Efficient multicriteria scheduling on large scale computing platforms

First year PhD student monitoring

Millian Poquet

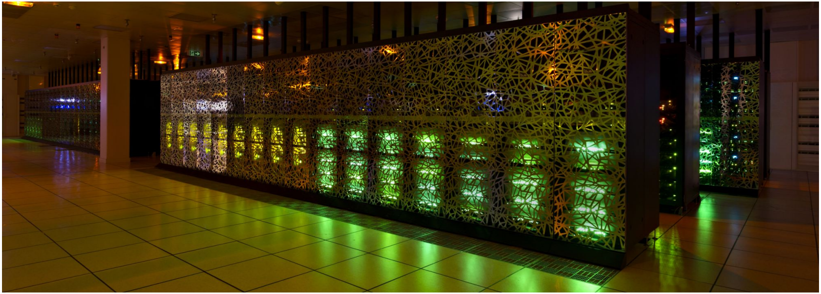
Supervisors: Denis Trystram and Pierre-François Dutot
Univ. Joseph Fourier, MOAIS, LIG, Inria

15 juin 2015

Table of contents

1. Topic
2. Batsim
3. First experiments
4. Future

HPC platforms



HPC platforms

- Computational power ↑
- Energy consumption ↑
- Exascale around 2023...

Our problem of interest

Input

- Jobs submitted by users at a given time
- A computing platform

Output

- A schedule : when and where the jobs should be executed

Optimization criteria

- User-oriented criterion *e.g.* average waiting time, bounded slowdown...
- Total energy consumption
- Power peak

Our problem of interest (2)

Realistic context

- Online
- Efficient
- Limited knowledge of the jobs
  - Uncertain runtime (but known upper bound)
  - Unknown computation and communication patterns

Table of contents

1 Topic

2 Batsim

3 First experiments

4 Future

Why simulate ?

Real-world experiments

- Long running time
- Access problems
- Unreasonable energy impact

Simulations

- Much faster
- Wide range of platforms
- Low energy impact
- Reproducibility

Batsim

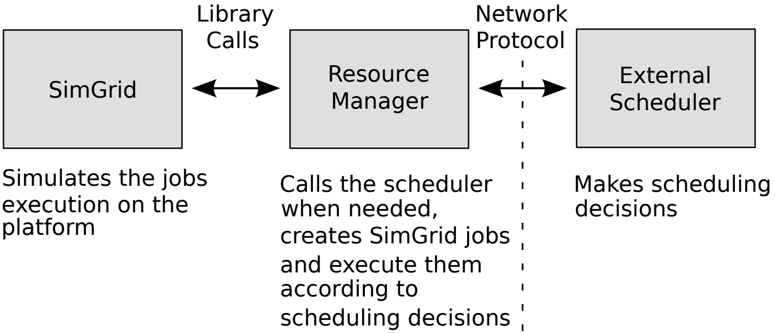
Needs

- Compare different scheduling algorithms
- Use existing scheduling algorithms, implemented in different languages
- Make easy the development of new algorithms
- *Realistic* results

Choices

- Use a simulation framework rather than a *from scratch* approach
- Strong separation of the scheduler

Batsim : overview



Simulates the jobs execution on the platform

Calls the scheduler when needed, creates SimGrid jobs and execute them according to scheduling decisions

Makes scheduling decisions

Table of contents

We now have a simulator prototype...

What can we do with it ?

- Add features to directly focus on the energy
- Shorter term : validate it with more classical scheduling problems

Locality constraints

Suppositions on the platform

- The resources are in a set M
- For each $m \in M$, let id_m be its unique identifier ($s.t.$ $id_m \geq 0$ and $id_m < |M|$)
- Hierarchical : resources are grouped in clusters

Contiguity

Allocations must be done on a contiguous subset of resources

Locality

A job fits entirely in 1 cluster → it must be allocated to only 1 cluster

Locality constraints : models vs simulations

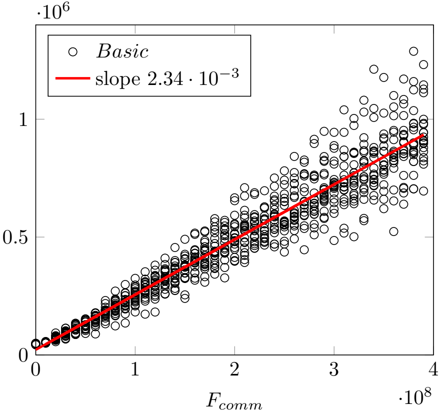
Theoretical results

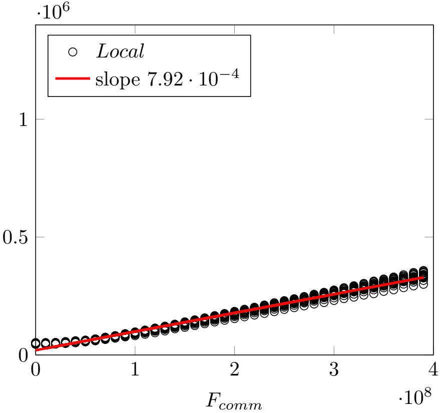
- Imposing such constraints reduces the solution space
- → might degrade performances (*e.g.* makespan up to a factor of 2)

Practical intuition

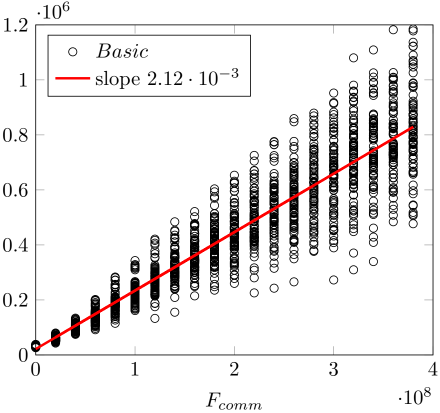
- Locality → faster job execution → reduced makespan...

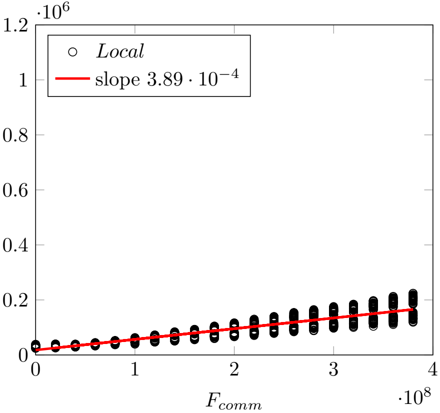
Results on homogeneous platforms





Results on heterogeneous platforms





Locality constraints : models vs simulations (2)

Conclusions

- As expected, increased amount of communication → increased gain with locality constraints
- Such constraints are beneficial for the schedules
- Models with parallel tasks whose internal communications are hidden are ill-suited to current architectures and should be reevaluated
- Paper submitted to HeteroPar'2015

Future

Batsim

- Energy
- IO
- OAR
- SLURM ?

Algorithms

- Energy
- Locality
- Topology ?