

Grid enabling MGAC (Modified Genetic Algorithms for Crystal and Cluster Structure Prediction) using Digital Sherpa

Mardi Gras Conference 2008 - Workshop on Grid-Enabling Applications

Ronald C. Price, Victor E. Bazterra, Wayne Bradford, Julio C. Facelli*

Center for High Performance Computing and Department of Biomedical Informatics,
The University of Utah

Partially funded by NSF ITR award # 0326027 and TeraGrid allocation TG-PHY070034N



Why global optimizations are important

- **Crystal Structure Prediction**
 - Polymorphism
 - Crystal Engineering
 - There are no “intuitive” guess structures
- **Atomic Clusters**
 - Micro to mezzo scale structural transitions
 - Non “motif” structures are important

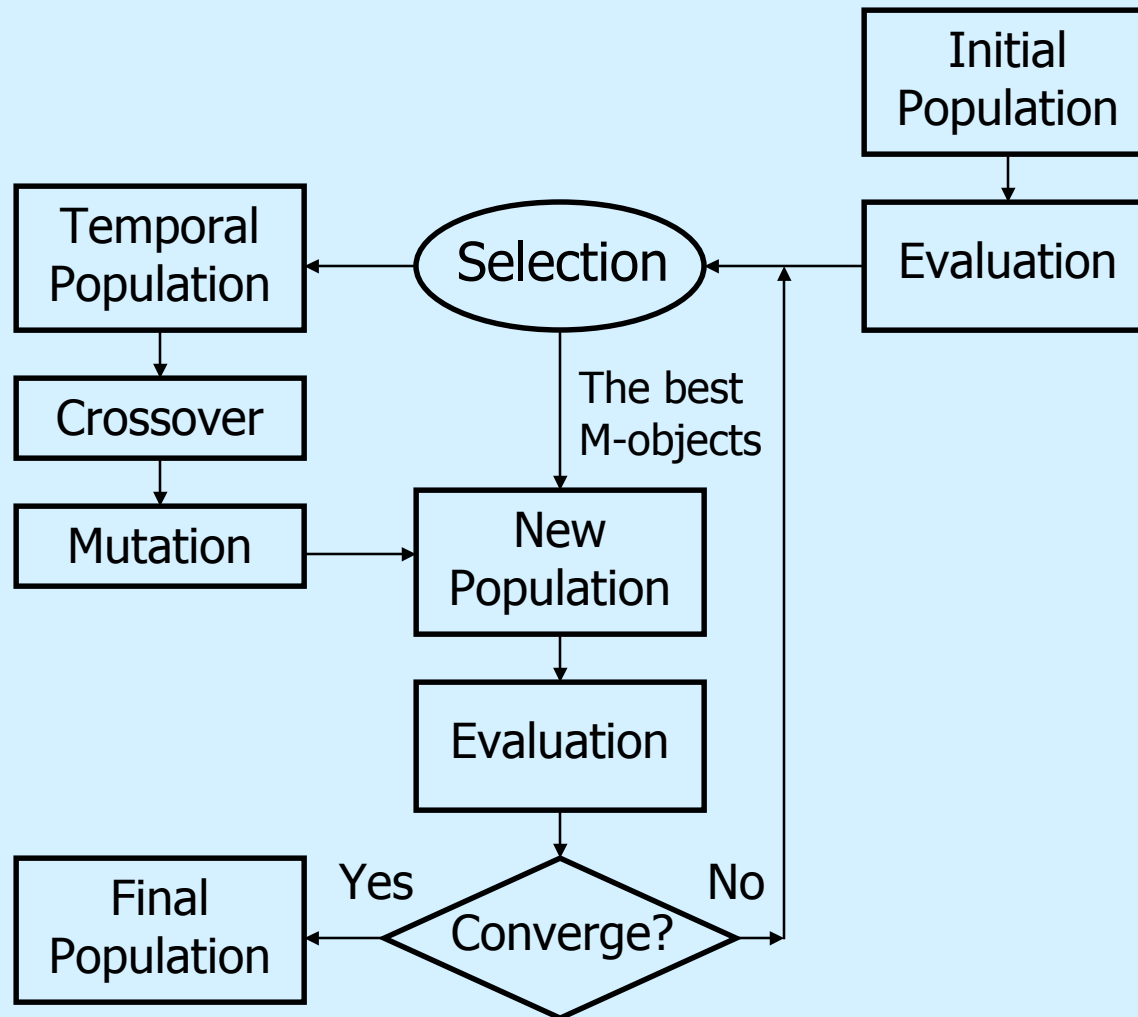
MGAC Crystal Structure Prediction Capabilities

Victor E. Bazterra, Matthew Thorley, Marta B. Ferraro, and Julio C. Facelli

J. Chem. Theory Comput. **2007**, 3, 201-209

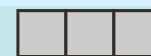
- Search for crystal structures within **any symmetry group** and with an **arbitrary number of molecules** and molecular types per asymmetric unit.
- Search structures using either the **rigid or flexible** molecule models.
- **Automatically** generate the molecule's **force field** using existing force field libraries.
- Increase the sampling power and the complexity of molecules amenable to CSP studies using the **parallel and distributed** computing capabilities of the system.
- **Automatically compare, sort and archive** the most relevant structures in a user database.

Genetic Algorithms



Molecular crystal genome

Unit Cell Parameters



$\vec{a}, \vec{b}, \vec{c}$

Position of the molecules



$\vec{\rho}_1$

$\vec{\rho}_2$

Orientation of the molecules



$\vec{\phi}_1$

$\vec{\phi}_2$

Dihedral angles



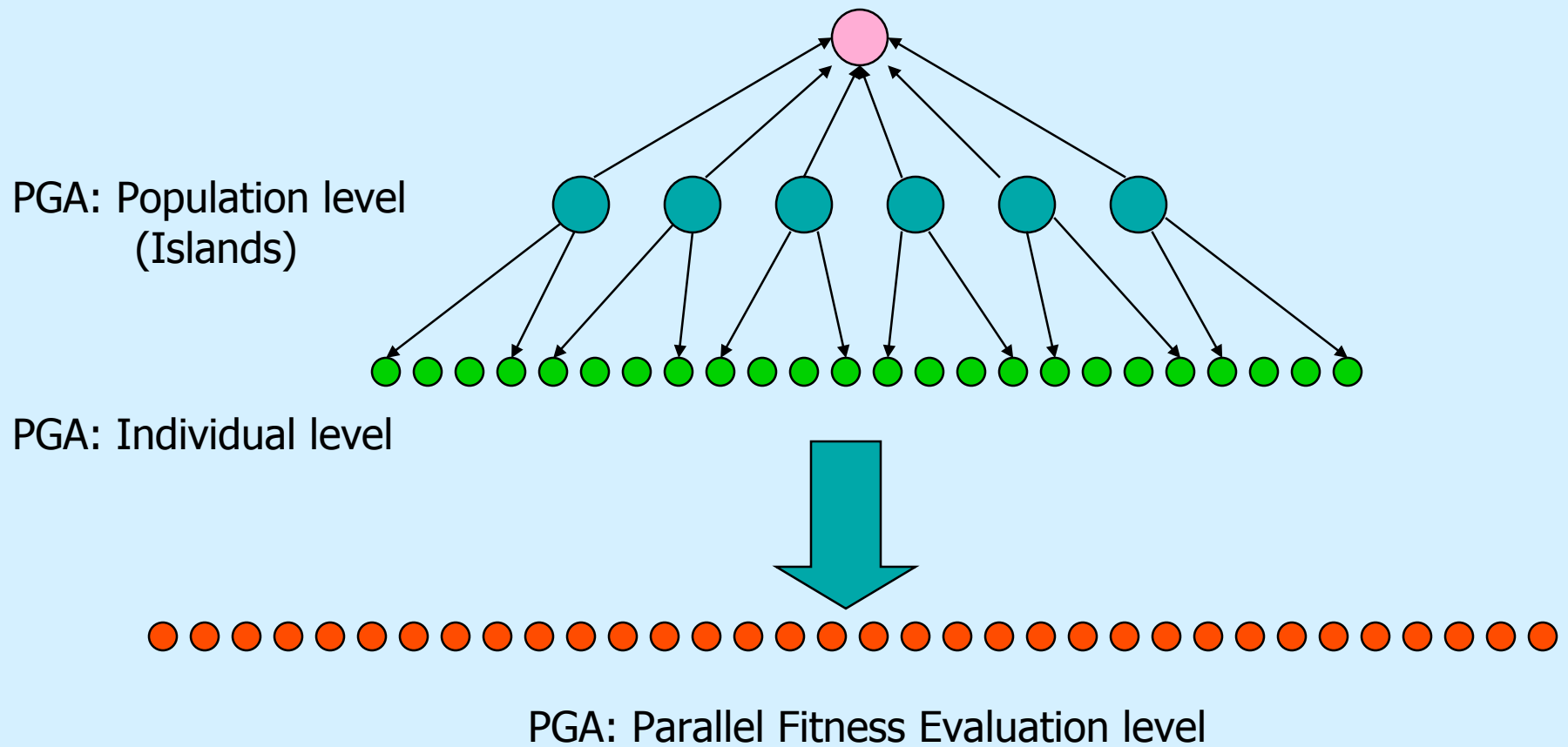
$\vec{\varphi}_1$

$\vec{\varphi}_2$

MGAC (Modified Genetic Algorithms for Crystals and Clusters)

- Local energy minimizations
 - Empirical potential methods: **CHARMM**
 - Semiempirical methods: **MSINDO**
 - DFT methods: **CPMD**
- Adaptive Parallel Genetic Algorithm (APGA)
 - V. E. Bazterra, M. Cuma, M. B. Ferraro and J. C. Facelli.
J. Parallel and Distributed Computing, 65, 48-57 (2005)

Parallelization in a GRID environment



Computational issues in MGAC

- **Very large number of runs**
- **Access to many different systems**
- **Load balancing issues due to stochastic initial conditions**
- Very large number of structures

BabySitter Scripts

- Resource manager and handler
- Proprietary states similar to the external states of the managed job services in WS-GRAM
- Not a general solution, scheduler specific

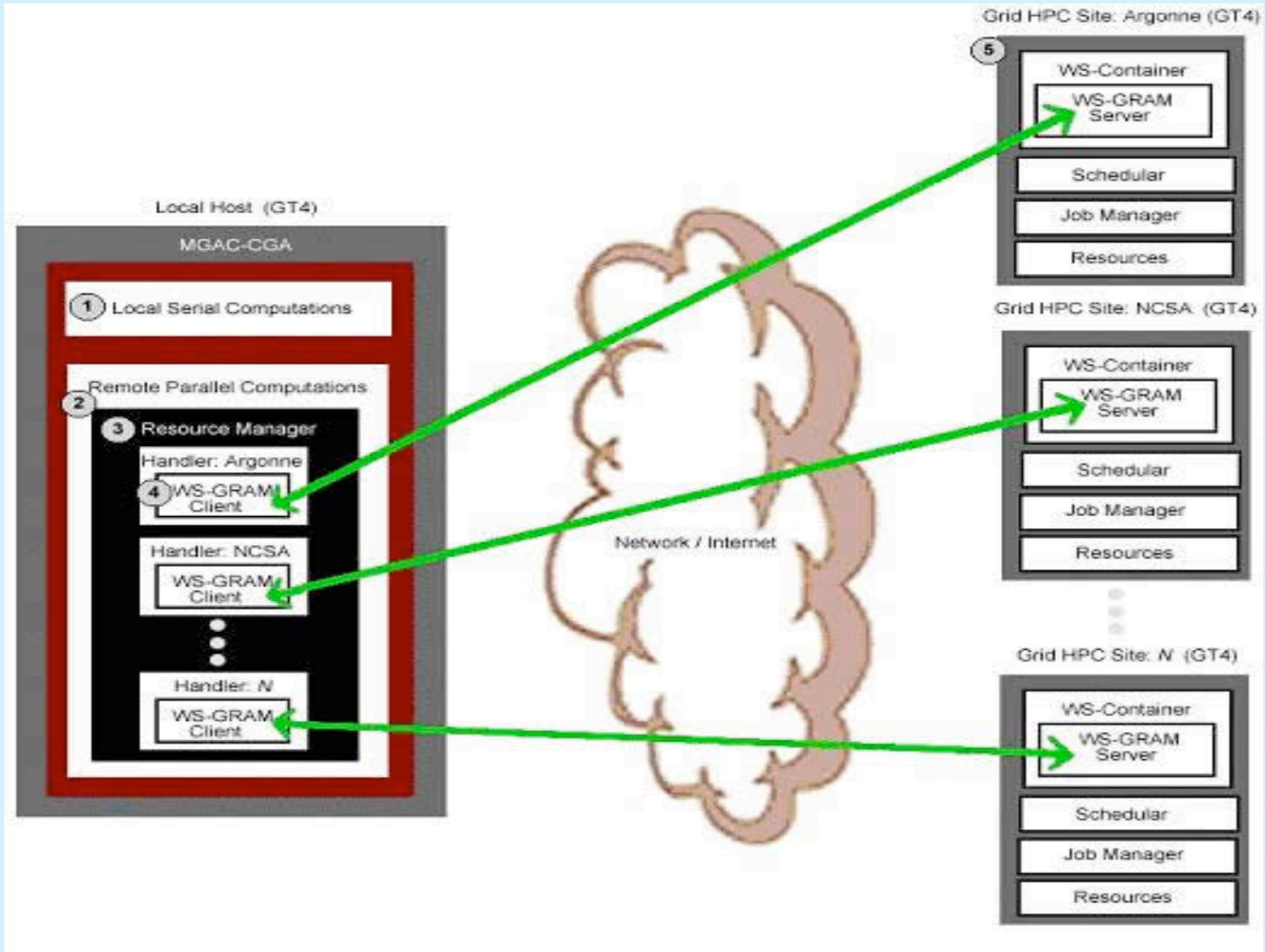
Workflows

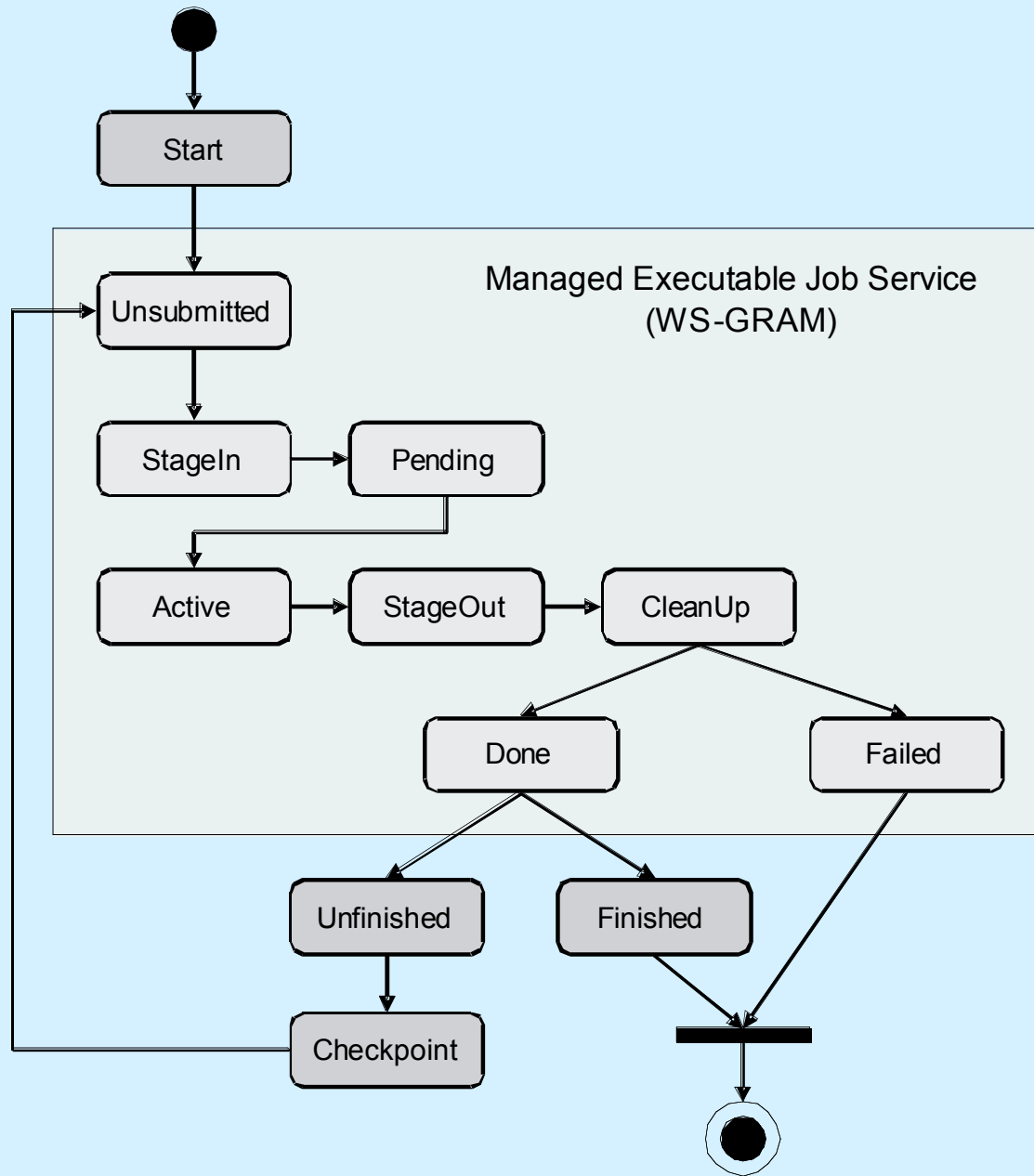
- Well establish technology in many applications
- No native to grid computing
- Several grid enabled, Kepler, Pegasus
- Difficult to use ?

Digital Sherpa (DS)

- Grid tool set:
 - Coordinate multiple job runs
 - Run jobs in multiple HPC resources
 - Monitor execution
 - Provides re-submission capabilities
 - Simpler than existing workflow tools

MGAC-CGA





Digital Sherpa Implementation

- GT4
 - WS-GRAM API
 - DS input uses WS-GRAM XML Schema
- WS-GRAM hides scheduler specific complexities

The “BLAH” Example: Test Jobs

- A test case for Sherpa: `***_blah.xml` corresponds to `***_blah.out` and `***_blahblah.xml` corresponds to `blahblah.out` ...
- Stage In:
 - Local `blahsrc.txt` -> remote server `blah.txt`
- Run:
 - `/bin/more blah.txt` (std out to: `blahtemp.out`)
- Stage Out:
 - Remote server `blahtemp.out` -> local `blah.out`
- Clean Up:
 - deletes `blahtemp.out` at remote HPC resource

“BLAH” on TeraGrid: Sherpa in Action

```
-bash-3.00$ java -DGLOBUS_LOCATION=$GLOBUS_LOCATION Sherpa argonne_blah.xml purdue_blahblahblah.xml
ncsamercury_blahblah.xml
Starting job in: argonne_blah.xml
Handler 1 Starting...argonne_blah.xml
Starting job in: purdue_blahblahblah.xml
Handler 2 Starting...purdue_blahblahblah.xml
Starting job in: ncsamercury_blahblah.xml
Handler 3 Starting...ncsamercury_blahblah.xml
Handler 3: StageIn
Handler 2: StageIn
Handler 1: StageIn
Handler 3: Pending
Handler 1: Pending
Handler 2: Pending
Handler 2: Active
Handler 2: StageOut
Handler 1: Active
Handler 2: CleanUp
Handler 2: Done
Handler 2 Complete.
Handler 3: Active
Handler 1: StageOut
Handler 3: StageOut
Handler 1: CleanUp
Handler 3: CleanUp
Handler 1: Done
Handler 1 Complete.
Handler 3: Done
Handler 3 Complete.
-bash-3.00$ hostname -f
watchman.chpc.utah.edu
```

Sherpa NCSA Mercury Results

-
blahblah.out
.....

Begin PBS Prologue Thu Apr 27 13:17:09 CDT 2006
Job ID: 612149.tg-master.ncsa.teragrid.org
Username: price
Group: oor
Nodes: tg-c421
End PBS Prologue Thu Apr 27 13:17:13 CDT 2006

BLAH BLAH

Begin PBS Epilogue Thu Apr 27 13:17:20 CDT 2006
Job ID: 612149.tg-master.ncsa.teragrid.org
Username: price
Group: oor
Job Name: STDIN
Session: 4042
Limits: ncpus=1,nodes=1,walltime=00:10:00
Resources: cput=00:00:01,mem=0kb,vmem=0kb,walltime=00:00:06
Queue: dque
Account: mud
Nodes: tg-c421

Killing leftovers...

End PBS Epilogue Thu Apr 27 13:17:24 CDT 2006

Sherpa UC/ANL Test Results

-
blah.out
.....

Begin PBS Prologue Thu Apr 27 13:16:53 CDT 2006
Job ID: 251168.tg-master.uc.teragrid.org
Username: rprice
Group: allocate
Nodes: tg-c061
End PBS Prologue Thu Apr 27 13:16:54 CDT 2006

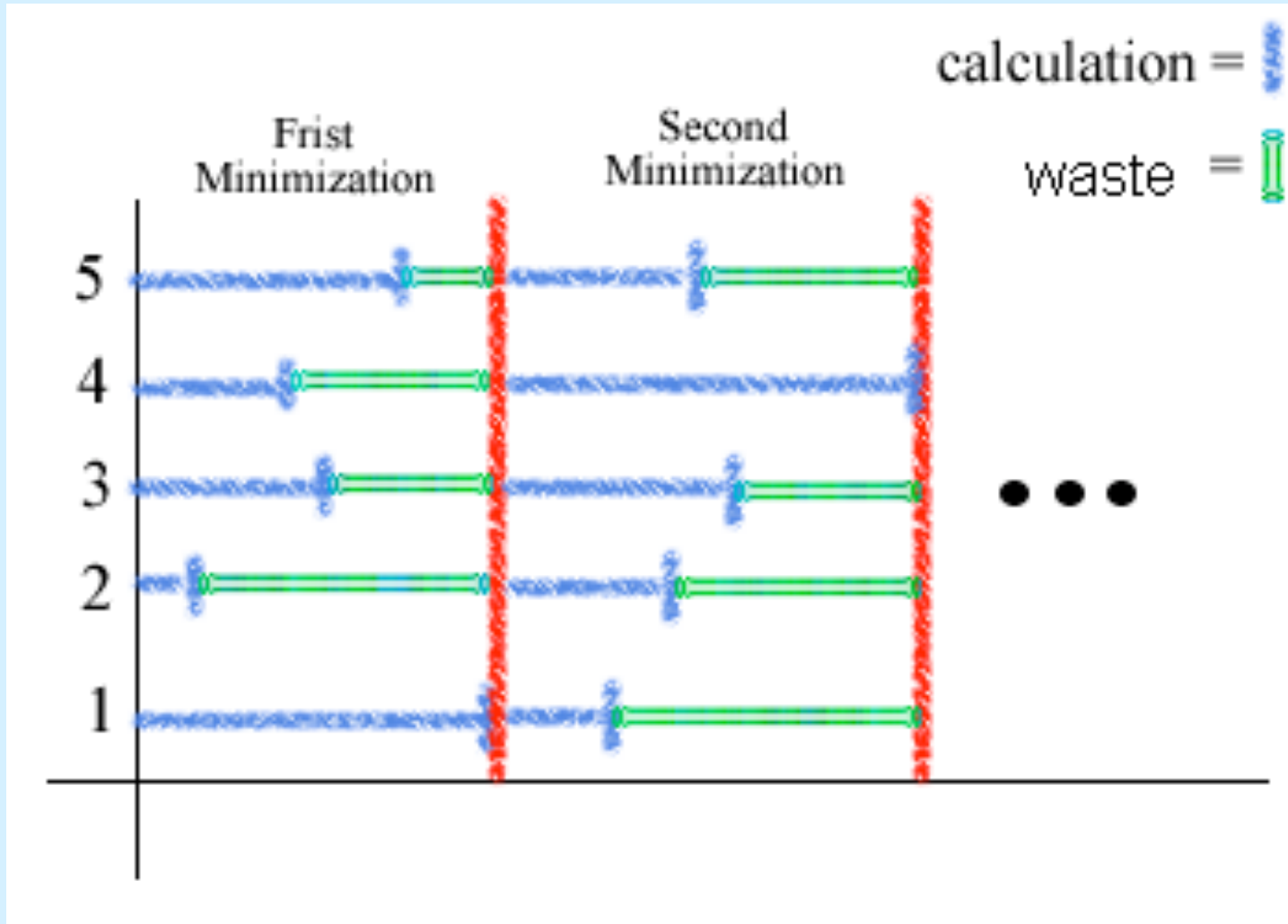
BLAH

Begin PBS Epilogue Thu Apr 27 13:17:00 CDT 2006
Job ID: 251168.tg-master.uc.teragrid.org
Username: rprice
Group: allocate
Job Name: STDIN
Session: 11367
Limits: nodes=1,walltime=00:15:00
Resources: cput=00:00:01,mem=0kb,vmem=0kb,walltime=00:00:02
Queue: dqe
Account: TG-MCA01S027
Nodes: tg-c061

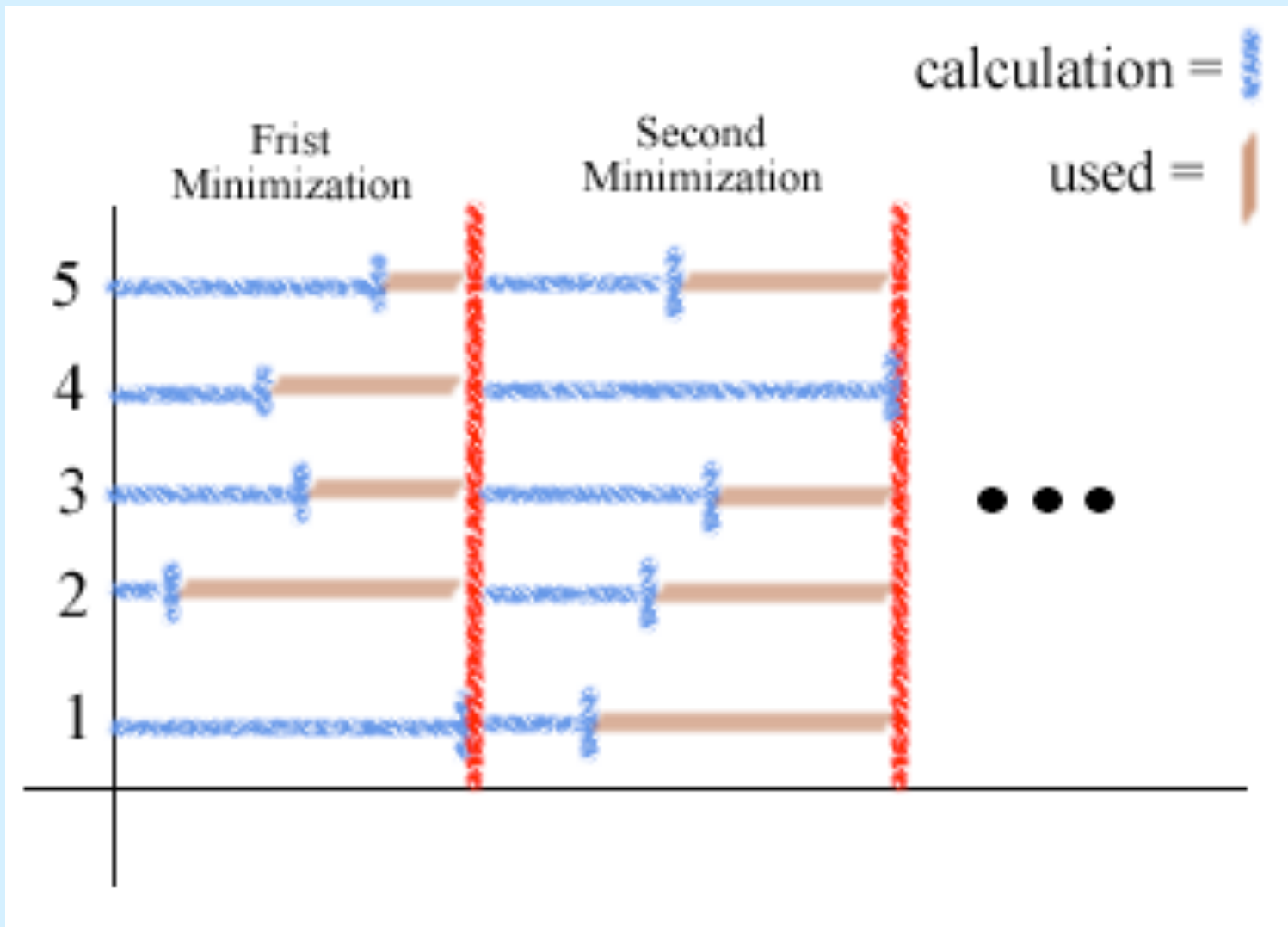
Killing leftovers...

End PBS Epilogue Thu Apr 27 13:17:16 CDT 2006

Single Job MGAC



MGAC-CGA



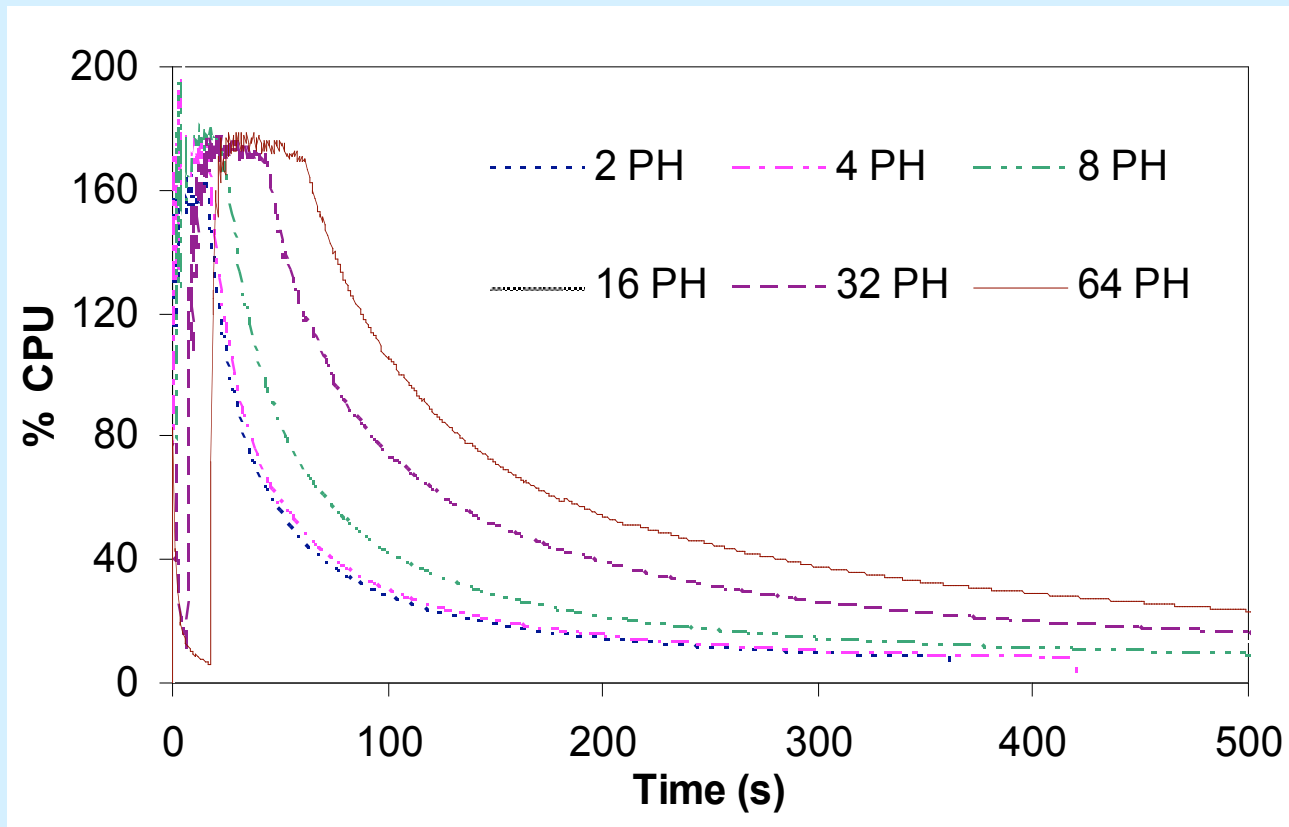
Efficiency and HPC Resources

- Our average MGAC run efficiency due to random initial conditions effects is: 46%
- Sherpa continuously submits one job at a time which reduces the scheduler side effect because multiple schedulers are involved and jobs are submitted in a more granular fashion
 - Improved Efficiency #1: increased granularity
- Necessary sharing policies prohibit large number of jobs from being submitted all at one HPC resource, queue times become too long
 - Improved Efficiency #2: access to more resources
- Total computational time (including queue time) reduced by 89%-60% in our initial testing.

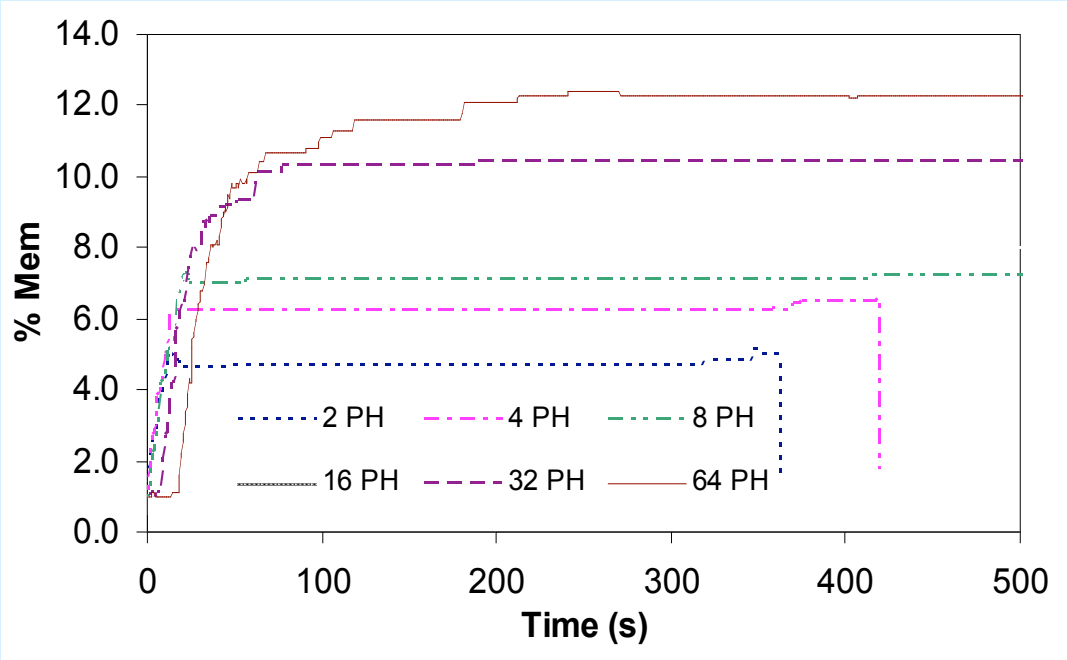
Sherpa Performance and Load Capability

- Performance:
 - DS is light weight because computationally intensive operations are done at HPC resource
 - Memory intensive
- Load Capability:
 - Successfully initiated 500 jobs
 - Emphasis on initiated, 500 jobs appeared in the test queue and although many ran to completion we did not have time to let them all run to completion

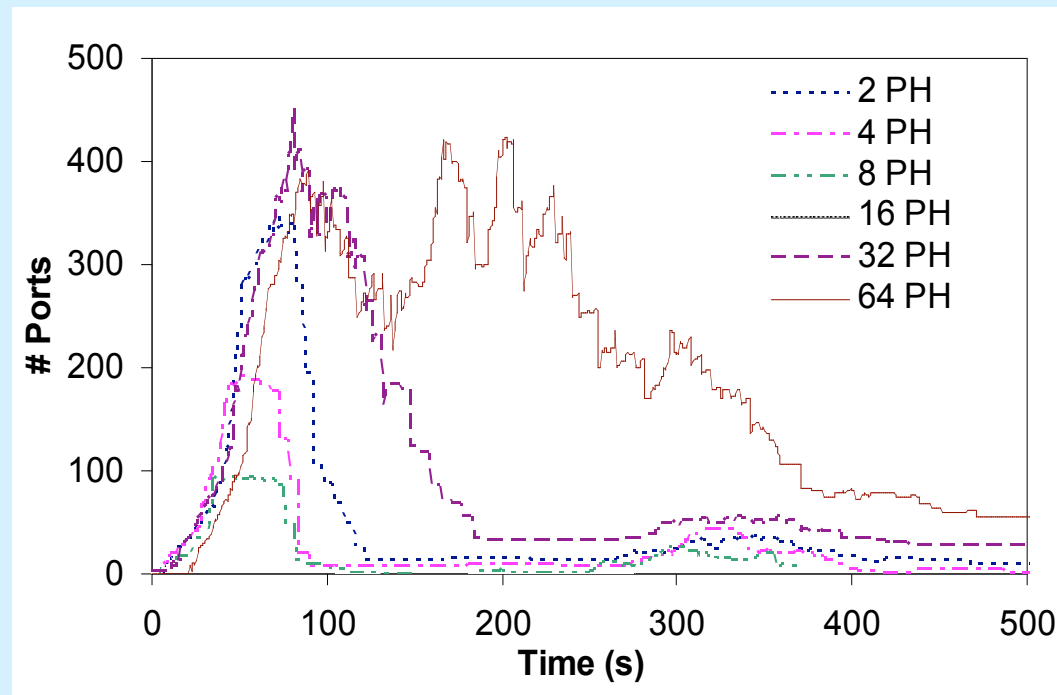
CPU usage for the first 500 s of execution of Digital Sherpa (DS) runs with different number of processor handlers (PH)



Percentage of memory usage for the first 500 s of execution of Digital Sherpa (DS) runs with different number of processor handlers (PH)



Number of ports used for the first 500 s of execution of Digital Sherpa (DS) runs with different number of processor handlers (PH)



Host Cert and Digital Sherpa

- Globus GSI:
 - Uses PKI to verify that users and hosts are who they claim to be, creates trust
 - User certs and host certs are different and they provide different functionality
- Digital Sherpa requires a host certificate
- ORNL granted us one but was revoked due policy changes
- NSF TeraGrid has difficulties in dealing with host certs

TeraGrid CA Caveats

- How do you allow your machines to fully interoperate with the TeraGrid without a host cert from a trusted CA?
- How do you get a host cert for the TeraGrid?
 - From least scalable to most scalable:
 - Work with site specific orgs to accept your CA's certs. (tedious for multiple sites)
 - Get TeraGrid security working groups approval for Local University CA (time consuming, not EDU scalable)
 - Get a TeraGrid trusted CA to issue you one. (unlikely as site policy seems to contradict this)
 - Become a TG member
 - Side Note: A satisfactory scalable solution does not seem to be currently in place and it's our understanding that Shibboleth and/or International Grid Trust Federation (IGTF) will eventually offer this service for EDU's in the future.

Not the End: Sherpa is Flexible

- Sherpa can work between any two machines that have GT4 installed and configured:
 - Flexible
 - Can work in many locations
 - Implicitly follows open standards

Future Features and Implementation

- Future efforts will be directed towards:
 - implementing monitoring and discovery client logic
 - polling feature that will help identify when system related issues have occurred (i.e. network down, scheduler unavailable)
- Implementation (move to a more general API)