



L I F C

IWJPDC'04

UNIVERSITÉ DE FRANCHE-COMTÉ



Combination of Java and asynchronism for the Grid : a comparative study based on a parallel power method.



FRE 2661

Jacques Bahi, Stephane Domas, Kamel Mazouzi
bahi, sdomas, mazouzi@iut-bm.univ-fcomte.fr

Young Team AND



L I F C

Overview

UNIVERSITÉ DE FRANCHE-COMTÉ



- ① Distributed computing and its problems,
- ① Asynchronous vs. synchronous iterative algorithms,
- ① Jace : an environment for asynchronous algorithms,
- ① Parallel Power Method,
- ① Experiments,
- ① Conclusion.



FRE 2661



L I F C

Distributed computing

UNIVERSITÉ DE FRANCHE-COMTÉ



- ① Classical hardware architecture:
 - clusters or LANs distributed over different geographical sites,
 - shared and variable bandwidth between sites.
- ① Classical software framework :
 - MPI-like environments,
 - meta-computing environments (GLOBUS, DIET,...).



FRE 2661



L I F C

Distributed computing : problems

UNIVERSITÉ DE FRANCHE-COMTÉ



- 🎯 Network :
 - restricted accesses (firewall, private IPs, ...),
 - costly synchronizations,
 - bandwidth between sites easily overloaded.
- 🎯 Heterogeneity:
 - load balancing often necessary,
 - different communication/computation libraries on sites,
 - fastidious compilation and installation.



FRE 2661



L I F C

Distributed computing : classical solutions

UNIVERSITÉ DE FRANCHE-COMTÉ



- ① Network : dedicated high speed network.
 - expensive.
- ① Heterogeneity : meta-computing environment.
 - a single site is chosen for each execution,
 - restrictions on the executed code,
 - user has little control over execution,
 - heavy environments.



FRE 2661

These solutions adapt the Grid to existing works !



L I F C

Distributed computing : innovating solutions

UNIVERSITÉ DE FRANCHE-COMTÉ



- ① Avoid synchronizations :
 - less bandwidth overload,
 - less idle times.
- ① Use a dedicated Java environment.
 - 100% portable,
 - lightweight environment,
 - user can execute any code he develops.



FRE 2661

Asynchronous algorithms implemented with Jace environment provide all of these !



L I F C

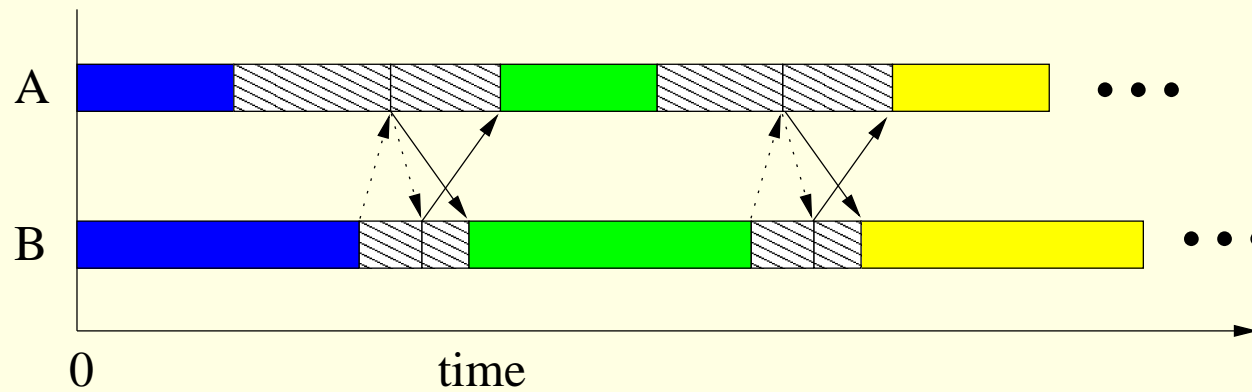
Synchronous vs. Asynchronous

UNIVERSITÉ DE FRANCHE-COMTÉ

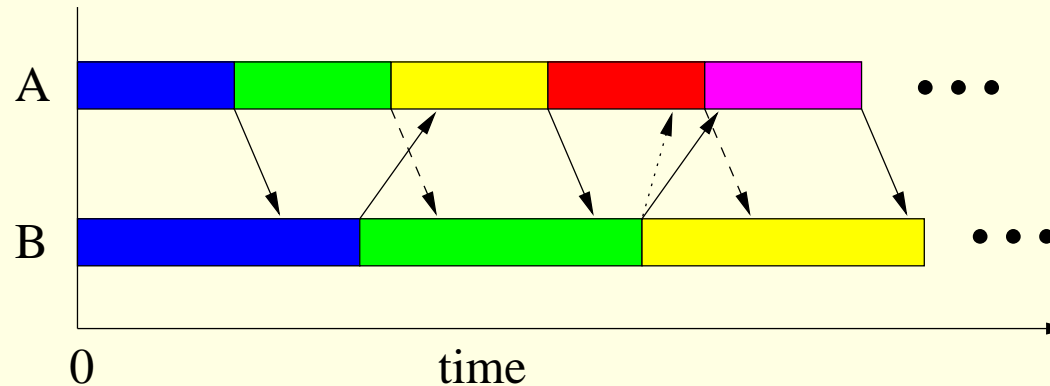


For an iterative algorithm :

Synchronous Iterations Asynchronous Communications



Asynchronous Iterations Asynchronous Communications



FRE 2661



L I F C

AIAC algorithms : summary (1)

UNIVERSITÉ DE FRANCHE-COMTÉ



- ① No synchronization phases :
 - no gather-scatter to test convergence at the end of each iteration,
 - exclusive use of non blocking communications.
- ① Message incoming tolerance :
 - no idle time to wait for needed messages,
 - messages can be delayed or even lost,
 - discarding "too old" messages is possible.



FRE 2661



L I F C

AIAC algorithms : summary (2)

UNIVERSITÉ DE FRANCHE-COMTÉ



- ① No idle times between iterations :
 - at each time, processors may compute different iterations,
 - heterogeneous workload has less impact.
 - ① Convergence under some conditions :
 - satisfied for a large class of scientific applications,
 - convergence acceleration mechanisms.
- ⇒ perfect for grid/distributed computing.



FRE 2661



L I F C

UNIVERSITÉ DE FRANCHE-COMTÉ



FRE 2661

Jace

- ⊙ Designed to manage communications for asynchronous algorithms,
- ⊙ Builds a parallel virtual machine by connecting a set of heterogeneous and distant machines,
- ⊙ Provides a simple programming interface to develop message passing codes.
- ⊙ Jace is composed of :
 - Daemon,
 - Task,
 - Spawner.



L I F C

Jace Architecture : Daemon

UNIVERSITÉ DE FRANCHE-COMTÉ



- 🎯 Launched on each computing node,
- 🎯 Provides authentication and task execution,
- 🎯 Each Daemon can run several tasks,
- 🎯 Manages communications :
 - messages are sent/received via RMI,
 - multi-threaded to implement non blocking communications,
 - older occurrences of the same message are **automatically** crushed.



FRE 2661



L I F C

Jace Architecture : Task

UNIVERSITÉ DE FRANCHE-COMTÉ



- ① The Task class is :
 - the computing unit,
 - a thread rather than a process (**unlike MPI**),
 - the programming interface to access to the communication layer (non-blocking send/receive, barrier,...).
- ① User's Task :
 - inherits from the Task Class,
 - defines a run method containing the computing code.



FRE 2661



L I F C

Jace : Raw Performances

UNIVERSITÉ DE FRANCHE-COMTÉ



Length in Kb	Jace	JPVM	MPI
0	7	8	0.11
10	8	12	2.23
100	24	27	21
500	108	113	100
800	183	211	170
1024	215	252	202

Table 1: Results of Ping-pong in ms



FRE 2661



L I F C

Jace : Current Extensions

UNIVERSITÉ DE FRANCHE-COMTÉ



- ① Hierarchical communication scheme :
 - access LANs or clusters with private addresses,
 - bypass firewall.
- ① Load balancing :
 - migrating user's tasks within a site,
 - "smart" spawner.



FRE 2661



L I F C

Parallel Power Method (1)

UNIVERSITÉ DE FRANCHE-COMTÉ



- ① Computes :
 - λ_1 : the eigenvalue of largest modulus,
 - its associated eigenvector.
- ① Assumptions :
 - $|\lambda_1|$ is a single eigenvalue,
 - p computing tasks,
 - $A_{n \times n}$ is cut in p horizontal strips.



FRE 2661



L I F C

Parallel Power Method (2)

UNIVERSITÉ DE FRANCHE-COMTÉ



receive A^{me} and x_0

$k = 1$

do

$$y_k^{me} = A^{me} \cdot x_{k-1}$$

send y_k^{me} to all other processors.

receive y_k^i for $i = 0, \dots, p - 1$ and $i \neq me$

$$\alpha_k = \|y_k\|$$

$$x_k = \frac{y_k}{\alpha_k}$$

test global convergence

receive stop / continue message.

$k+ = 1$

while no stop message is received



FRE 2661



L I F C

Experiments

UNIVERSITÉ DE FRANCHE-COMTÉ



- ① Framework :
 - 8 heterogeneous machines,
 - 3 distant sites,
 - heterogeneous bandwidth (ADSL, 10Mb, 100Mb).
- ① Tests :
 - sparse stochastic matrix of 4096 order.
 - comparison between synchronous and asynchronous versions,
 - increasing precision,
 - with and without static load balancing.



FRE 2661



L I F C

Experiments : no load balancing

UNIVERSITÉ DE FRANCHE-COMTÉ



precision	siac	aiac	ratio
10^{-3}	82 (18)	60 (43 - 239)	1.36
10^{-4}	136 (30)	92 (67 - 357)	1.47
10^{-5}	196 (42)	113 (82 - 429)	1.73
10^{-6}	255 (55)	132 (95 - 536)	1.93
10^{-7}	300 (67)	170 (116 - 641)	1.76

Table 2: Comparison between synchronous and asynchronous versions without load balancing.



FRE 2661



L I F C

Experiments : static load balancing

UNIVERSITÉ DE FRANCHE-COMTÉ



precision	siac	aiac	ratio
10^{-3}	86 (18)	28 (53 - 98)	3.07
10^{-4}	140 (30)	54 (84 - 203)	2.59
10^{-5}	196 (42)	74 (147 - 266)	2.64
10^{-6}	258 (55)	91 (174 - 313)	2.83
10^{-7}	316 (67)	115 (225 - 412)	2.74

Table 3: Comparison between synchronous and asynchronous versions with static load-balancing.



FRE 2661



L I F C

Experiments : Summary

UNIVERSITÉ DE FRANCHE-COMTÉ



- ① AIAC implies more iterations,
- ① AIAC is more efficient if synchronous version does a significant number of iterations,
- ① static load balancing has a real impact on AIAC,
- ① results greatly depend on the ratio computation/communication.



FRE 2661



L I F C

Conclusion

UNIVERSITÉ DE FRANCHE-COMTÉ



- ① Asynchronism + Java :
 - a new environment : Jace,
 - an innovating way to adapt scientific computation to the Grid,
 - portable codes,
 - efficient executions on heterogeneous multi-sites architectures.

- ① In the future:
 - a general study of AIAC algorithms using Jace,
 - extending Jace specifications to integrate classical grid services.



FRE 2661