The Costs of HPC-Based Science in the Exascale Era

Prof. Dr. Thomas Ludwig

German Climate Computing Centre & University of Hamburg
Hamburg, Germany
ludwig@dkrz.de
Many science fields base their knowledge gaining process on high performance computing. Constant exponential increase in performance allows in particular natural sciences to run more and more sophisticated numerical simulations. However, one may wonder, does the quality of results correlate to the increase in costs? In particular with the advent of the Exascale era and with Big Data we are confronted with possibly prohibitive energy costs. In addition, our installations grow in size and we typically replace them every 4-6 years. The talk will analyze the cost-benefit ratio of HPC-based science and consider economic and ecological aspects. We will have a closer look onto different science fields and evaluate the impact of their research results on society.
Costs
Benefits
Quantifications
Optimizations
Conclusions
Costs

Benefits

Quantifications

Optimizations

Conclusions
Cost model for total cost of ownership (TCO)

• Investment cost
  – Computer hardware and software
  – Data center facility
  – ...

• Operational costs
  – Human resources (brainware)
  – Electricity
  – ...
Terascale and Petascale Era of Computing

= 

Megascale Era of Costs
Investment costs

- 2002: Earth Simulator (Yokohama): $600 million
- 2010: Tianhe-1A (Tjanin): $88 million
- 2011: K computer (Kobe): around $1 billion
- 2011: Sequoia (Livermore): $250 million
- 2012: SuperMUC (Munich): €135 million

- often including data center facility
- sometimes including power and/or power station
Costs in the Petascale Era...

Scalable Cluster-Computing
Operational costs: electricity

1 MW 24/7 for one year is 8,760,000 kWh/y

$0.1 per kWh results in $876,000 per year
Operational costs: electricity

- 2002: Earth Simulator (Yokohama): $600 million
  - 3 MW → $2.5 million/year
- 2010: Tianhe-1A (Tianjin): $88 million
  - 4 MW → $3.5 million/year
- 2011: K computer (Kobe): around $1 billion
  - 12 MW → $10 million/year
- 2011: Sequoia (Livermore): $250 million
  - 8 MW → $7 million/year
- 2012: SuperMUC (Munich): €135 million
  - 3 MW → €5 million/year
Exascale Era of Computing

= 

Gigascale Era of Costs
Research and development costs

- Exascale programs to build an Exaflops computer with Exabyte storage systems
- USA, Japan, Europe, China, Russia
  - multi-billion investment in R&D

Investment cost

- First EFLOPS-computer: $500-$1500 million

Operational costs

- 20 MW $20 million/year
Costs in the Exascale Era…

Scalable Cluster-Computing
Operational costs: electricity

1 MW 24/7 for one year is 8,760,000 kWh/y
20 MW 24/7 for one year is 175,200,000 kWh/y
Greenhouse Gas Equivalencies Calculator

UPDATED May 2011. New NYUP sub region and national average non-baseload emissions rates updated. See the revision history page for more details.

Did you ever wonder what reducing carbon dioxide (CO₂) emissions by 1 million metric tons means in everyday terms? The greenhouse gas equivalencies calculator can help you understand just that, translating abstract measurements into concrete terms you can understand, such as "equivalent to avoiding the carbon dioxide emissions of 183,000 cars annually."

This calculator may be useful in communicating your greenhouse gas reduction strategy, reduction targets, or other initiatives aimed at reducing greenhouse gas emissions.
1 kWh corresponds to 0.00069 metric tons of CO₂ (around 1.5 lb)
Equivalency Results

Click on the question mark ? link to read the explanation of that particular calculation.

The information you entered above is equivalent to one of the following statements:

- Annual greenhouse gas emissions from \( \underline{23,688} \) passenger vehicles (click for calculation)
- \( \underline{13,543,739} \) gallons of gasoline consumed
- \( \underline{280,954} \) barrels of oil consumed
- \( \underline{1,593} \) tanker trucks’ worth of gasoline
- \( \underline{15,064} \) homes for one year
• Costs of current HPC are in the range of Megadollars

• Costs of Exascale HPC will be in the range of Gigadollars
Costs

Benefits

Quantifications

Optimizations

Conclusions
HPC enhances theory and experiment

- Provides numerical simulation as a means of knowledge gaining
- Indispensable for modern science and engineering

HPC enables competitive science and engineering for its users
HPC and Science

- Climate research
  - Understand clouds
- Life sciences
  - Understand the brain and simulate it
  - Understand genes
- Physics
  - Understand the universe
  - Understand the smallest particles
- etc.
HPC and Engineering

• Automotive
  – Develop more efficient engines
  – Optimize tires
• Aviation
  – Develop safe and efficient airplanes
• Oil and gas industry
  – Reservoir detection
• etc.
Cooperation of Boing and ORNL
Boeing airplane design

- Model aeroelasticity
- Lighter composites for wing design and performance
- 11 physical wing designs for 787 Dreamliner
  - Instead of 77 physical wings for 767
  - Construction of real wings heavily reduced
  - Tremendous cost saving!
HPC in Science and Engineering

TOP500
June 2012

system share

Academia
Research
Industry
HPC in Science and Engineering...

TOP500
June 2012
performance share

Academia
Research
Industry

1995
2000
2005
2010

Vendor
Government
Industry
Classified
Research
Academic

Share
• HPC enables unprecedented science

• HPC enables unprecedented engineering

• HPC is a key factor to the development of industrialized societies
Costs

Benefits

Quantifications

Optimizations

Conclusions
Research Questions

• How can we quantify the costs?
• How can we quantify the benefits?
• How can we define a benefit-cost ratio?

• What are potential consequences...
  ... for academia?
  ... for industry?
  ... for society?
There is not much research available to answer these questions.

In fact: almost no research

Approach here:

• Show practical example
• Report on analytical approaches
• Show more examples 😊
DKRZ in Hamburg

Deutsches Klimarechenzentrum (DKRZ)

German Climate Computing Centre
IBM Power6 Computer System

- Rank 232 in TOP500/Nov12
- 8,064 cores, 115 TFLOPS Linpack
- 6PB disks
Sun StorageTek Tape Library

- 100 PB storage capacity
- 90 tape drives
- HPSS HSM system
Energy Costs at DKRZ

• 2 MW for computer, storage, cooling, building

• Annual budget for power > €2 million

• Currently we use certified renewable energy
  – i.e. CO$_2$ free energy
  – Otherwise ca. 10,000t CO$_2$/y
Energy Cost History at DKRZ

[Graph showing energy cost history with two lines: Peak Performance and Operating Costs, with data points from 1988 to 2010.]
Cost-Benefit Modell at DKRZ

€2 M

2 MW

IPCC

Deutsches Klimarecherzentrum
5th IPCC status report:

- German part uses ca. 30M corehours at DKRZ
- DKRZ offers ca. 60M corehours/y
- Energy costs for the German IPCC contribution: **ca. €1 m**
  - **9,000,000 kWh to solution** with DKRZ’s Blizzard system
  - 4,500 metric tons of CO₂ with regular German electricity

Climate researchers should predict the climate change...

... and not produce it!
TCO of DKRZ per year: approximately €16M

- €8M hardware
- €2M electricity
- €3M brainware

Publications per year: let’s assume 100

Mean price per publication: €160,000

+ costs for scientists 😊

It is tax money – society expects a benefit
Finally: cloud computing

200km
Typical resolution of IPCC AR4 models

25km
Upper limit of climate models with cloud parameterizations

1km
Cloud system resolving models are a transformational change
1. Analytical Approach

Suzy Tichenor (Council of Competitiveness) and Albert Reuther (MIT Lincoln Laboratory)

Making the Business Case for High Performance Computing: A Benefit-Cost Analysis Methodology

CTWatchQuarterly, November 2006

• Boardrooms in U.S. industry see HPC only as a cost of doing business
• Try to quantify benefits and costs in academia and industry
• Give assistance to convince decision makers
Quantitative Approach

- Benefit-cost ratio BCR (bcr = benefit / cost)  
  [also: BCR = ROI / TCO]
- Internal rate of return IRR (IRR=BCR-1)
- Needs a collection of accurate data
- Evaluations conducted for one year periods
Quantitative approach

For research oriented organizations

\[
\text{productivity (BCR)} = \frac{(\text{time saved by users on system})}{(\text{time to parallelize}) + (\text{time to train}) + (\text{time to launch}) + (\text{time to administrate}) + (\text{system cost})}
\]

For industry environments

\[
\text{productivity (BCR)} = \frac{\sum (\text{Profit gained or maintained by project})}{(\text{Cost of software}) + (\text{Training cost}) + (\text{Admin cost}) + (\text{System cost})}
\]

(cf. Jeremy Kepner, MIT Lincoln Laboratory, HPCS Productivity Team member)
MIT Lincoln Laboratory: 600 processor cluster, 200 users, average full burdened salary of $200,000 per year

- 36,000 hours of user time saved
- Time to parallelize 200 user codes: 6,200 hours
- Total training time of 800 hours
- System administrator needs 2,000 hours per year
- HPC system costs $500,000 (equals 5,000 staff hours)
Example...

\[
BCR = \frac{[\text{Salary}] \times 36000}{[\text{Salary}] \times (6200 + 800 + 37.8 + 2000 + 5000)} = \frac{36000}{14028} = 2.6,
\]

\[
\text{IRR}_{1\text{ year}} = BCR - 1 = 1.6 = 160\%.
\]

Saved time for all the 200 users

Typical chancellor: “Why save time for scientist?—they get payed anyway!”

(Why pay for taxis when there are busses?)
Benefit-Cost Example in Industry

- $260 M
- 260 MW
- $8,500 M
IRR (=BCR-1)

Academia
HR-bound

Industry
market-bound

costs
2. Analytical Approach

Amy Apon (University of Arkansas),
Stan Ahalt (University of North Carolina) et al.

High Performance Computing Instrumentation and Research Productivity in U.S. Universities

Journal of Information Technology Impact, Vol. 10/2, 2010

• Research institutes with powerful HPC systems are more successful with their science

• Results are economically and statistically significant
Apon/Ahalt study the following variables

- dRankSum  Sum of derived ranks (500...1)
- Counts    #lists in which institution appeared
- NSF       Sum of NSF funding for institution
- Pubs      Sum of publications
- FF        Sum of federal funding
- DOE       Sum of DOE funding
- DOD       Sum of DOD funding
- NIH       Sum of NIH funding
- USNews    US News and World Report ranking
## Correlation Analysis

<table>
<thead>
<tr>
<th></th>
<th>Counts</th>
<th>NSF</th>
<th>Pubs</th>
<th>All Fed</th>
<th>DOE</th>
<th>DOD</th>
<th>NIH</th>
<th>USNews</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>dRankSum</strong></td>
<td>0.8198</td>
<td>0.6545</td>
<td>0.2643</td>
<td>0.2566</td>
<td>0.2339</td>
<td>0.1418</td>
<td>0.1194</td>
<td>-0.243</td>
</tr>
<tr>
<td><strong>Counts</strong></td>
<td>0.6746</td>
<td>0.4088</td>
<td></td>
<td>0.3601</td>
<td>0.3486</td>
<td>0.1931</td>
<td>0.2022</td>
<td>-0.339</td>
</tr>
<tr>
<td><strong>NSF</strong></td>
<td></td>
<td></td>
<td>0.7123</td>
<td>0.6542</td>
<td>0.5439</td>
<td>0.2685</td>
<td>0.4830</td>
<td>-0.540</td>
</tr>
<tr>
<td><strong>Pubs</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.8665</td>
<td>0.4846</td>
<td>0.3960</td>
<td>0.8218</td>
<td>-0.588</td>
</tr>
<tr>
<td><strong>All Fed</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.543</td>
</tr>
<tr>
<td><strong>DOE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.384</td>
</tr>
<tr>
<td><strong>DOD</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.252</td>
</tr>
<tr>
<td><strong>NIH</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.500</td>
</tr>
</tbody>
</table>

cf. slides by Apon, Ahalt on “Investment in High Performance Computing”
Authors test two relationships

- NSF funding as a function of contemporaneous and lagged appearance on the TOP500 list count and publication count
- Publication count as a function of contemporaneous and lagged appearance on the TOP500 list count and NSF funding

[endogeneity between Pubs and NFS was tested and corrected]
According to the authors

- An entry in a list results in an increase of yearly NSF funding of $2.4M
- An entry in a list results in an increase in yearly publications of 60
- Rank has a positive impact on competitiveness, but with reduced confidence
- HPC investments suffer from fast depreciation over a 2 year horizon
- Consistent investments in HPC, even at modest level [at least TOP500!], are strongly related to research competitiveness
Apon/Ahal’s work is a typical example for data driven science – not yet data intensive

- The Fourth Paradigm
- Combine existing data and derive new insight
- I would call it secondary level science
- We will see much more of it

This talk is third level science...
Quantification Summary

- Quantification is possible! 😊
- We need more research on quantification
- You can only control what you can measure

- Benefit is difficult to quantify
- It is not necessary to quantify benefit as it is always very high
Benefit Considerations

2 more examples
Higgs Boson
aka The God Particle

• Large Hadron Collider construction costs $4.75 billion
• Electricity costs per year $23.5 million
• Total operating budget per year of the LHC runs to about $1 billion

• Total costs of finding the Higgs Boson $13.25 billion
What have the Romans physicists ever done for us?
Costs

Benefits

Quantifications

Optimizations

Conclusions
Available money is often decided upon by politicians

Benefits of HPC are always very high

Question

Can we spend the financial resources more efficiently in order to have even higher benefit?
How to increase BCR?

General approach

– Increase benefit and/or **decrease costs**

In detail

– Invest in human resources *(use intellectual capital)*
– Tune programs (sequential and parallel)
– Increase application performance
– Thus increase scientific productiveness

Hardware, software, brainware
In detail: shift expenses and reduce costs

- Invest in human resources
- Tune programs (sequential and parallel)

  Costs measured in salary for person months

- Increase application performance

  Cost savings effectuated by energy savings

- Thus increase scientific productiveness

  Do more science with your (fixed) energy budget
Example IPCC AR5 production runs

• Remember

Energy costs for the German IPCC contribution: ca. 1 M€

– **9,000,000 kWh to solution** with DKRZ’s system

– 4,500,000 kg of CO₂ with regular German electricity

• Approach: Tune program and save 10% runtime

  • Saves 900,000 kWh

  • Saves €100,000 (is one person year)

  • Saves 450 metric tons CO₂
HECToR is the UK National Supercomputing Service

- dCSE programme has a focus helping users to improve their code
- There are many published success stories with quantifications

E.g.

- Oceanography code NEMO: better speed and I/O
  - 6 PMs effort, saves £96K per year
- Key materials science code CASTEP: 4x speed, 4x scalability
  - 8 PM effort, saves £320K - £480K per year
- Plus: protecting the environment
Invest in people!

We need more HPC specialists

- Co-design and code development
- Tuning of applications
- many other things...

Gigadollars for iron and electricity
will not be the solution!
Costs
Benefits
Quantifications
Optimizations

Conclusions
There is a proven positive correlation between costs and benefits for science and engineering.

BCR in science: most results are only possible just because of HPC
- Costs are investments in a better future
  - Therefore no cost calculation

BCR in industry: many products are only possible just because of HPC
- At the moment benefits exceed costs dramatically
  - Therefore no real cost calculation
**BUT**: With Exascale costs will be **much** higher!

And financial resources are always limited...

Therefore:

- Optimize the usage of your financial resources
  - measure – evaluate – optimize
- Use people and their intellectual capacities
- Invest in brainware – not just hardware/software

**Tell the story to your political representative**
Perhaps see you again at...

EnA-HPC 2013

Fourth International Conference on Energy-Aware High Performance Computing

September 2-3, 2013
Dresden, Germany

www.ena-hpc.org
References

• Suzy Tichenor, Albert Reuther: Making the Business Case for High Performance Computing: A Benefit-Cost Analysis Methodology
  In: CTWatchQuarterly, November 2006

• Amy Apon, Stan Ahalt et al.: High Performance Computing Instrumentation and Research Productivity in U.S. Universities