Stay Tuned
Computational Science Team @ NeSI

Jordi Blasco
(jordi.blasco@nesi.org.nz)
Outline

1. About NeSI CS Team
   Who we are?
2. Identify the Bottlenecks
   Identify the Most Popular Apps
   Profile and Debug
3. Tuning
   Increase the performance
   Increase the scalability
   Increase the efficiency
4. Back to the future
   Moore’s Law
   Sequencing cost
   Comparing results
About NeSI CS Team

Computational Science Team

• NeSI’s team support researchers to get the most out of our platforms and services.
• The CS Team has a lot of experience in HPC that spans many science domains.
• Collaboratively enhance the performance of research software codes.
  • Troubleshoot memory and other or I/O bottlenecks.
  • Connect researchers and scientific software experts.
  • The team is available to support researchers across any research institution in New Zealand.
• More information at
About NeSI CS Team

Computational Science Team

Basically, it means that we are the Researchers best friends :-)

Jordi Blasco (jordi.blasco@nesi.org.nz) | Stay Tuned
Detect the Most Popular Apps

Top users list

- Previously based on the top user list
- Now we are using snoopy
Identify the bottlenecks

We use several tools for instrument and profile the code in order to identify potential bottlenecks and explore potential tuning opportunities.
Identify the bottlenecks with Intel Trace Analyzer
Identify the bottlenecks with Intel Vtune Amplifier
Profile and Debug

Identify the bottlenecks with ParaProf (new)
Profile and Debug

Available software

- Intel Vtune Amplifier
- Intel Trace Analyzer
- DDT
- Score-P
- HPC Toolkit
- Scalasca
- Cube
- PAPI
- TAU
- Parallel Profile Visualization (ParaProf)
There are several ways to tune an HPC Application

- Most obvious: tune the algorithm.
- Choose the right Libraries + Compilers + MPI "Flavour".
- Choose the right Options and Environment.
- Work in the work-flow.
- Explore the scalability (i.e. MPI Collectives).
- Check if your Benchmark results are good enough.
PhyML Case Study

**PhyML** is a software that estimates maximum likelihood phylogenies from alignments of nucleotide or amino acid sequences. The main strength of PhyML lies in the large number of substitution models coupled to various options to search the space of phylogenetic tree topologies, going from very fast and efficient methods to slower but generally more accurate approaches. The right compilers and optimization options for an specific architecture can increase the performance quite a lot!
Increase the performance

PhyML Case Study: Speed Up

PhyML-20120412 Nucleic results for Intel SandyBridge (E5-2680)

Jordi Blasco (jordi.blasco@nesi.org.nz)

Stay Tuned
Increase the performance

PhyML Case Study: Performance

PhyML-20120412 Nucleic results for Intel SandyBridge (E5-2680)

Performance improvement (%) vs. Number of cores

- GCC+ompi + IB
- ICS 2013 + GB
- ICS 2013 + IB
- ICS 2013 + OPT
Increase the scalability

Theoretical limit : Amdahl’s Law

- There are theoretical limits but the reality can be really surprising!
- The benchmarks can help to discover the real scalability limits.
- With this information you can get the results faster and save computational resources for other jobs.
Increase the scalability

![Amdahl's Law Diagram]

- Parallel Portion
  - 50%
  - 75%
  - 90%
  - 95%

Number of Processors vs. Speedup
Increase the scalability

Migrate-n Case Study

Migrate estimates effective population sizes and past migration rates between n population assuming a migration matrix model with asymmetric migration rates and different subpopulation sizes.
Increase the scalability

Migrate-n Case Study

Sarah J. Knight Benchmark results for Intel SandyBridge (E5-2680)

SpeedUp (x)

Number of cores

1 8 16 32 48 64 80 96 112

Intel Cluster Studio XE 2013
Increase the efficiency

Your knowledge really matters

The work-flow can save many HPC resources and in some cases you can achieve a linear scalability.
NCBI-BLAST+ Case Study

BLAST (Basic Local Alignment Search Tool) command line applications developed at the National Center for Biotechnology Information (NCBI).
Increase the efficiency

NCBI-BLAST+ Case Study: Speed Up

![Graph showing speedup for NCBI-BLAST+/2.2.28 Benchmark 2013 results for Intel SandyBridge (E5-2680)]
Increase the efficiency

NCBI-BLAST+ Case Study: Efficiency

NCBI-BLAST+/2.2.28 Benchmark 2013 results for Intel SandyBridge (E5-2680)

Number of threads

Jobs/h

1 4 8 16

GNU
Intel Cluster Studio 2013 + Tuning

Jordi Blasco (jordi.blasco@nesi.org.nz)
Stay Tuned
Back to the future

Moore’s Law

Figure 1. Transistors, frequency, power, performance, and processor cores over time. The original Moore’s law projection of increasing transistors per chip remains unabated even as performance has stalled.

Jordi Blasco (jordi.blasco@nesi.org.nz)  
Stay Tuned
Back to the future

Sequencing Cost per Genome

Cost per Genome

$100M
$10M
$1M
$100K
$10K
$1K


Moore's Law

National Human Genome Research Institute
genome.gov/sequencingcosts

Jordi Blasco (jordi.blasco@nesi.org.nz) Stay Tuned
Back to the future
Back to the future

Sequencing Cost per Genome

Cost per Genome

NIH National Human Genome Research Institute
genome.gov/sequencingcosts
Back to the future

Are you still living in the 80's?
Comparing results

Is your performance as good as you expect?

APOA1 Benchmark results for Intel SandyBridge (E5-2680)

Share your benchmark results to the community.

Jordi Blasco (jordi.blasco@nesi.org.nz)

Stay Tuned
Questions & Answers