

REPRODUCIBILITY CRISIS, OPEN SCIENCE, AND COMPUTER SCIENCE

Arnaud Legrand



December 2020



PUBLIC EVIDENCE FOR A LACK OF REPRODUCIBILITY

- J.P. Ioannidis. *Why Most Published Research Findings Are False* PLoS Med. 2005.
- *Lies, Damned Lies, and Medical Science*, The Atlantic. Nov, 2010
- *Reproducibility: A tragedy of errors*, Nature, Feb 2016.
- Steen RG, *Retractions in the scientific literature: is the incidence of research fraud increasing?*, J. Med. Ethics 37, 2011

Los Angeles Times | BUSINESS

LOCAL U.S. WORLD BUSINESS SPORTS ENTERTAINMENT HEALTH STYLE TRAVEL

Science has lost its way, at a big cost to humanity

Researchers are rewarded for splashy findings, not for double-checking accuracy. So many scientists looking for cures to diseases have been building on ideas that aren't even true.

Science AAAS-ORG FEEDBACK HELP LIBRARAINS All Science Journals ▾ Issue Search ▾

AAAS NEWS SCIENCE JOURNALS CAREERS MULTIMEDIA COLLECTIONS

Science The World's Leading Journal of Original Scientific Research, Global News, and Commentary.

Science Home Current Issue Previous Issues Science Express Science Products My Science About the Journal

Home > SCIENCE Magazine > 12 JUNE 2014 > MONDAY, 30 MAY (160) : 229

Article Views Science 17 January 2014; Vol. 343 no. 6168 p. 229 DOI: 10.1126/science.1250475

Summary Full Text Full Text (PDF)

EDITORIAL

Reproducibility

Marcia McNutt

► Marcia McNutt is editor-in-chief of *Science*.

Science advances on a foundation of trusted data. But the lack of reproducibility in science is an approach that scientists used to gain confidence in their results. Now, the scientific community is shaken by reports that a trend toward irreproducibility is real. Because confidence in results is essential to progress, we are calling for new measures to combat this trend. For example, the National Institutes of Health has issued a set of recommendations of the U.S. National Institute of General Medical Sciences. Authors will indicate how they handled data (such as how to deal with outliers), whether they ensured a sufficient signal-to-noise ratio, whether the experimenter was blind to the conduct of the experiments, and more.

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

Rights & Permissions

Commercial Reprints and E-Prints

View Public Citation

Related Content

Save to My Folders

Download Citation

Alert Me When Article Is Cited

Post to CiteULike

E-mail This Page

</

NEWSWORTHY STORIES ABOUT SCIENTIFIC MISCONDUCT

Dong-Pyou Han Assistant professor, Biomedical sciences, Iowa State University, 2013

Falsified blood results to make it appear as though a vaccine exhibited anti-HIV activity

- Han and his team received \approx \$19 million from NIH
- 1 retracted publication and resignation of university. Sentenced in 2015 to 57 months imprisonment for fabricating and falsifying data in HIV vaccine trials He was also fined US \$7.2 million!

NEWSWORTHY STORIES ABOUT SCIENTIFIC MISCONDUCT

Dong-Pyou Han Assistant professor, Biomedical sciences, Iowa State University, 2013

Falsified blood results to make it appear as though a vaccine exhibited anti-HIV activity

- Han and his team received \approx \$19 million from NIH
- 1 retracted publication and resignation of university. Sentenced in 2015 to 57 months imprisonment for fabricating and falsifying data in HIV vaccine trials He was also fined US \$7.2 million!

Dieterik Stapel Professor, Social Psychology, Univ. Amsterdam, 2011

I failed as a scientist. I adapted research data and fabricated research. Not once, but several times, not for a short period, but over a longer period of time. [...] I am aware of the suffering and sorrow that I caused to my colleagues... I did not withstand the pressure to score, to publish, the pressure to get better in time. I wanted too much, too fast. In a system where there are few checks and balances, where people work alone, I took the wrong turn.

58 retracted publications

NEWSWORTHY STORIES ABOUT SCIENTIFIC MISCONDUCT

Dong-Pyou Han Assistant professor, Biomedical sciences, Iowa State University, 2013

Falsified blood results to make it appear as though a vaccine exhibited anti-HIV activity

- Han and his team received \approx \$19 million from NIH
- 1 retracted publication and resignation of university. Sentenced in 2015 to 57 months imprisonment for fabricating and falsifying data in HIV vaccine trials He was also fined US \$7.2 million!

Dieterik Stapel Professor, Social Psychology, Univ. Amsterdam, 2011

I failed as a scientist. I adapted research data and fabricated research. Not once, but several times, not for a short period, but over a longer period of time. [...] I am aware of the suffering and sorrow that I caused to my colleagues... I did not withstand the pressure to score, to publish, the pressure to get better in time. I wanted too much, too fast. In a system where there are few checks and balances, where people work alone, I took the wrong turn.

58 retracted publications

Brian Wansink Professor, Psychological Nutrition, Cornell, 2016

I gave her a data set of a self-funded, failed study which had null results. I said "This cost us a lot of time and our own money to collect. There's got to be something here we can salvage because it's a cool (rich & unique) data set." I told her what the analyses should be. [...] Every day she came back with puzzling new results, and every day we would scratch our heads, ask "Why," and come up with another way to reanalyze the data with yet another set of plausible hypotheses

17 retracted publications

SCIENTIFIC MISCONDUCT? WHAT ARE THE CONSEQUENCES ?

Bad science is deleterious

- It is used to backup stupid politics, it affects people's life, ...
- It blurs the frontier between scientists and crooks

Media attention inflates conspiracy opinions 🙄

- *Scientific result are worthless.*
- *Scientists cannot even agree with each others on economy/climate/vaccine/5G/...*
- *Stop the scientific dictatorship/lobby!*

A CREDIBILITY CRISIS?

Scientific misconduct is obviously wrong but it's **not new!**

- Every domain has its black sheep
- The publish or perish pressure is a huge pain

The Battle against Scientific Fraud

CNRS International Magazine

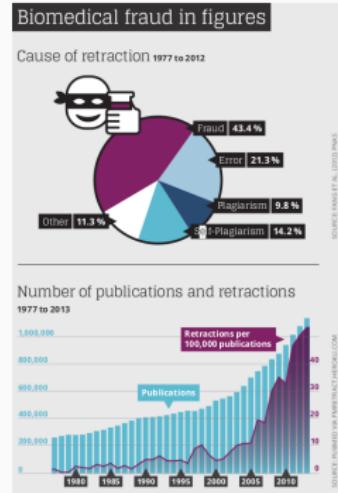


A CREDIBILITY CRISIS?

Scientific misconduct is obviously wrong but it's **not new!**

- Every domain has its black sheep
- The publish or perish pressure is a huge pain

Fraud is the (**uninteresting**) visible part of the iceberg



The Battle against Scientific Fraud

CNRS International Magazine



- **Failing to reproduce the results of others is common**

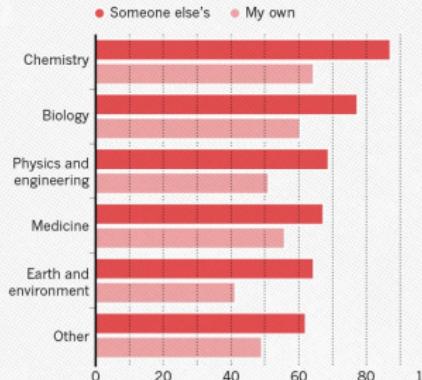
1,500 scientists lift the lid on reproducibility

Nature, May 2016

- How so? **Why now? Why is this important?** What can we do about it?

HAVE YOU FAILED TO REPRODUCE AN EXPERIMENT?

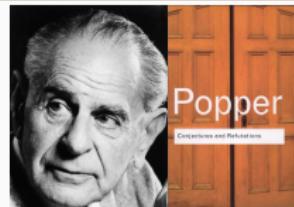
Most scientists have experienced failure to reproduce results



REPRODUCIBILITY OF EXPERIMENTAL RESULTS: THE HALLMARK OF SCIENCE

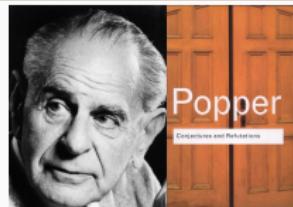
1934: Karl Popper puts the notions of **falsifiability** and **crucial experiment** as the **hallmark of science**

- If no experiment can be set up to **disprove** your theory, it is not science
- Good experiments **discriminate** good theories from bad ones
- **Non-reproducible** single occurrences are of no significance to science



REPRODUCIBILITY OF EXPERIMENTAL RESULTS: THE HALLMARK OF SCIENCE

1934: Karl Popper puts the notions of **falsifiability** and **crucial experiment** as the **hallmark of science**



- If no experiment can be set up to **disprove** your theory, it is not science
- Good experiments **discriminate good theories from bad ones**
- **Non-reproducible** single occurrences are of no significance to science

An ideal rather than the norm

Popper's proposal works well for Physics from the 18th century but is not so simple for many other domains:

- Theory of evolution
- Spotting a SuperNova
- Particle Physics (a single LHC)
- Biology (every animal does not behave in the same way)
- Anthropology (impact on people from a remote culture)

REPRODUCIBILITY: A CORE VALUE OF SCIENCE

1. Universality: Science aims for objective findings, accessible to anyone

Reproducibility acts as a Universality/Robustness control

2. Incremental: We build on each others work but everybody makes mistakes

Methods, biases, ... How to discriminate sound theories experiments from bad ones? 😊

Reproducibility acts as a Quality control

REPRODUCIBILITY: A CORE VALUE OF SCIENCE

1. Universality: Science aims for objective findings, accessible to anyone

Reproducibility acts as a Universality/Robustness control

2. Incremental: We build on each others work but everybody makes mistakes

Methods, biases, ... How to discriminate sound theories experiments from bad ones? 😊

Reproducibility acts as a Quality control

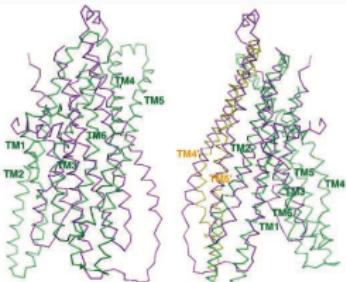
But, scientific practices have greatly evolved, in particular since we rely on computers



How computers broke science – and what we can do about it

– Ben Marwick, The conversation, 2015

How COMPUTERS BROKE SCIENCE



Geoffrey Chang (Scripps, UCSD) works on crystallography and studies the structure of cell membrane proteins.

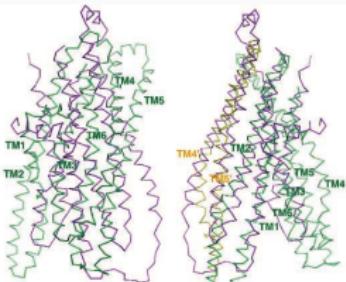
He specialized in structures of **multidrug resistant transporter proteins in bacteria**: MsbA de Escherichia Choli (Science, 2001), Vibrio cholera (Mol. Biology, 2003), Salmonella typhimurium (Science, 2005)

2006: Inconsistencies reveal **a programming mistake**

A homemade data-analysis program had flipped two columns of data, inverting the electron-density map from which his team had derived the protein structure.

5 retractions that motivate **improved software engineering practices** in comp. biology

How COMPUTERS BROKE SCIENCE



Geoffrey Chang (Scripps, UCSD) works on crystallography and studies the structure of cell membrane proteins.

He specialized in structures of **multidrug resistant transporter proteins in bacteria**: MsbA de Escherichia Choli (Science, 2001), Vibrio cholera (Mol. Biology, 2003), Salmonella typhimurium (Science, 2005)

2006: Inconsistencies reveal **a programming mistake**

A homemade data-analysis program had flipped two columns of data, inverting the electron-density map from which his team had derived the protein structure.

5 retractions that motivate **improved software engineering practices** in comp. biology

There is **worse!**

- The generalized and intensive use of **spreadsheets** (**COVID tracing**)
- Relying on **black box** statistical methods is infinitely easier than understanding them (**Learning and Data Analytics frameworks are nuclear weapons**)
- **Numerical errors** and **software environment unawareness**

DIFFERENT REPRODUCIBILITY CONCERN IN MODERN SCIENCE

Social Sciences, Oncology, ... methodology, statistics

Genomics software engineering, computational reproducibility, provenance

Computational fluid dynamics numerical issues

The processing steps between raw observations and findings have gotten increasingly numerous and complex

Authors



Data

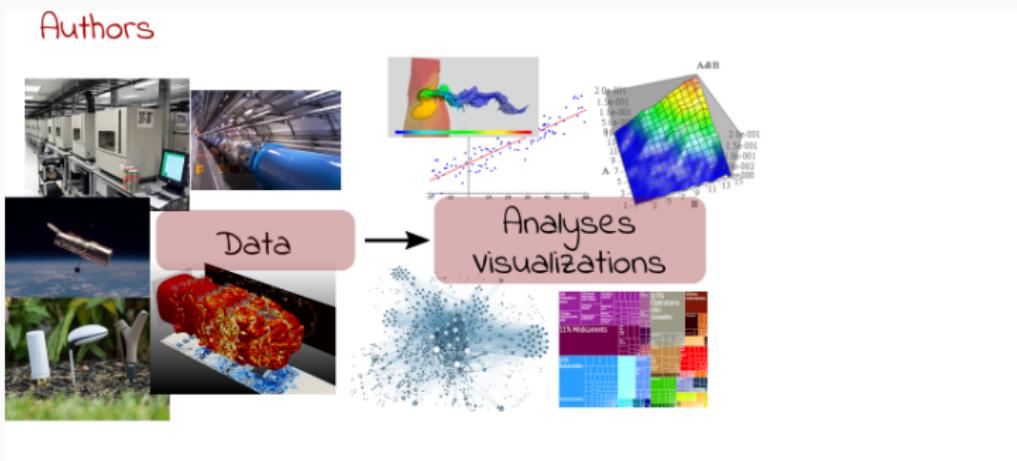
DIFFERENT REPRODUCIBILITY CONCERN IN MODERN SCIENCE

Social Sciences, Oncology, ... methodology, statistics

Genomics software engineering, computational reproducibility, provenance

Computational fluid dynamics numerical issues

The processing steps between raw observations and findings have gotten increasingly numerous and complex



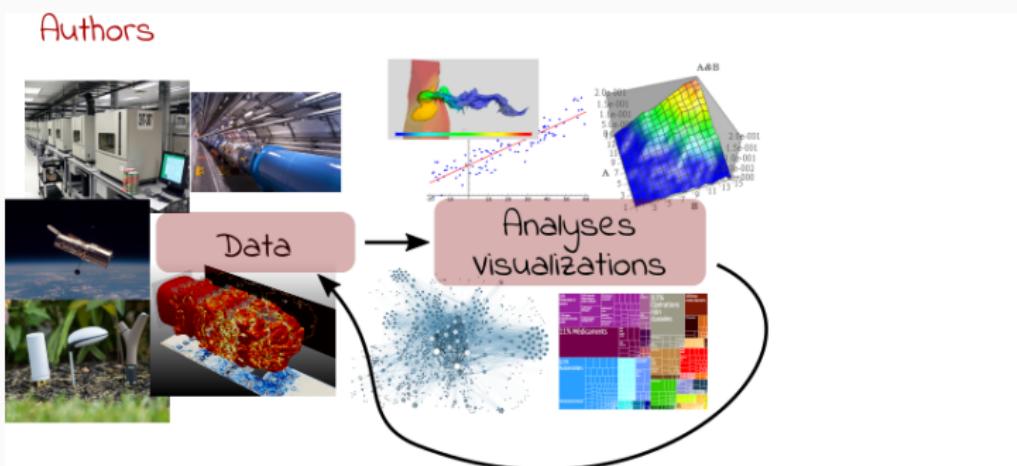
DIFFERENT REPRODUCIBILITY CONCERN IN MODERN SCIENCE

Social Sciences, Oncology, ... methodology, statistics

Genomics software engineering, computational reproducibility, provenance

Computational fluid dynamics numerical issues

The processing steps between raw observations and findings have gotten increasingly numerous and complex



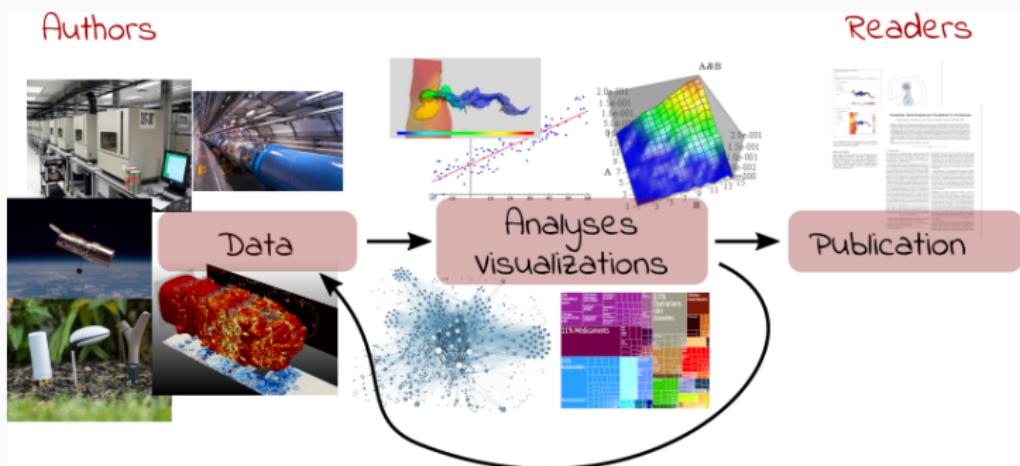
DIFFERENT REPRODUCIBILITY CONCERN IN MODERN SCIENCE

Social Sciences, Oncology, ... methodology, statistics

Genomics software engineering, computational reproducibility, provenance

Computational fluid dynamics numerical issues

The processing steps between raw observations and findings have gotten increasingly numerous and complex



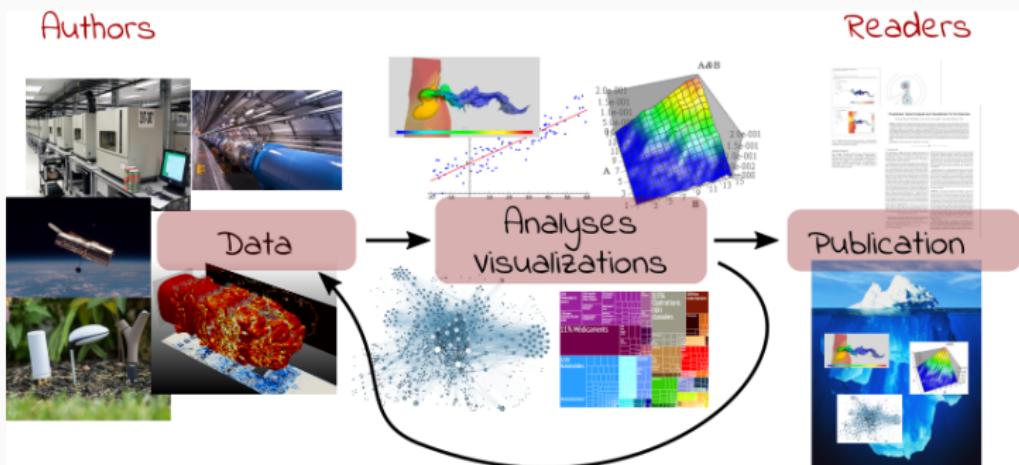
Different Reproducibility Concerns in Modern Science

Social Sciences, Oncology, ... methodology, statistics

Genomics software engineering, computational reproducibility, provenance

Computational fluid dynamics numerical issues

The processing steps between raw observations and findings have gotten increasingly numerous and complex



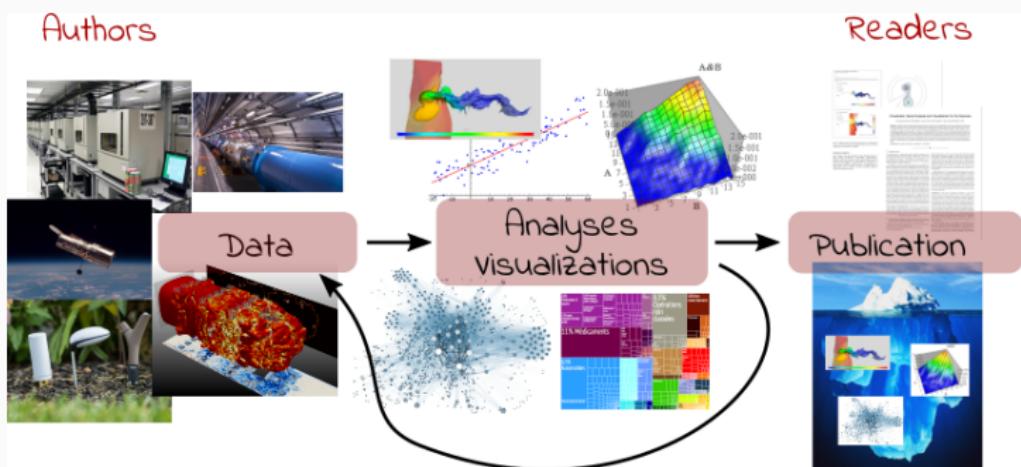
DIFFERENT REPRODUCIBILITY CONCERN IN MODERN SCIENCE

Social Sciences, Oncology, ... methodology, statistics

Genomics software engineering, computational reproducibility, provenance

Computational fluid dynamics numerical issues

The processing steps between raw observations and findings have gotten increasingly numerous and complex



Reproducible Research = Bridging the Gap by working Transparently

REPRODUCIBLE RESEARCH PRACTICES

"REPRODUCIBLE RESEARCH": FIRST APPEARANCE

Claerbout & Karrenbach, meeting of the Society of Exploration Geophysics, 1992

Electronic Documents Give Reproducible Research a New Meaning

RE1.3

Jon F. Claerbout and Martin Karrenbach, Stanford Univ.

SUMMARY

A revolution in education and technology transfer follows from the marriage of word processing and software command scripts. In this marriage an author attaches to every figure caption a pushbutton or a name tag usable to recalculate the figure from all its data, parameters, and programs. This provides a new meaning of reproducibility in computer documents.

In 1990, we set this sequence of goals:

- Learn how to merge a publication with its underlying computational analysis.
- Teach researchers how to prepare a document in a form where they themselves can reproduce their own research results a year or more later by "pressing a single button".
- Learn how to leave finished work in a condition where coworkers can reproduce the calculation including the final illustration by pressing a button in its caption.
- Prepare a complete copy of our local software environment so that graduating students can take their work away with them to other sites, press a button, and reproduce their Stanford work.
- Merge electronic documents written by multiple authors (SEP reports).

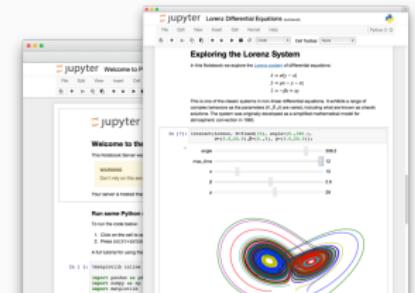
- make incremental improvements in electronic-document software
- seek partners for broadening standards (and making incremental improvements).

Our basic goal is reproducible research. The electronic document is our means to this end. In principle, reproducibility in research can be achieved without electronic documents and that is how we started. Our first nonelectronic reproducible document was a textbook in which the paper document contained the name of a program script in every figure caption. The program scripts were organized by book chapter and section so they could be correlated to an accompanying magnetic tape dump of the file system. The magnetic tape also contained all the necessary data to feed the program script.

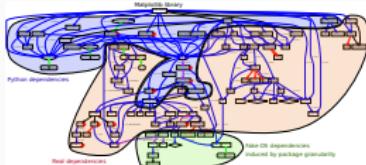
Now that we have begun using CD-ROM publication, we can go much further. Every figure caption contains a pushbutton that jumps to the appropriate science directory (folder) and initiates a figure rebuild command and then displays the figure, possibly as a movie or interactive program. We normally display seismic images of the earth's interior, but to reach wider audiences, Figure 1 shows a satellite weather picture which the pushbutton will animate as seen on commercial television. We include all our plot software as well as freely available software from many sources, including compilers and the L^AT_EX word processing systems. Naturally we cannot include licensed software, but with the exception

EXISTING TOOLS, EMERGING STANDARDS

Notebooks and workflows



Software environments



Sharing platforms



TOOL 1: COMPUTATIONAL NOTEBOOKS/LITTERATE PROGRAMMING

Un document computationnel

Mon ordinateur m'indique que π vaut approximativement

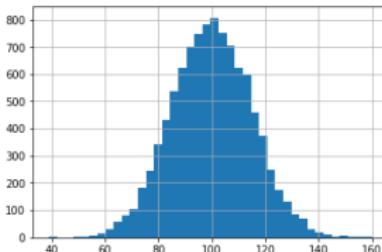
3.141592653589793

Mais calculé avec la méthode des [aiguilles de Buffon](#), on obtiendrait comme approximation :

```
import numpy as np
N = 1000000
x = np.random.uniform(size=N, low=0, high=1)
theta = np.random.uniform(size=N, low=0, high=pi/2)
2/(sum((x+np.sin(theta))>1)/N)
```

3.1437198694098765

On peut inclure des formules mathématiques comme $\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ et des dessins qui n'ont rien à voir avec π (si ce n'est une constante de normalisation... ☺).



TOOL 1: COMPUTATIONAL NOTEBOOKS/LITTERATE PROGRAMMING

Document initial dans son environnement

The screenshot shows a Jupyter Notebook interface with the following details:

- Title:** # Un document computationnel
- Cell 1:** In [1]:

```
from math import *
print(pi)
3.141592653589793
```

Output: Mais calculé avec la [méthode des aiguilles de Buffon](#) (https://fr.wikipedia.org/wiki/Aiguille_de_Buffon), on obtient d'abord comme approximation :
- Cell 2:** In [2]:

```
import numpy as np
N = 1000000
x = np.random.uniform(size=N, low=0, high=1)
theta = np.random.uniform(size=N, low=0, high=pi/2)
2/(sum((x+np.sin(theta))>1))/N
```

Output: Out[2]: 3.1437198694998765
- Cell 3:** In [3]:

```
%matplotlib inline
import matplotlib.pyplot as plt
mu, sigma = 100, 15
x = mu + sigma*np.random.randn(10000)
plt.hist(x,40)
plt.grid(True)
plt.show()
```

Output: A histogram showing a normal distribution centered at 100.

Document final

Un document computationnel

Mon ordinateur m'indique que π vaut approximativement

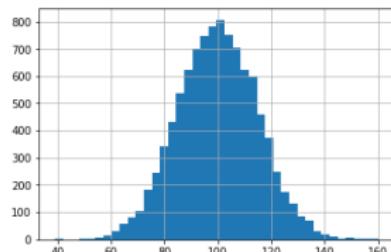
3.141592653589793

Mais calculé avec la [méthode des aiguilles de Buffon](#), on obtiendrait comme approximation :

```
import numpy as np
N = 1000000
x = np.random.uniform(size=N, low=0, high=1)
theta = np.random.uniform(size=N, low=0, high=pi/2)
2/(sum((x+np.sin(theta))>1))/N
```

3.1437198694998765

On peut inclure des formules mathématiques comme $\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ et des dessins qui n'ont rien à voir avec π (si ce n'est une constante de normalisation... ☺).



TOOL 1: COMPUTATIONAL NOTEBOOKS/LITTERATE PROGRAMMING

Document initial dans son environnement

```
# Un document computationnel

from math import *
print(pi)
3.141592653589793

Mais calculé avec la méthode des aiguilles de Buffon (https://fr.wikipedia.org/wiki/Aiguille\_de\_Buffon), on obtient comme approximation : 3.1437198694998765

import numpy as np
N = 1000000
x = np.random.uniform(size=N, low=0, high=1)
theta = np.random.uniform(size=N, low=0, high=pi/2)
2/(sum((x+np.sin(theta))>1))/N

Out[2]: 3.1437198694998765

On peut inclure des formules mathématiques comme $ \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) $ et des dessins qui n'ont rien à voir avec  $\pi$  (si ce n'est une constante de normalisation... ☺).

In [3]:
%matplotlib inline
import matplotlib.pyplot as plt

mu, sigma = 100, 15
x = mu + sigma*np.random.randn(10000)

plt.hist(x,40)
plt.grid(True)
plt.show()
```

Document final

Un document computationnel

Mon ordinateur m'indique que π vaut approximativement

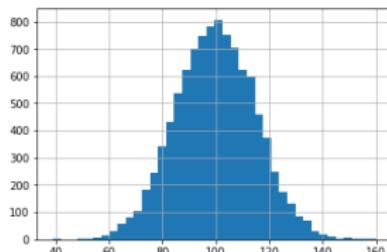
3.141592653589793

Mais calculé avec la [méthode des aiguilles de Buffon](#), on obtient comme approximation :

```
import numpy as np
N = 1000000
x = np.random.uniform(size=N, low=0, high=1)
theta = np.random.uniform(size=N, low=0, high=pi/2)
2/(sum((x+np.sin(theta))>1))/N
```

3.1437198694998765

On peut inclure des formules mathématiques comme $\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ et des dessins qui n'ont rien à voir avec π (si ce n'est une constante de normalisation... ☺).



TOOL 1: COMPUTATIONAL NOTEBOOKS/LITTERATE PROGRAMMING

Document initial dans son environnement

The screenshot shows a Jupyter Notebook interface with three code cells:

- In [1]:** Prints the value of pi (3.141592653589793) and includes a note about calculating pi with Buffon's needle method.
- In [2]:** Generates a uniform distribution of points (x, theta) and calculates an approximation of pi based on the ratio of points where x >= sin(theta).
- In [3]:** Plots a histogram of x values, showing a bell-shaped distribution centered around 100.

Annotations with red arrows point from the text "Code" in the first section to the code blocks in the notebook, and from the text "Document initial dans son environnement" to the top of the screenshot.

Document final

Un document computationnel

Mon ordinateur m'indique que π vaut approximativement

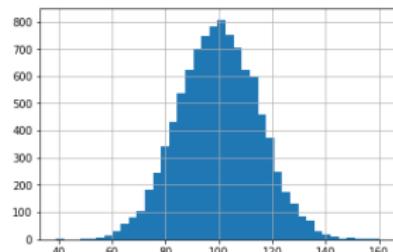
3.141592653589793

Mais calculé avec la méthode des [aiguilles de Buffon](#), on obtiendrait comme approximation :

```
import numpy as np
N = 1000000
x = np.random.uniform(size=N, low=0, high=1)
theta = np.random.uniform(size=N, low=0, high=pi/2)
2*(sum((x+np.sin(theta))>1))/N
```

3.14371986949098765

On peut inclure des formules mathématiques comme $\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ et des dessins qui n'ont rien à voir avec π (si ce n'est une constante de normalisation... ☺).



TOOL 1: COMPUTATIONAL NOTEBOOKS/LITTERATE PROGRAMMING

Document initial dans son environnement

A screenshot of a Jupyter Notebook interface. The top bar shows 'jupyter example_pi' and 'Python 3'. The notebook contains three code cells:

- In [1]:** Prints the value of pi: `print(pi)` resulting in `3.141592653589793`. A note below says: "Mais calculé avec la `_methode_ des aiguilles de Buffon` (https://fr.wikipedia.org/wiki/Aiguille_de_Buffon), on obtiendrait comme approximation :".
- In [2]:** Imports numpy and generates random numbers between 0 and 1. It calculates the ratio of points inside a unit circle to the total number of points to estimate pi.
- In [3]:** Imports matplotlib and plots a histogram of a random sample of 100 values. The plot shows a bell-shaped curve centered around 100.

Annotations in red:

- A large red arrow labeled "Résultats" points from the output of In [1] to the final document.
- Two smaller red arrows point from the notes in In [1] and In [2] to the explanatory text in the final document.

Document final

Un document computationnel

Mon ordinateur m'indique que π vaut approximativement

3.141592653589793

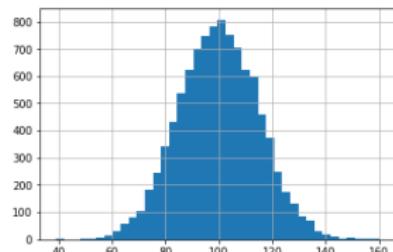
Mais calculé avec la **méthode des aiguilles de Buffon**, on obtiendrait comme approximation :

```
import numpy as np
N = 1000000
x = np.random.uniform(size=N, low=0, high=1)
theta = np.random.uniform(size=N, low=0, high=pi/2)
2*(sum((x+np.sin(theta))>1))/N
```

3.14371986949098765

On peut inclure des formules mathématiques comme $\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ et

des dessins qui n'ont rien à voir avec π (si ce n'est une constante de normalisation... ☺).



TOOL 1: COMPUTATIONAL NOTEBOOKS/LITTERATE PROGRAMMING

Document initial dans son environnement

The screenshot shows a Jupyter Notebook interface with three code cells:

- In [1]:** Prints the value of pi (3.141592653589793) and includes a note about calculating pi with Buffon's needle method.
- In [2]:** Generates random points (x, theta) and calculates the ratio of points where x <= mu + sigma * sin(theta) to the total number of points, which approximates pi.
- In [3]:** Plots a histogram of x values, showing a bell-shaped distribution centered around 100.

Document final

Un document computationnel

Mon ordinateur m'indique que π vaut approximativement

3.141592653589793

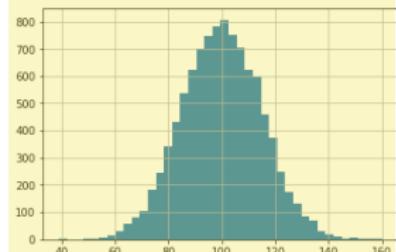
Mais calculé avec la méthode des [aiguilles de Buffon](#), on obtiendrait comme approximation :

```
import numpy as np
N = 1000000
x = np.random.uniform(size=N, low=0, high=1)
theta = np.random.uniform(size=N, low=0, high=pi/2)
2/(sum((x+np.sin(theta))>1))/N
```

3.1437198694098765

Export

On peut inclure des formules mathématiques comme $\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ et des dessins qui n'ont rien à voir avec π (si ce n'est une constante de normalisation... ☺).



TOOL 1: COMPUTATIONAL NOTEBOOKS/LITTERATE PROGRAMMING

Document initial dans son environnement

The screenshot shows a Jupyter Notebook interface with three code cells:

- In [1]:** Prints the value of pi (3.141592653589793) and includes a note about approximating pi with Buffon's needle method.
- In [2]:** Generates random points (x, theta) and calculates the ratio of points where x >= sin(theta) to the total number of points, which approximates pi/2.
- In [3]:** Plots a histogram of x values, showing a bell-shaped distribution centered around 100.

Document final

Un document computationnel

Mon ordinateur m'indique que π vaut approximativement

3.141592653589793

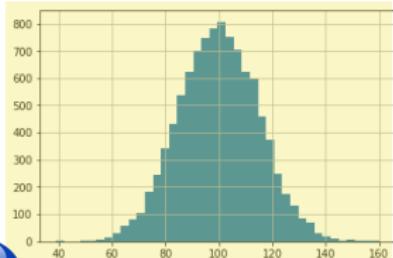
Mais calculé avec la **méthode des aiguilles de Buffon**, on obtiendrait comme approximation :

```
import numpy as np
N = 1000000
x = np.random.uniform(size=N, low=0, high=1)
theta = np.random.uniform(size=N, low=0, high=pi/2)
2/(sum((x+np.sin(theta))>1))/N
```

3.1437198694098765

Export

On peut inclure des formules mathématiques comme $\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ et des dessins qui n'ont rien à voir avec π (si ce n'est une constante de normalisation... ☺).



TOOL 2: FIGHTING SOFTWARE ENVIRONMENTS NIGHTMARE

What is hiding behind a simple

```
import matplotlib
```

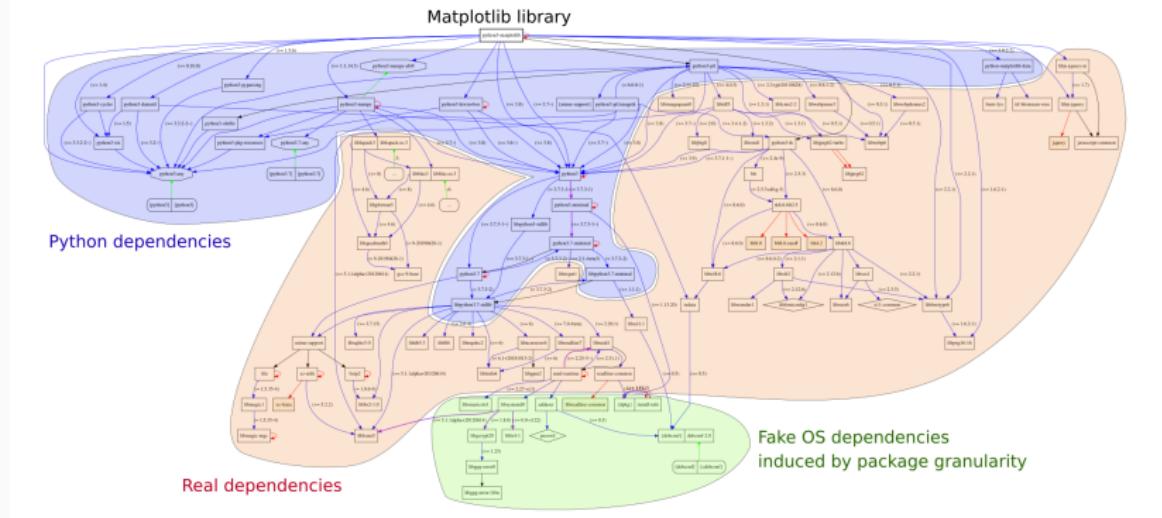
Package: python3-matplotlib
Version: 2.1.1-2
Depends: python3-dateutil, python-matplotlib-data (>= 2.1.1-2),
python3-pyparsing (>= 1.5.6), python3-six (>= 1.10), python3-tz,
libjs-jquery, libjs-jquery-ui, python3-numpy (>= 1:1.13.1),
python3-numpy-abi9, python3 (<< 3.7), python3 (>= 3.6~),
python3-cycler (>= 0.10.0), python3:any (>= 3.3.2-2~), libc6 (>= 2.14), libfreetype6 (>= 2.2.1), libgcc1 (>= 1:3.0), libpng16-16 (>= 1.6.2-1), libstdc++6 (>= 5.2), zlib1g (>= 1:1.1.4)

TOOL 2: FIGHTING SOFTWARE ENVIRONMENTS NIGHTMARE

What is hiding behind a simple

```
import matplotlib
```

Package: python3-matplotlib



TOOL 2: FIGHTING SOFTWARE ENVIRONMENTS NIGHTMARE

Python and its rapidly evolving environment

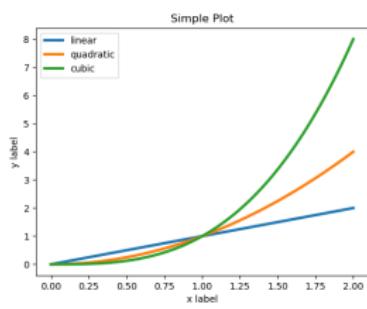
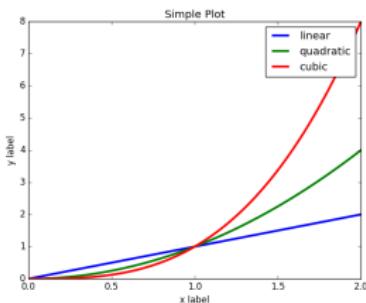
```
python2 -c "print(10/3)"  
python3 -c "print(10/3)"
```

3
3.333333333333335

TOOL 2: FIGHTING SOFTWARE ENVIRONMENTS NIGHTMARE

Python and its rapidly evolving environment

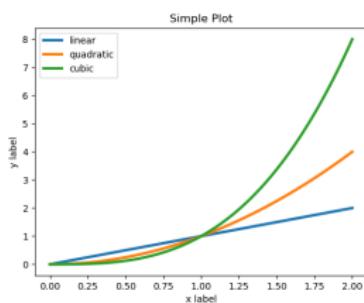
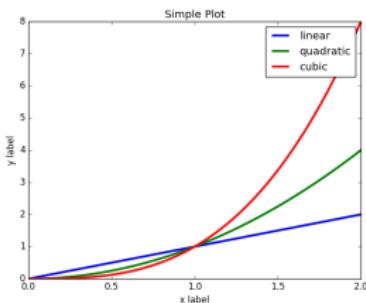
```
python2 -c "print(10/3)"  
python3 -c "print(10/3)"
```



TOOL 2: FIGHTING SOFTWARE ENVIRONMENTS NIGHTMARE

Python and its rapidly evolving environment

```
python2 -c "print(10/3)"  
python3 -c "print(10/3)"
```

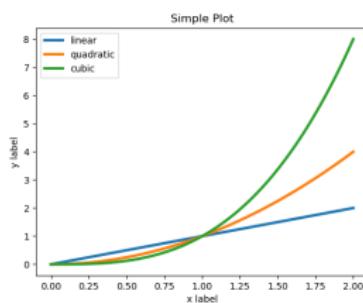
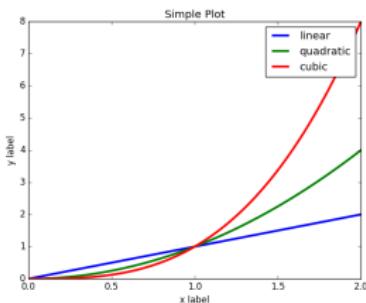


Cortical Thickness Measurements (PLOS ONE, June 2012): *FreeSurfer: differences were found between the Mac and HP workstations and between Mac OSX 10.5 and OSX 10.6.*

TOOL 2: FIGHTING SOFTWARE ENVIRONMENTS NIGHTMARE

Python and its rapidly evolving environment

```
python2 -c "print(10/3)"  
python3 -c "print(10/3)"
```



Cortical Thickness Measurements (PLOS ONE, June 2012): *FreeSurfer: differences were found between the Mac and HP workstations and between Mac OSX 10.5 and OSX 10.6.*



TOOL 3: FIGHTING INFORMATION LOSS WITH ARCHIVES

D. Spinellis. *The Decay and Failures of URL References*. CACM, 46(1), Jan 2003.

The half-life of a referenced URL is approximately 4 years from its publication date.

P. Habibzadeh. *Decay of References to Web sites in Articles Published in General Medical Journals: Mainstream vs Small Journals*. Applied Clinical Informatics. 4 (4), 2013

half life ranged from 2.2 years in EMHJ to 5.3 years in BMJ

TOOL 3: FIGHTING INFORMATION LOSS WITH ARCHIVES

D. Spinellis. *The Decay and Failures of URL References*. CACM, 46(1), Jan 2003.

The half-life of a referenced URL is approximately 4 years from its publication date.

P. Habibzadeh. *Decay of References to Web sites in Articles Published in General Medical Journals: Mainstream vs Small Journals*. Applied Clinical Informatics. 4 (4), 2013

half life ranged from 2.2 years in EMHJ to 5.3 years in BMJ

Article archives arXiv.org HAL
archives-ouvertes.fr

Data archives figshare zenodo

Software Archive Software Heritage



or



= awesome collaborations ≠ archive

CHANGING RESEARCH PRACTICES

Soft. Engineering, Statistics, and Reproducible Research in the curricula



- Book on RR *Vers une recherche
reproductible: Faire évoluer ses pratiques*
- MOOC on RR (3rd edition, ongoing)
- A new "Advanced RR" MOOC (2021?)
 - Software environment control
 - Scientific workflow
 - Managing data

Manifesto: "*I solemnly pledge*" (**WSSSPE, Lorena Barba, FAIR**)

1. I will teach my graduate students about reproducibility
2. All our research code (and writing) is under version control
3. We will always carry out verification and validation
4. We will share data, plotting script & figure under CC-BY
5. We will upload the preprint to arXiv at the time of submission of a paper
6. We will release code at the time of submission of a paper
7. We will add a "Reproducibility" declaration at the end of each paper
8. I will keep an up-to-date web presence

WHAT ABOUT COMPUTER SCIENCE ?

ALL THIS IS ABOUT COMPUTATIONAL SCIENCES. SHOULD WE CARE ?

Computer Science is young and inherits from Mathematics,
Engineering, Nat. Sciences, Linguistic, ...

Purely theoretical scientists whose practice is close to mathematics
may not be concerned (can't publish a math article without releasing
the proofs).

- Have a look at talk by Vladimir Voevodsky in 2014 at Princeton 😊

ALL THIS IS ABOUT COMPUTATIONAL SCIENCES. SHOULD WE CARE ?

Computer Science is young and inherits from Mathematics,
Engineering, Nat. Sciences, Linguistic, ...

Purely theoretical scientists whose practice is close to mathematics
may not be concerned (can't publish a math article without releasing
the proofs).

- Have a look at talk by Vladimir Voevodsky in 2014 at Princeton 😊

Computer science is not more related to computers than Astronomy to telescopes

– Dijkstra (mis-attributed)

Right, why should we care about computers? They are **deterministic**
machines after all, right? 😊

Model ≠ Reality. Although designed and built by human beings,
computer systems are **so complex** that mistakes easily slip in...

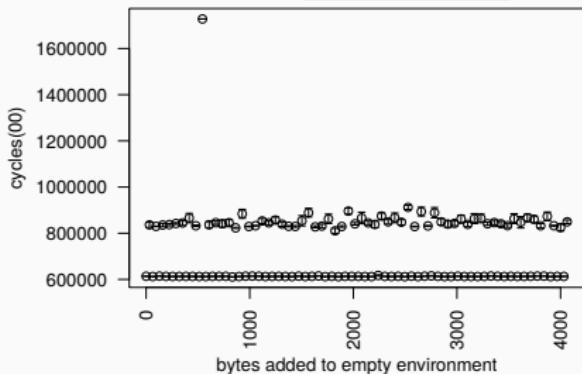
EXPERIMENTING WITH COMPUTERS

Machines are real!



Brendan Gregg: Shouting in the data center

Machines are complicated



Mytkowicz et al. *Producing wrong data without doing anything obviously wrong!*
ACM SIGPLAN Not. 44(3), March 2009

Our reality evolves!!! The hardware keeps evolving so most results on old platforms quickly become obsolete (although, we keep building on such results 😊).

We need to regularly revisit and allow others to build on our work!

COMPUTER PERFORMANCE ? WELL, I DESIGN ALGORITHMS!

- "Real" problems are all NP-hard, Log-APX, etc.
- Real workload = ~~NP-completeness proof widgets~~, regularities and properties (difficult to formally state but that should be exploited)

Algorithms are evaluated on particular **workloads** that impact both their running time and the quality of the solutions

COMPUTER PERFORMANCE ? WELL, I DESIGN ALGORITHMS!

- "Real" problems are all NP-hard, Log-APX, etc.
- Real workload = ~~NP-completeness proof~~ widgets, regularities and properties (difficult to formally state but that should be exploited)

Algorithms are evaluated on particular **workloads** that impact both their running time and the quality of the solutions

Image Processing: **True horror stories**, E. Meinhardt-Llopis, CANUM 2016

- *The proposed multigrid algorithm converges to the solution of the problem in $O(N)$ using biharmonic functions*
- Surprisingly, our naive multi-scale Gauss-Seidel converges much faster

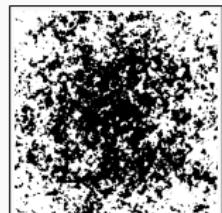
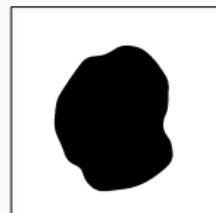
COMPUTER PERFORMANCE ? WELL, I DESIGN ALGORITHMS!

- "Real" problems are all NP-hard, Log-APX, etc.
- Real workload = ~~NP-completeness proof~~ widgets, regularities and properties (difficult to formally state but that should be exploited)

Algorithms are evaluated on particular **workloads** that impact both their running time and the quality of the solutions

Image Processing: True horror stories, E. Meinhardt-Llopis, CANUM 2016

- *The proposed multigrid algorithm converges to the solution of the problem in $O(N)$ using biharmonic functions*
- Surprisingly, our naive multi-scale Gauss-Seidel converges much faster





Machine Learning: Trouble at the lab, The Economist 2013

According to some estimates, three-quarters of published scientific papers in the field of machine learning are bunk because of this "overfitting".

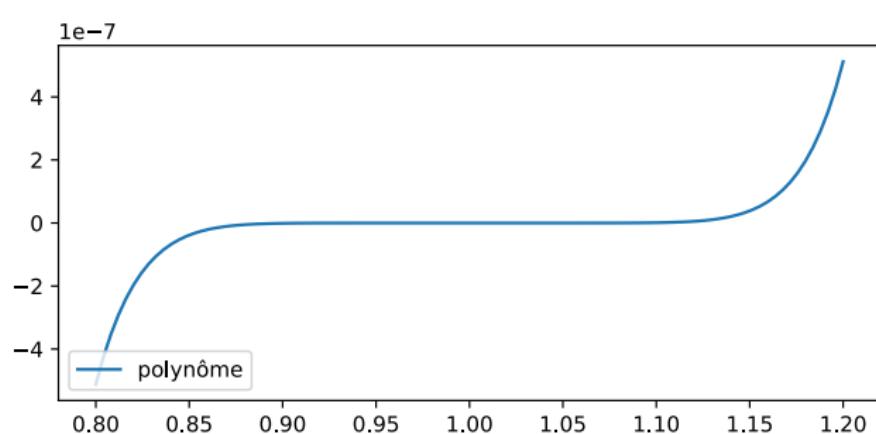
– Sandy Pentland (MIT)

Every month in CACM, there is an article about the ethical consequences of Machine Learning on:

- Car driving, Autonomous guns, Law enforcement (risk assessment, predictive policing), ...
It's Not the Algorithm, It's the Data (CACM, Feb. 2017)
- Advertising, Loan attribution, Selection at University, Organ transplant

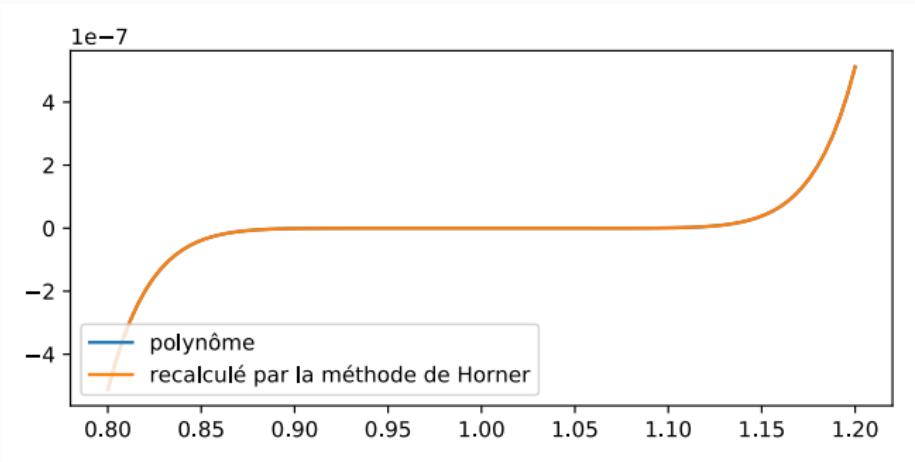
Increasing society concern about **fairness** and **transparency**

ALL I CARE ABOUT IS THE ALGORITHM OUTPUT (FP)



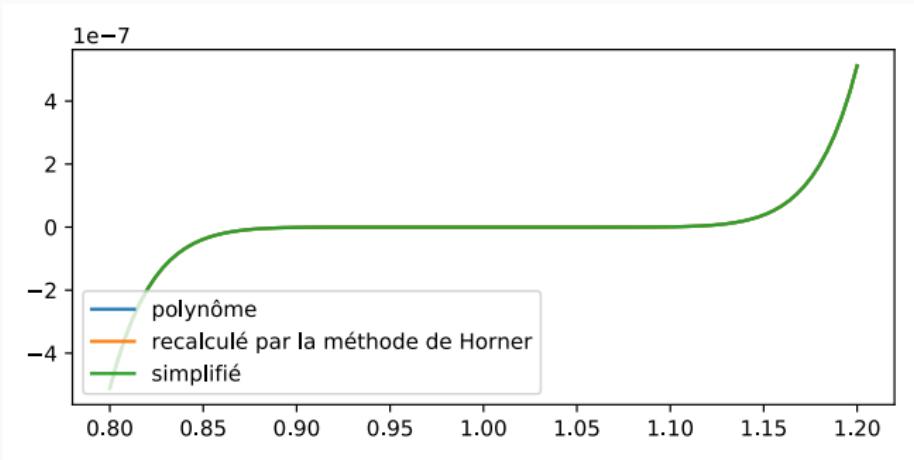
```
def polynome(x):
    return x**9 - 9.*x**8 + 36.*x**7 - 84.*x**6 + 126.*x**5 \
           - 126.*x**4 + 84.*x**3 - 36.*x**2 + 9.*x - 1.
```

FLOATING-POINT ARITHMETIC



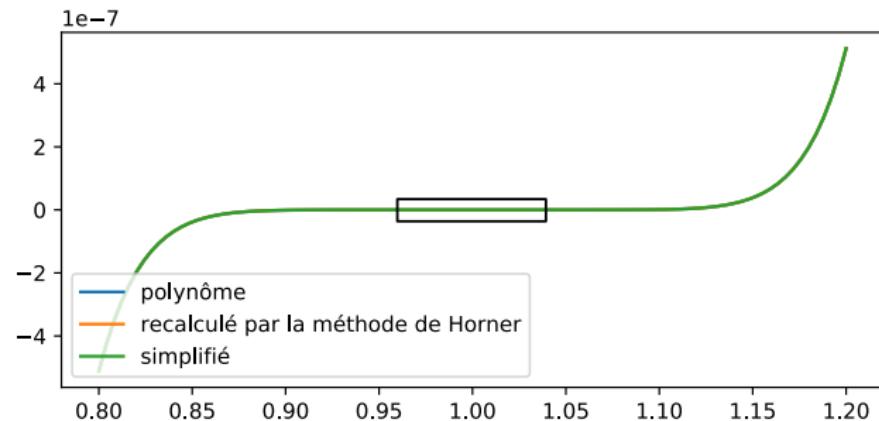
```
def horner(x):
    return x*(x*(x*(x*(x*(x*(x*(x - 9.) + 36.) - 84.) + 126.) \
        - 126.) + 84.) - 36.) + 9.) - 1.
```

FLOATING-POINT ARITHMETIC

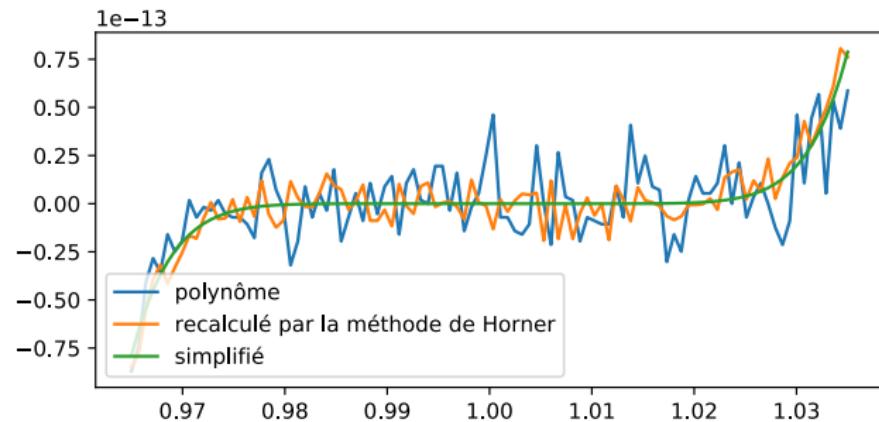


```
def simple(x):
    return (x-1.)**9
# Easy! ;)
```

FLOATING-POINT ARITHMETIC



FLOATING-POINT ARITHMETIC



ROUNDING

- Every operation includes implicit rounding.
- $a+b$ is actually `round(a+b)`.
- Unfortunately:

$$\text{round}(\text{round}(a+b)+c) \neq \text{round}(a+\text{round}(b+c)).$$

- Operation order therefore matters.

For a reproducible computation, operation order must be preserved!!!

HOW TO EXPLAIN IT TO MY COMPILER?

To speed up computations, compilers may change operation order, and thus results.

Two options for computing reproducibly:

1. Insist on the preservation of operation order,
 - if the language permits it.
 - Example: Module 'ieee_arithmetic' in Fortran 2003
2. Make compilation reproducible:
 - Record the precise compiler version
 - Record all compilation options

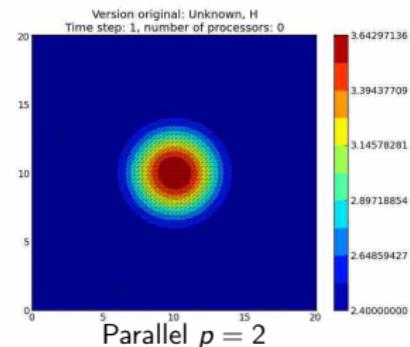
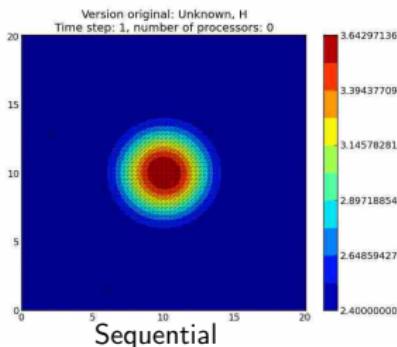
DID I MENTION WE HAVE PARALLEL MACHINES NOWADAYS?

Telemac2D: the simplest gouttedeo simulation

The gouttedeo test case

- 2D-simulation of a water drop fall in a square bassin
- Unknown: water depth for a 0.2 sec time step
- Triangular mesh: 8978 elements and 4624 nodes

Expected numerical reproducibility (time step = 1, 2, ...)

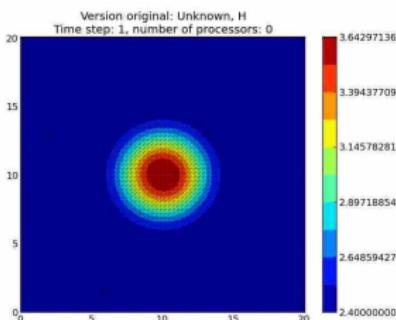


DID I MENTION WE HAVE PARALLEL MACHINES NOWADAYS?

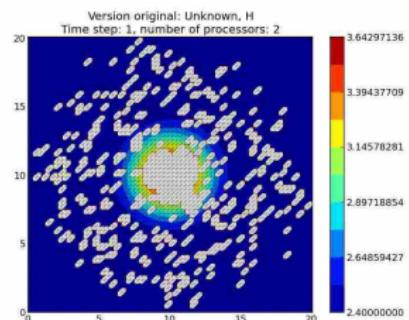
A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 1



Sequential



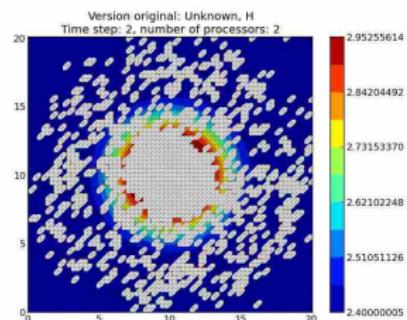
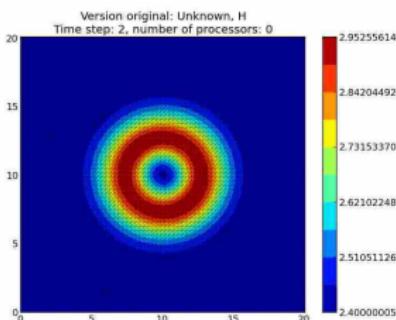
Parallel $p = 2$

DID I MENTION WE HAVE PARALLEL MACHINES NOWADAYS?

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 2

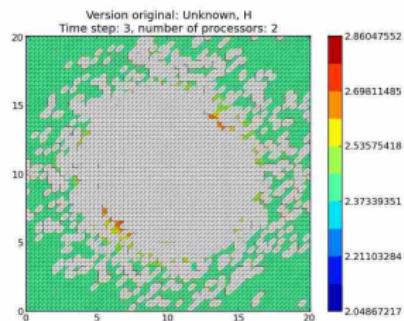
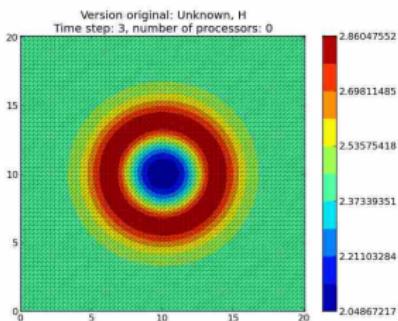


DID I MENTION WE HAVE PARALLEL MACHINES NOWADAYS?

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 3

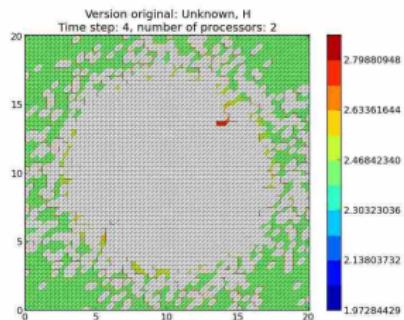
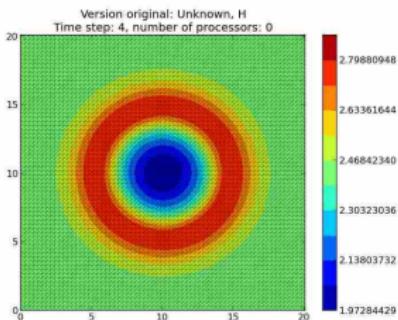


DID I MENTION WE HAVE PARALLEL MACHINES NOWADAYS?

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 4

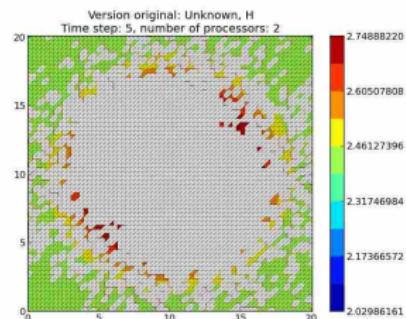
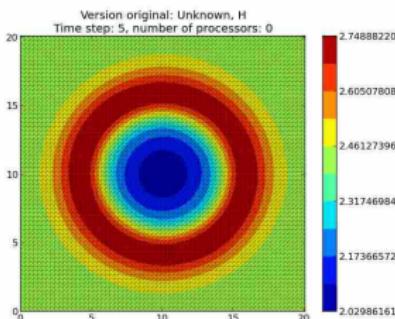


DID I MENTION WE HAVE PARALLEL MACHINES NOWADAYS?

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 5

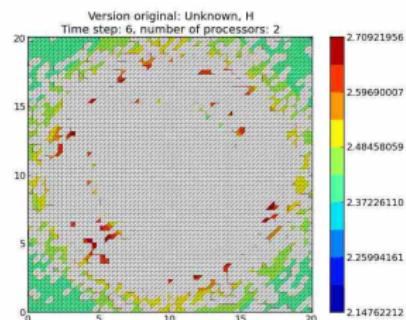
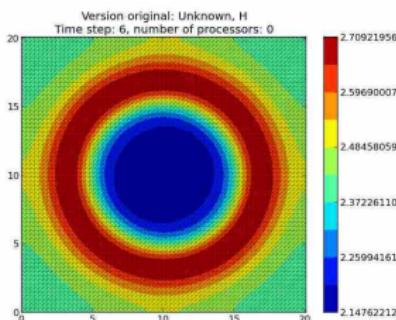


DID I MENTION WE HAVE PARALLEL MACHINES NOWADAYS?

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 6

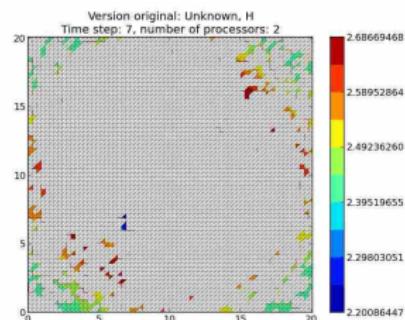
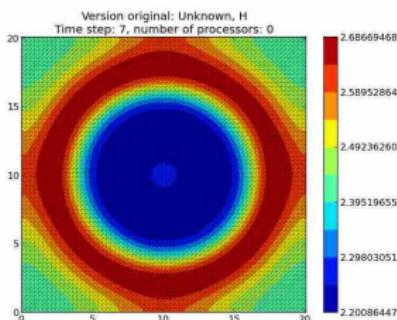


DID I MENTION WE HAVE PARALLEL MACHINES NOWADAYS?

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 7

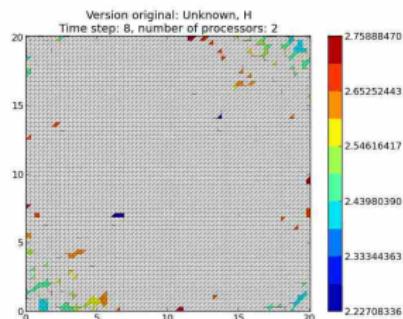
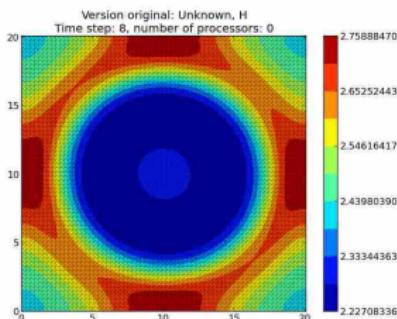


DID I MENTION WE HAVE PARALLEL MACHINES NOWADAYS?

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 8

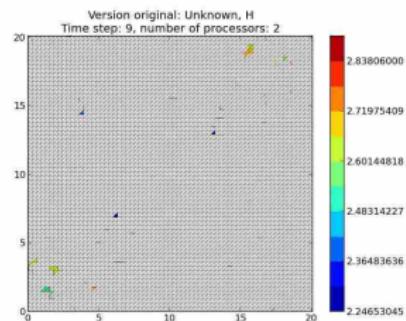
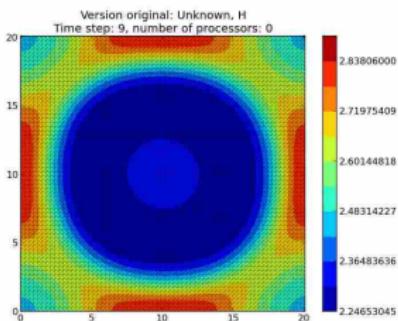


DID I MENTION WE HAVE PARALLEL MACHINES NOWADAYS?

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 9

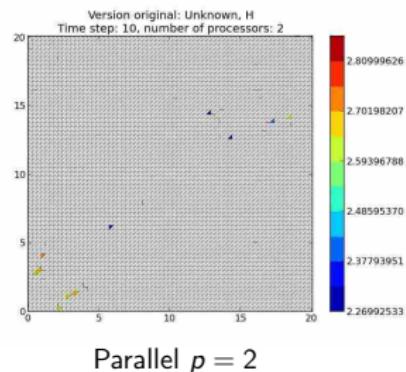
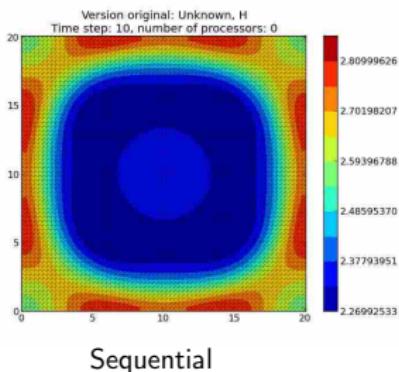


DID I MENTION WE HAVE PARALLEL MACHINES NOWADAYS?

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 10

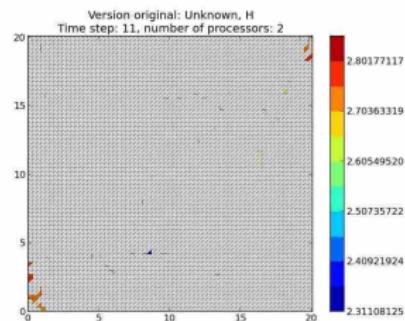
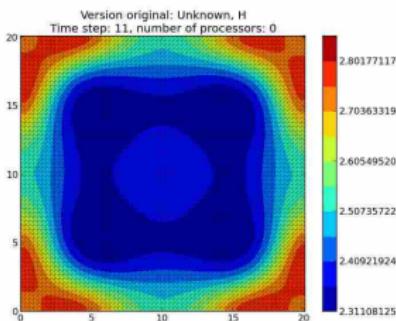


DID I MENTION WE HAVE PARALLEL MACHINES NOWADAYS?

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 11

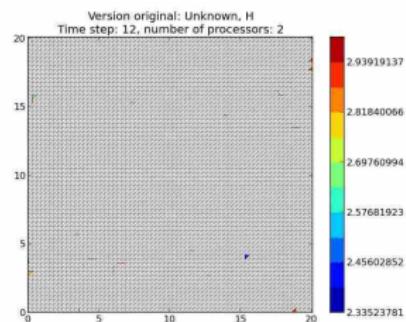
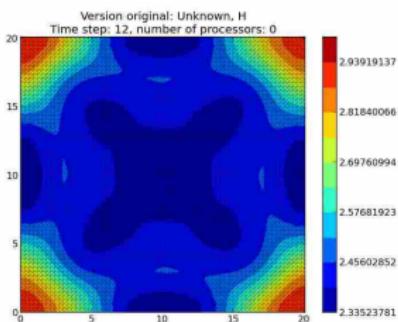


DID I MENTION WE HAVE PARALLEL MACHINES NOWADAYS?

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 12

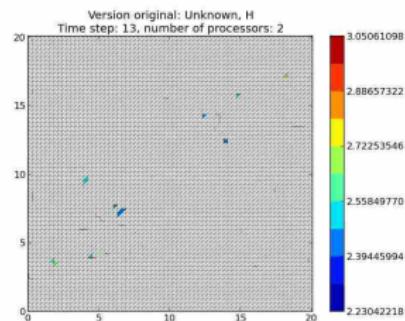
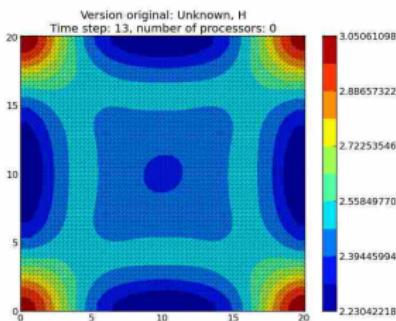


DID I MENTION WE HAVE PARALLEL MACHINES NOWADAYS?

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 13

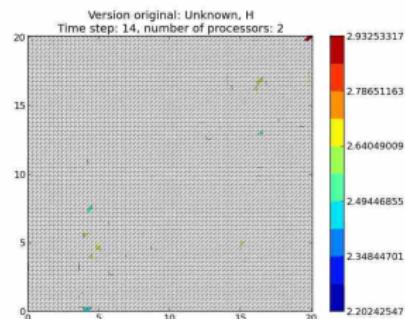
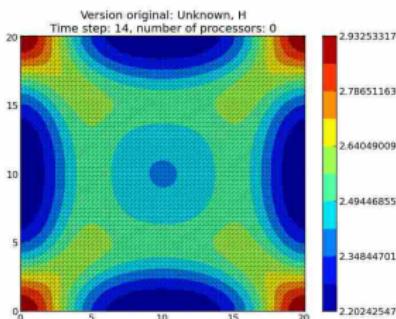


DID I MENTION WE HAVE PARALLEL MACHINES NOWADAYS?

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 14

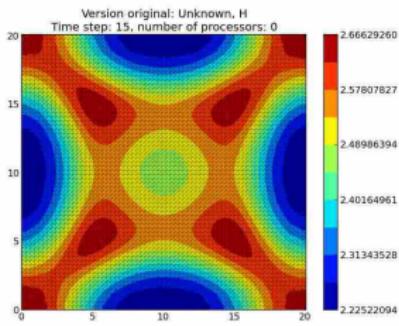


DID I MENTION WE HAVE PARALLEL MACHINES NOWADAYS?

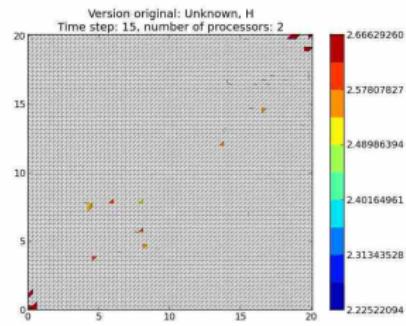
A white plot displays a non-reproducible value

NO numerical reproducibility!

time step = 15

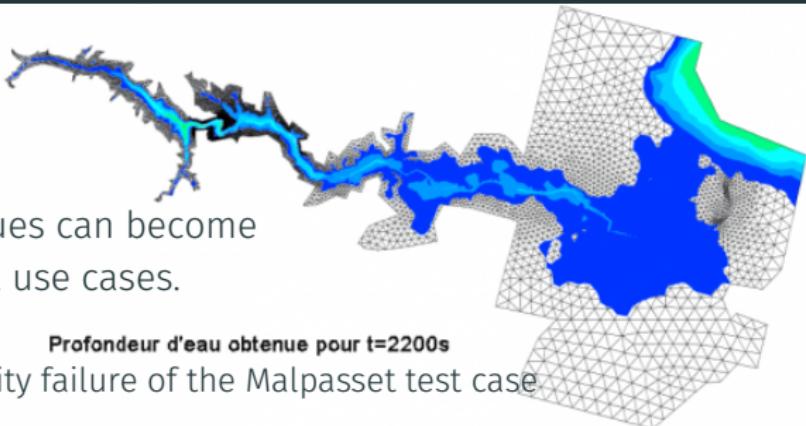


Sequential



Parallel $p = 2$

DID I MENTION WE HAVE PARALLEL MACHINES NOWADAYS?



These numerical issues can become quite harmful in real use cases.

TABLE 1.1: Reproducibility failure of the Malpasset test case

	The sequential run	a 64 procs run	a 128 procs run
depth H	0.3500122E-01	0.2748817E-01	0.1327634E-01
velocity U	0.4029747E-02	0.4935279E-02	0.4512116E-02
velocity V	0.7570773E-02	0.3422730E-02	0.7545233E-02

Numerical reproducibility: Approximations in the model, in the algorithm, in its implementation, in its execution.

The whole chain needs to be revisited.

Courtesy of P. Langlois and R. Nheili

Artifact evaluation and ACM badges



Major conferences

- Supercomputing: Artifact Description (AD) mandatory, Artifact Evaluation (AE) still optional, Double blind vs. RR
- NeurIPS, ICLR: open reviews, reproducibility challenge



Joelle Pineau @ NeurIPS'18

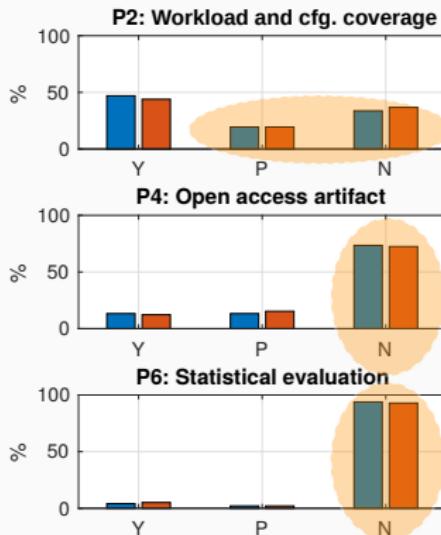
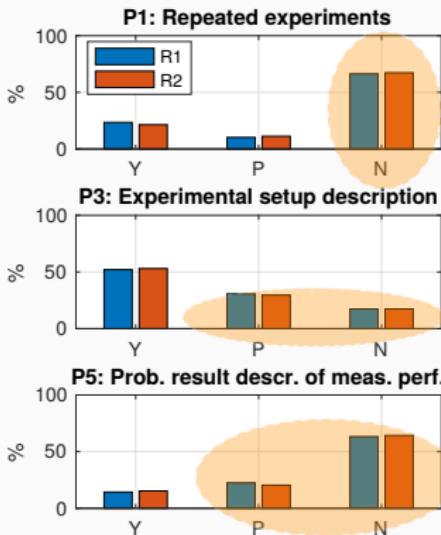
- ACM SIGMOD 2015-2019, Most Reproducible Paper Award...

Mentalities are evolving people care, make stuff available, errors are found and fixed* Resources and Acknowledgments

TOWARD REPRODUCIBLE COMPUTER SCIENCE EXPERIMENTS ?

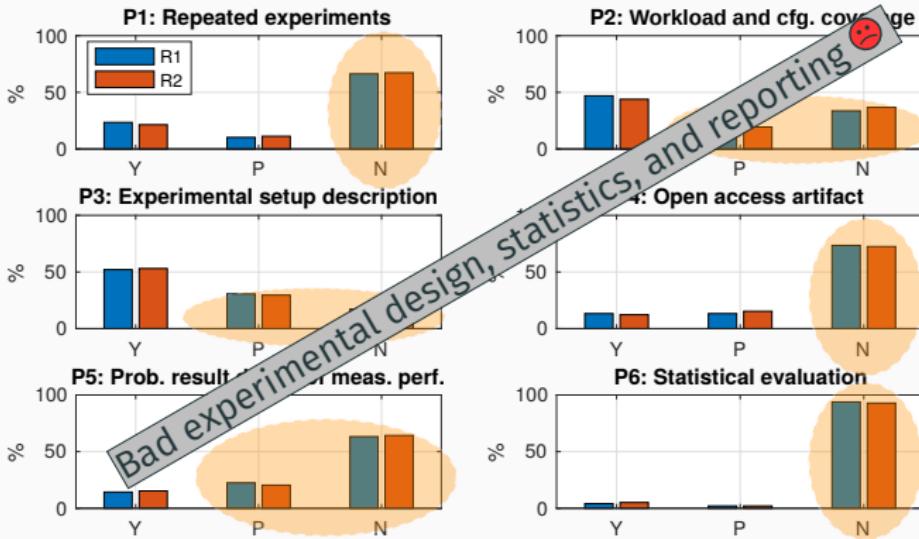
KEY CONCERN FOR OUR COMMUNITY (ROOM FOR IMPROVEMENT)

How are cloud performance currently obtained and reported?, Methodological Principles for Reproducible Performance Evaluation in Cloud Computing, IEEE Trans. on Soft. Eng., July 2019



KEY CONCERN FOR OUR COMMUNITY (ROOM FOR IMPROVEMENT)

How are cloud performance currently obtained and reported?, Methodological Principles for Reproducible Performance Evaluation in Cloud Computing, IEEE Trans. on Soft. Eng., July 2019



Key DoE principles:

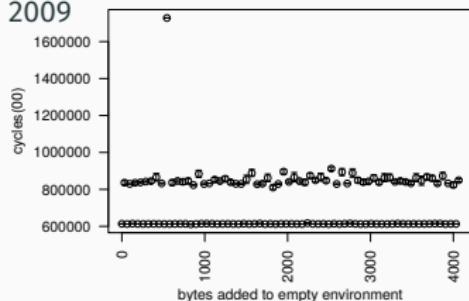
1. Replicate to increase reliability.
2. Randomize to reduce bias \rightsquigarrow Evaluate statistical confidence.

MEASURING PERFORMANCE IS DIFFICULT

Producing wrong data without doing anything obviously wrong!

Mytkowicz et al. in ACM SIGPLAN Not. 44(3), March 2009

changing the size of environment variables can trigger performance degradation as high as 300%; simply changing the link order of object files can cause performance to decrease by as much as 57%.

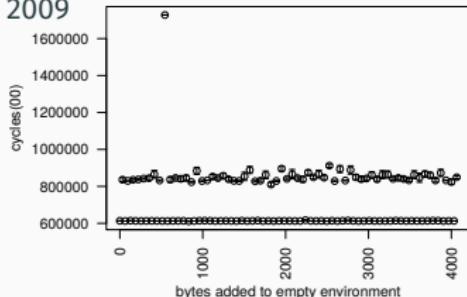


MEASURING PERFORMANCE IS DIFFICULT

Producing wrong data without doing anything obviously wrong!

Mytkowicz et al. in ACM SIGPLAN Not. 44(3), March 2009

changing the size of *environment variables* can trigger performance degradation as high as 300%; simply changing the *link order* of object files can cause performance to decrease by as much as 57%.



Taming the Influence of Memory Layout.

STABILIZER: *Statistically Sound Performance Evaluation*, C. Curtsinger and E. Berger in ASPLOS 2013

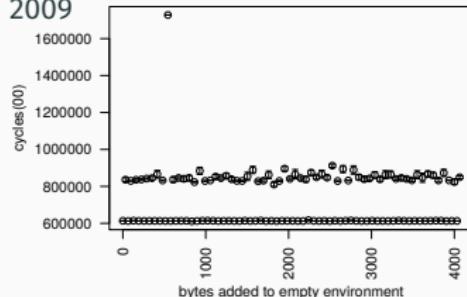
STABILIZER forces executions to sample the space of memory configurations by *repeatedly rerandomizing* layouts of code, stack, and heap objects at run-time. [...] Re-randomization ensures that layout effects *follow a Gaussian distribution*, enabling the use of statistical tests like ANOVA.

MEASURING PERFORMANCE IS DIFFICULT

Producing wrong data without doing anything obviously wrong!

Mytkowicz et al. in ACM SIGPLAN Not. 44(3), March 2009

changing the size of *environment variables* can trigger performance degradation as high as 300%; simply changing the *link order* of object files can cause performance to decrease by as much as 57%.



Taming the Influence of Memory Layout.

STABILIZER: Statistically Sound Performance Evaluation, C. Curtsinger and E. Berger in ASPLOS 2013

STABILIZER forces executions to sample the space of memory configurations by *repeatedly rerandomizing* layouts of code, stack, and heap objects at run-time. [...] Re-randomization ensures that layout effects *follow a Gaussian distribution*, enabling the use of statistical tests like ANOVA.

Randomization helps fighting bias incurred by:

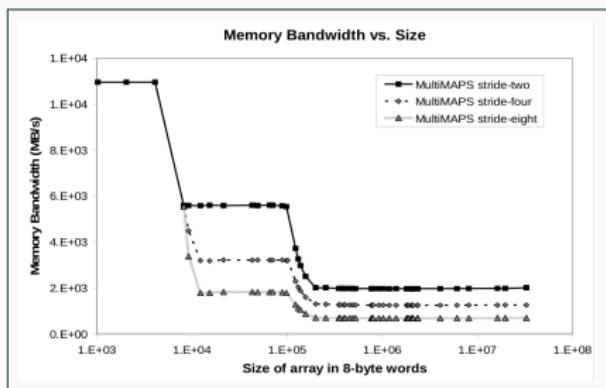
1. specific configurations AA...A → A₁A₂...A_n (*pseudo-replication*)
2. temporary perturbations AA...A BB...B → ABBAAAAB...

IMPACT OF WORKING SET SIZE ON EFFECTIVE BANDWIDTH

- Cache hierarchy (L1, L2, L3, RAM) with different bandwidth
- LRU, pre-fetching for linear access
- An array fits in a cache level \Rightarrow operate at the corresponding bandwidth
- Stride access decrease bandwidth

IMPACT OF WORKING SET SIZE ON EFFECTIVE BANDWIDTH

- Cache hierarchy (L1, L2, L3, RAM) with different bandwidth
- LRU, pre-fetching for linear access
- An array fits in a cache level \Rightarrow operate at the corresponding bandwidth
- Stride access decrease bandwidth



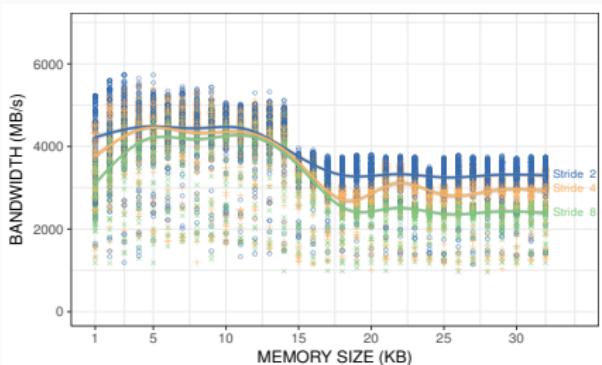
MultiMAPS on an Opteron

Genetic Algorithms Approach to Modeling the Performance of Memory-bound Computations,
Tikir et. al. in SC'07

```
MultiMAPS(size, stride, nloops) {  
    allocate buffer[size];  
    timer_start();  
    for rep in (1..nloops)  
        for i in (0..size/stride)  
            access buffer[stride*i];  
    timer_stop();  
    bandwidth = nb_access /  
               elapsed_time;  
    deallocate buffer;  
}
```

IMPACT OF WORKING SET SIZE ON EFFECTIVE BANDWIDTH

- Cache hierarchy (L1, L2, L3, RAM) with different bandwidth
- LRU, pre-fetching for linear access
- An array fits in a cache level \Rightarrow operate at the corresponding bandwidth
- Stride access decrease bandwidth

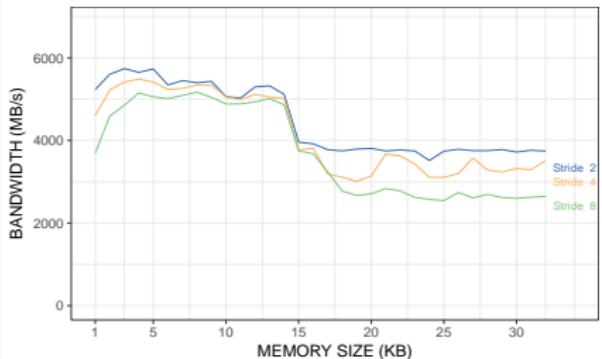


Our first attempt on a Pentium 4...

```
MultIMAPS(size, stride, nloops) {  
    allocate buffer[size];  
    timer_start();  
    for rep in (1..nloops)  
        for i in (0..size/stride)  
            access buffer[stride*i];  
    timer_stop();  
    bandwidth = nb_access /  
               elapsed_time;  
    deallocate buffer;  
}
```

IMPACT OF WORKING SET SIZE ON EFFECTIVE BANDWIDTH

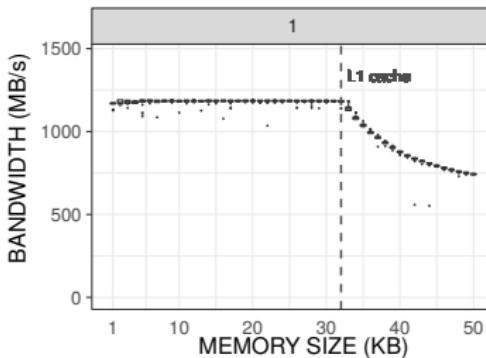
- Cache hierarchy (L1, L2, L3, RAM) with different bandwidth
- LRU, pre-fetching for linear access
- An array fits in a cache level \Rightarrow operate at the corresponding bandwidth
- Stride access decrease bandwidth



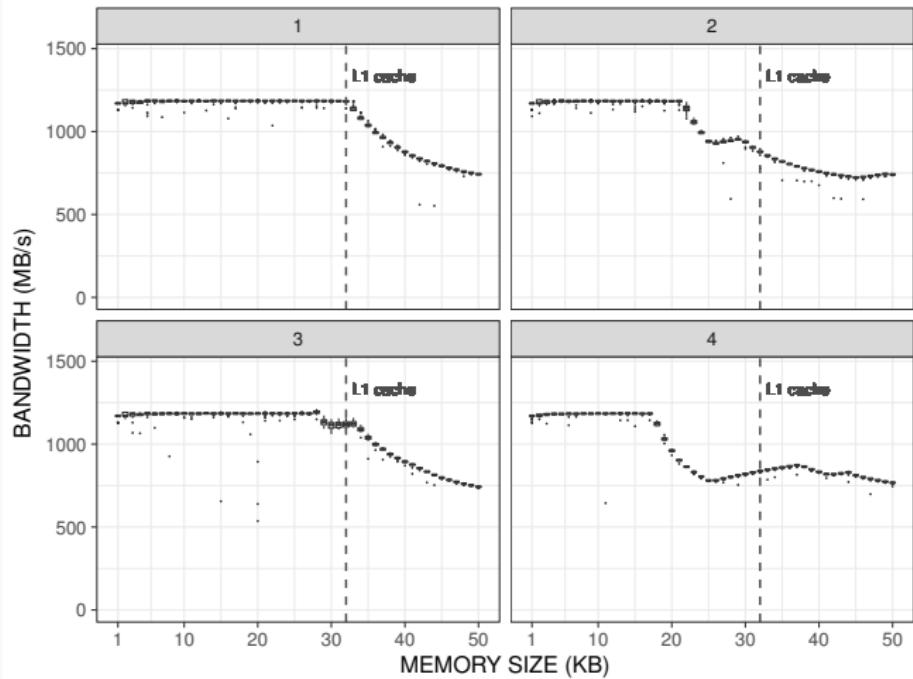
Our first attempt on a Pentium 4...

```
MultIMAPS(size, stride, nloops) {  
    allocate buffer[size];  
    timer_start();  
    for rep in (1..nloops)  
        for i in (0..size/stride)  
            access buffer[stride*i];  
    timer_stop();  
    bandwidth = nb_access /  
               elapsed_time;  
    deallocate buffer;  
}
```

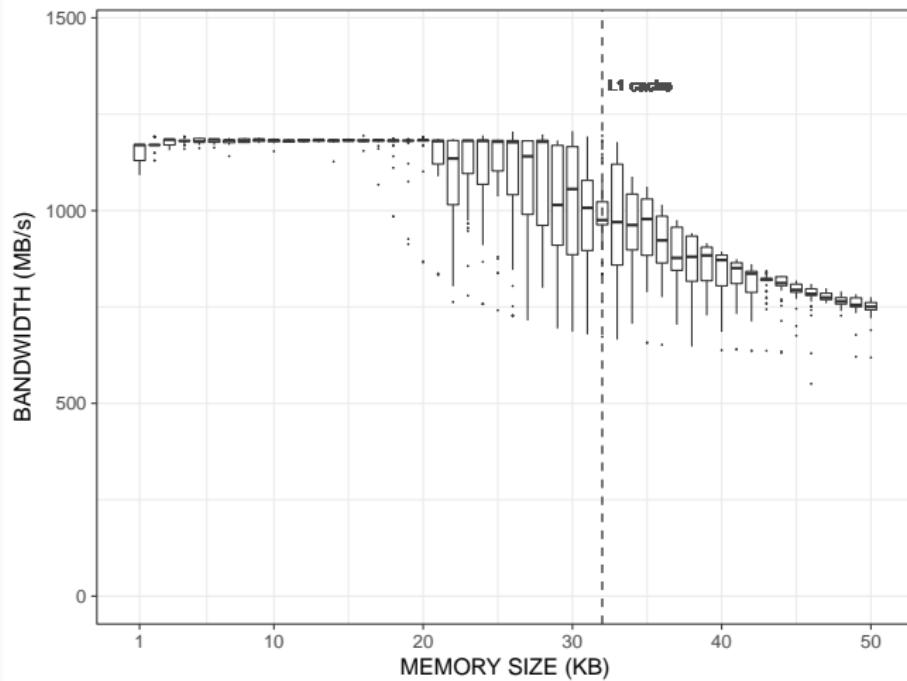
IMPACT OF ARCHITECTURE



IMPACT OF ARCHITECTURE



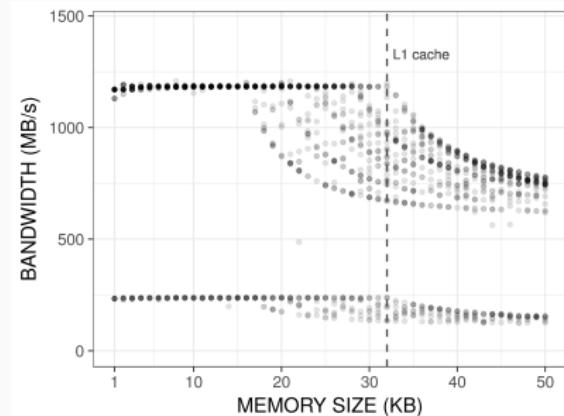
IMPACT OF ARCHITECTURE (THE ARM ASSOCIATIVITY ISSUE)



Randomize physical address start!

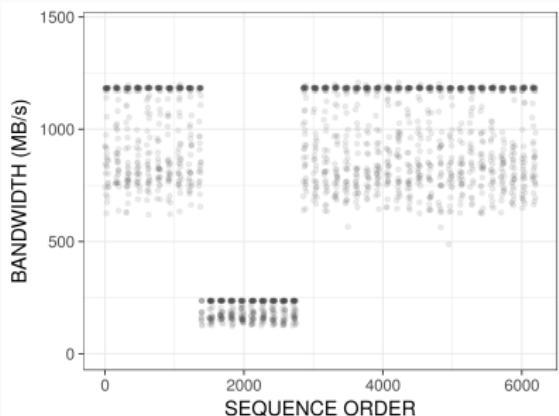
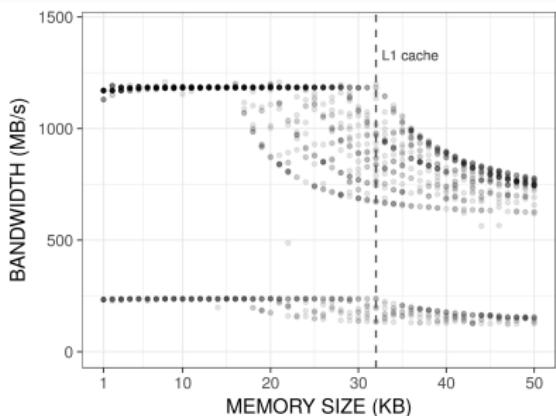
IMPACT OF OPERATING SYSTEM SCHEDULER

- Activating real-time kernel scheduler



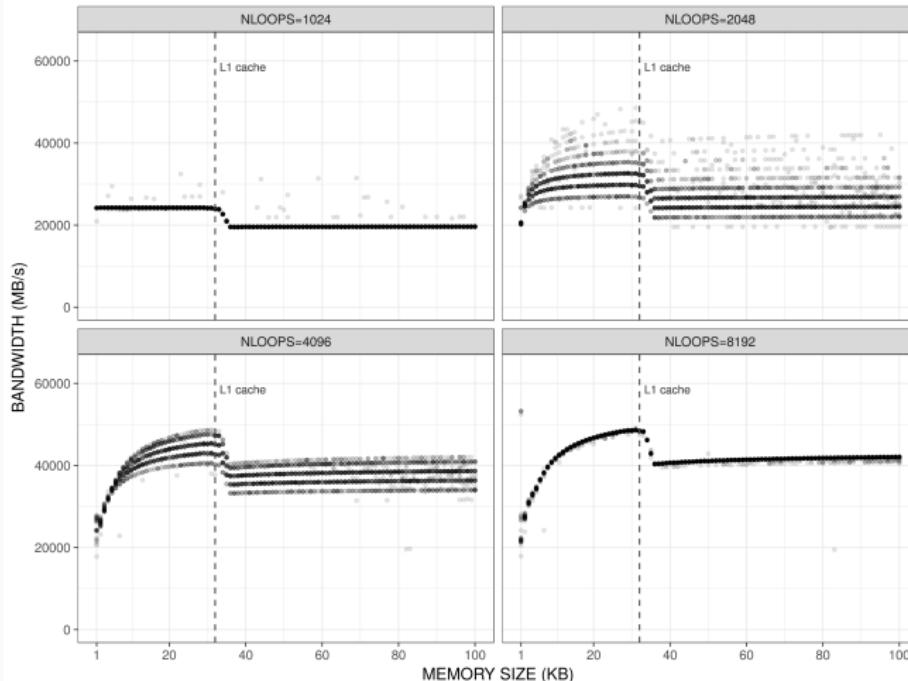
IMPACT OF OPERATING SYSTEM SCHEDULER

- Activating real-time kernel scheduler



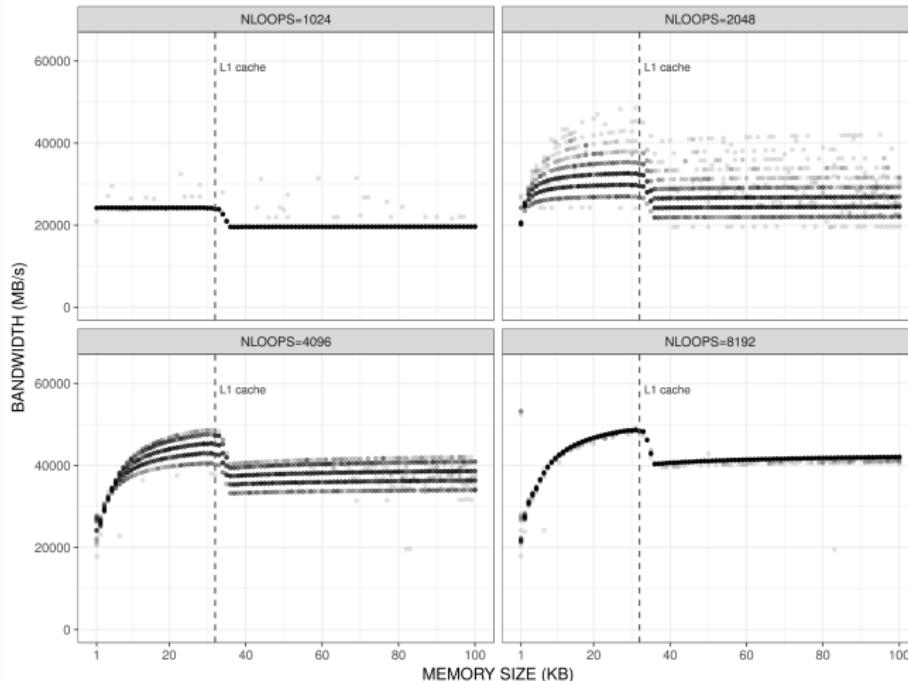
IMPACT OF REPETITIONS

- Remember `nloops` ?



IMPACT OF REPETITIONS DVFS

- Remember nloops ?



MPI PERFORMANCE CHARACTERIZATION: THE MANY BIASES

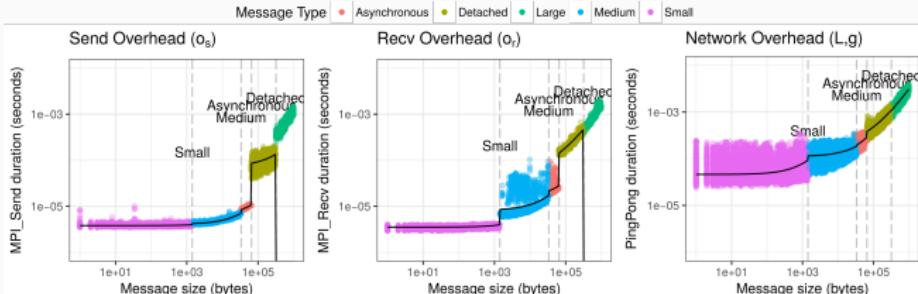
- Powers of two for message sizes ? Linear ?
- Sensitivity to temporal perturbations (in order N_{rep})
- Breakpoint detection (increasing message size, minimal range length, ...)
- Outlier removal (assumes normality, ignores uncertainty of previous measurements)

Measurement proposal:

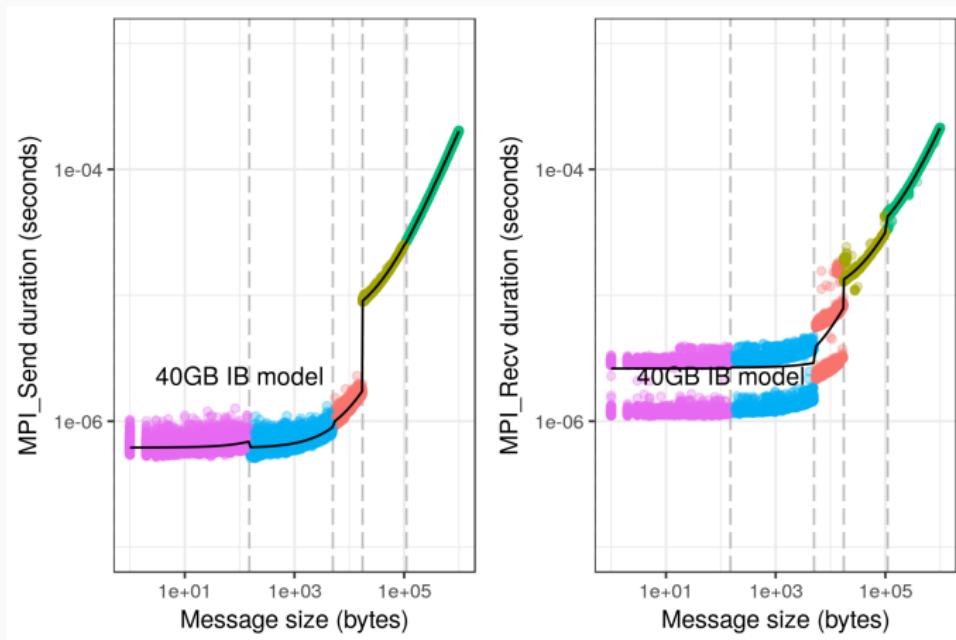
- Ping-Pong `data_size` = 10^X , where $X \sim \text{Unif}(\log_{10}(a), \log_{10}(b))$.
- Record the time taken in every `MPI_Send` and `MPI_Receive` operation

Analysis proposal:

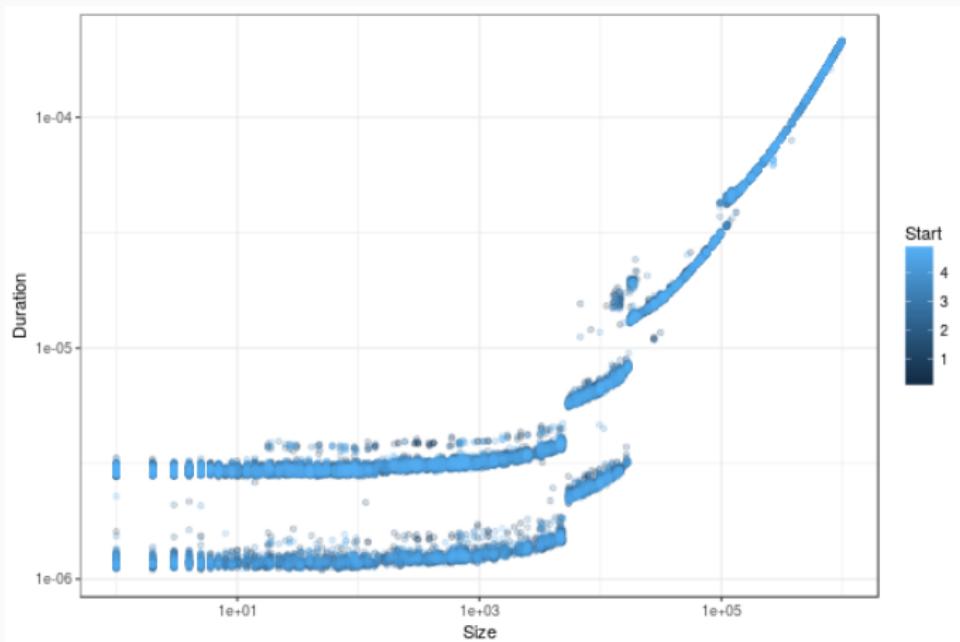
- Manually provided breakpoints
- Regression in R
- Inspect regression output and hypothesis (linearity, noise, "outliers")



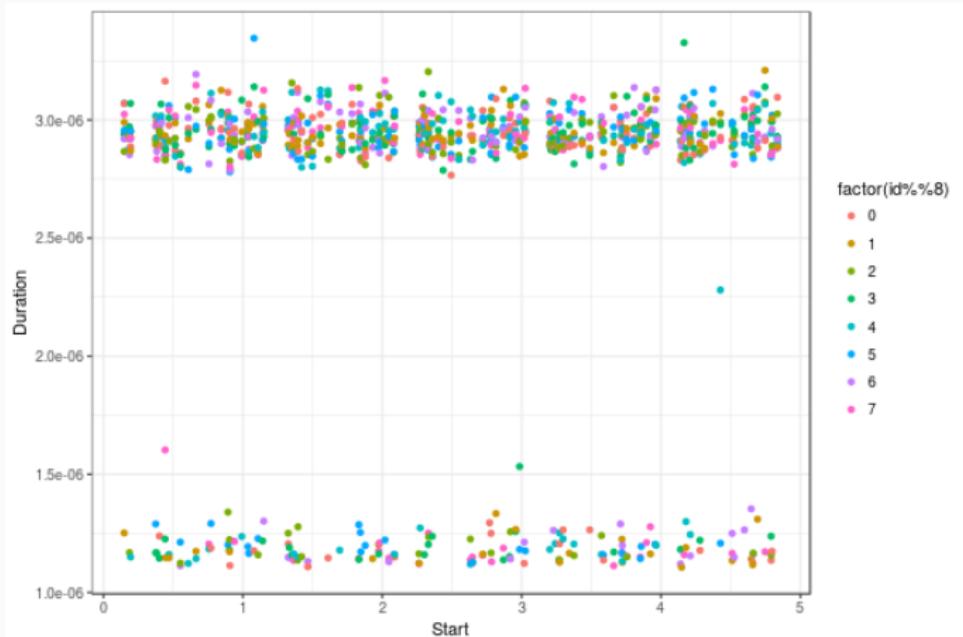
MPI MEASUREMENT: RANDOMIZATION IN ACTION (STAMPEDE@TACC)



MPI MEASUREMENT: RANDOMIZATION IN ACTION (STAMPEDE@TACC)

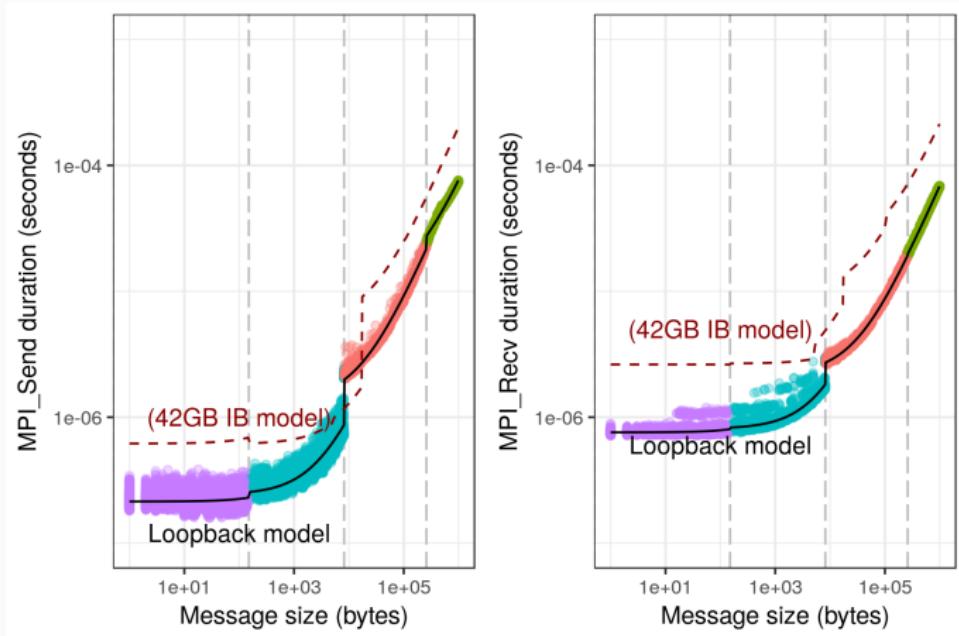


MPI MEASUREMENT: RANDOMIZATION IN ACTION (STAMPEDE@TACC)

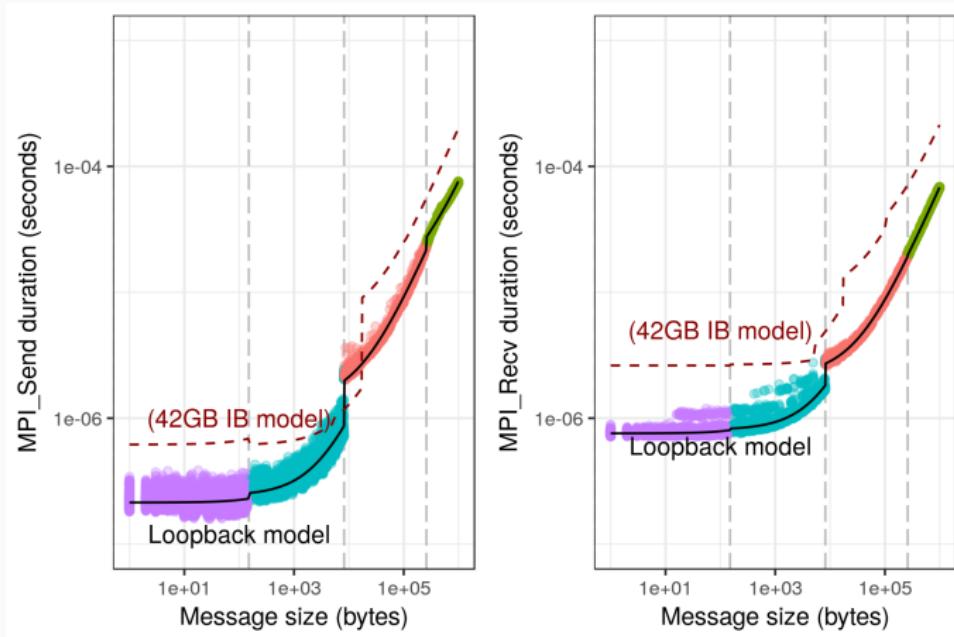


c558-[203,304]

MPI MEASUREMENT: RANDOMIZATION IN ACTION (STAMPEDE@TACC)



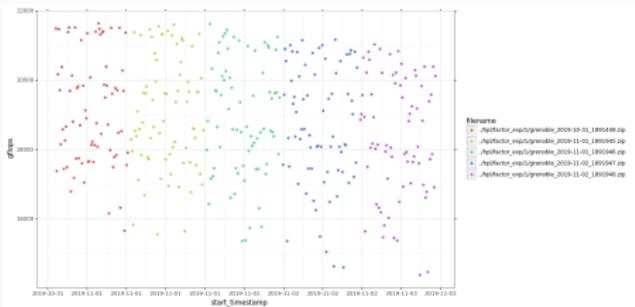
MPI MEASUREMENT: RANDOMIZATION IN ACTION (STAMPEDE@TACC)



Even the simpler physical quantities can be very tricky to measure because our models and protocols are often naive

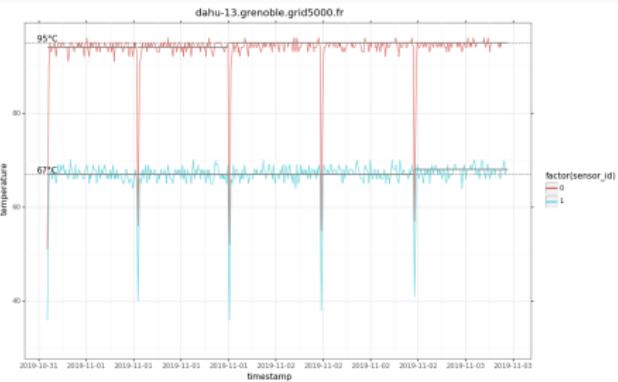
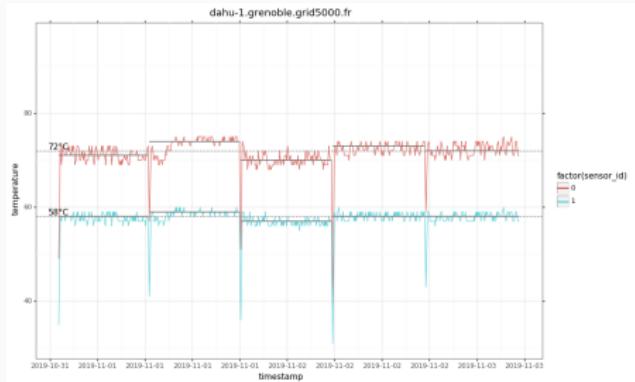
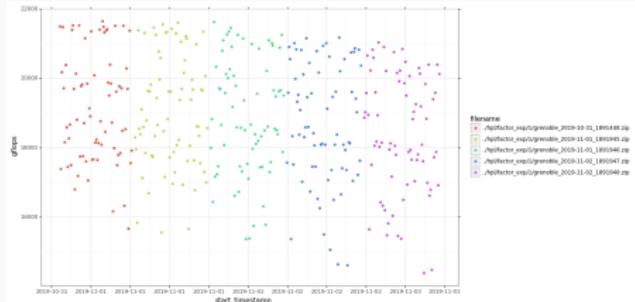
AVOIDING "TEMPORARY" PERTURBATIONS (RANDOMIZING A FACTORIAL DESIGN)

- HPL performance (32 nodes, 70 cfg., 5 repetitions) • Time scale = 3 days



AVOIDING "TEMPORARY" PERTURBATIONS (RANDOMIZING A FACTORIAL DESIGN)

- HPL performance (32 nodes, 70 cfg., 5 repetitions) • Time scale = 3 days



AVOIDING "TEMPORARY" PERTURBATIONS (RANDOMIZING A FACTORIAL DESIGN)

- HPL performance (32 nodes, 70 cfg., 5 repetitions) • Time scale = 3 days



Factor:
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
5510
5511
5512
5513
5514
5515
5516
5517
5518
5519
5520
5521
5522
5523
5524
5525
5526
5527
5528
5529
5530
5531
5532
5533
5534
5535
5536
5537
5538
5539
55310
55311
55312
55313
55314
55315
55316
55317
55318
55319
55320
55321
55322
55323
55324
55325
55326
55327
55328
55329
55330
55331
55332
55333
55334
55335
55336
55337
55338
55339
55340
55341
55342
55343
55344
55345
55346
55347
55348
55349
55350
55351
55352
55353
55354
55355
55356
55357
55358
55359
55360
55361
55362
55363
55364
55365
55366
55367
55368
55369
55370
55371
55372
55373
55374
55375
55376
55377
55378
55379
55380
55381
55382
55383
55384
55385
55386
55387
55388
55389
55390
55391
55392
55393
55394
55395
55396
55397
55398
55399
553100
553101
553102
553103
553104
553105
553106
553107
553108
553109
553110
553111
553112
553113
553114
553115
553116
553117
553118
553119
553120
553121
553122
553123
553124
553125
553126
553127
553128
553129
553130
553131
553132
553133
553134
553135
553136
553137
553138
553139
553140
553141
553142
553143
553144
553145
553146
553147
553148
553149
553150
553151
553152
553153
553154
553155
553156
553157
553158
553159
553160
553161
553162
553163
553164
553165
553166
553167
553168
553169
553170
553171
553172
553173
553174
553175
553176
553177
553178
553179
553180
553181
553182
553183
553184
553185
553186
553187
553188
553189
553190
553191
553192
553193
553194
553195
553196
553197
553198
553199
553200
553201
553202
553203
553204
553205
553206
553207
553208
553209
553210
553211
553212
553213
553214
553215
553216
553217
553218
553219
553220
553221
553222
553223
553224
553225
553226
553227
553228
553229
553230
553231
553232
553233
553234
553235
553236
553237
553238
553239
553240
553241
553242
553243
553244
553245
553246
553247
553248
553249
553250
553251
553252
553253
553254
553255
553256
553257
553258
553259
553260
553261
553262
553263
553264
553265
553266
553267
553268
553269
553270
553271
553272
553273
553274
553275
553276
553277
553278
553279
553280
553281
553282
553283
553284
553285
553286
553287
553288
553289
553290
553291
553292
553293
553294
553295
553296
553297
553298
553299
5532100
5532101
5532102
5532103
5532104
5532105
5532106
5532107
5532108
5532109
5532110
5532111
5532112
5532113
5532114
5532115
5532116
5532117
5532118
5532119
5532120
5532121
5532122
5532123
5532124
5532125
5532126
5532127
5532128
5532129
5532130
5532131
5532132
5532133
5532134
5532135
5532136
5532137
5532138
5532139
5532140
5532141
5532142
5532143
5532144
5532145
5532146
5532147
5532148
5532149
5532150
5532151
5532152
5532153
5532154
5532155
5532156
5532157
5532158
5532159
5532160
5532161
5532162
5532163
5532164
5532165
5532166
5532167
5532168
5532169
5532170
5532171
5532172
5532173
5532174
5532175
5532176
5532177
5532178
5532179
5532180
5532181
5532182
5532183
5532184
5532185
5532186
5532187
5532188
5532189
5532190
5532191
5532192
5532193
5532194
5532195
5532196
5532197
5532198
5532199
5532200
5532201
5532202
5532203
5532204
5532205
5532206
5532207
5532208
5532209
5532210
5532211
5532212
5532213
5532214
5532215
5532216
5532217
5532218
5532219
5532220
5532221
5532222
5532223
5532224
5532225
5532226
5532227
5532228
5532229
55322210
55322211
55322212
55322213
55322214
55322215
55322216
55322217
55322218
55322219
55322220
55322221
55322222
55322223
55322224
55322225
55322226
55322227
55322228
55322229
553222210
553222211
553222212
553222213
553222214
553222215
553222216
553222217
553222218
553222219
553222220
553222221
553222222
553222223
553222224
553222225
553222226
553222227
553222228
553222229
5532222210
5532222211
5532222212
5532222213
5532222214
5532222215
5532222216
5532222217
5532222218
5532222219
5532222220
5532222221
5532222222
5532222223
5532222224
5532222225
5532222226
5532222227
5532222228
5532222229
55322222210
55322222211
55322222212
55322222213
55322222214
55322222215
55322222216
55322222217
55322222218
55322222219
55322222220
55322222221
55322222222
55322222223
55322222224
55322222225
55322222226
55322222227
55322222228
55322222229
553222222210
553222222211
553222222212
553222222213
553222222214
553222222215
553222222216
553222222217
553222222218
553222222219
553222222220
553222222221
553222222222
553222222223
553222222224
553222222225
553222222226
553222222227
553222222228
553222222229
5532222222210
5532222222211
5532222222212
5532222222213
5532222222214
5532222222215
5532222222216
5532222222217
5532222222218
5532222222219
5532222222220
5532222222221
5532222222222
5532222222223
5532222222224
5532222222225
5532222222226
5532222222227
5532222222228
5532222222229
55322222222210
55322222222211
55322222222212
55322222222213
55322222222214
55322222222215
55322222222216
55322222222217
55322222222218
55322222222219
55322222222220
55322222222221
55322222222222
55322222222223
55322222222224
55322222222225
55322222222226
55322222222227
55322222222228
55322222222229
553222222222210
553222222222211
553222222222212
553222222222213
553222222222214
553222222222215
553222222222216
553222222222217
553222222222218
553222222222219
553222222222220
553222222222221
553222222222222
553222222222223
553222222222224
553222222222225
553222222222226
553222222222227
553222222222228
553222222222229
5532222222222210
5532222222222211
5532222222222212
5532222222222213
5532222222222214
5532222222222215
5532222222222216
5532222222222217
5532222222222218
5532222222222219
5532222222222220
5532222222222221
5532222222222222
5532222222222223
5532222222222224
5532222222222225
5532222222222226
5532222222222227
5532222222222228
5532222222222229
55322222222222210
55322222222222211
55322222222222212
55322222222222213
55322222222222214
55322222222222215
55322222222222216
55322222222222217
55322222222222218
55322222222222219
55322222222222220
55322222222222221
55322222222222222
55322222222222223
55322222222222224
55322222222222225
55322222222222226
55322222222222227
55322222222222228
55322222222222229
553222222222222210
553222222222222211
553222222222222212
553222222222222213
553222222222222214
553222222222222215
553222222222222216
553222222222222217
553222222222222218
553222222222222219
553222222222222220
553222222222222221
553222222222222222
553222222222222223
553222222222222224
553222222222222225
553222222222222226
553222222222222227
553222222222222228
553222222222222229
5532222222222222210
5532222222222222211
5532222222222222212
5532222222222222213
5532222222222222214
5532222222222222215
5532222222222222216
5532222222222222217
5532222222222222218
5532222222222222219
5532222222222222220
5532222222222222221
5532222222222222222
5532222222222222223
5532222222222222224
5532222222222222225
5532222222222222226
5532222222222222227
5532222222222222228
5532222222222222229
55322222222222222210
55322222222222222211
55322222222222222212
55322222222222222213
55322222222222222214
55322222222222222215
55322222222222222216
55322222222222222217
55322222222222222218
55322222222222222219
55322222222222222220
55322222222222222221
55322222222222222222
55322222222222222223
55322222222222222224
55322222222222222225
55322222222222222226
55322222222222222227
55322222222222222228
55322222222222222229
553222222222222222210
553222222222222222211
553222222222222222212
553222222222222222213
553222222222222222214
553222222222222222215
553222222222222222216
553222222222222222217
553222222222222222218
553222222222222222219
553222222222222222220
553222222222222222221
553222222222222222222
553222222222222222223
553222222222222222224
553222222222222222225
553222222222222222226
553222222222222222227
553222222222222222228
553222222222222222229
5532222222222222222210
5532222222222222222211
5532222222222222222212
5532222222222222222213
5532222222222222222214
5532222222222222222215
5532222222222222222216
5532222222222222222217
5532222222222222222218
5532222222222222222219
5532222222222222222220
5532222222222222222221
5532222222222222222222
5532222222222222222223
5532222222222222222224
5532222222222222222225
5532222222222222222226
5532222222222222222227
5532222222222222222228
5532222222222222222229
55322222222222222222210
55322222222222222222211
55322222222222222222212
55322222222222222222213
55322222222222222222214
55322222222222222222215
55322222222222222222216
55322222222222222222217
55322222222222222222218
55322222222222222222219
55322222222222222222220
55322222222222222222221
55322222222222222222222
55322222222222222222223
55322222222222222222224
55322222222222222222225
55322222222222222222226
55322222222222222222227
55322222222222222222228
55322222222222222222229
553222222222222222222210
553222222222222222222211
553222222222222222222212
553222222222222222222213
553222222222222222222214
553222222222222222222215
553222222222222222222216
553222222222222222222217
553222222222222222222218
5532

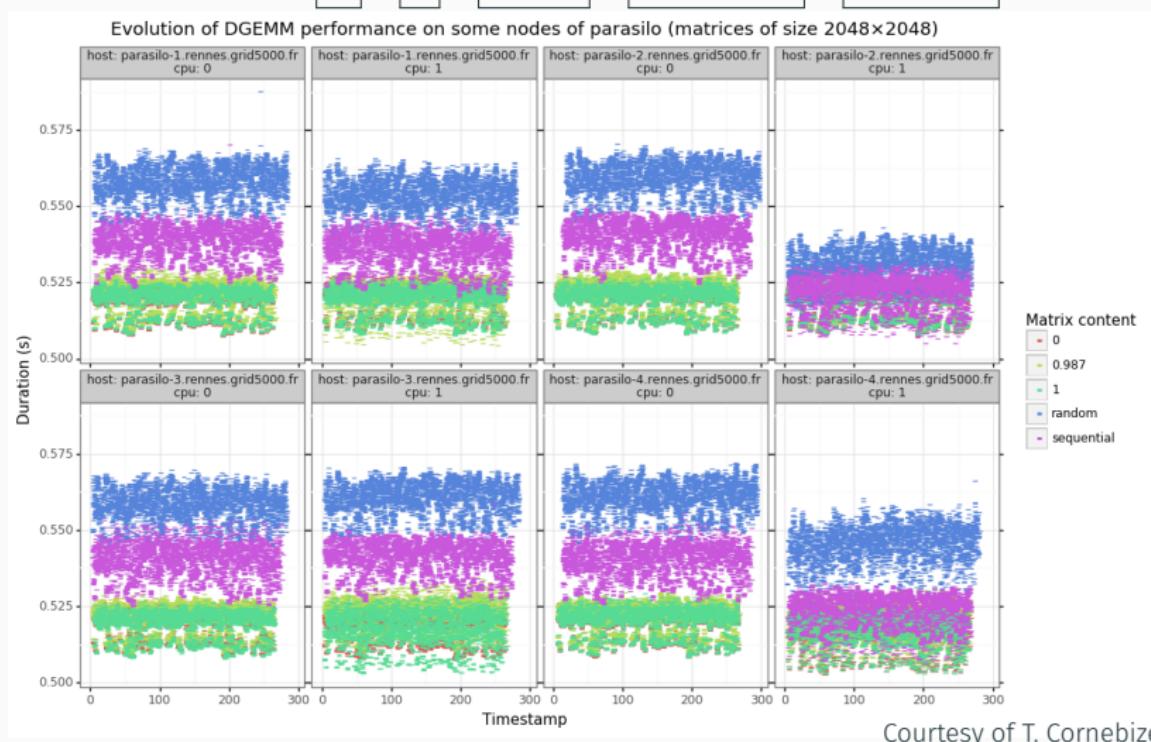
ON THE IMPORTANCE OF CONTENT INITIALIZATION

- $C = A \times A$ (2048×2048), independant
- Time scale = 5 minutes
- A initialized with

ON THE IMPORTANCE OF CONTENT INITIALIZATION

- $C = A \times A$ (2048×2048), independant
- A initialized with

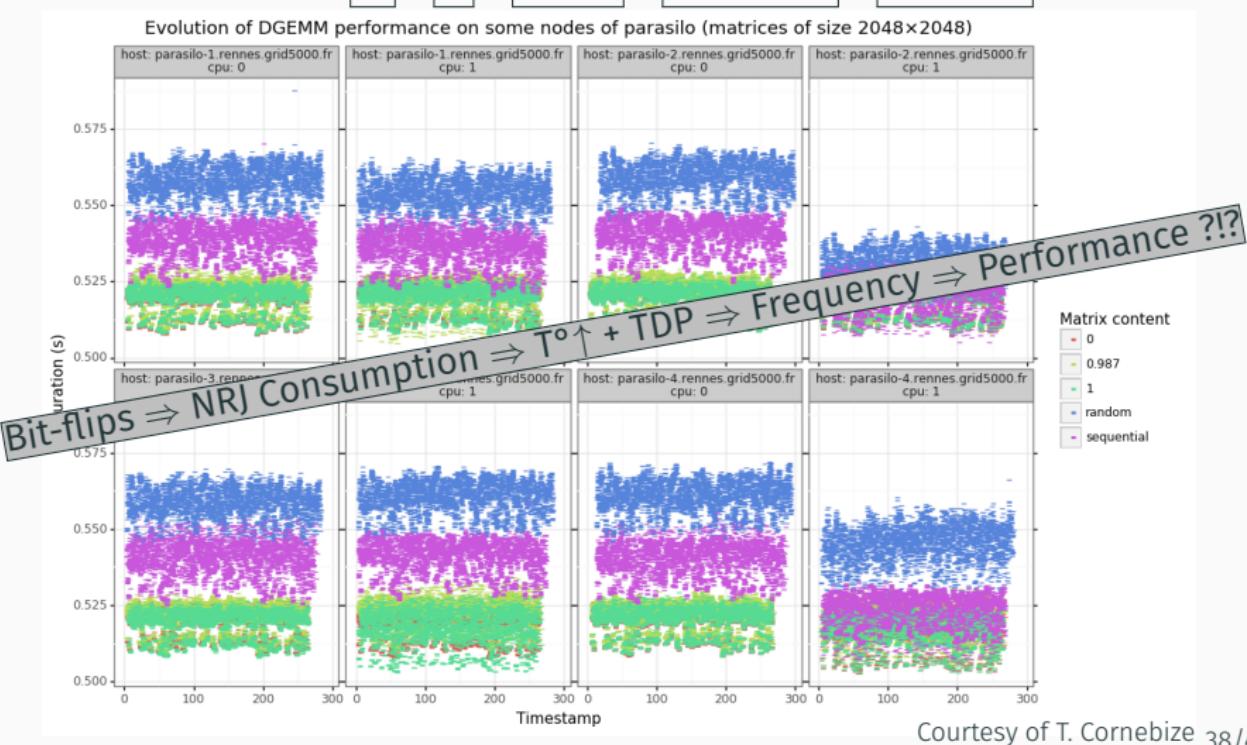
0	1	0.987	1, 2, 3, ...	random ?
---	---	-------	--------------	----------
- Time scale = 5 minutes



ON THE IMPORTANCE OF CONTENT INITIALIZATION

- $C = A \times A$ (2048×2048), independant
- A initialized with

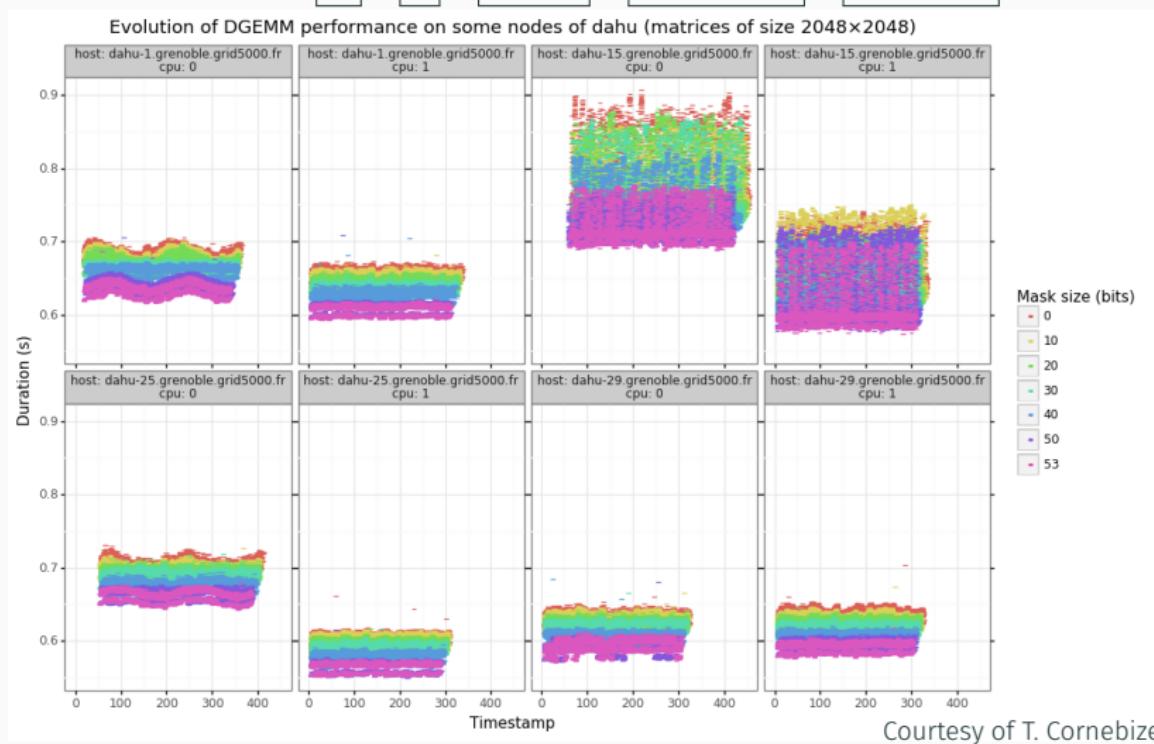
0	1	0.987	1, 2, 3, ...	random ?
---	---	-------	--------------	----------
- Time scale = 5 minutes



ON THE IMPORTANCE OF CONTENT INITIALIZATION

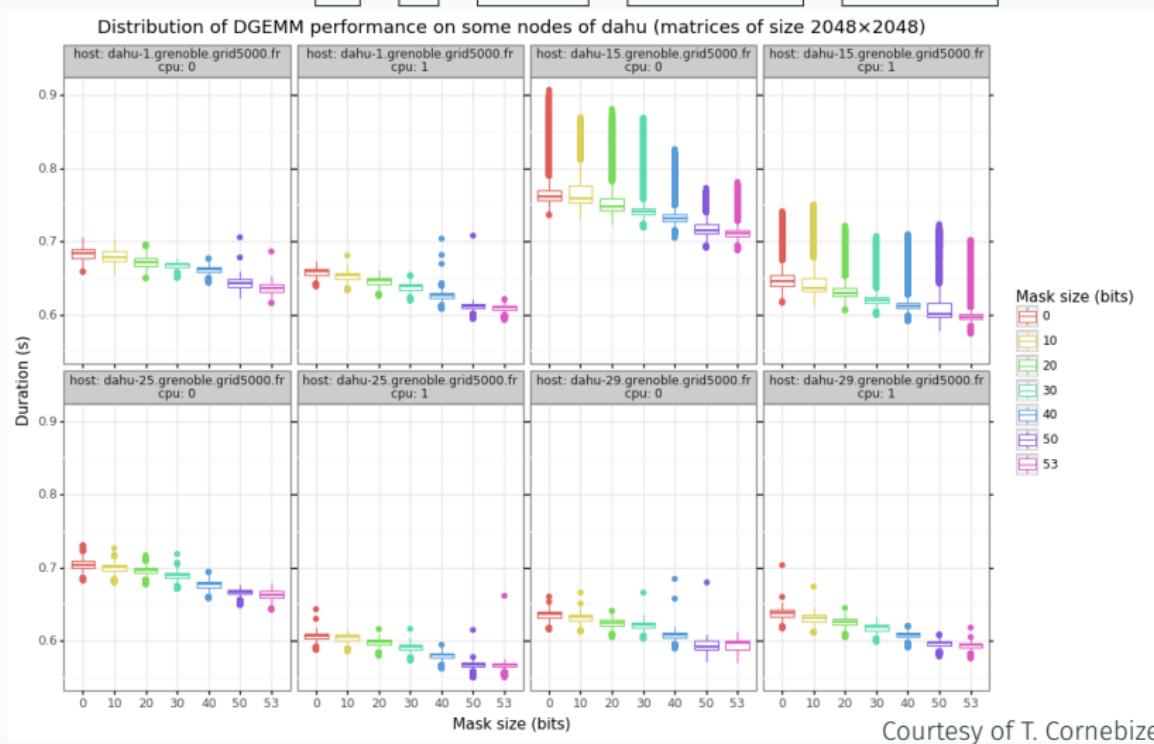
- $C = A \times A$ (2048×2048), independant
- A initialized with

0	1	0.987	1, 2, 3, ...	random ?
---	---	-------	--------------	----------
- Time scale = 5 minutes



ON THE IMPORTANCE OF CONTENT INITIALIZATION

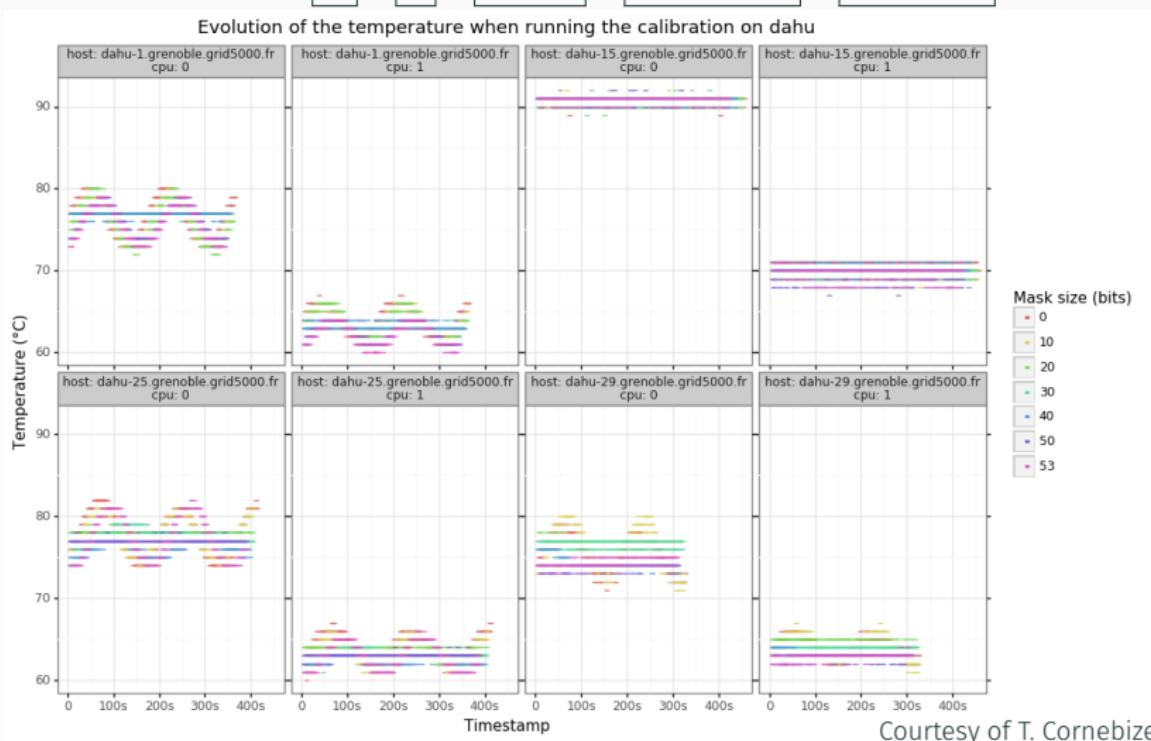
- $C = A \times A$ (2048×2048), independant
- A initialized with 0 1 0.987 $1, 2, 3, \dots$ $random?$
- Time scale = 5 minutes



ON THE IMPORTANCE OF CONTENT INITIALIZATION

- $C = A \times A$ (2048×2048), independant
- A initialized with

0	1	0.987	1, 2, 3, ...	random ?
---	---	-------	--------------	----------
- Time scale = 5 minutes



ON THE IMPORTANCE OF CONTENT INITIALIZATION

- $C = A \times A$ (2048×2048), independant
- A initialized with

0	1	0.987	1, 2, 3, ...	random ?
---	---	-------	--------------	----------
- Time scale = 5 minutes



CONCLUSION

REPRODUCIBLE RESEARCH = RIGOR AND TRANSPARENCY

To err is human.

Good research requires time and resources

1. Train yourself and your students: RR, statistics, experiments
 - Beware of checklists and norms
 - Understand what's at stake
2. Change the norm: make publication practices evolve
3. Incentive: consider RR/open science when hiring/promoting

REPRODUCIBLE RESEARCH = RIGOR AND TRANSPARENCY

To err is human.

Good research requires time and resources

1. Train yourself and your students: RR, statistics, experiments
 - Beware of checklists and norms
 - Understand what's at stake
2. Change the norm: make publication practices evolve
3. Incentive: consider RR/open science when hiring/promoting
4. Prepare the Future: Toward **literate experimentation?**
 - Reuse, reuse, reuse!
 - Shared and controlled testbeds (e.g., Grid'5000/SILECS)
 - How to share Experiments ?



SOME ADVERTISING

TEN YEARS REPRODUCIBILITY CHALLENGE

RESCIENCE SPECIAL ISSUE
FREE TO READ - FREE TO PUBLISH



Would you dare to run the code from your past self?
(the one that does not answer mail)

SUBMISSION DEADLINE 01/04/2020
<http://rescience.github.io/ten-years>
In association with Inria, CMS, Software Heritage, Rescience, Comité pour la Science Ouverte, DRFST Bordeaux & Mission de la pédagogie et du numérique pour l'enseignement supérieur.



3rd Edition: ongoing

A new MOOC: "Advanced RR" (2021)

- Software environment control
- Scientific workflow
- Managing data