# MASSIVIZING COMPUTER SYSTEMS

## MASSIVIZING HIGH PERFORMANCE COMPUTING FOR AI AND ML:
VU ON THE SCIENCE, DESIGN, AND ENGINEERING OF AI AND ML ECOSYSTEMS

@Large Research
Massivizing Computer Systems

http://atlarge.science

bit.ly/MassivizingHPC22

Massivizing =
Rich challenge of computer science → high societal impact!

Sponsored by:

NWO

@Alosup

VU VRIJE UNIVERSITEIT AMSTERDAM

Contributions from the MCS/AtLarge teams. Many thanks!
Many thanks to our collaborators, international working groups, authors of all images included here.
Also to Gianfranco Bilardi for invitation, ScalPerf for discussion.

Prof.dr.ir. Alexandru
IOSUP

# US IN 1 MINUTE

WE'RE MASSIVIZING COMPUTER SYSTEMS!

# VU AMSTERDAM < SCHIPHOL < THE NETHERLANDS < EUROPE

Amsterdam
founded 10th century
pop: 850,000


VU
founded 1880
pop: 23,500

The Netherlands


Europe


VU UNIVERSITY AMSTERDAM

# http://atlarge.science

VU VRIJE UNIVERSITEIT AMSTERDAM

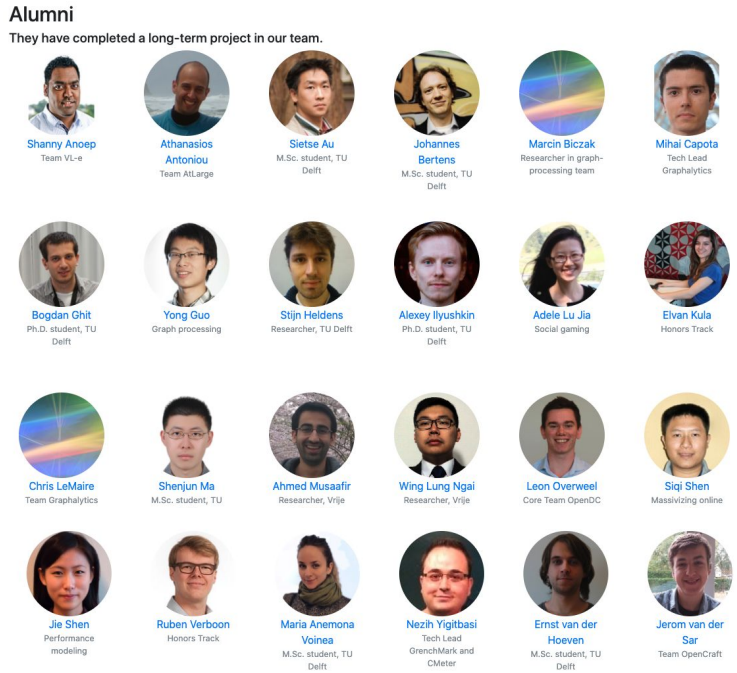
TU Delft

CURRENT TEAM

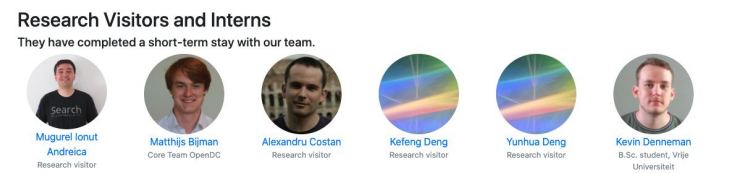

WE ARE HIRING
A NEW ASST. PROF.!

- Professor
- Assistant Prof.
- Teacher
- Visitor/P.-doc
- Ph.D. student
- Early Scientist

ALUMNI




VISITORS




WE ARE A FRIENDLY, DIVERSE GROUP, OF DIFFERENT RACES AND ETHNICITIES, GENDERS AND SEXUAL PREFERENCES, AND VIEWS OF CULTURE, POLITICS, AND RELIGION. YOU ARE WELCOME TO JOIN!

# WHO AM I?
# PROF. DR. IR. ALEXANDRU IOSUP

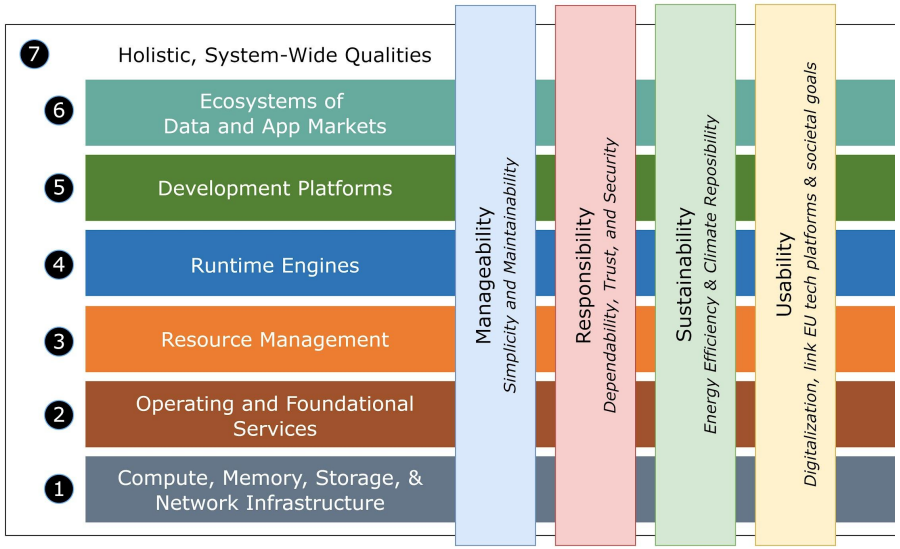
- Education, my courses:
  - > Honours Programme, Computer Org. (BSc)
  - > Distributed Systems, Cloud Computing (MSc)
- Research, 15 years in DistribSys:
  - > Massivizing Computer Systems
- About me:
  - > Worked in 7 countries, NL since 2004
  - > I like to help… I train people in need
  - > VU University Research Chair + Group Chair
  - > NL ICT Researcher of the Year
  - > NL Higher-Education Teacher of the Year
  - > NL Young Royal Academy of Arts & Sciences
  - > Knighted in 2020


VU

# WE ARE ALIGNED WITH COMMUNITY CONCERNS…

## The Manifesto on Computer Systems and Networking Research

### Clear vision for the field in the NL, 2021-2035




### Signed

50+ PIs / Leads
7 universities
5 relevant societal stakeholders

### Available

Full version (40+ pages) https://arxiv.org/pdf/2206.03259
Who's Who in CompSysNL? https://bit.ly/CompSysNLWhosWho

# ONE PROJECT TO MENTION…

## Big Graph Processing: Used in AI/ML, FinTech, Sci/Pharma, Industry 4.0, Energy Mgmt.*, etc.

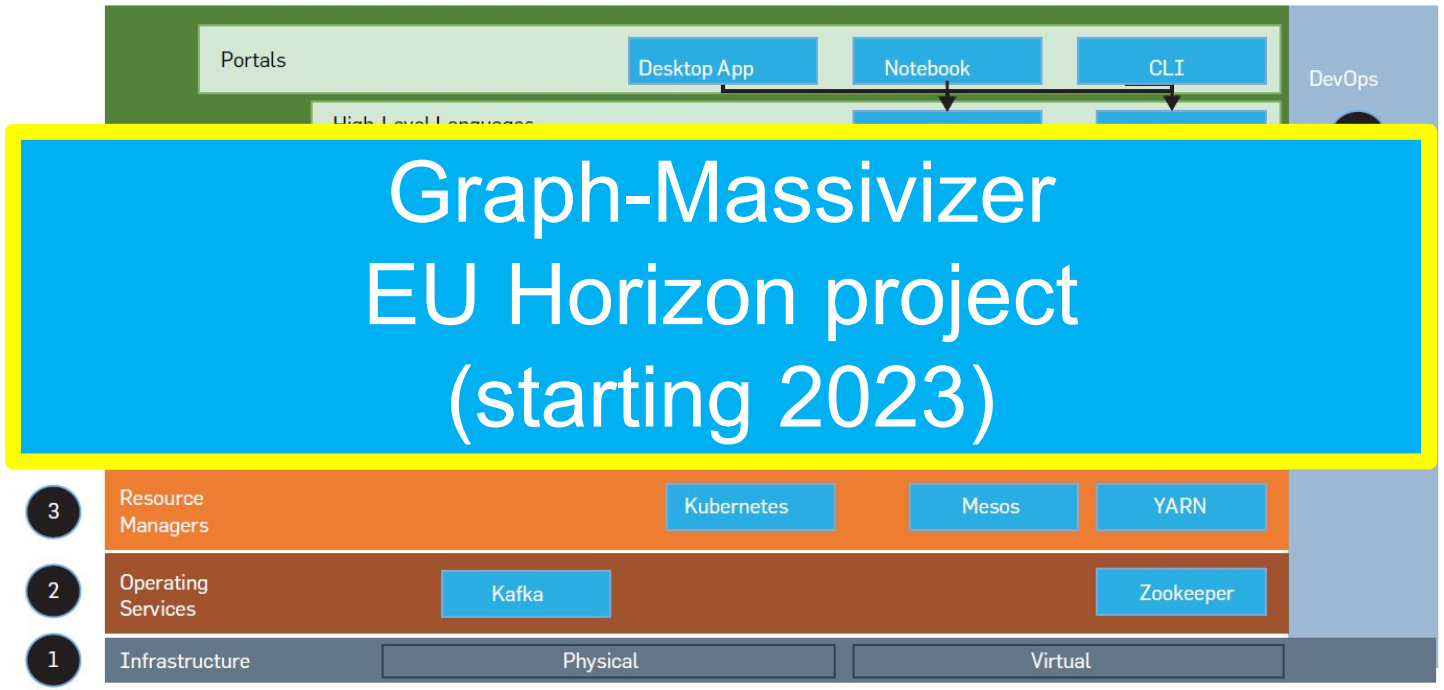
**Vision: Massivizing computer systems approaches are key to enable big graph ecosystems**





CACM Cover/Featured article, Sep 2021

Sakr, Bonifati, Voigt, Iosup, et al. (2021) The Future Is Big Graphs! CACM.




(*) Digital twin for datacenters, with partners CINECA, UniBo, etc.

Prodan et al. (2022) The GraphMassivizer project: Towards Extreme and Sustainable Graph Processing for Urgent Societal Challenges in Europe.

# THIS IS THE GOLDEN AGE OF COMPUTER ECOSYSTEMS

1

# THIS IS THE GOLDEN AGE OF MASSIVE COMPUTER ECOSYSTEMS

+ AI
+ ML

Education for Everyone (Online)

Business Services

Big Data

Edge / IoT / Fog / … Computing

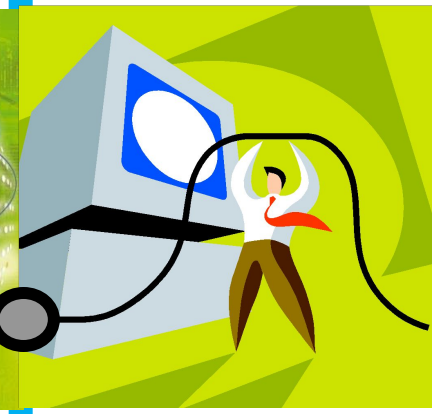
Big Science

Internet

5G/6G/*

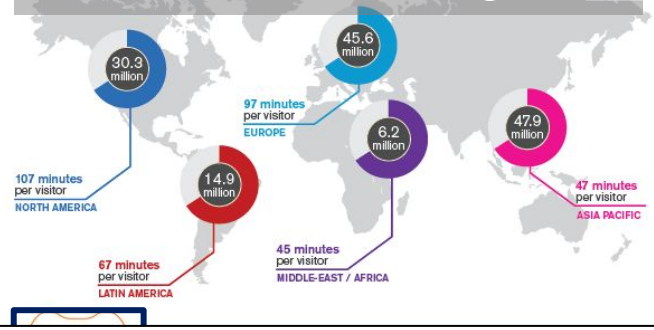
THz

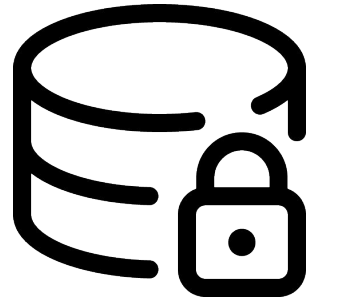
VLC



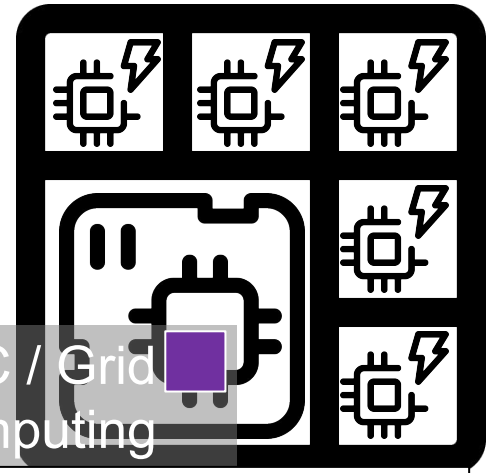
Cloud Computing

Online Gaming



Datacenter



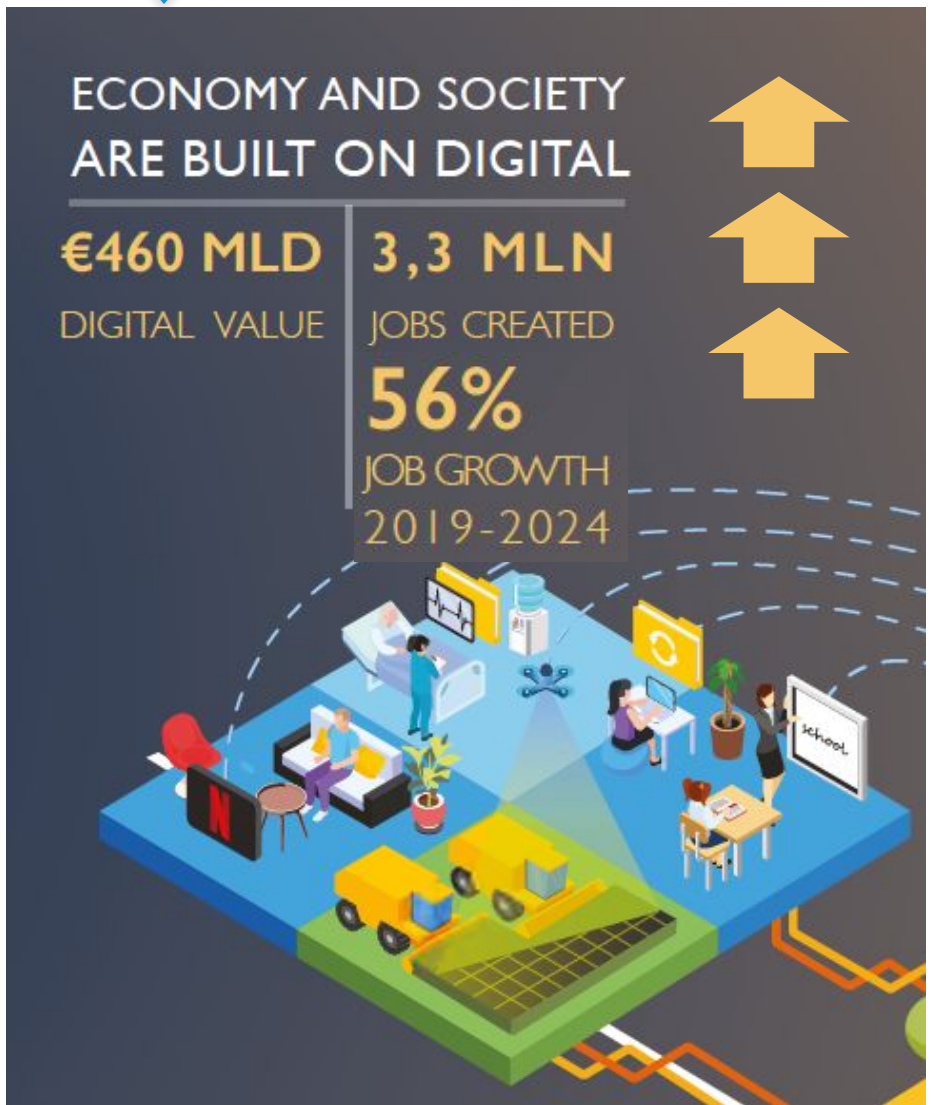
HPC / Grid Computing

ABN·AMRO

Daily Life

Iosup et al., Massivizing Computer Systems, ICDCS 2018. [Online] Hesselman, Grosso, Kuipers, et al. (2020) A Responsible Internet to increase Trust in the Digital World. INSM [Online]

# THE ECONOMIC IMPACT OF MASSIVE COMPUTER ECOSYSTEMS




DIVERSE SERVICES FOR ALL

EVERY €1 → €15 ADDED VALUE

**Impacting ≥60% of the NL GDP (1 trillion EUR/y)**

**Attracting ≥20% of all foreign direct investments in NL**

Sources: Iosup et al., Massivizing Computer Systems, ICDCS 2018 [Online] / Dutch Data Center Association, 2020 [Online] / Growth: NL Gov't, Flexera, Binx 2020. Gartner 2019. IA 2017.

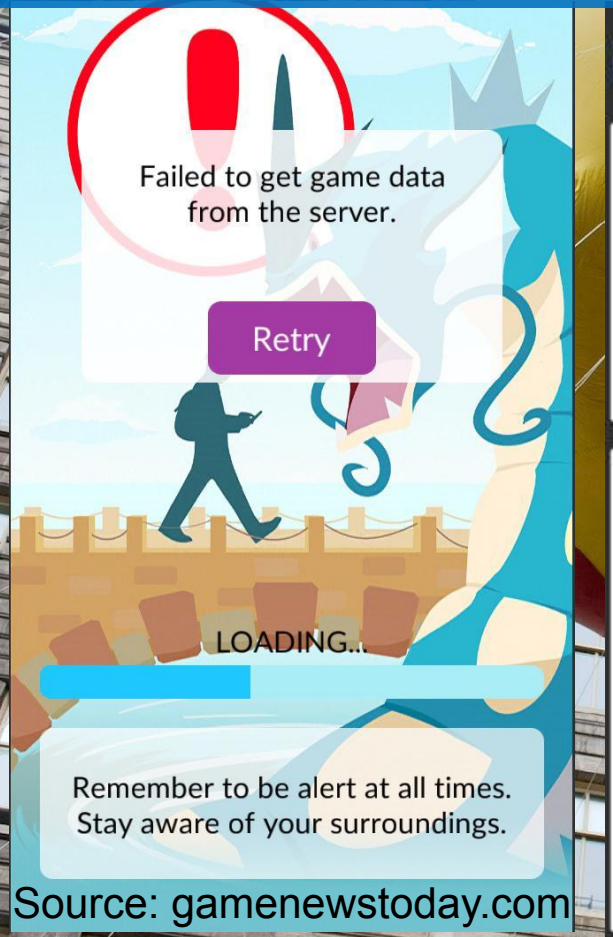
# BUT WE CANNOT TAKE THIS TECHNOLOGY FOR GRANTED

2

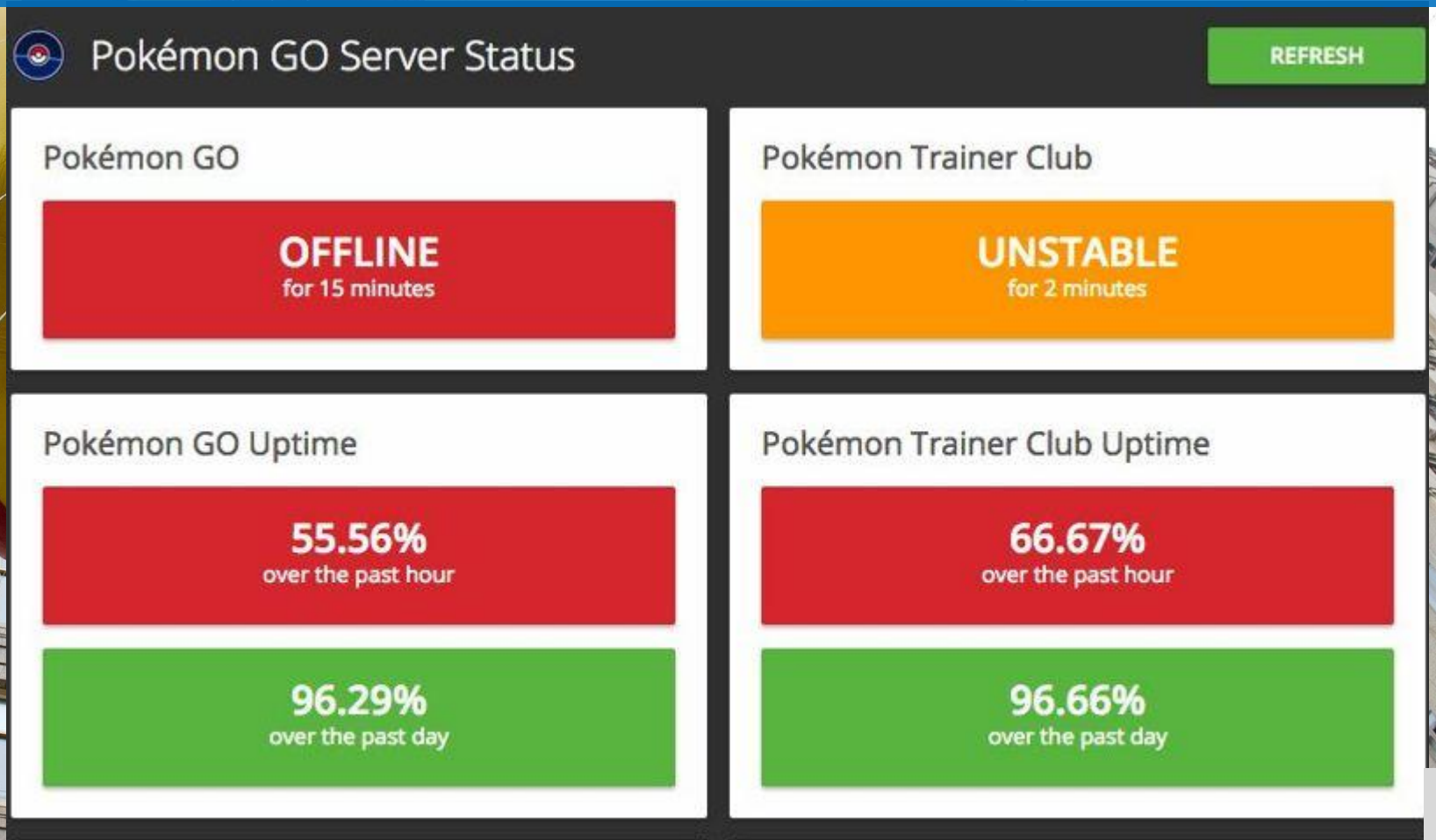
(So, this is why I am giving this talk)

# PHENOMENON: FAILURES IN CLOUD SERVICES

## UNCOVERING THE PRESENCE OF FAILURES




Source: gamenewstoday.com



# PHENOMENON: PERFORMANCE IN CLOUD SERVICES

## UNCOVERING THE PRESENCE OF PERFORMANCE ISSUES, EVEN LEADING TO CRASHES

Polygon

Source: http://bit.ly/EveOnline21Crash

NEWS

**Players in Eve Online broke a world record — and then the game itself**

*Developers said they're not 'able to predict the server performance in these kinds of situations'*

By Charlie Hall | @Charlie_L_Hall | Jan 5, 2021, 2:54pm EST

Source: Razorien/CCP Games

# PHENOMENON: CLOUD DATACENTER SUSTAINABILITY

## UNCOVERING THE USE OF ENERGY AND WATER, THE IMPACT ON CLIMATE

Power consumption of datacenters:
**>1% of global electricity**

Source: Nature, 2018 [Online]

Water consumption of datacenters in the US:
**>625Bn. l/y (0,1%)**

Source: Energy Technologies Area, 2016 [Online]

Power consumption of datacenters in the Netherlands:
**1→3% of national electricity**

Source: NRC, 2019 [Online]

Other greenhouse emissions:
**Largely unknown**

Source: Nature Climate Change, 2020 [Online]

Source: NASA Earth Observatory

# THIS TALK: MASSIVIZING = LET'S THINK ECOSYSTEMS!

WE TAKE A HOLISTIC VIEW, BASED ON COMPUTER ECOSYSTEMS

Technology **not ready, many issues***

A Why does this* happen?

B What to do about it*?

* In modern computer systems, issues are often linked.

Source: Alexandru's personal library.

# A new science, of complex, smart computer ecosystems

3

(operational simplicity for the user)

# AN ANALOGY: MASSIVIZING CLIMATE SCIENCE

## TAKE A HOLISTIC VIEW, BASED ON COUPLED NATURAL SYSTEMS

Can be understood only with coupled models

Climatologist Bjorn Stevens. . Source: HBOWire

* In climate science, issues are often linked.
The same occurs in massive computer (eco)systems.

# ECOSYSTEM VS. SYSTEMS, A PRIMER

Structure: composites of smaller assemblies, but for ecosystems some assemblies are produced elsewhere, by teams with different practices

Operation: ecosystems exhibit many unknown, possibly emergent, phenomena, dynamics, and socio-technical issues

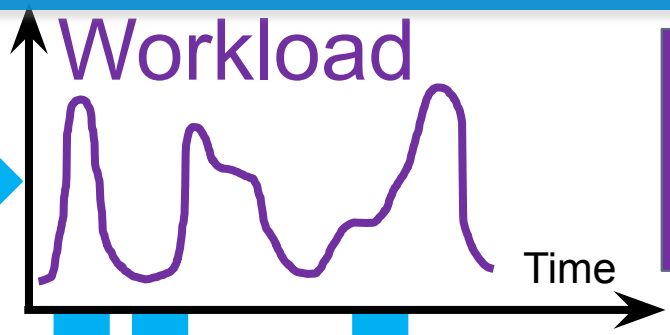
Lifecycle: some ecosystem constituents will perish, or be replaced with others that may not actually fulfill the needs

In plain English: in modern computer systems, several or all issues may be linked. Thus, looking at any single issue for an isolated system is no longer sufficient. We are now creating the science that explains this—Massivizing ...

# ECOSYSTEM = SERVICES + COMPUTING + SMARTS + GOALS

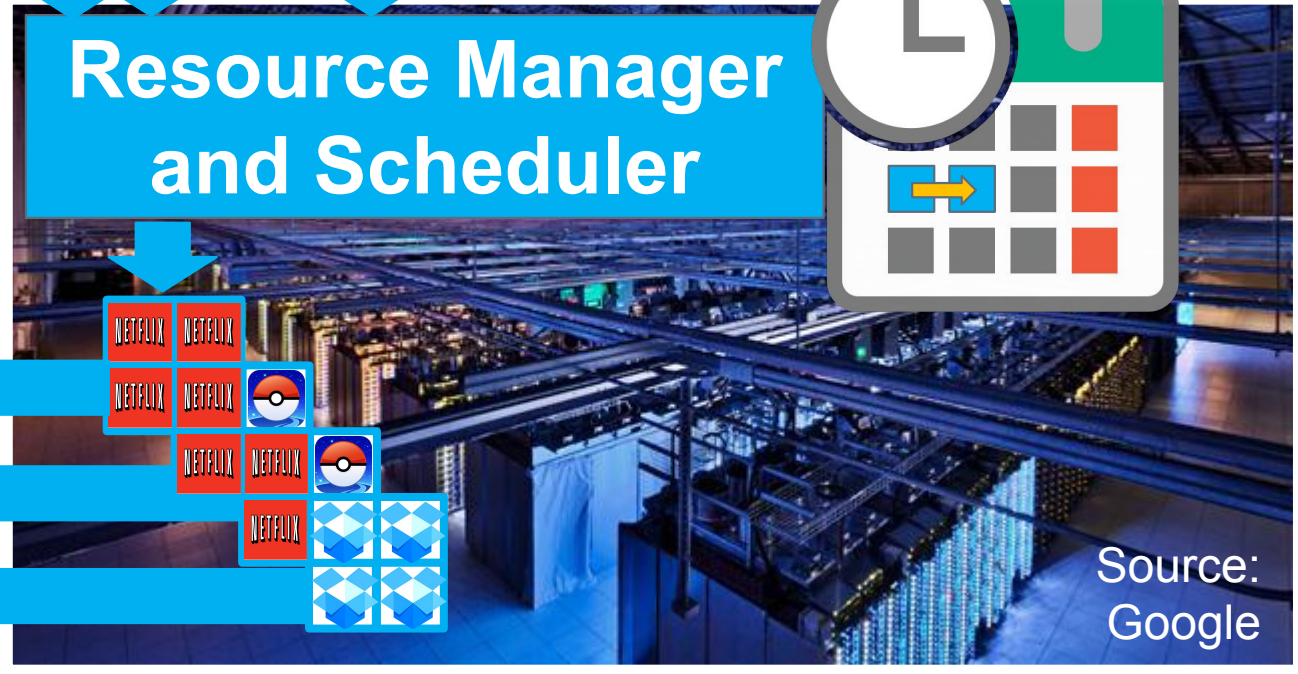
Creators


Digital Services




**100s of services … an ecosystem**

**Resource Manager and Scheduler**




Computing Continuum

Extreme Automation, Performance, Dependability, Sustainability

# DISTRIBUTED ECOSYSTEMS, OUR DEFINITION

1. Set of 2+ constituents, often heterogeneous
2. Each constituent is a system or an ecosystem (recursively)
3. Constituents are autonomous, cooperative or in competition
4. Ecosystem structure and organization ensure responsibility
    1. Completing functions and providing services
    2. Providing desirable non-functional properties
    3. Fulfill agreements with both operators and clients, clients in the loop
5. Long and short-term dynamics occur in the ecosystem

| |
|---|
| Iosup et al., Lecture Notes in Distributed Systems, Section 1.1.1 |
| Iosup et al., Massivizing Computer Systems, ICDCS 2018. [Online] |

# 8 PRINCIPLES OF DISTRIBUTED SYSTEMS & ECOSYSTEMS

P1: The Golden Age

P2: Design for massive scale, focus on scalability, elasticity

P3: Phenomena to discover, seed innovation

P4: Inherent functional requirements

P5: Non-functional requirements

P6: Resource management and scheduling

P7: Interaction programming model - system architecture

P8: Super-distribution

FIND THEM HERE → Iosup et al. Distributed Sytems and Ecosystems,

# High-level, theoretical approaches

∧

(our theories are often frameworks, designs, etc.)

# SERVERLESS AI/ML/DL OPERATIONS

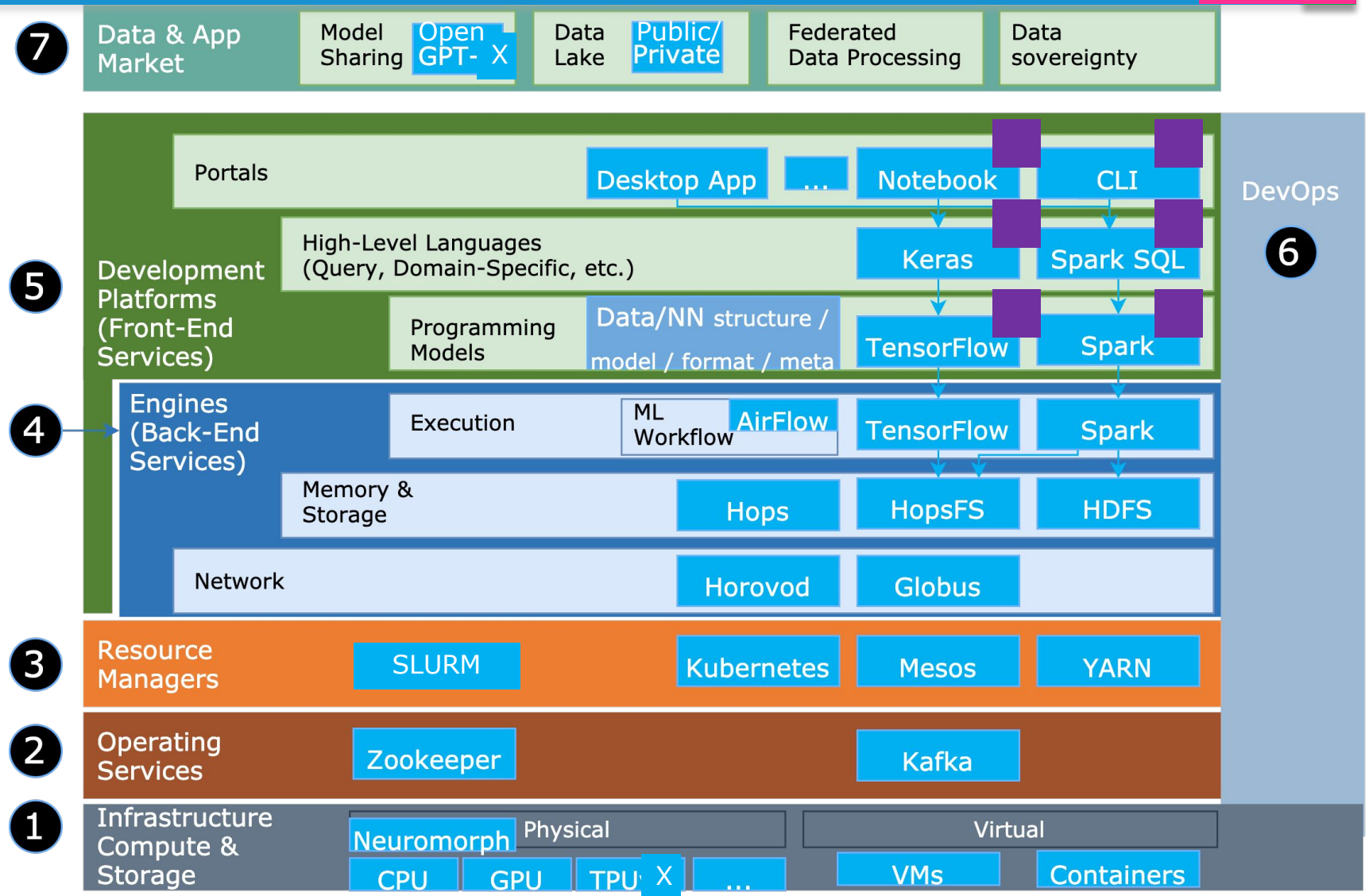
ISSUES: COMPLEXITY, NON-TECHNICAL

Actual ML app is a very small part!

Rest is systems, HW+SW, including HPC

Adapted from:

Sakr, Bonifati, Voigt, Iosup, et al. (2021) The Future Is Big Graphs! CACM



# SERVERLESS AI/ML/DL OPERATIONS

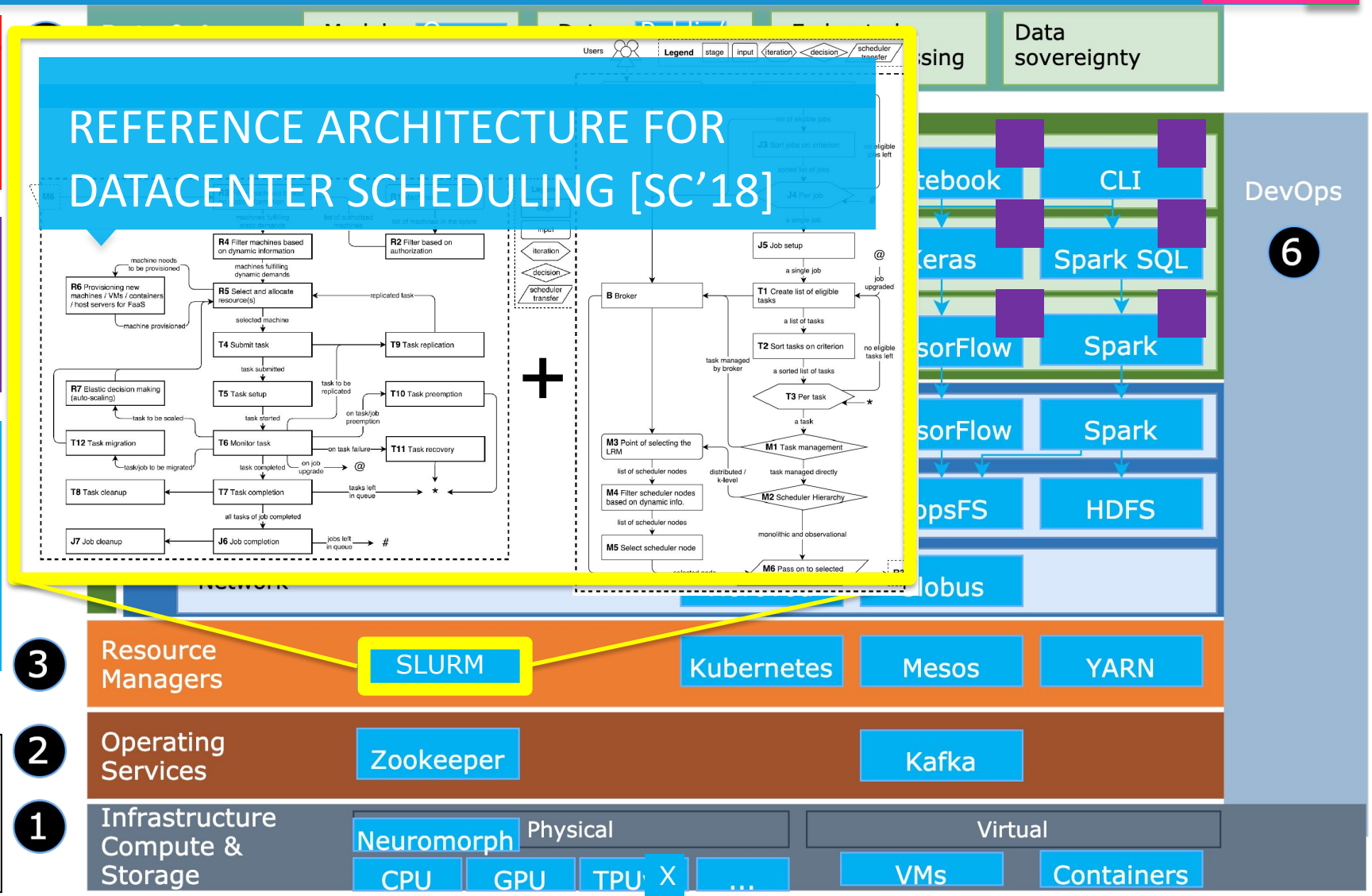
ISSUES: COMPLEXITY, NON-TECHNICAL

Actual ML app is a very small part!

Rest is systems, HW+SW, including HPC

Adapted from:

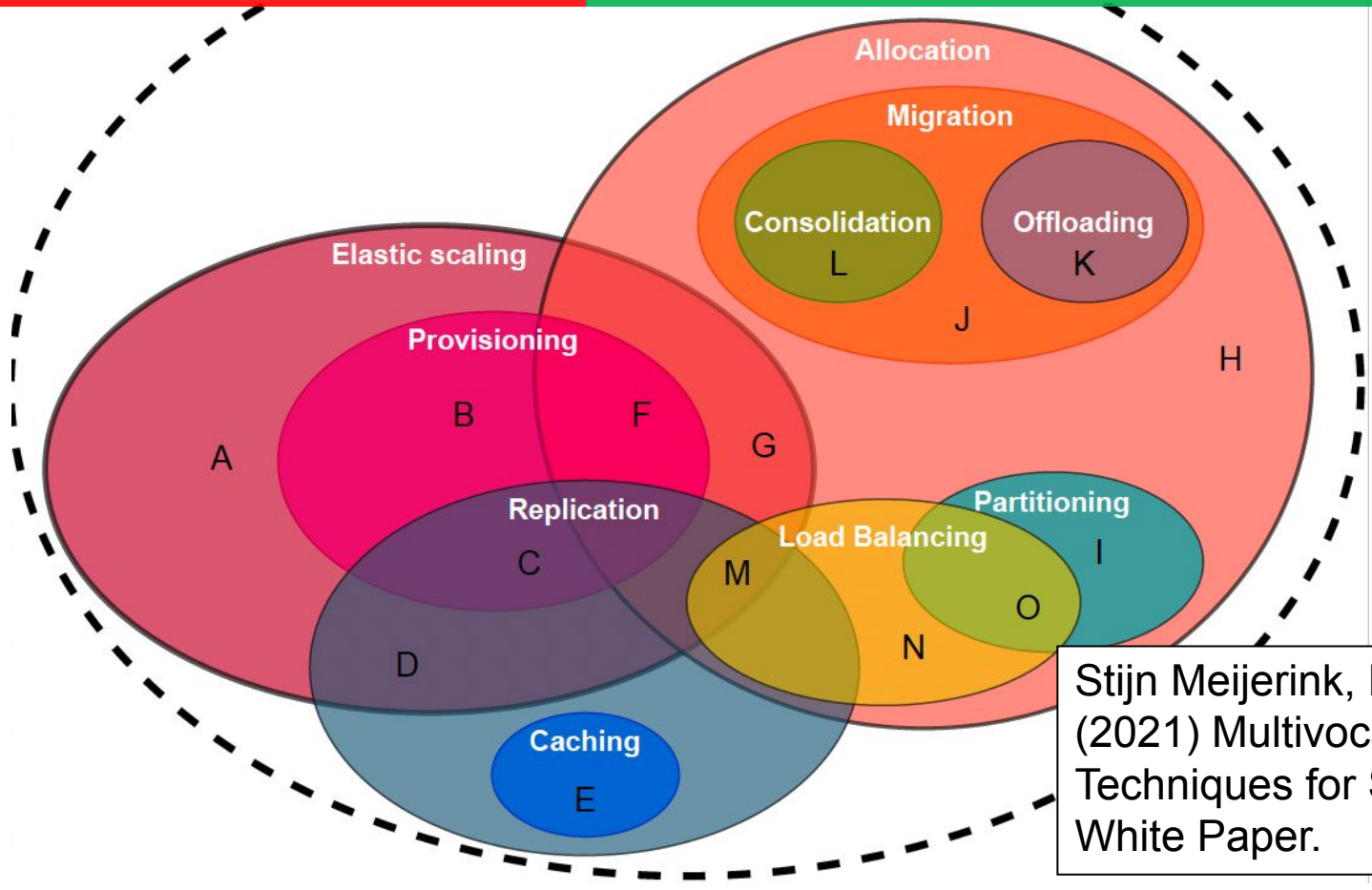
Sakr, Bonifati, Voigt, Iosup, et al. (2021) The Future Is Big Graphs! CACM



# HOW TO DO RM&S ACROSS THE ECOSYSTEM?

## IT'S OPERATIONS!

## REFERENCE VIEW ON OPERATIONAL TECHNIQUES



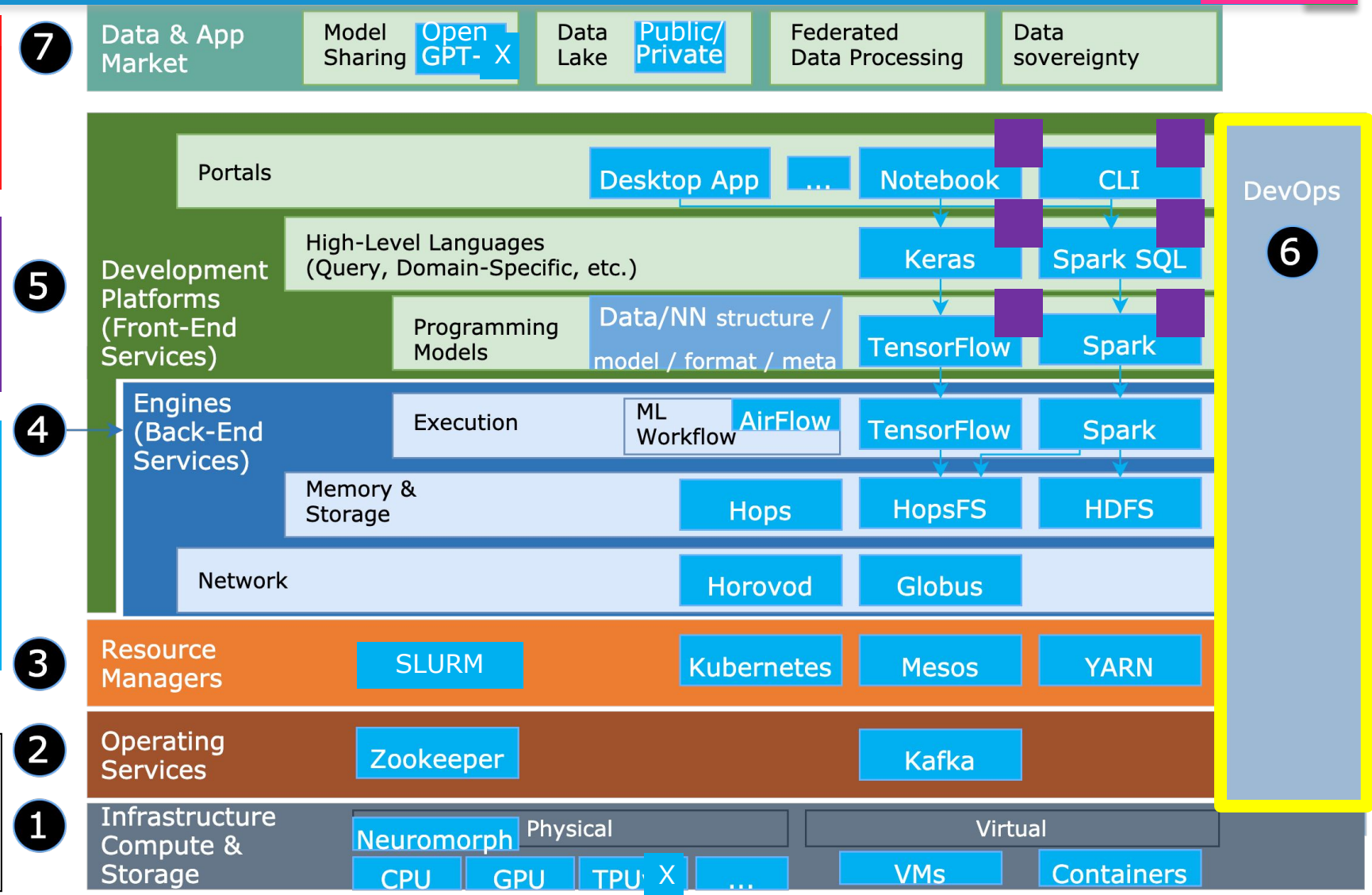

Stijn Meijerink, Erwin van Eyk, Alexandru Iosup (2021) Multivocal Survey of Operational Techniques for Serverless Computing. White Paper.

# SERVERLESS AI/ML/DL OPERATIONS

ISSUES: COMPLEXITY, NON-TECHNICAL

Actual ML app is a very small part!

Rest is systems, HW+SW, including HPC



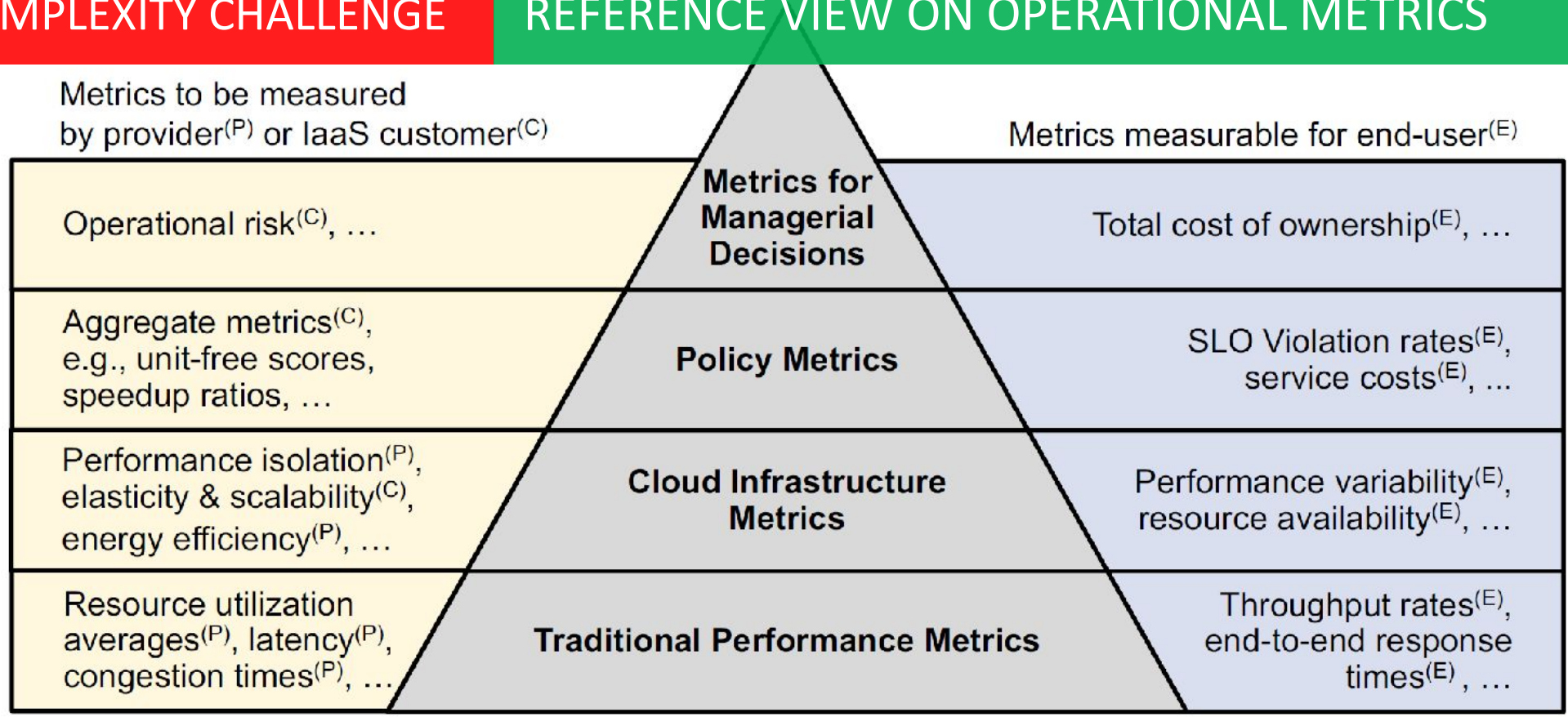

Adapted from:

Sakr, Bonifati, Voigt, Iosup, et al. (2021) The Future Is Big Graphs! CACM

# HOW TO REPORT PERFORMANCE?

THE COMPLEXITY CHALLENGE

REFERENCE VIEW ON OPERATIONAL METRICS




N. Herbst, E. Van Eyk, C. L. Abad, A. Iosup, et al. (2018) Quantifying Cloud Performance and Dependability: Taxonomy, Metric Design, and Emerging Challenges. TOMPECS 3(4): 19:1-19:36

# SERVERLESS AI/ML/DL OPERATIONS

```
cwlVersion: v1.0
class: CommandLineTool

doc: Spoa is a partial order aligment…

inputs:
  readsFA:
    type: File
    format: edam:format_1929
    doc: FASTA file containing a set of sequences.

requirements:
  InlineJavascriptRequirement: {}
hints:
  DockerRequirement:
    dockerPull: "quay.io/biocontainers/spoa:3.4.0--hc9558a2_0"
  ResourceRequirement:
    ramMin: $(15 * 1024)
    outdirMin: $(Math.ceil(inputs.readsFA.size/(1024*1024*1024) + 20))

baseCommand: spoa

arguments: [ $(inputs.readsFA), -G, -g, '-6' ]

stdout: $(inputs.readsFA.nameroot).g6.gfa

outputs:
  spoaGFA:
    type: stdout
    format: edam:format_3976
    doc: result in Graphical Fragment Assembly (GFA) format

$namespaces:
  edam: http://edamontology.org
```

1. Community Maintained File Format Identifier
2. Software Container
3. Dynamic Resource Requirements

Crusoe et al. (2022) Methods Included: Standardizing Computational Reuse and Portability with the Common

Model Sharing: Open GPT- X | Data Lake: Public/Private | Federated Data Processing | Data sovereignty

Desktop App | … | Notebook | CLI

High-Level Languages (Query, Domain-Specific, etc.)

Keras | Spark SQL

DevOps 6

Local execution on Linux, macOS, and MS Windows via the CWL reference implementation (`cwltool`) and Docker/uDocker/Singularity/podman/...

Authors of CWL tool and workflow descriptions

CWL

BATCH HTC

CLOUD

HTCondor | IBM Spectrum LSF | gridengine | slurm workload manager | PBS Works | Apache MESOS | kubernetes | Google Cloud | openstack. | amazon web services EC2 | Azure

Backends supported by various F/OSS CWL implementations

2 Operating Services

1 Infrastructure Compute & Storage

Neuromorph | Physical | CPU | GPU | TPU | X | … | Virtual | VMs | Containers

# Detailed, practical approaches

+

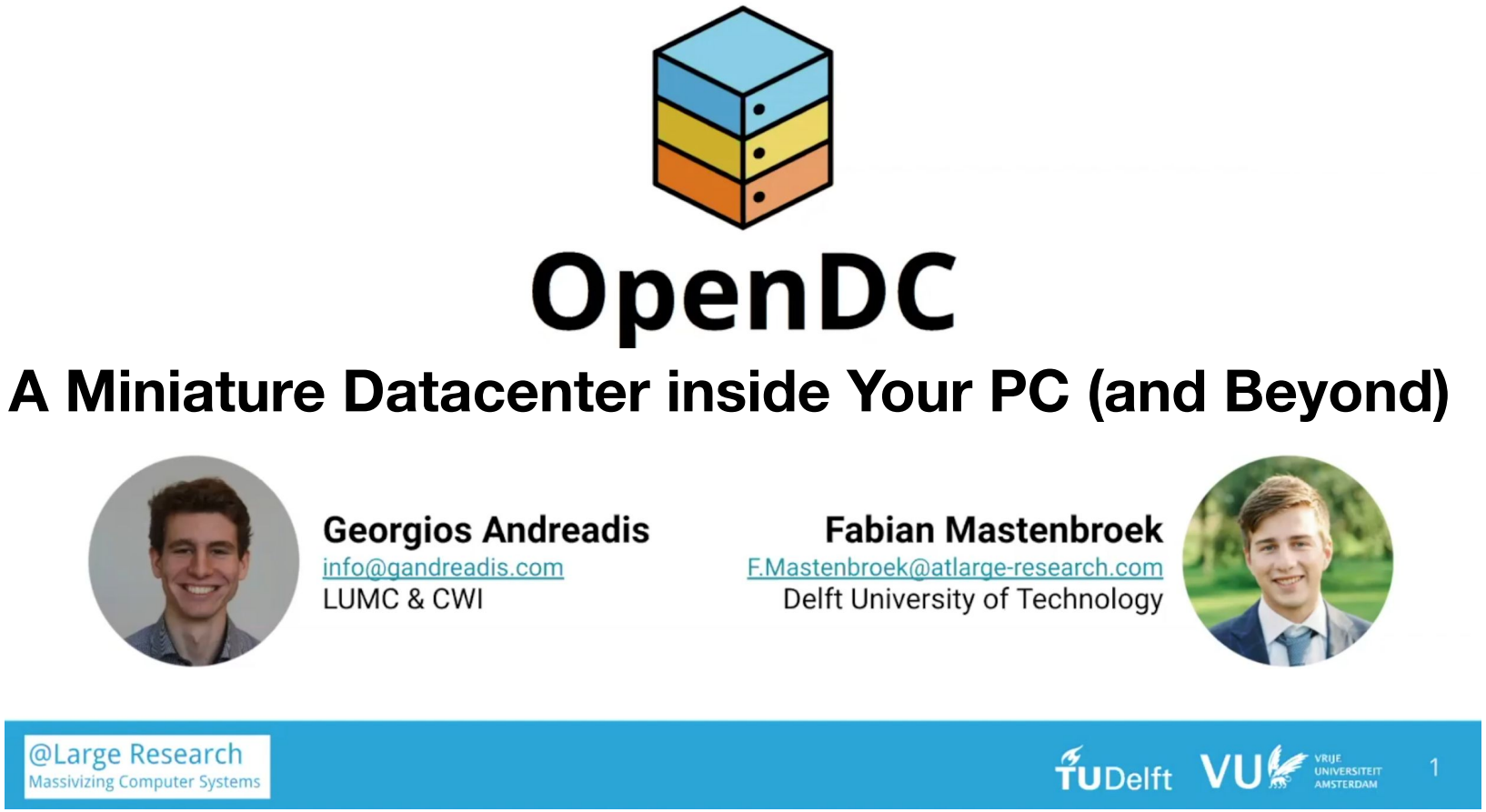
(our practical approaches are often instruments and datasets)

# CHALLENGE: CONDUCT EXPERIMENTS EFFICIENTLY

## THE CHALLENGE OF REAL-WORLD AND ANALYTICAL APPROACHES

- Testing in the lab costs 6-9 months of effort → real-world experiments are costly
- Analytical approaches rely on limiting assumptions, and on the existence of a (valid, calibrated) model → analytical approaches rarely easy to do for new technology

# OPENDC: SIMULATE DATACENTERS, TOGETHER



# … CAN WE AFFORD A? WHAT IF B? X vs. Y … vs. Z?

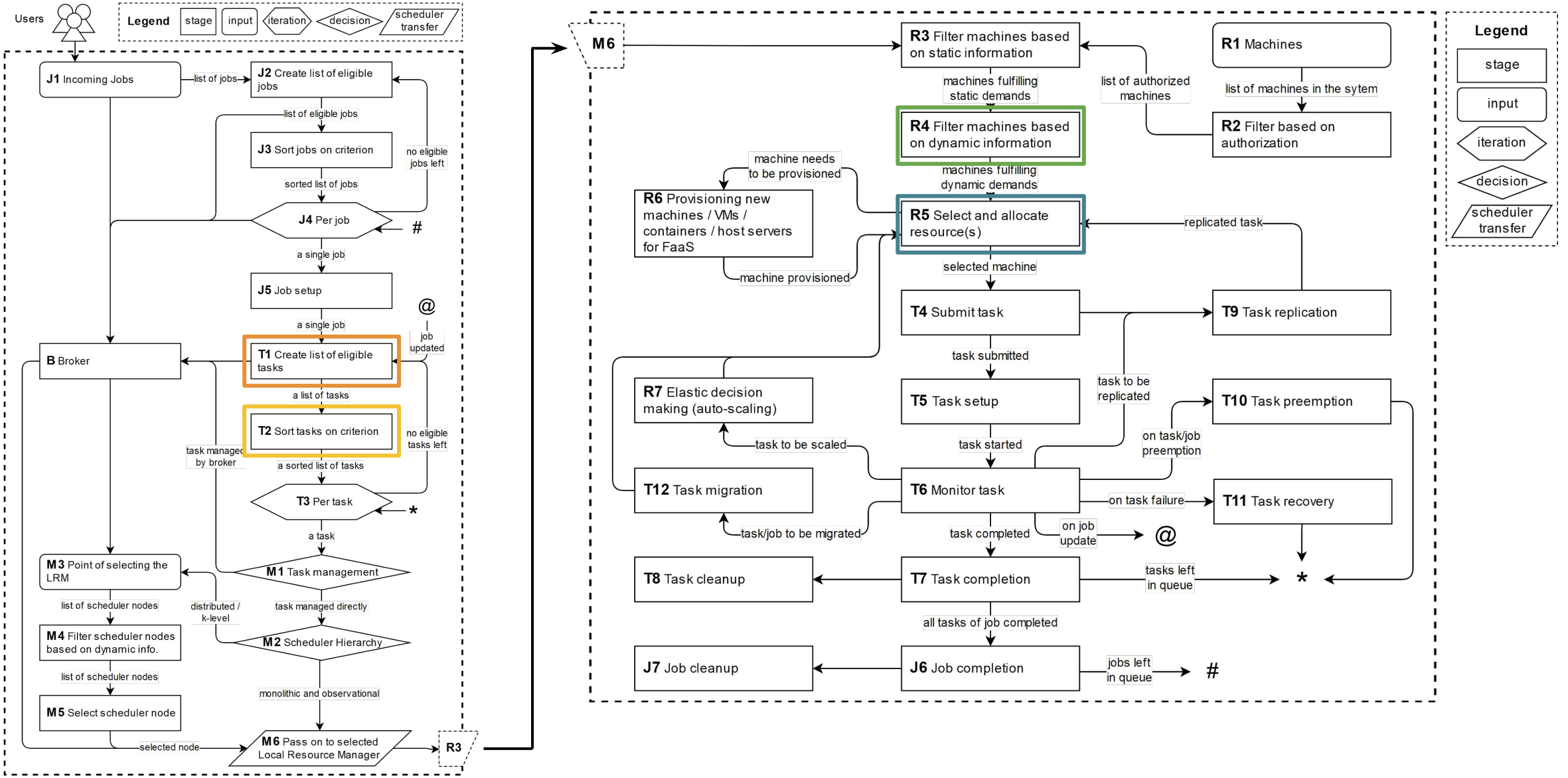
## TOO COSTLY TO CONDUCT REAL-WORLD EXPERIMENTS, SO USE A SIMULATOR

OpenDC
simulator

Learn more:
opendc.org

- Short-term resource management
- Long-term capacity planning
- Sophisticated model  many Qs, goals
- Supports many kinds of workloads
- Supports many kinds of resources
- Validated for various scenarios
- Work with major NL hoster
- Used in training, education, research

*and more…*


Users
Legend
stage
input
iteration
decision
scheduler transfer
J1 Incoming Jobs
list of jobs
J2 Create list of eligible jobs
list of eligible jobs
J3 Sort jobs on criterion
no eligible jobs left
sorted list of jobs
J4 Per job
#
a single job
J5 Job setup
@
job updated
a single job
B Broker
T1 Create list of eligible tasks
a list of tasks
T2 Sort tasks on criterion
no eligible tasks left
task managed by broker
a sorted list of tasks
T3 Per task
*
a task
M3 Point of selecting the LRM
M1 Task management
list of scheduler nodes
distributed / k-level
task managed directly
M4 Filter scheduler nodes based on dynamic info.
M2 Scheduler Hierarchy
list of scheduler nodes
monolithic and observational
M5 Select scheduler node
selected node
M6 Pass on to selected Local Resource Manager
R3
M6
R3 Filter machines based on static information
machines fulfilling static demands
list of authorized machines
R1 Machines
list of machines in the sytem
R2 Filter based on authorization
R4 Filter machines based on dynamic information
machine needs to be provisioned
machines fulfilling dynamic demands
R6 Provisioning new machines / VMs / containers / host servers for FaaS
R5 Select and allocate resource(s)
replicated task
machine provisioned
selected machine
T4 Submit task
T9 Task replication
task submitted
R7 Elastic decision making (auto-scaling)
T5 Task setup
task to be replicated
T10 Task preemption
task to be scaled
task started
on task/job preemption
T12 Task migration
T6 Monitor task
on task failure
T11 Task recovery
task/job to be migrated
task completed
on job update
@
T8 Task cleanup
T7 Task completion
tasks left in queue
*
all tasks of job completed
J7 Job cleanup
J6 Job completion
jobs left in queue
#
Legend
stage
input
iteration
decision
scheduler transfer

# IMPACT OF UNDERSPECIFICATION ON PERFORMANCE RESULTS

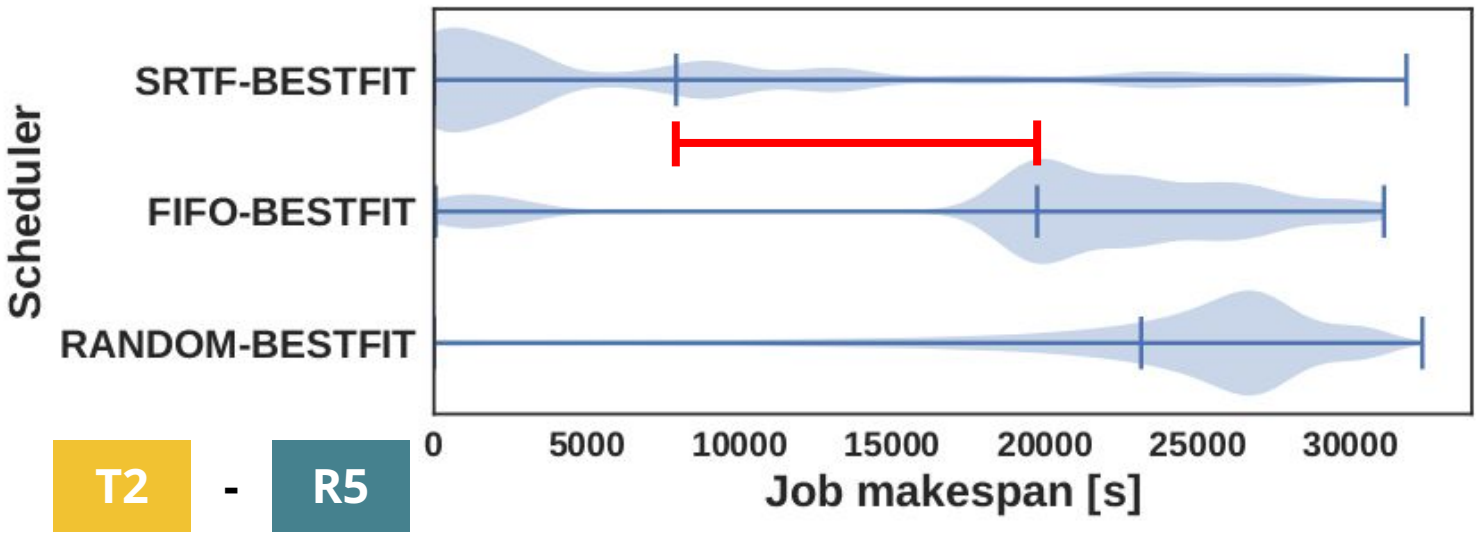
- From the highly cited article on the scheduler of a Big Tech company
- Selected two underspecified stages
- Compared different credible policies that could have been used

**Significant difference in performance!**

💡 Different policy combinations lead to different results

**T2**: Sort tasks

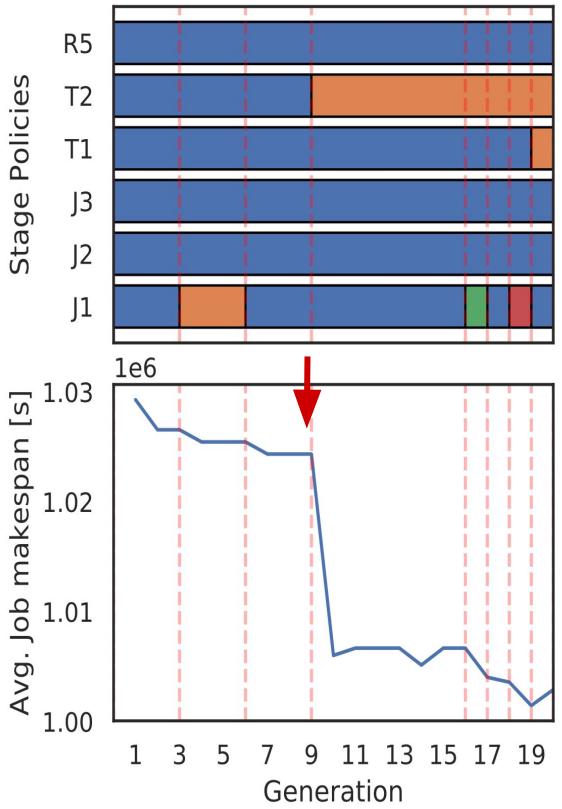
**R5**: Select resources



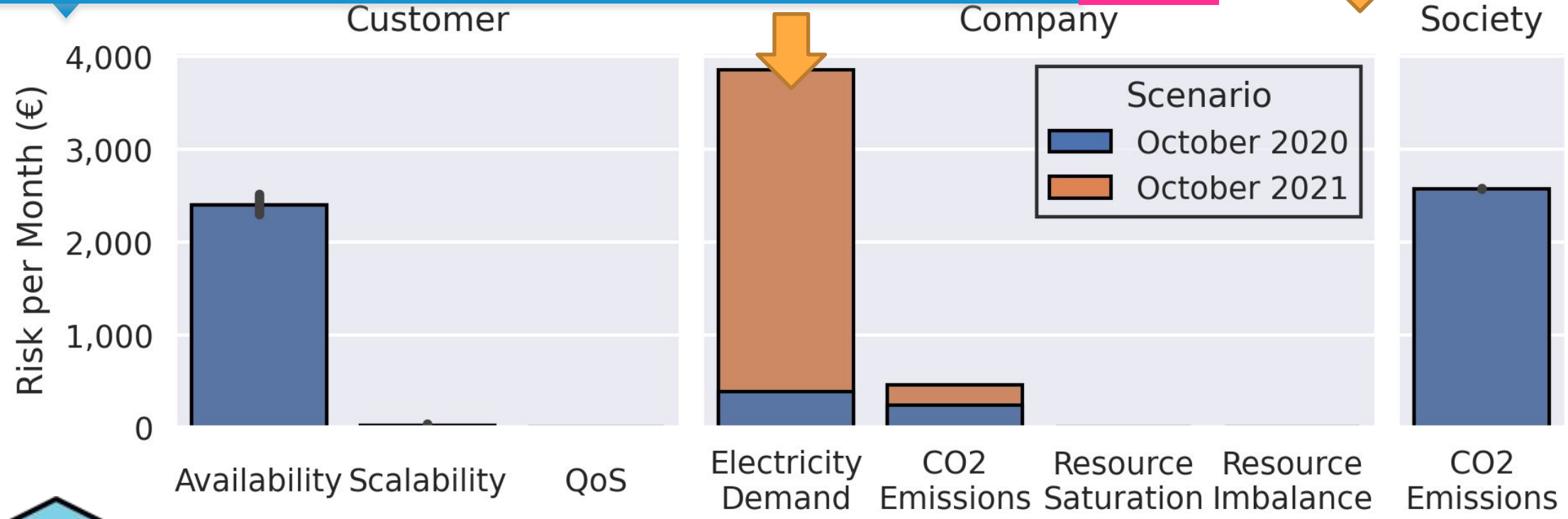
# Follow-Up:
# Automatic Scheduler Design Exploration

- A novel approach for designing better schedulers
- We decompose schedulers into 33 stages
- Systematically construct new schedulers using building blocks
- Trying every combination takes over **$10^{18}$ years!**
- A smarter exploration approach based on an automated process of natural selection
- Algorithm is able to adapt scheduler to workload patterns

Fabian Mastenbroek



# Electricity expenses become High Risk


Electricity expenses could become **primary risk** in datacenters

# Experiments are also Expensive!
# An Environmental Perspective

**112**

Core-hours

**2.1**

kWh

Delft →  → Rotterdam

**In Simulation**

**8,600,000,000**

Core-hours

**116,000,000,000**

kWh

 **x 55 billion!**

**In Reality**

*Exact numbers confidential, depend on topology*

# Summary:
# Simulation-based experiments are very useful.

1. Simulators are very useful for many lines of inquiry
2. Especially exploratory research can benefit □ simulation-based experiments are very efficient
3. We are building a datacenter simulator with exceptional capabilities
4. Already published results in top-tier conferences

# CHALLENGE: MEANINGFUL REAL-WORLD EXPERIMENTS

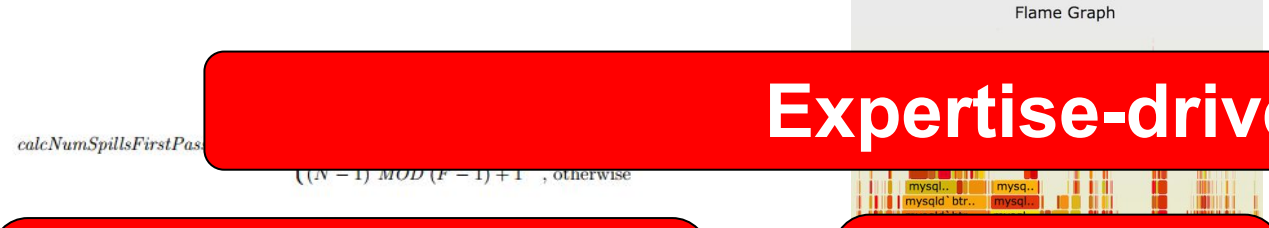
## THE CHALLENGE OF REAL-WORLD AND ANALYTICAL APPROACHES

Objective: analyse performance of (graph-based) AI/ML systems

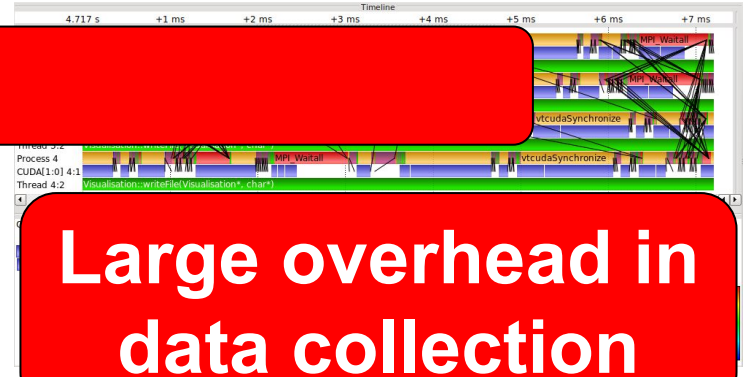
Analytical modeling

Profiling

Tracing






**Expertise-driven**

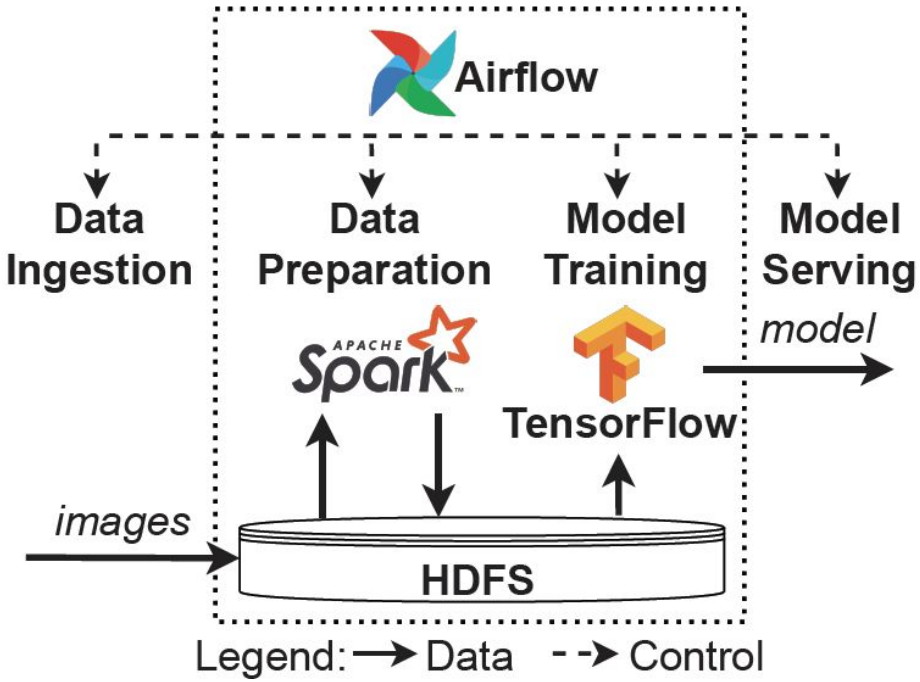
**Challenging for complex systems**

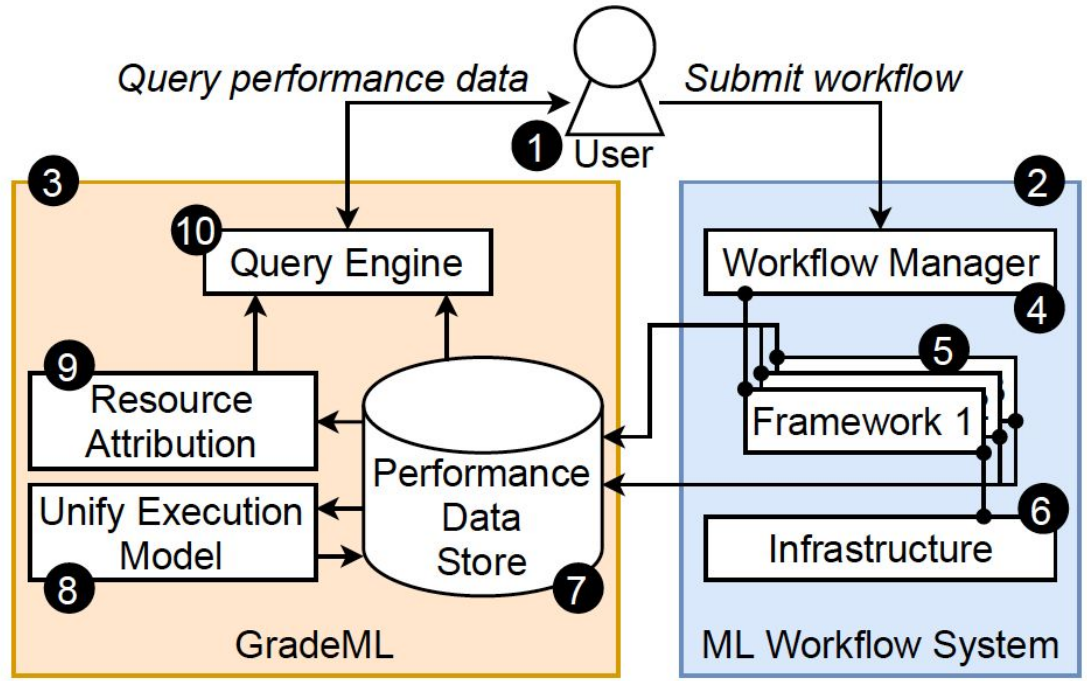
**Limited insight**

**Large overhead in data collection**

# GradeML: PERFORMANCE MEASUREMENT AND ANALYSIS FOR COMPLEX ML WORKFLOWS

## HIGHLY EFFICIENT MEASUREMENT, ADVANCED AND VERSATILE ANALYSIS

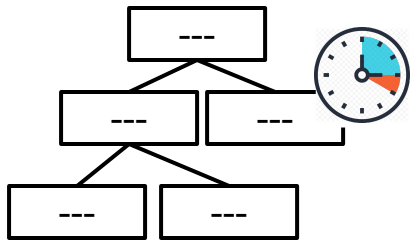


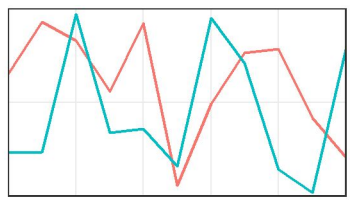




- Also easy to use, low overhead in terms of modeling

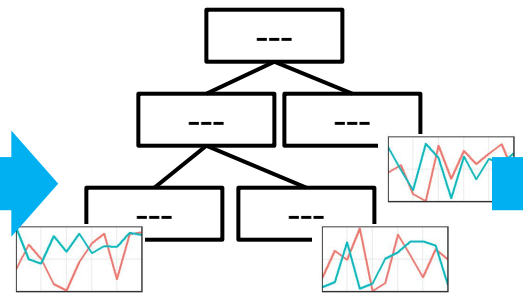
# GradeML: AUTOMATED BOTTLENECK DETECTION AND PERFORMANCE ISSUE IDENTIFICATION FOR AI/ML

System under test

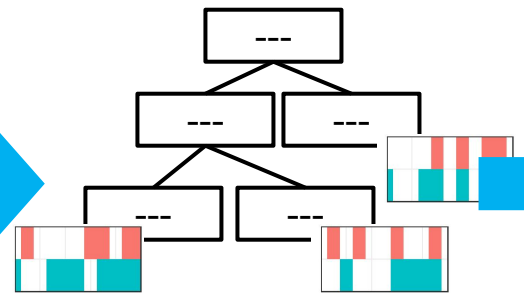


Monitoring



Resource attribution



Bottleneck detection

Top bottlenecks:

---

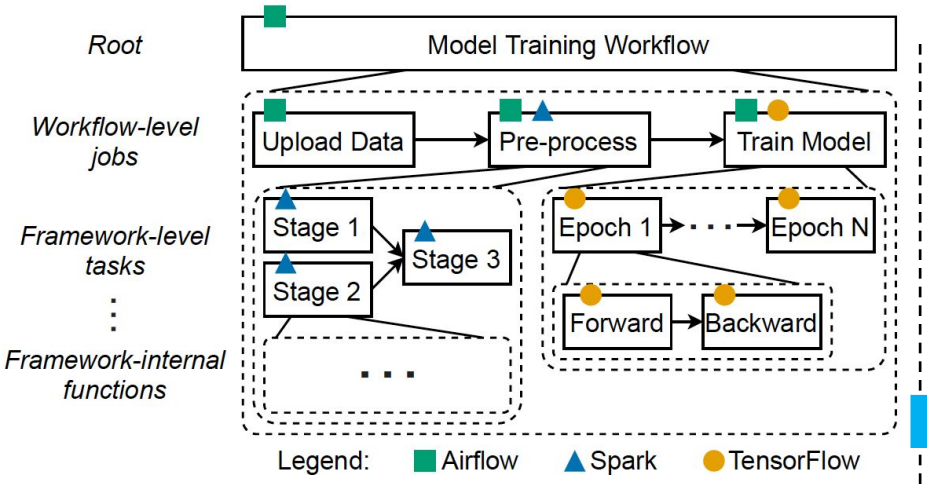
---

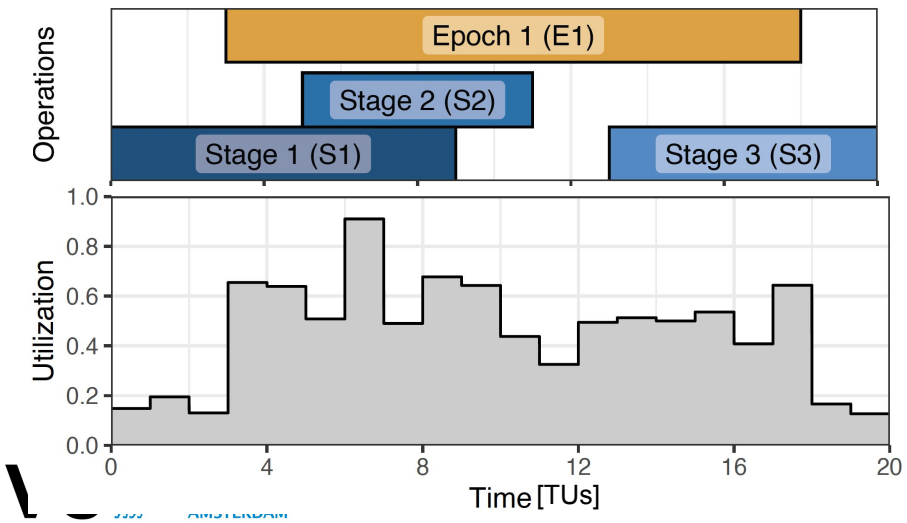
Performance issue identification

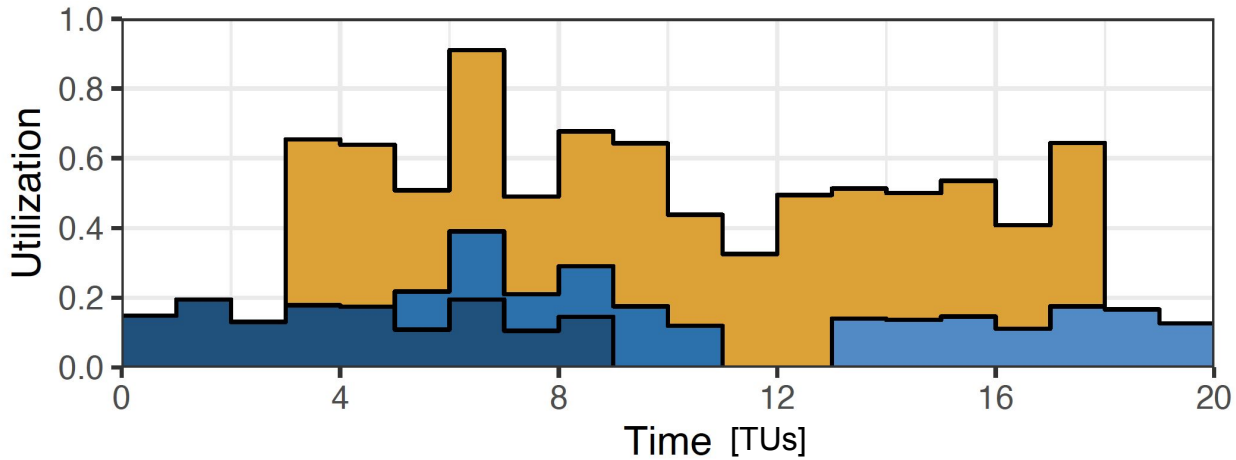
**Semantic gap**
**System-level vs. Workflow-level data**

Grade10 analysis engine

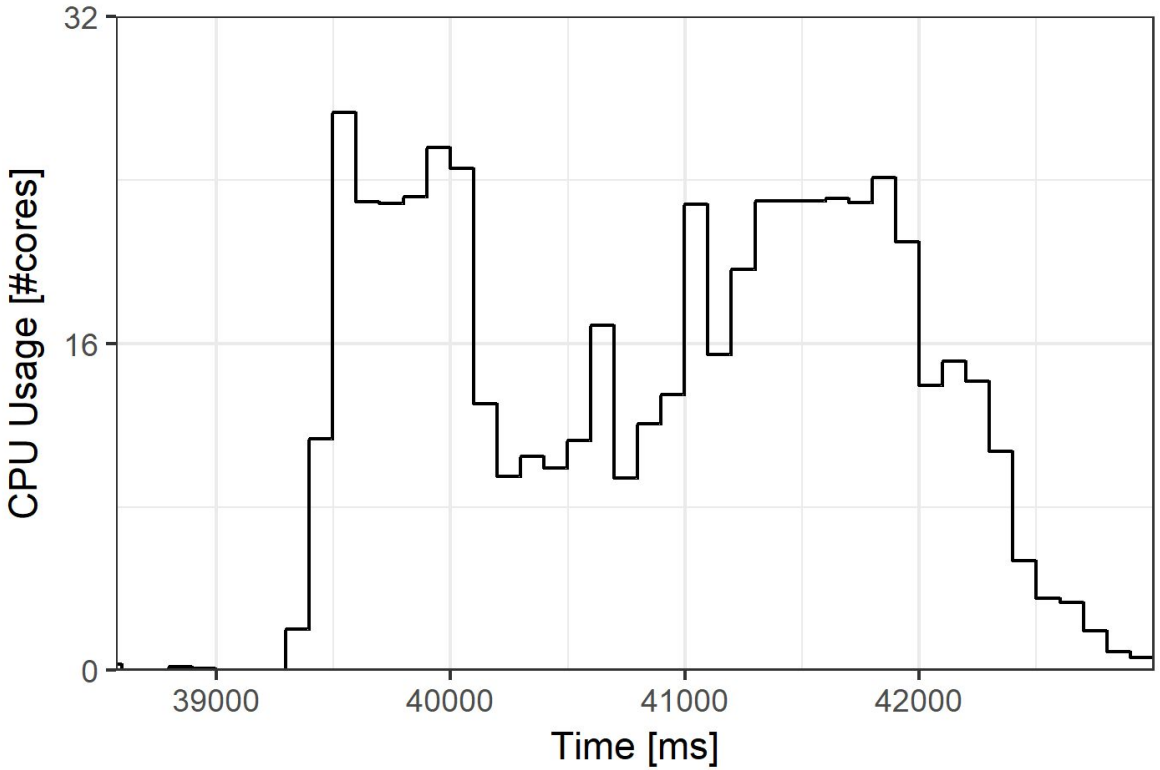
# GradeML: AUTOMATED BOTTLENECK DETECTION AND PERFORMANCE ISSUE IDENTIFICATION FOR AI/ML



# Grade10 Result: Analysing a Giraph Job




WorkerSuperstep

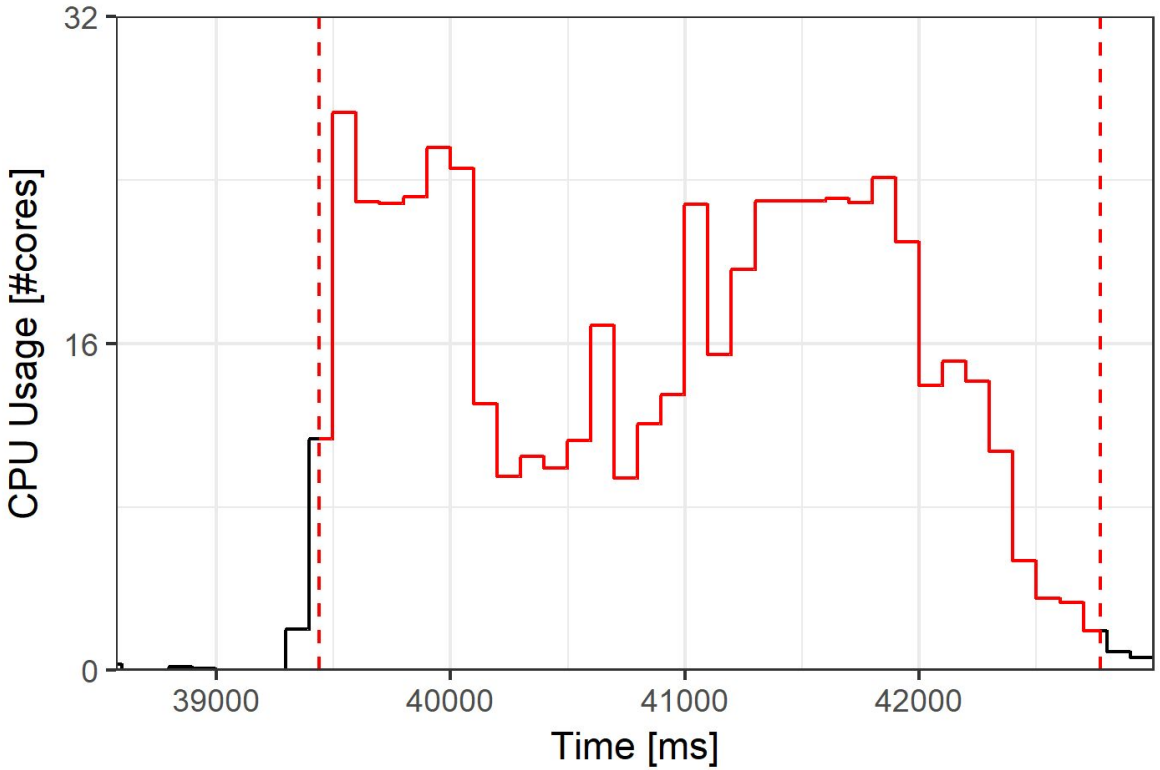
**CPU usage < 32 cores (100%), so no bottleneck**

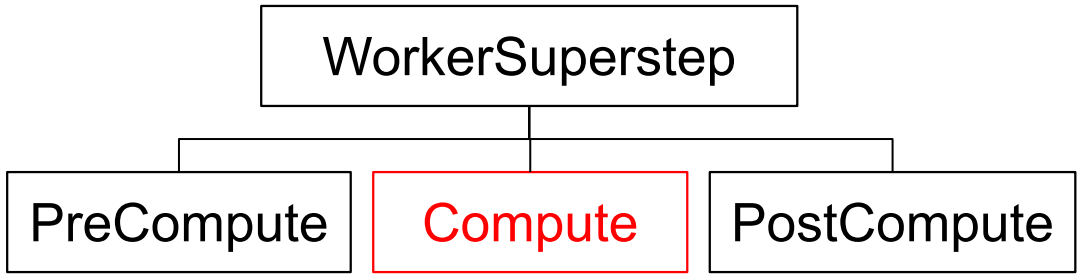
**… Unless …**



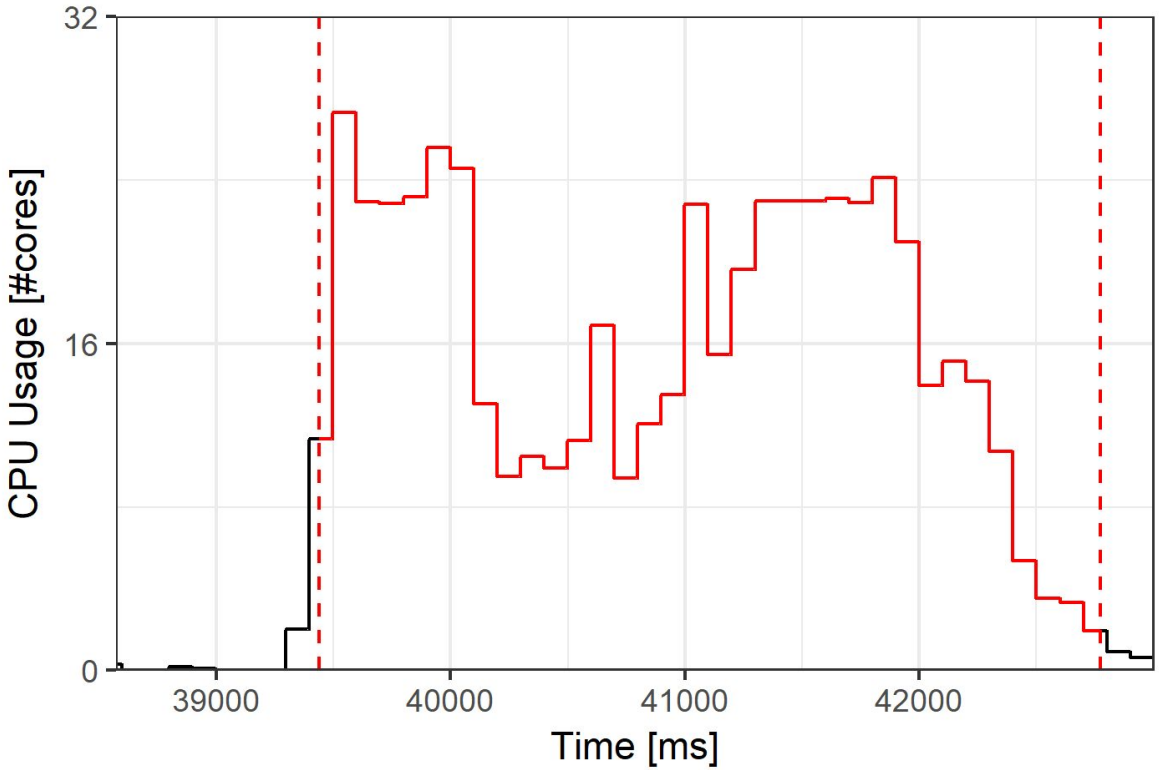
Hegeman et al., Grade10: A Framework for Performance Characterization of Distributed Graph Processing, IEEE Cluster

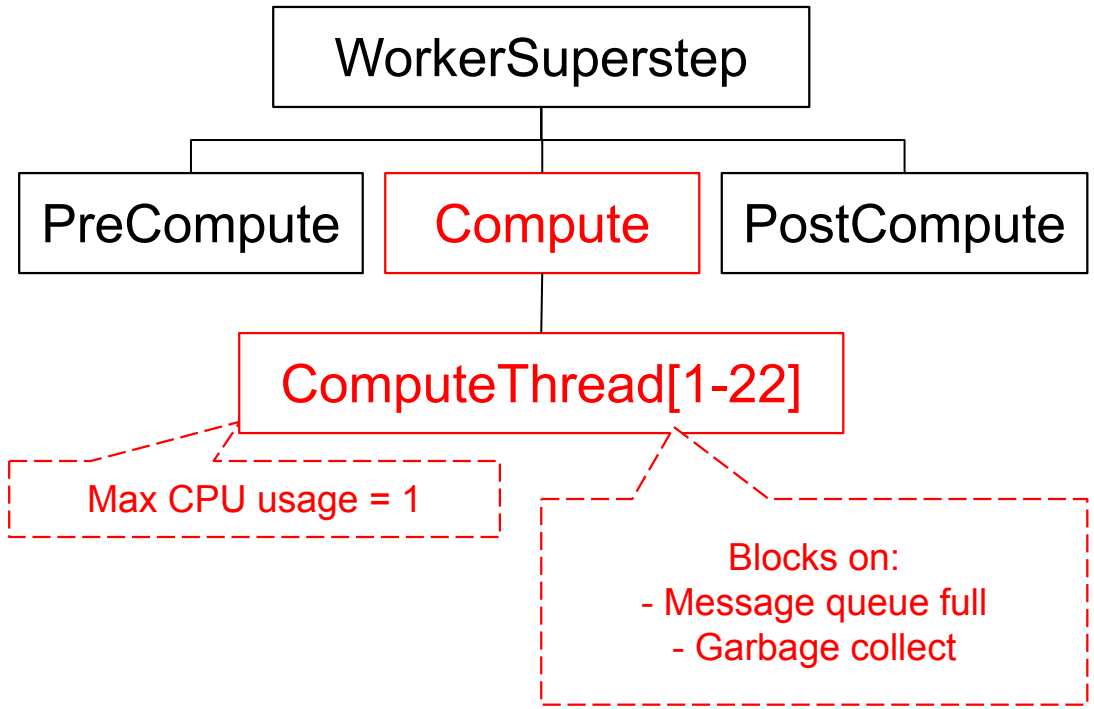
# Grade10 Result: Analysing a Giraph Job



# Grade10 Result: Analysing a Giraph Job

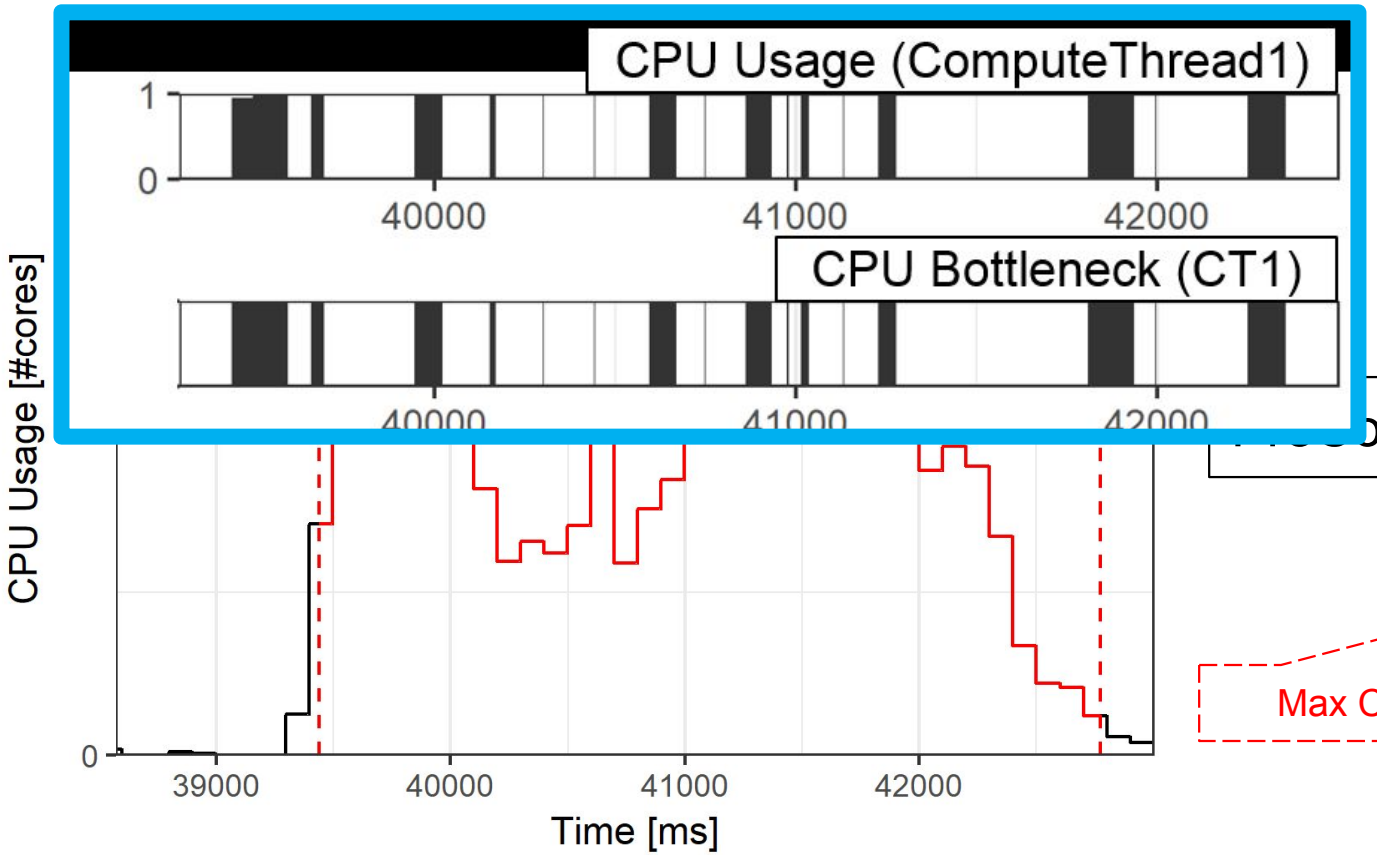


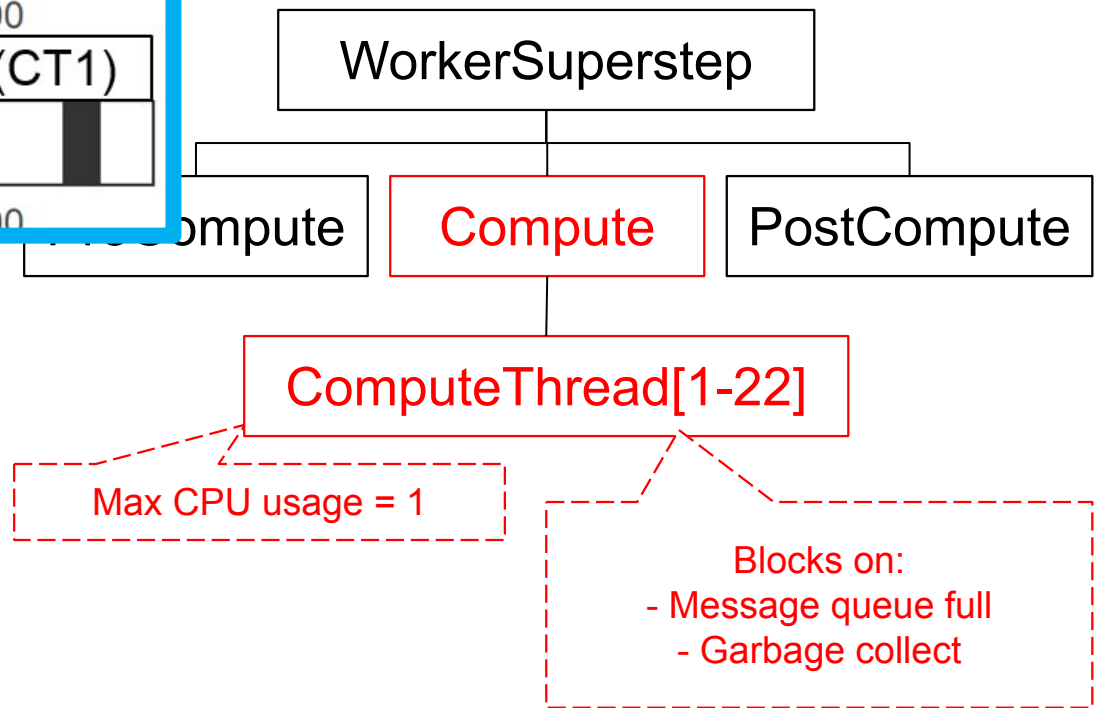





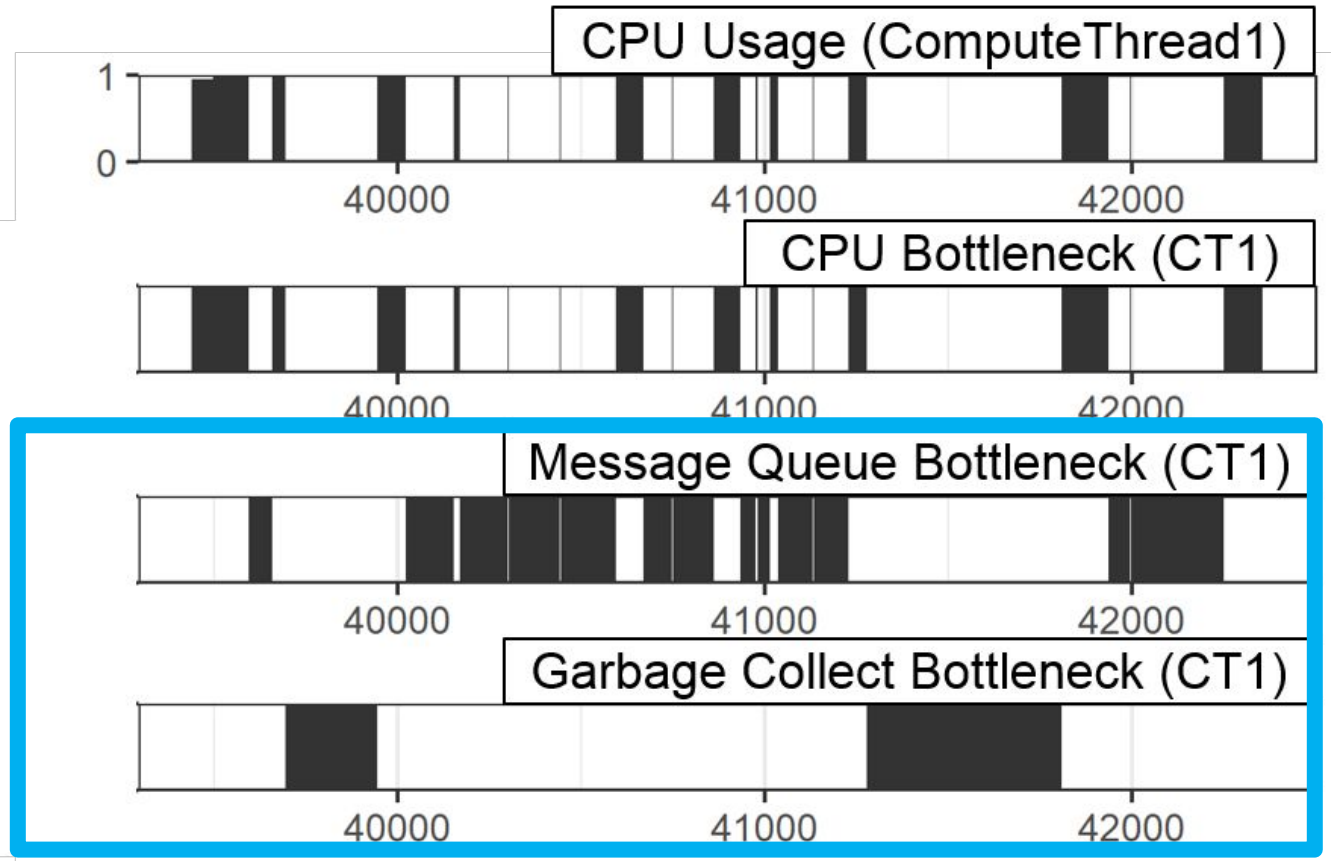

Hegeman et al., Grade10: A Framework for Performance Characterization of Distributed Graph Processing, IEEE Cluster

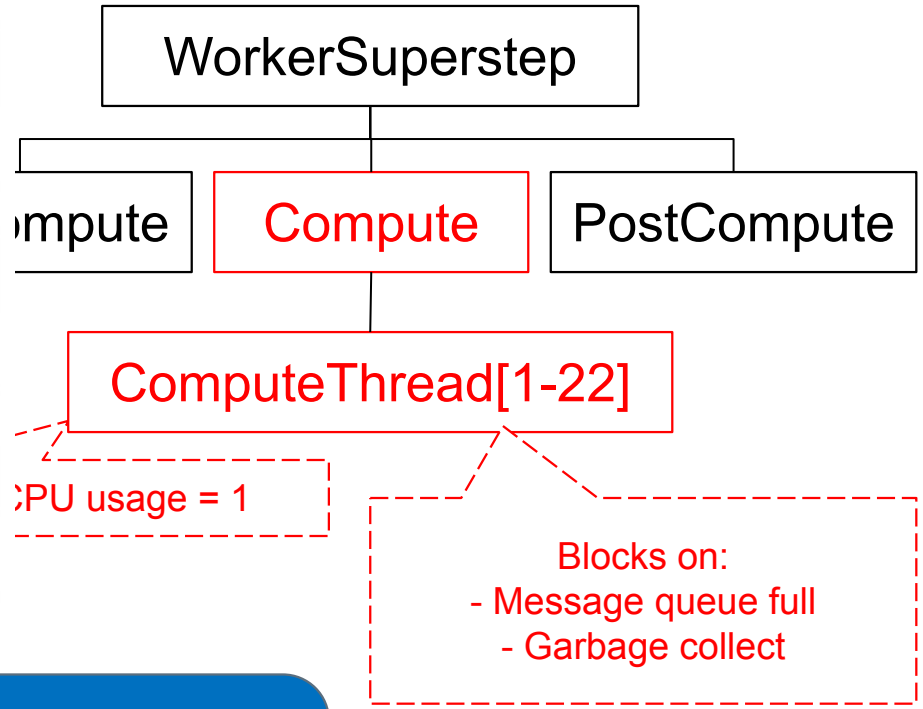
# Grade10 Result: Analysing a Giraph Job



# Grade10 Result: Analysing a Giraph Job




**Plenty of bottlenecks to discover!**

# CHALLENGE: FALSE SENSE OF KNOWING HOW THINGS WORK

## THE CHALLENGE OF TOO MUCH EXPERTISE

An anecdote

- For grid computing, we built them from scratch, so we though we knew all about how they work
- The community leads thought the future is "Big, parallel jobs"
- Whether we actually have parallel jobs or not really matters
- Unfortunately, the belief was not true

Iosup and Epema: Grid Computing Workloads. IEEE Internet Computing 15(2): 19-26 (2011)

# OUR VISION: THE DISTRIBUTED SYSTEMS MEMEX

Bush (1945) As we may think. The Atlantic, Jul 1945.

+

## UNCOVERING THE MYSTERIES OF OUR UNIVERSE, PHYSICAL AND DIGITAL

Find and eradicate performance issues

Get quantitative evidence

Enable new designs and automation

Consider culture and ethics

Understand how entire ecosystems behave and evolve

| Cloud, Grid, Edge, Fog, etc. | One aspect: BigData, P2P | Sci.&Eng. Apps+Sys. | Consumer Apps+Sys. | Enterprise Sys. | Systems, Ecosystems | Performance, Availability, etc. |
|---|---|---|---|---|---|---|
|   | P2P T.A. | [Versluis et al. [PDS'21] Groups, Workflows | Game Trace Archive | Business -Critical |  | [Versluis et al. [FGCS'22] |
| [Iosup et al. FGCS'08] | [Zhang et al. CoNext'10] | [Iosup et al. IEEE IC'11] | [Guo et al. NETGAMES'12] | [Shen et al. CCGRID'15] | [Ghit et al. CCGRID'14] | [Iosup et al. CCGRID'10] |

# DISCOVERY = LARGE-SCALE, LONG-TERM STUDY

+

## UNCOVERING THE MYSTERIES OF OUR PHYSICAL UNIVERSE

GEORGE SMOOT
NOBEL PRIZE 2006

SKA FUNDING: 500+ FTE, EUR 1.5B

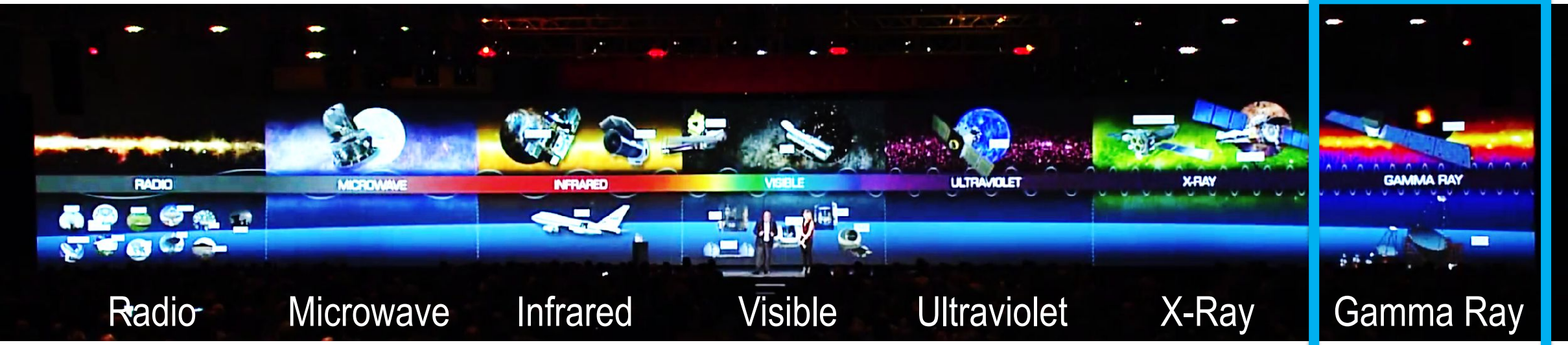




| |
|---|
| James Cordes, The Square Kilometer Array, Project Description, 2009 [Online] |
| The Square Kilometer Array Factsheet, How much will it cost?, 2012 [Online] |
| Phil Diamond and Rosie Bolton, Life, the Universe & Computing: The story of the SKA Telescope, SC17 Keynote. [Online] |

# DISCOVERY = LARGE-SCALE, LONG-TERM STUDY

## UNCOVERING THE MYSTERIES OF OUR UNIVERSE, PHYSICAL AND DIGITAL




| Radio | Microwave | Infrared | Visible | Ultraviolet | X-Ray | Gamma Ray |
|---|---|---|---|---|---|---|
| Cloud, Grid, Edge, Fog, etc. | One aspect: BigData, P2P | Sci.&Eng. Apps+Sys. | Consumer Apps+Sys. | Enterprise Sys. | Systems, Ecosystems | Performance, Availability, etc. |
| [Iosup et al. FGCS'08] | [Zhang et al. CoNext'10] | [Iosup et al. IEEE IC'11] | [Guo et al. NETGAMES'12] | [Shen et al. CCGRID'15] | [Ghit et al. CCGRID'14] | [Iosup et al. CCGRID'10] |

THE GRID WORKLOADS ARCHIVE

P2P T.A.

[Versluis et al. FGCS'22]

[Versluis et al. [PDS'21]

Groups, Workflows

Game Trace Archive

Business -Critical

# UNKNOWN PHENOMENA: INTER-, ADAPT-, EXAPTATION +

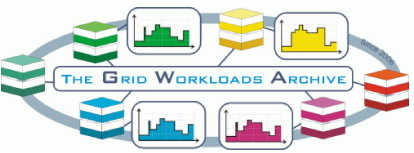
## UNCOVERING THE MYSTERIES OF OUR UNIVERSE, PHYSICAL AND DIGITAL

### SOME OF OUR DISCOVERIES

... ... ... ... ...

**BOTS, NOT PARALLEL JOBS**

Cloud, Grid, Edge, Fog, etc.




[Iosup et al. FGCS'08]

**GROUPS NOT RARE, DOMINANT**

One aspect: BigData, P2P

P2P T.A.

[Zhang et al. CoNext'10]

Sci.&Eng. Apps+Sys.

[Versluis et al. [PDS'21]

Groups, Workflows

[Iosup et al. IEEE IC'11]

**COMMUNITY DYNAMICS**

Consumer Apps+Sys.

Game Trace Archive

[Guo et al. NETGAMES'12]

**SYSTEMIC VARIABILITY**

Enterprise Sys.

Business -Critical

[Shen et al. CCGRID'15]

**CORRELATED, NOT IID FAILURES**

Systems, Ecosystems

[Ghit et al. CCGRID'14]

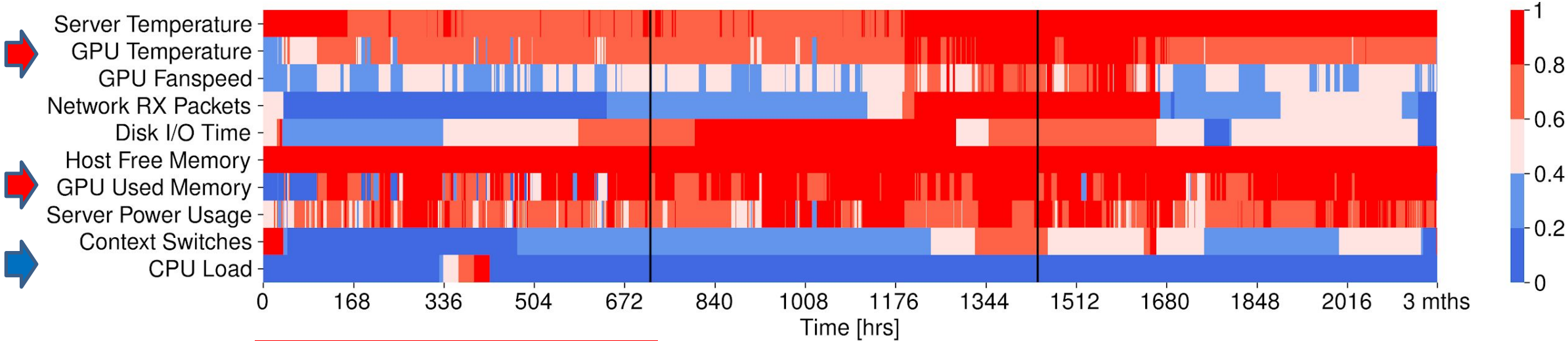
Performance, Availability, etc.

[Versluis et al. FGCS'22]

[Iosup et al. CCGRID'10]

# One Basic Result

Uta et al., Beneath the SURFace: An MRI-like View into the Life of a 21st-Century Datacenter. USENIX; login 2020.




**Hot GPUs**

**Cold CPUs***

* when GPUs

**Could use less powerful, less expensive CPUs**

# TAKE-HOME MESSAGE

Massivizing → computer ecosystems with good functional and non-functional properties, for all

The ecosystem is vast, challenging: many apps, many platforms, may goals, many approaches

Many modern, open challenges: resource management and scheduling, telemetry, analysis, simulation, experimentation, etc.

# MASSIVIZING COMPUTER SYSTEMS

## FURTHER READING

https://atlarge-research.com/publications.html

1. Iosup et al. Massivizing Computer Systems. ICDCS 2018 ← start here
2. Andreadis et al. A Reference Architecture for Datacenter Scheduling, SC18
3. Van Eyk et al. Serverless is More: From PaaS to Present Cloud Computing, IEEE IC Sep/Oct 2018
4. Uta et al. Exploring HPC and Big Data Convergence: A Graph Processing Study on Intel Knights Landing, IEEE Cluster 2018

5. Talluri et al. Big Data Storage Workload in the Cloud. ACM/SPEC ICPE 2019.
6. Toader et al. Graphless. IEEE ISPDC'19.
7. Jiang et al. Mirror. CCPE 2018.
8. Ilyushkin et al. Autoscalers. TOMPECS 2018.
9. Versluis et al. Autoscaling Workflows. CCGRID'18.
10. Uta et al. Elasticity in Graph Analytics? IEEE Cluster 2018.

11. Herbst et al. Ready for rain? TOMPECS 2018.
12. Guo et al. Streaming Graph-partitioning. JPDC'18.
13. Iosup et al. The OpenDC Vision. ISPDC'17.
14. Iosup et al. Self-Aware Computing Systems book.
15. Iosup et al. LDBC Graphalytics. PVLDB 2016.

Etc.

# MASSIVIZING COMPUTER SYSTEMS

## FURTHER READING

https://atlarge-research.com/publications.html


1. Crusoe, Iosup, et al. (2022) Methods Included: CWL. CACM
2. Sakr, Bonifati, Voigt, Iosup, et al. (2021) The Future Is Big Graphs! CACM
3. Andreadis et al. (2021) Capelin: Data-Driven Capacity Procurement for Cloud Datacenters using Portfolios of Scenarios. TPDS, under review.
4. Versluis et al. The Workflow Trace Archive: Open-Access Data From Public and Private Computing Infrastructures. TPDS 2020.
5. Uta et al. (2020) Beneath the SURFace: An MRI-like View into the Life of a 21st-Century Datacenter. login USENIX
6. Iosup, Hegeman, et al. (2020) The LDBC Graphalytics Benchmark. CoRR. https://arxiv.org/abs/2011.15028
7. Hegeman et al. (2021) GradeML. HotCloudPerf.
8. Abad, Iosup, et al. An Analysis of Distributed Systems Syllabi With a Focus on Performance-Related Topics. WEPPE 2021. https://arxiv.org/abs/2103.01858

Etc.

# MASSIVIZING COMPUTER SYSTEMS

## FURTHER READING

https://atlarge-research.com/publications.html

1. Iosup et al. The AtLarge Vision on the Design of Distributed Systems and Ecosystems. ICDCS 2019 ← Start here
2. Uta et al. Is big data performance reproducible in modern cloud networks? NSDI 2020
3. Van Eyk et al. The SPEC-RG Reference Architecture for FaaS: From Microservices and Containers to Serverless Platforms, IEEE IC 2019
4. Papadopoulos et al. Methodological Principles for Reproducible Performance Evaluation in Cloud Computing. TSE 2019 and (journal-first) ICSE 2020

5. van Beek et al. Portfolio Scheduling for Managing Operational and Disaster-Recovery Risks in Virtualized Datacenters Hosting Business-Critical Workloads. ISPDC 2019
6. van Beek et al. A CPU Contention Predictor for Business-Critical Workloads in Cloud Datacenters. HotCloudPerf19

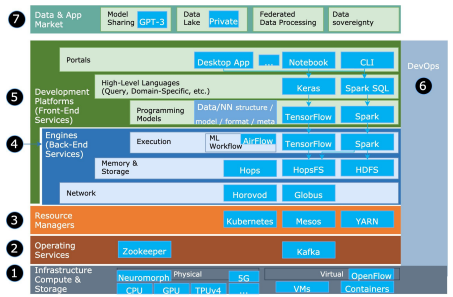
\+ Iyushkin et al. Performance-Feedback Autoscaling with Budget Constraints for Cloud-based Workloads of Workflows. Under submission

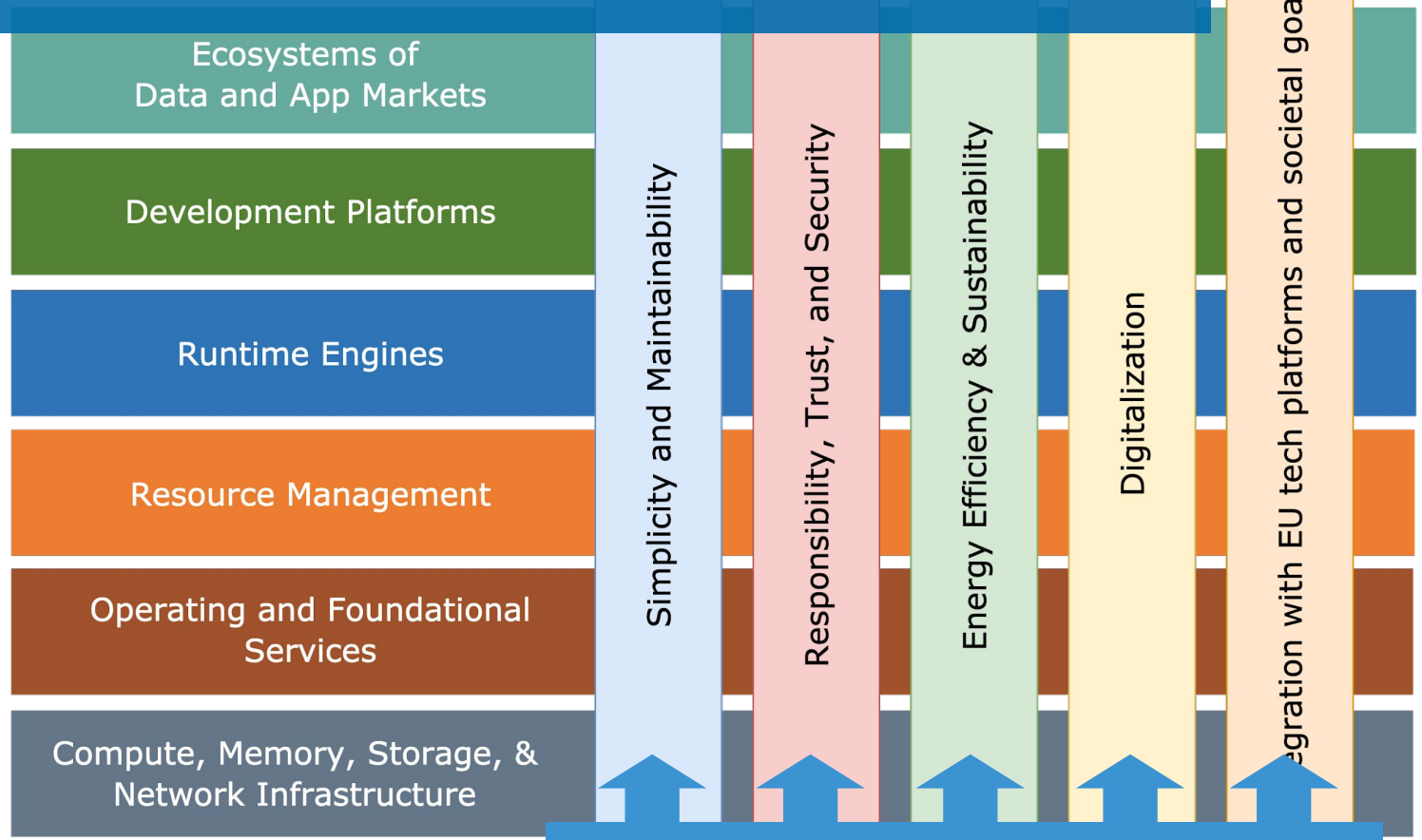
Etc.

# EXTRAS

# MASSIVIZING COMPUTER SYSTEMS

## A LARGER VISION OF HOW COMPUTING WILL HELP OUR SOCIETY



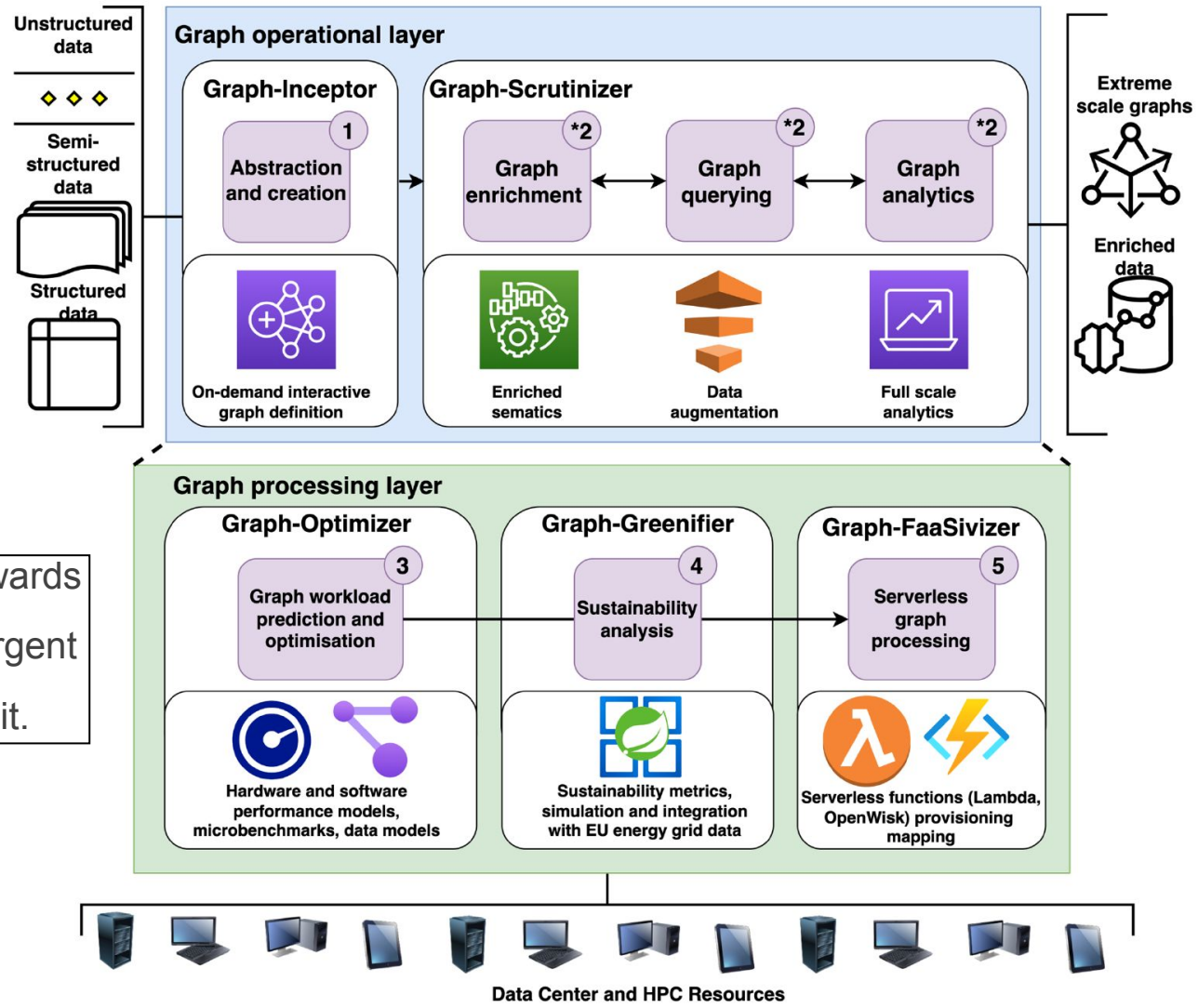




A.Iosup@vu.nl
http://atlarge.science

# ONE PROJECT TO MENTION…

## GraphMassiziver

**Big Graph Processing: Used in AI/ML, FinTech, Sci/Pharma, Industry 4.0, Energy Mgmt.*, etc.**

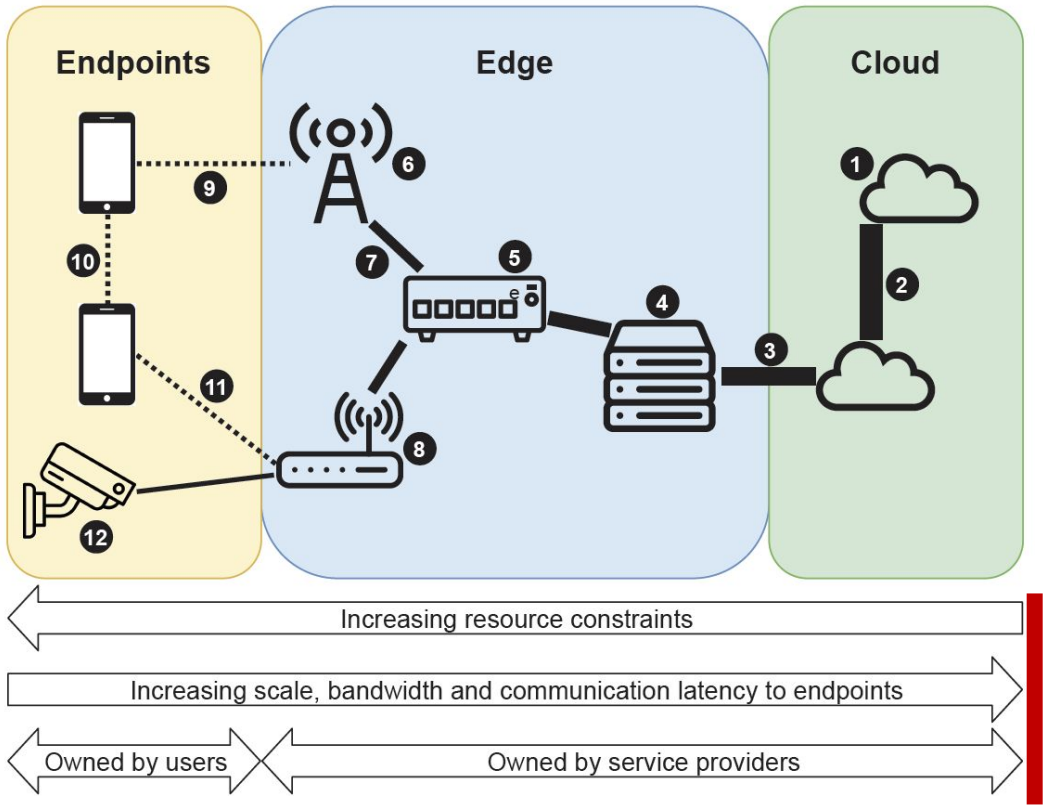


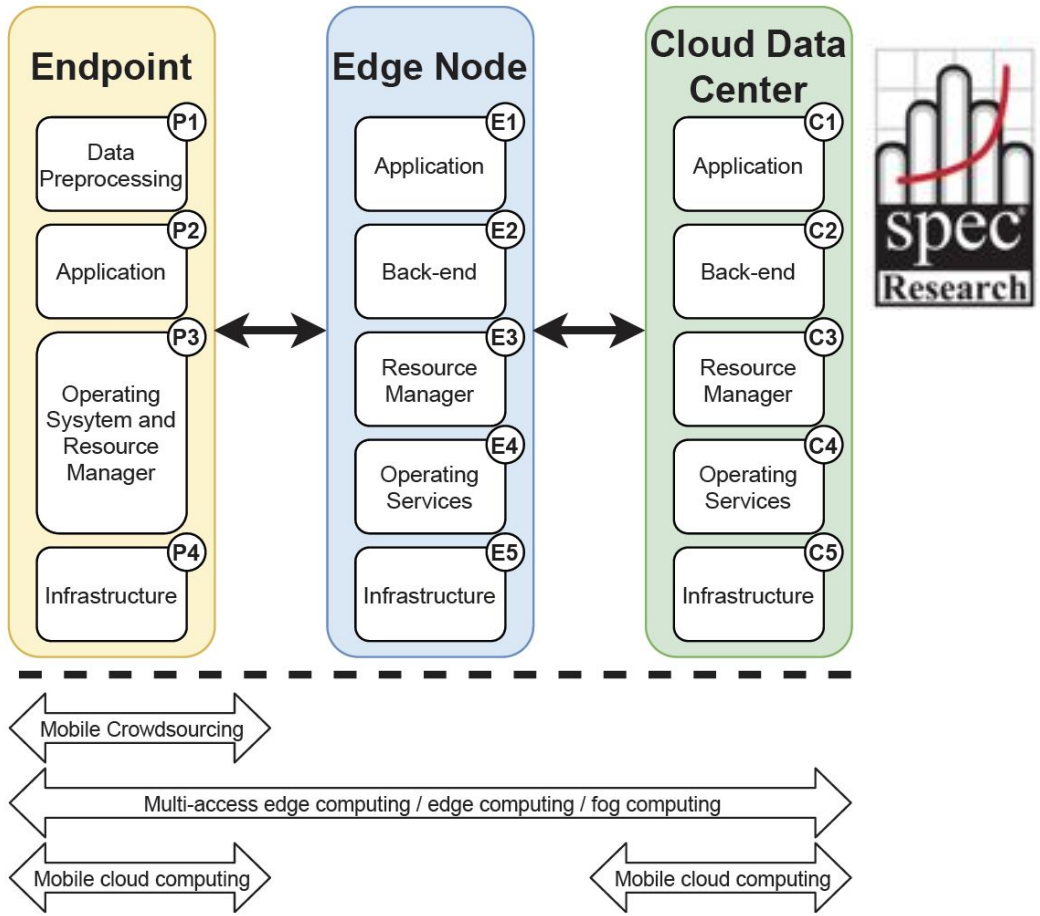

Prodan et al. (2022) The GraphMassivizer project: Towards Extreme and Sustainable Graph Processing for Urgent Societal Challenges in Europe. IEEE Cloud Summit.

# Computational Continuum
# DC → Endpoint-edge-cloud

**Vision: Massivizing computer systems approaches enable holistic understanding and management in the computational continuum**

APPROACH: WITH BROAD PARTICIPATION, UNIFY CURRENT COMPUTING MODELS, FORM A COMPUTATIONAL CONTINUUM




Trivedi, Wang, Bal, Iosup (2021) Sharing and Caring of Data at the Edge. HotEdge.

Jansen et al. (2022) The SPEC-RG Reference Architecture for the Edge Continuum. CoRR abs/2207.04159