Trends in Computing Power used by Various BOINC Communities

Yuri Gordienko
G.V.Kurdyumov Institute for Metal Physics, National Academy of Sciences (Kyiv, Ukraine)

10th BOINC Workshop – 30 September 2014 – Budapest

Under partial support and in cooperation with EU FP7 SCI-BUS project, contract no. RI-283481
Moore’s Law: CPU transistors versus dates of introduction. The line corresponds to exponential growth with transistor count doubling every two years.
Kurzweil's extension of Moore's law: calculations per second versus time - from integrated circuits to earlier transistors, vacuum tubes, relays and electromechanical computers.
Let’s Measure Progress of Computing Power in BOINC Projects

Motivation:
- Can we measure the actual (not theoretical like in Moore’s law) increase of global computing power using ...
  ... BOINC projects are some subsets of the global PC computing community?
- Can we consider them as “statistically representative samples”?

Aims:
- to explore the progress of computing power in BOINC projects
- to compare it among various BOINC projects
- to check and extend Moore’s Law?

Methods:
- use available open statistics on hosts in BOINC-projects
- use closed statistics on actually working machines (not available at the moment)
Data Used

• FLOPs (FPOPs) – the FLoating Point OPerations per second
  "p_flops"
• IOPs – the Integer OPerations per second
  "p_iops"
• CPUs – the number of CPUs in the host
  "ncpus"
  ...
  ...
• Others like:
  “p_membw”
• or
  “d_total”
• and more ...
Example 1:
SLinCA@Home IMP Desktop Grid

size: small (even **nano**)
history: short (2009-2014-...)
hosts: CPU
Scaling Laws in Cluster Aggregation - SLinCA@Home

Our team in G.V.Kurdyumov Institute for Metal Physics (IMP), National Academy of Sciences (Kyiv, Ukraine) maintains DG BOINC infrastructure on the basis of BOINC SZTAKI Desktop Grid technology at the premises of SLinCA@Home IMP Desktop Grid (http://dg.imp.kiev.ua/slinca).
SLinCA@Home – Overview

The current status of IMP Desktop Grid infrastructure:
- ~4000 workers;
- ~20 000 in-progress workunits;
- ~150 GFLOPs (average performance)
- ~300-550 GFLOPs (weekly peak performance) and 1.3 TFLOPs (max).

Typical performance timeline

Number of in-progress workunits
SLinCA@Home – Volunteer Community

IMP team mostly cooperate with a national-wide Ukrainian user community on the premises of the public site “Distributed Computing Team ‘Ukraine’” (http://distributed.org.ua), discussed and contributed the best practices in DCI operations at their special fora.
Use Case 1: Stress relaxation phenomena in Al/Cu/Si crystals

Application: LAMMPS-over-DCI (LAMMPS wrapped in DC-API by SZTAKI)

Science Community: physics, materials science – 2 user groups (6 end users): IMP + KNU (Taras Shevchenko Kyiv National University, Kyiv, Ukraine).

Scientific Aim: MD simulation of relaxation behavior of stresses in nanocrystals.

The evolution of the defect substructure in Al nanocrystal (after 15 ps, defects only are shown: gray color — point defects, red color — stacking faults).

Oscillation of the internal stress $P_{xx}$ for different nanocrystal sizes.
Use Case 2: Nanoindentation of graphene membrane

Application: LAMMPS-over-DCI (LAMMPS wrapped in DC-API by SZTAKI)

Science Community: nanotechnologies – 2 user groups (5 end users): IMP + SPM&RS-Centre (Centre of scanning probe microscopy and resonance spectroscopy, Kyiv, Ukraine).

Scientific Aim: MD simulation of nanoindentation was performed for monolayer graphene membrane in an atomic force microscope.

Visualization of nanoindentation: red color — mobile graphene atoms, yellow color — fixed graphene atoms, gray color – nanoindentor atoms.

Stress-strain dependencies for different speeds (A/pm) of the nanoindentor.
Use Case 3: Thermal stability of boron nitride nanotubes

Application: LAMMPS-over-DCI (LAMMPS wrapped in DC-API by SZTAKI)

Science Community: nanotechnologies, materials science – 2 user groups (4 end users): IMP + IPMS (Frantsevich Institute for Problems in Materials Science, Kiev, Ukraine).

Scientific Aim: MD simulation of boron nitride nanotubes (BNNT) with exceptional physical properties, which are a prerequisite for their wide practical applications in the future.

example of boron nitride nanotube (TEM-image)³

Collapse of NTNB (after 0.1 ns, nitrogen atoms are shown by red color, boron atoms — by blue color).

Dependence of decay temperature on the vacancies concentration and location (cap or side) in BNNT
Detachment of m-CNTs after application of driving force per atom $F=0.17 \text{ eV}/\text{A}$ and usage of the second Si-substrate ("stamp") in the presence of s-CNTs: two m-CNT c(6,6); two s-CNT c(7,5), two s-CNT c(9,2), and two m-CNT c(10,0) (from left to right).
Publications


8. Application of the Science Gateway Portal on the Basis of WS-PGRADE Technology for Simulation of Aggregation Kinetics and Molecular Dynamics Simulations of Metal-Organic Nanostructure, International Workshop on Science Gateways (Zurich, Switzerland), 03-05 June, 2013, PDF.


Example 1:
SLinCA@Home IMP Desktop Grid
size: small (even nano)
history: short (2009-2014-...)
hosts: CPU

Let’s measure
its averaged (per host)
performance characteristics...
IMP SLinCA@Home DG - CPUs/host
IMP DG - Performance - IOPs

IOPs - Mean

Year

Mean (GIOPS)
IMP DG - Performance - FLOPs
Distribution of Performances

Floating Point Operations Per second (FPOPs) in linear-probabilistic coordinates among the registered hosts (black), the worked hosts (red), and the hosts used for LAMMPS-application (green)

It seems to be normal, BUT it is more close to Weibull!

And Weibull was observed also at Ibercivis BOINC project from private communication with Alejandro Rivero (in 2012), manager of Volunteer Computing at Ibercivis (2007 to 2011)

Distribution of CPUs per Host

It seems to be log-normal

CPUs per host in the logarithmic-probabilistic coordinates (right) among the registered hosts (black), the worked hosts (red), and the hosts used for LAMMPS-application (green)

Distribution of RAM sizes

It seems to be log-normal... at least for low values

RAM sizes for the registered hosts (black), the worked hosts (red), and the hosts used for LAMMPS-application (green)

Distribution of HDD sizes

It seems to be log-normal... at least for low values

HDD sizes for the registered hosts (black), the worked hosts (red), and the hosts used for LAMMPS-application (green)

Example 2:
SZTAKI Desktop Grid

size: large,
history: long,
hosts: CPU
About SZTAKI Desktop Grid

SZTAKI Desktop Grid is operated by the Laboratory of Parallel and Distributed Systems at the Hungarian Academy of Sciences, Budapest, Hungary.

The SZTAKI Desktop Grid and its applications are partly supported by the ongoing IDGF-SP project. The work leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n° RI-312297. The experts of the International Desktop Grid Federation provide further support for the SZTAKI Desktop Grid infrastructure, its applications, and its integration into the European Grid Infrastructure.

Join SZTAKI Desktop Grid

http://szdg.lpds.sztaki.hu/szdg/
SZTAKI DG - Performance - IOPs

IOPs - Mean

Year

Mean (GIOPS)
SZTAKI DG - Performance - FLOPs

- Mean (GFLOPs)
- Year
- Mean values range from 1.4 to 2.6
Example 3:
Ibercivis Desktop Grid

size: large,
history: long,
hosts: CPU
Ibercivis is a research project that uses Internet-connected computers to do research in physics, material science and biomedicine.

You can participate by downloading and running a free program on your computer.

Ibercivis is based at several institutes and universities, like Zaragoza, CETA-CIEMAT, CSIC, Coimbra. More info about the different projects at:

http://registro.ibercivis.es/
Ibercivis - CPUs/host

CPUs - Mean

<table>
<thead>
<tr>
<th>Year</th>
<th>Mean (CPUs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>2</td>
</tr>
<tr>
<td>2009</td>
<td>3</td>
</tr>
<tr>
<td>2010</td>
<td>4</td>
</tr>
<tr>
<td>2011</td>
<td>5</td>
</tr>
<tr>
<td>2012</td>
<td>6</td>
</tr>
<tr>
<td>2013</td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td></td>
</tr>
</tbody>
</table>
Ibercivis - Performance - IOPs
Ibercivis - Performance - FLOPs
Example 4:
EDGeS Desktop Grid

size: medium,
history: short,
hosts: CPU
EDGeS

Project Performance:
Number of users: 36764
Number of hosts: 28911
Estimated performance of last 48 hours: 1257.723 GFlop/s

About EDGeS@Home
The aim of the EDGeS@Home project is to support the execution of selected and validated scientific applications developed by the IDGF and EGI communities.

This umbrella BOINC project supports several communities with applications: AutoDock - Molecular docking simulations used by Chemists/Biologists, Biome - BGC and Biome-BGC MuSo is executed by scientists from BIOVEL, Riemann Zeta Research application supports the work of mathematicians from ELTE, GBAC is used by the WeNMR community. DSP is used for functional testing.

The EDGeS@Home Desktop Grid and its applications are partly supported by the ongoing IDGF-SP project. The work leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-

Join EDGeS@Home

- Прочтите наши правила и политику
- This project uses BOINC. If you're already running BOINC, select Attach to Project.
- If not, download BOINC.
- When prompted, enter http://home.edges-grid.eu/home/
- If you're running a command-line or pre-5.0 version of BOINC, create an account first.
- If you have any problems, get help here.

Proud participant of
International Desktop Grid Federation

http://home.edges-grid.eu/home/
EDGeS DG - CPUs/host

CPUs - Mean

Year

Mean (CPUs)
EDGeS DG - Performance - IOPs

IOPs - Mean

Year

Mean (GIOPS)
EDGeS DG - Performance - FLOPs
Example 5:
GPUGRID

size: medium,
history: medium,
hosts: GPU
GPUGRID

Do real science, at home.

http://www.gpugrid.net/
GPUGRID - Performance - IOPs

IOPs - Mean

Year

Mean (GIOps)
Summary
as to the examples
CPUs/host

- SZDG
- Ibercivis
- EDGeS
- GPUGRID
- IMP
## Progress of FPOPs and IOPs

<table>
<thead>
<tr>
<th>Community</th>
<th>Increase</th>
<th>Time Range</th>
<th>Speedup</th>
</tr>
</thead>
<tbody>
<tr>
<td>SZDG</td>
<td>1,27</td>
<td>8</td>
<td>0,159</td>
</tr>
<tr>
<td>EDGeS</td>
<td>0,58</td>
<td>4</td>
<td>0,145</td>
</tr>
<tr>
<td>Ibercivis</td>
<td>0,91</td>
<td>6</td>
<td>0,152</td>
</tr>
<tr>
<td>GPUGRID</td>
<td>0,46</td>
<td>4</td>
<td>0,115</td>
</tr>
</tbody>
</table>

**MEAN:** $0,14\pm0,02$ (GFPOPs/year)

<table>
<thead>
<tr>
<th>Community</th>
<th>Increase</th>
<th>Time Range</th>
<th>Speedup</th>
</tr>
</thead>
<tbody>
<tr>
<td>SZDG</td>
<td>5,2</td>
<td>8</td>
<td>0,65</td>
</tr>
<tr>
<td>EDGeS</td>
<td>3.36</td>
<td>4</td>
<td>0,84</td>
</tr>
<tr>
<td>Ibercivis</td>
<td>4,17</td>
<td>6</td>
<td>0,695</td>
</tr>
<tr>
<td>GPUGRID</td>
<td>3,96</td>
<td>5</td>
<td>0,792</td>
</tr>
</tbody>
</table>

**MEAN:** $0,74\pm0,09$ (GFPOPs/year)
Summary on Progress of Computing Power in BOINC Projects

- Several BOINC projects of different kinds (big-small, long-short, CPU-GPU, geography) demonstrate the similar progress of CPUs/host, FPOPs, IOPs, therefore they can be considered as some subsets of the global PC computing community.

- The longest (>6 years) and largest (>10000 hosts) BOINC projects are very similar in these aspects and can be considered as “statistically representative samples”

- The actual increase of global computing power (in IOPS/FPOPs) averaged over communities is very slow (0.74/0.14 per year) in comparison to Moore’s law (2x each 2 years)

- This metrics – let’s title it like BOINC Law (or Anderson’s 😊 Law) - can be considered as extension of Moore’s Law (?)

BUT

- These results are previous and should be confirmed by future ...
Future Actions – other (much bigger) communities

• e.g. ABC (big size >160 MB)

... then go to

... then go to

• WCG (huge size >2.5 GB)
• SETI (biggest size >4 GB, longest history >11 years)
• LHC (biggest hype, because of Higgs bozon)...
Future Actions – It should be better (much deeper) analysis

To take into account:

• actually **working** hosts,
• host subsets **for some applications** (low sense, but it can be interesting for load balancing, planning, scaling, ...),
• **weights** on credits earned (if it has sense?)
• “finer-grain” time series (from **years** to **months** and weeks?)
Future Actions – other (much wider) metrics

• Higher moment ($2^{nd}$ – std, $3^{rd}$ – skewness, $4^{th}$ - kurtosis, ...) analysis,

• Bootstrapping analysis,

• Scaling analysis
  • ...
Thank you for your attention!

Under partial support and in cooperation with EU FP7 SCI-BUS project, contract no. RI-283481