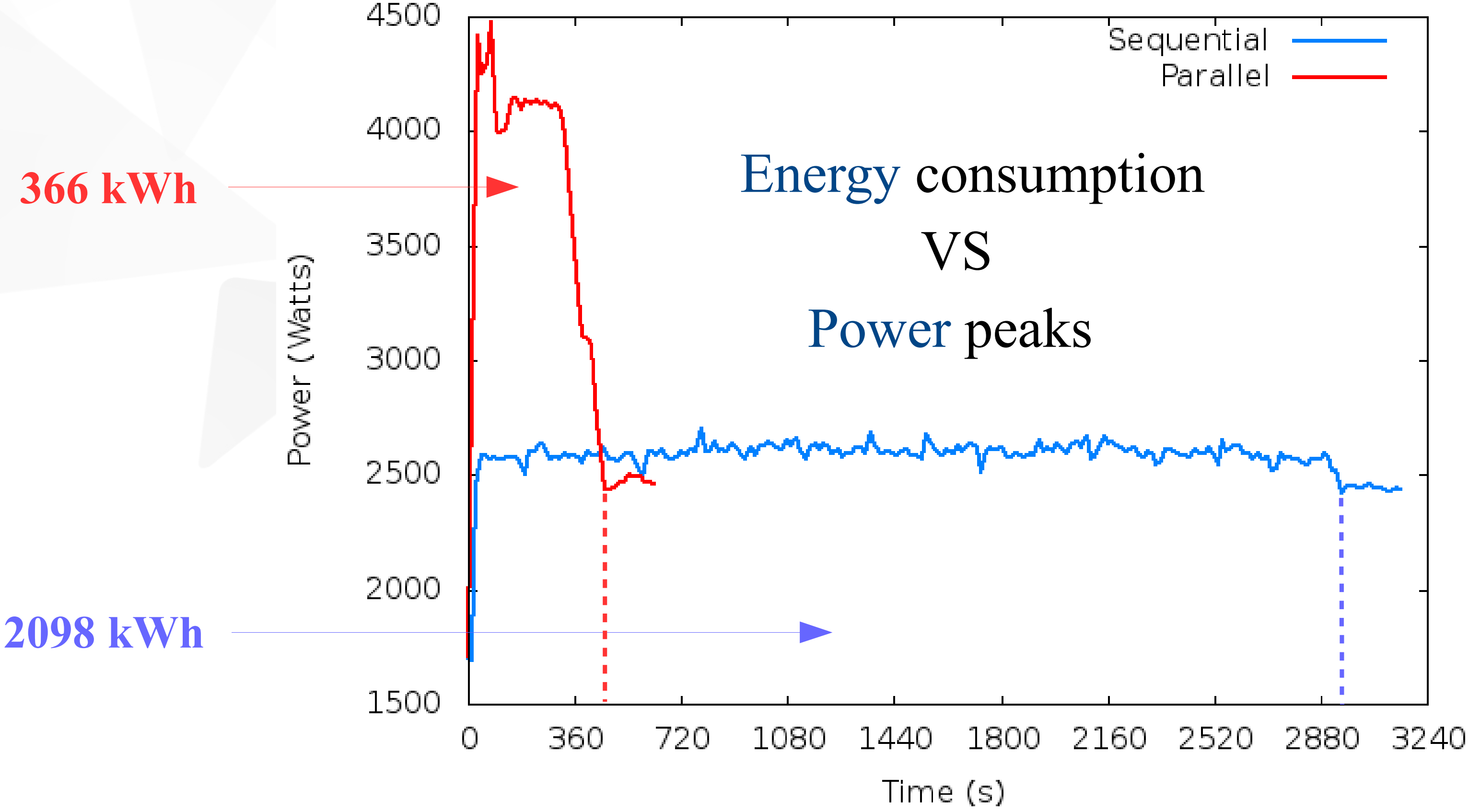
Energy consumption

VS

Power peaks

Energy consumption

- **Blade-center maintenance**



366 kWh

2098 kWh

Energy consumption

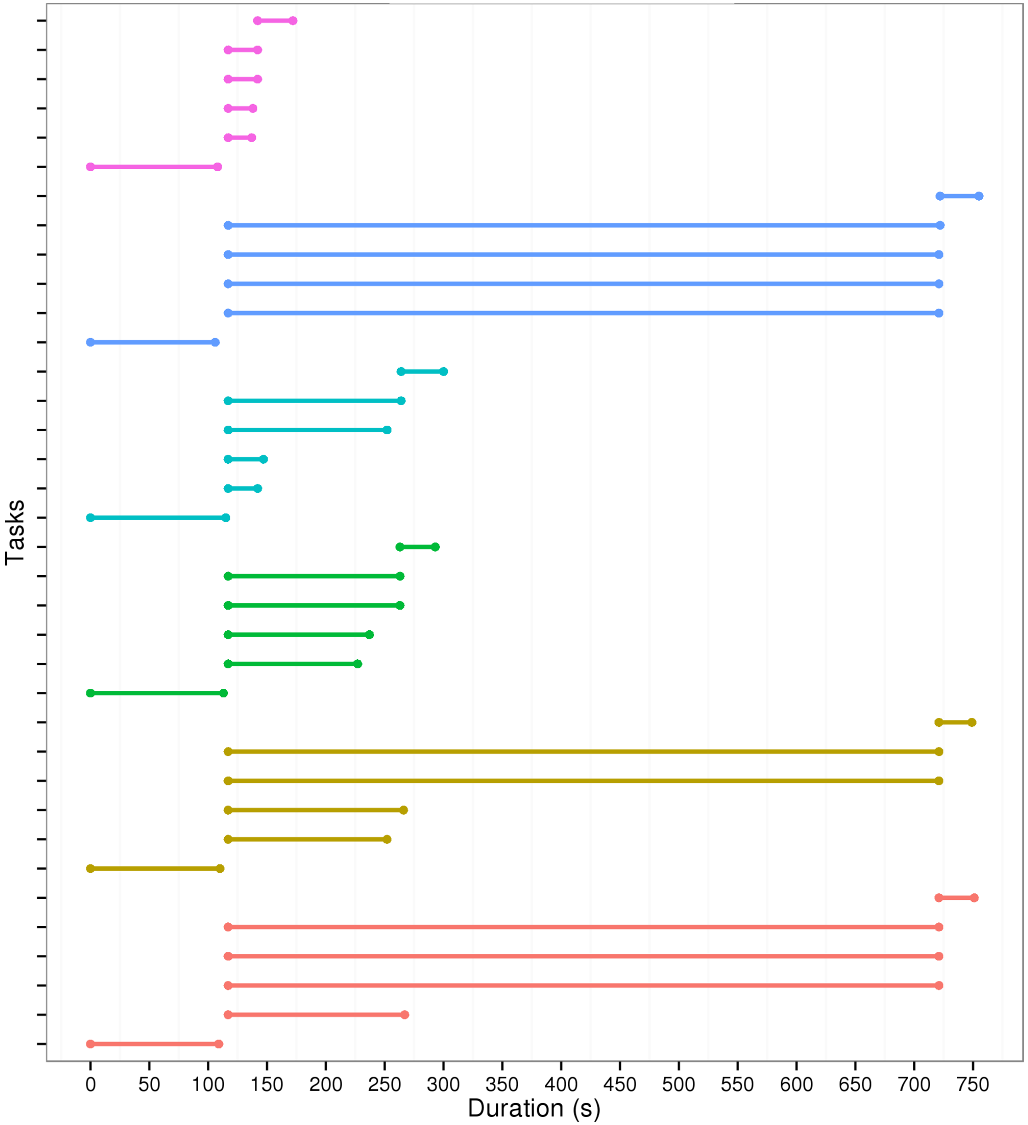
VS

Power peaks

Neither of the two strategies is efficient

- No generic solution
  - It is a matter of trade-offs
  - Parallel / Sequential migrations
- Adaptive to the environment peculiarities
  - Network links capacity / topology
  - Workload specificities

Adaptation to the interlink peculiarities



6 to 6 servers, 4 VMs per server

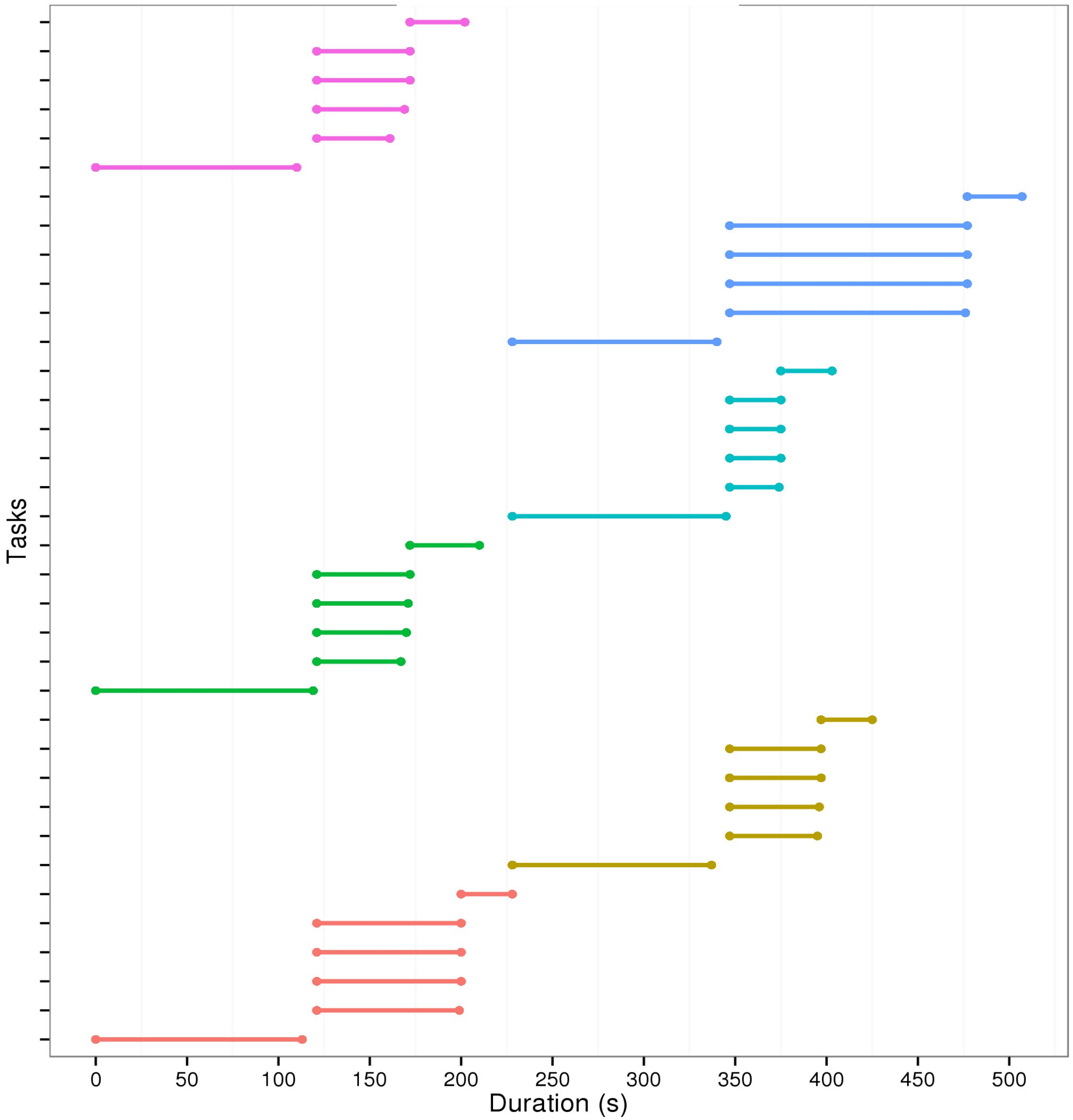
24 migrations

- Full parallel scenario

- Timeout: 10min

9 cold migrations

Improvements – Interlink peculiarities



Nodes replaced 3 by 3

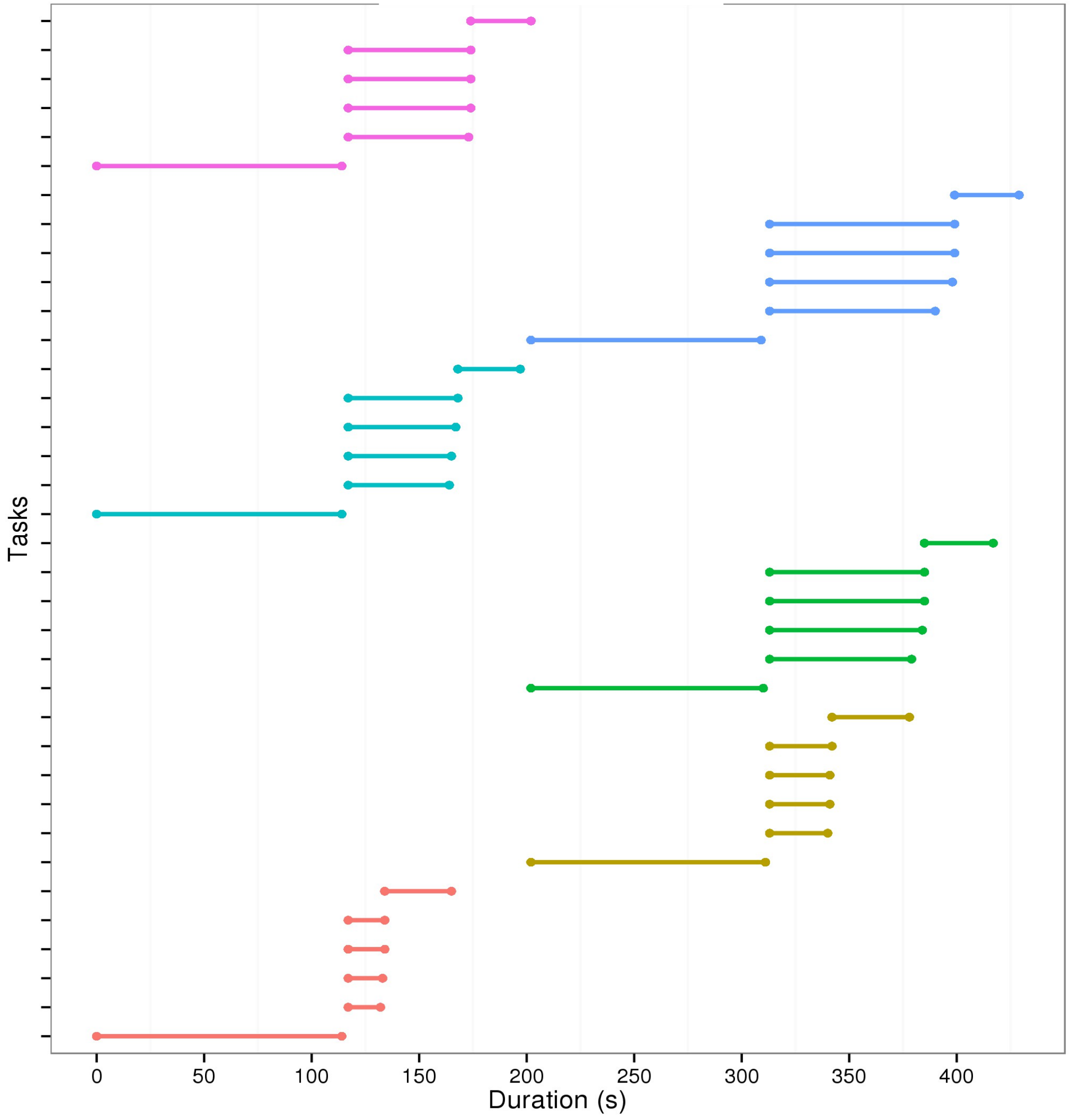
- Completion time reduced by 33%
- Migration times 4 times lower

- Better interlink usage

No cold migration

- High std. deviation: 33.15 ?
- **Aggregation not fair !**

Improvements – Interlink peculiarities



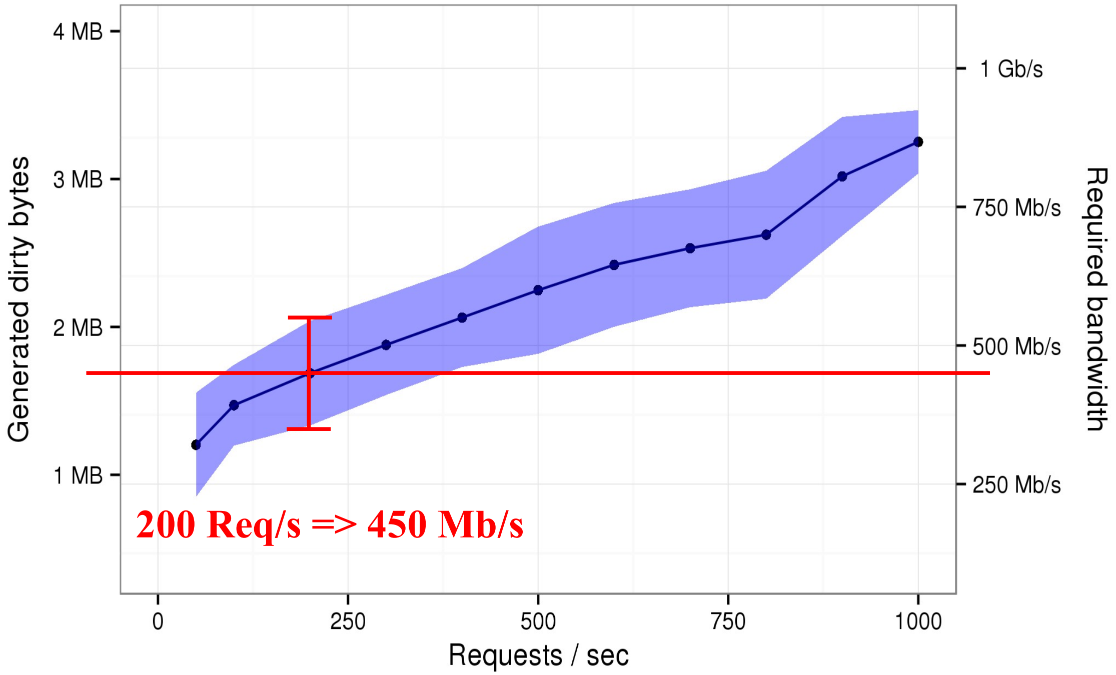
Ensuring fair aggregation

- Completion time reduced by 44%
- Migration times 5 times lower
- Still high std .deviation: 23.17

1Gb/s links saturated

Adaptation to the workload peculiarities

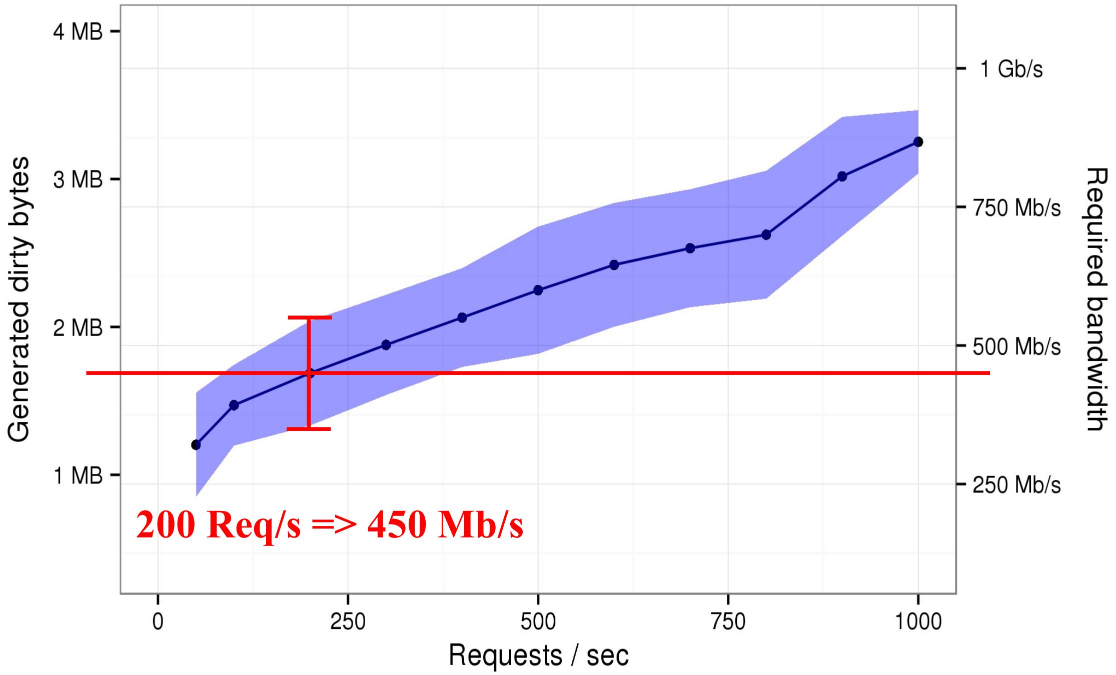
- How many migrations in parallel ?
- Wrt. the VM's dirty page rate:
  - KVM downtime: 30ms
  - Httperf workload



200 Req/s => 450 Mb/s

Adaptation to the workload peculiarities

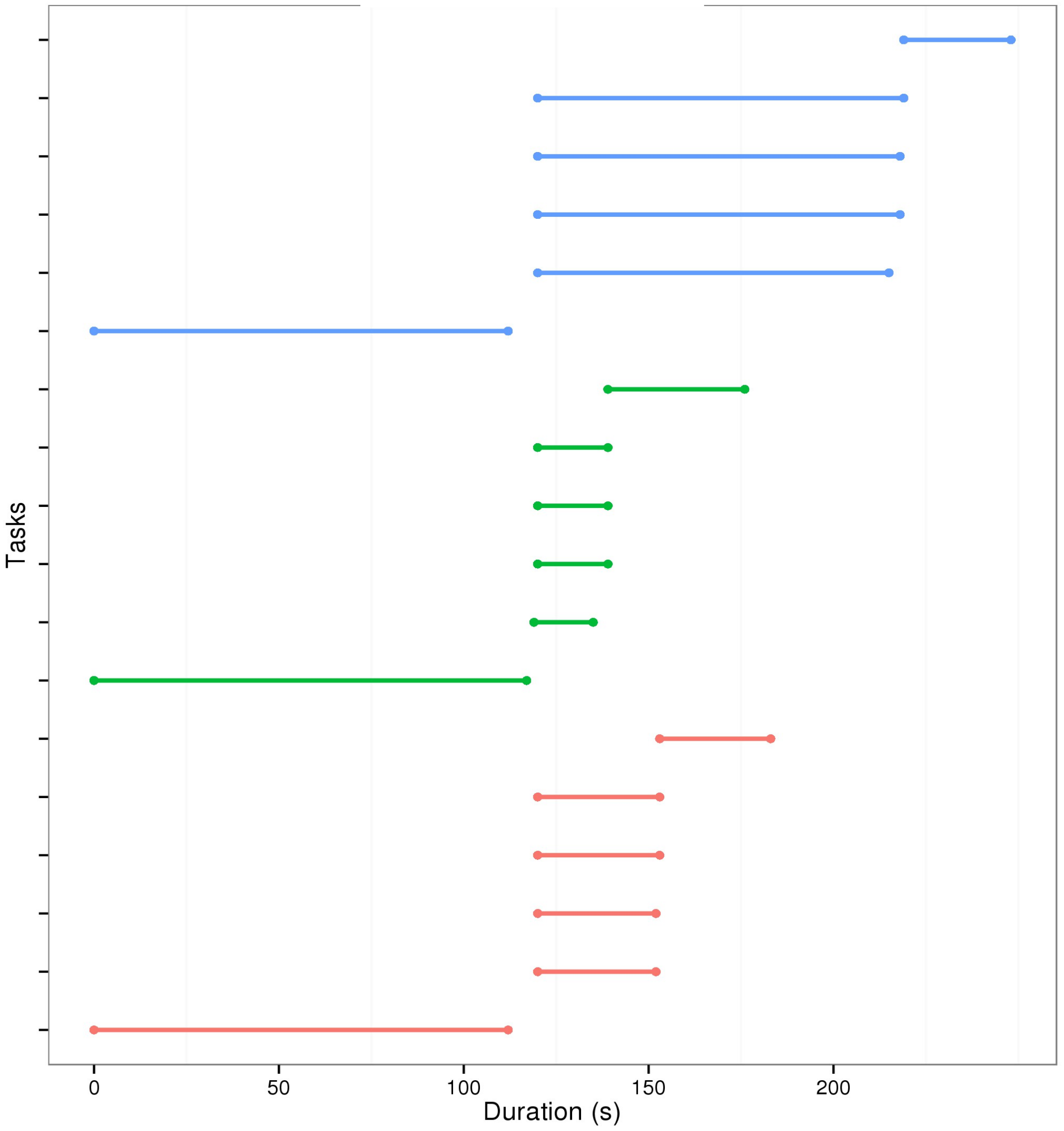
- How many migrations in parallel ?
- Wrt. the VM's dirty page rate:
  - KVM downtime: 30ms
  - Httperf workload



200 Req/s => 450 Mb/s

- **2 migrations max per 1Gb link**

Improvements – Workload peculiarities



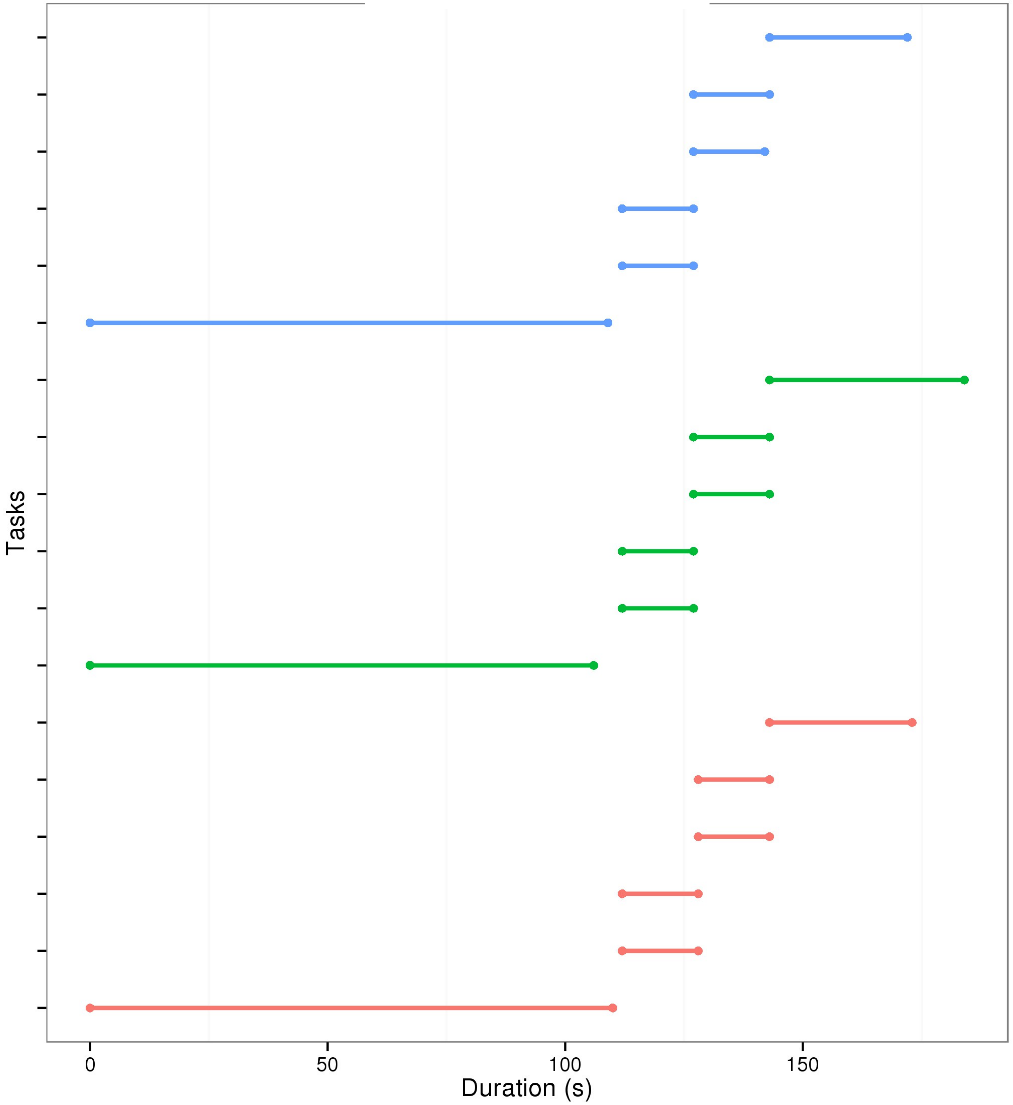
- 3 to 3 servers, 4 VMs per server
  12 migrations

- Full parallel scenario

- Mean migration duration :
  49.51 sec.

  Std.dev: 34.51

Improvements – Workload peculiarities



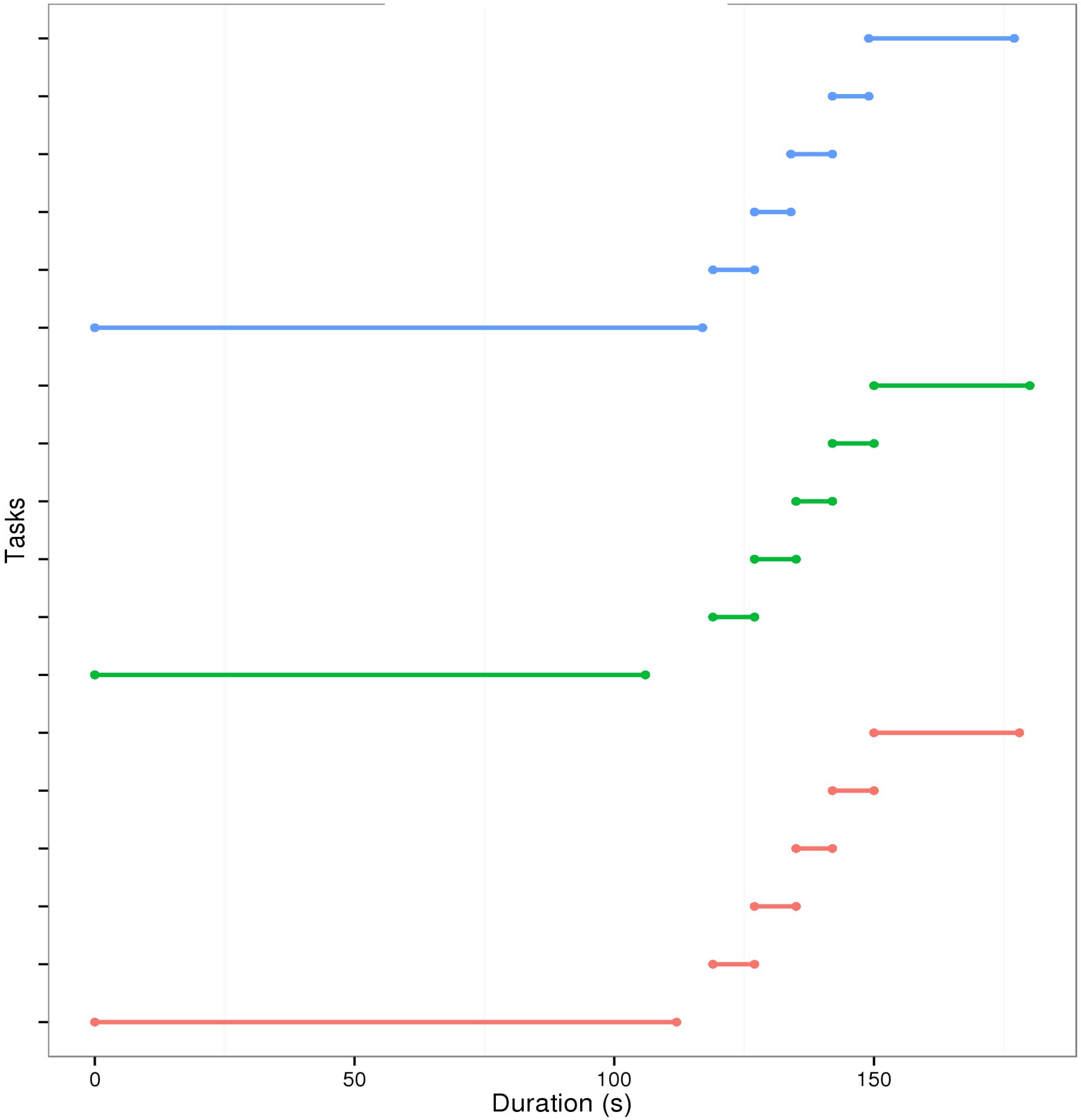
VMs migrated 2 by 2

- Mean migration duration : 15.41 sec.

  Std. deviation: 0.47

- Completion time reduced by 31%

Improvements – Workload peculiarities



VMs migrated 1 by 1

- Mean migration duration : 7.66 sec.

  Std.dev: 0.45

- Same completion time than previous

- But no one is better !
  - Depend on application context

Conclusion

- Based on good environment knowledge, we greatly improved the maintenance preparation efficiency.
- Prepare efficient migrations plans to maintenance preparation is complicated !
- To automate we need to be adaptive to the peculiarities:
  - Knowledge of the environment
  - Highly dependent of specific metrics

Future works

- Model the aspects of a migration plan
  - Dirty page rate
  - Network topology
  - Estimated migration durations

- Model interaction with external side constraints
  - Power budget, Completion deadline, Licensing policy, ...

- Implement the model over the VM manager
  - Composable VM placement algorithm
  - Support side constraints expressed by operators

BtrPlace

Planning live-migrations to prepare servers for maintenance

Vincent Kherbache*, Fabien Hermenier,*
Eric Madelaine

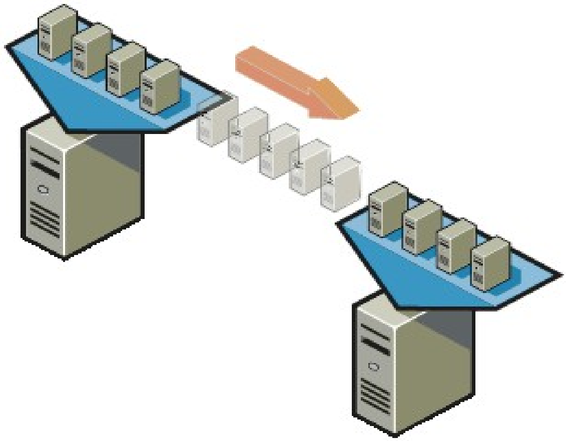
Servers Upgrading

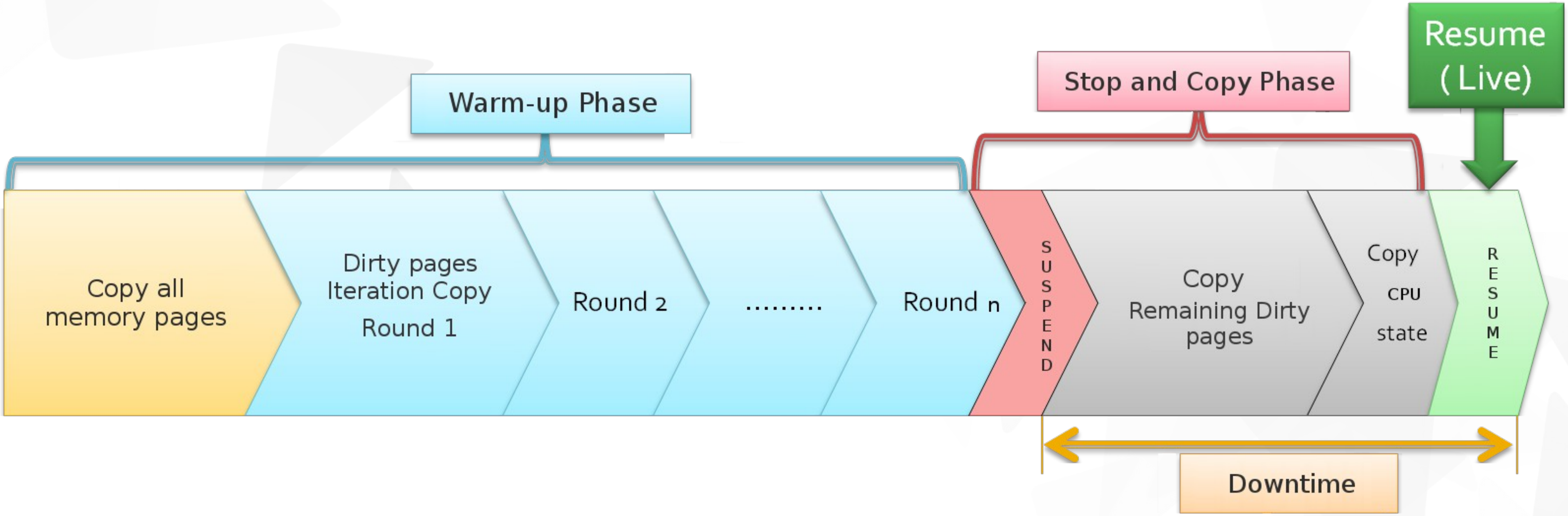
Plan

- Introduction
  - Virtualized Data-centers
  - Maintenance tasks
  - Live migrations
- Planning live migrations
  - Intuitive migrations plan analysis
  - Toward smarter migration plans
- Conclusion & Future works
  - Efficient plans automation
  - VM manager BtrPlace

Introduction - Live-migration

- Migration with negligible downtime
- Pre-copy algorithm (KVM, XEN, ..)
  - 1. Warm-up phase
  - 2. Stop-and-copy phase





Scenarios Comparison

	Maintenance		***Upgrading***	
Metrics	**Sequential**	**Parallel**	**Sequential**	**Parallel**
Time to completion (sec.)	2871	446	3467	384
Mean migration duration (sec.)	12.2	192.9	11.2	158
standard deviation	5.41	45.12	4.81	52.97
Server boot duration (sec.)	113.1	116.5	114.9	115
Server shutdown duration (sec.)	29.5	28.8	32.2	32.1

Energy consumption (kWh)	2098.4	366.4	3317.5	548.1
Max. peak power (kW)	2.7	4.47	4.24	6.05