# Massivizing = Scalable, High Performance, Reliable, Efficient Graph Processing Systems

@AIosup

dr. ir. Alexandru Iosup
Parallel and Distributed Systems Group

TUDelft

DELFT DATA SCIENCE

# Societal Challenges


The quadruple helix: prosperous society & blooming economy & inventive academia & wise governance depend on datacenters

- Enable data access & processing as a fundamental right in Europe
- Enable big science and engineering (2020: €100 bn., 1 mil. jobs)
- “To out-compute is to out-compete”, but with energy footprint <5%
- Keep Internet-services affordable yet high quality in Europe
- The Schiphol of computation: Netherlands as a world-wide ICT hub

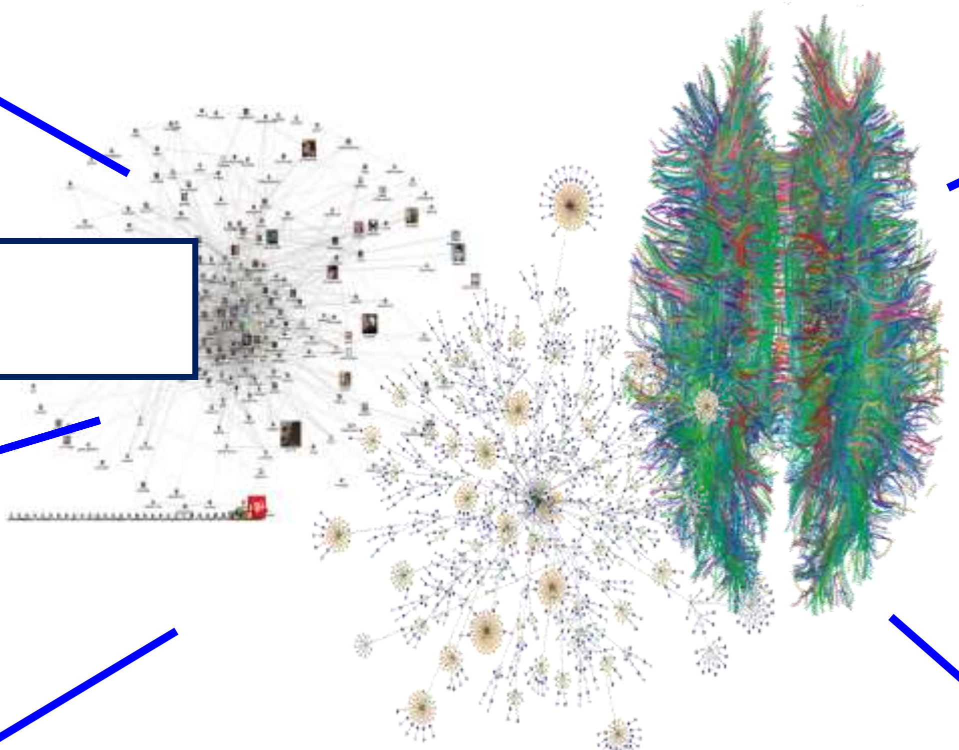
KIVI Engineering Society

DELFT DATA SCIENCE

# Societal Challenges, Concretely: Graph Processing for Everyone

Oracle 1.2M followers,
132k employees

company/day:
40-60 posts, 500-700 comments



Source: Smith, CHI'10; Blog webpage; Gigandet et al., PLoS ONE 3(12)]

270M MAU
200+ avg followers

>54B edges

1.2B MAU 0.8B DAU
200+ avg followers

>240B edges

friendster (SME*)

XFIRE (SME*)

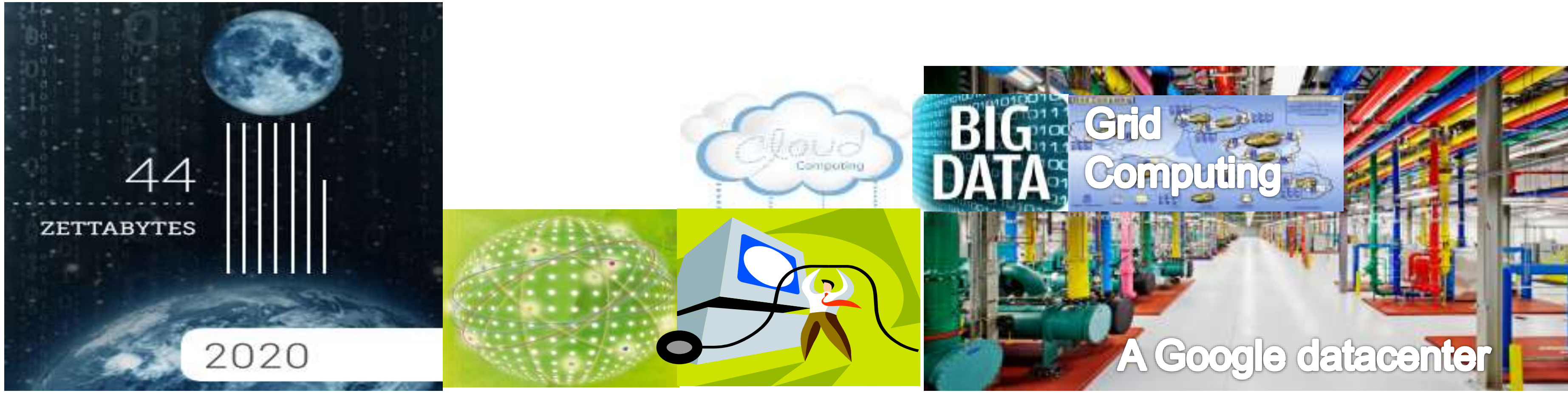
* SMEs in EU/NL= 60% gross value added, little to no ICT expertise



TU Delft

Delft Data Science

# Scientific Challenges




## How to massivize graph processing?

- Super-scalable, super-flexible, yet efficient graph-processing infrastructure
- End-to-end automation of large-scale graph processing
- Dynamic, compute- and data-intensive graph-processing workloads
- Evolving, heterogeneous hardware and software
- Strict performance, cost, energy, reliability, and fairness requirements

# Massivizing Graph-Processing Systems

Interactive

5' — Pitch on Massivizing Graph-Processing Systems

5' — Two Exemplary Steps Forward

- Benchmarking distributed *or* heterogeneous graph-processing systems
- Designing distributed *and* heterogeneous graph-processing systems

5' — Towards a Taskforce on Data Science as a Service

# What does a benchmark consist of?

- Four main elements:
  - data schema: defines the structure of the data
  - workloads: defines the set of operations to perform
  - performance metrics: used to measure (quantitatively) the performance of the systems
  - execution rules: defined to assure that the results from different executions of the benchmark are valid and comparable

- Software as Open Source (GitHub)
  - data generator, query drivers, validation tools, ...



Y. Guo, M. Biczak, A. L. Varbanescu, A. Iosup, C. Martella, and T. L. Willke, "How Well Do Graph-Processing Platforms Perform? An Empirical Performance Evaluation and Analysis", IPDPS, 2014

# Graphalytics, in a nutshell

- An LDBC benchmark
- Advanced benchmarking harness
- Diverse real and synthetic datasets
- Many classes of algorithms
- Granula for manual choke-point analysis
- Modern software engineering practices
- Supports many distributed/heterogeneous platforms

**http://graphalytics.ewi.tudelft.nl/**
**https://github.com/tudelft-atlarge/graphalytics/**

# What is the performance of

TOTEM

medusa-gpu
Medusa: Simplified Graph Processing on GPUs

mapgraph Beta
Massively Parallel Graph processing on GPU

**General Challenges**

Performance Metrics + Graph Diversity + Algorithm Diversity

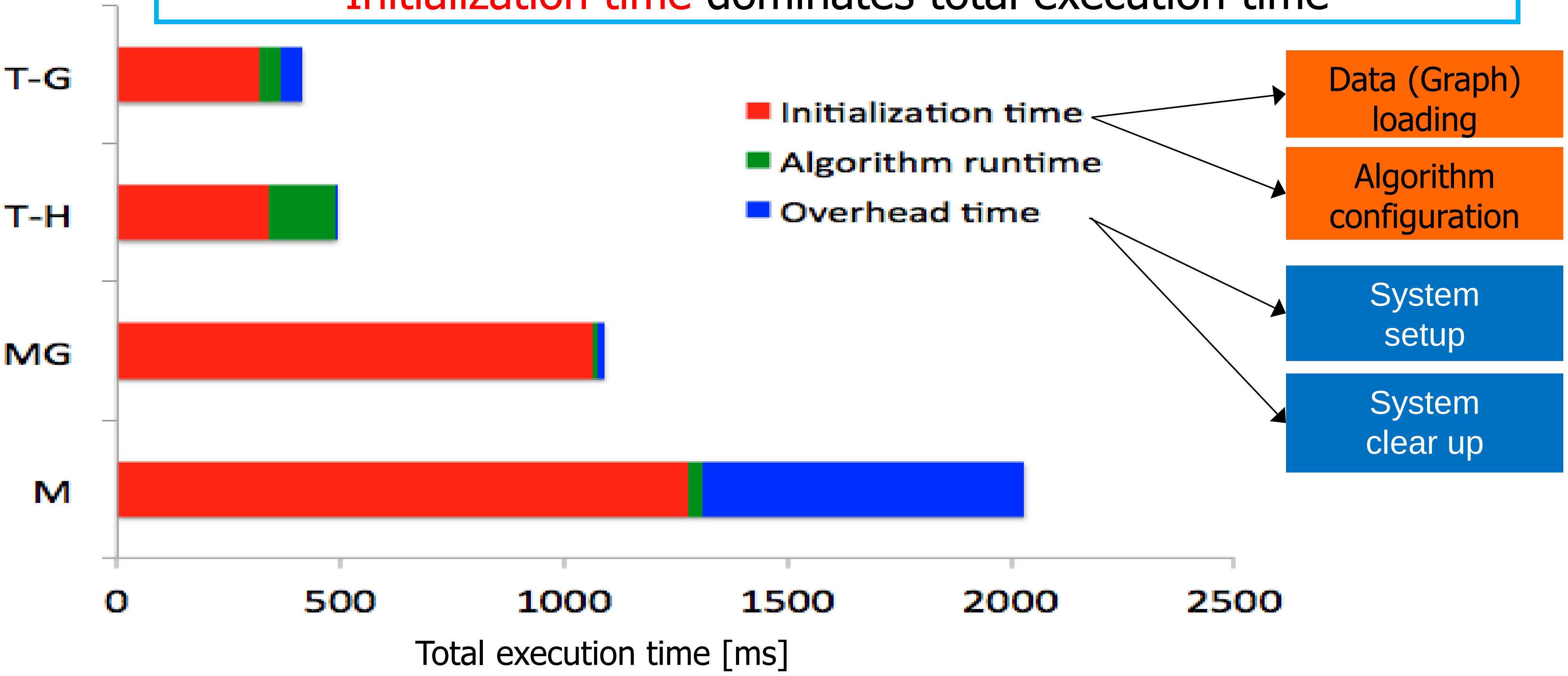
**Challenges for evaluating GPU-enabled systems**

In-memory graph formats + Optimization techniques + GPU generations

Y. Guo, A. L. Varbanescu, A. Iosup, and D. Epema, "An Empirical Performance Evaluation of GPU-Enabled Graph-Processing Systems," *CCGrid, 2015.*

# Sample Result:
# BFS Algo on Amazon Data for all systems

Initialization time dominates total execution time




Y. Guo, A. L. Varbanescu, A. Iosup, and D. Epema, "An Empirical Performance Evaluation of GPU-Enabled Graph-Processing Systems," *CCGrid, 2015.*

# Massivizing Graph-Processing Systems

Interactive

5' — Pitch on Massivizing Graph-Processing Systems

5' — Two Exemplary Steps Forward

- Benchmarking distributed *or* heterogeneous graph-processing systems
- Designing distributed *and* heterogeneous graph-processing systems

5' — Towards a Taskforce on Data Science as a Service

# Existing Graph-Processing Systems: *Either* Distributed *or* Heterogeneous

- Distributed CPU-based systems cannot use additional computational power of accelerators

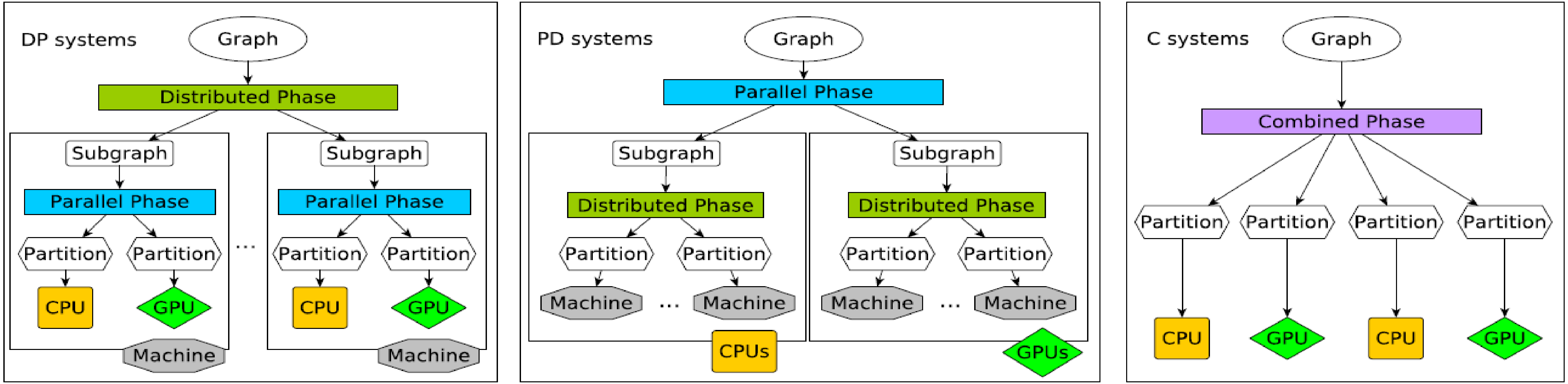
Oracle Labs PGX · Apache Giraph · GraphLab · GraphX · YARN

- GPU-enabled systems are (mostly) single-machine systems, cannot handle large-scale graphs

NetSysLab · medusa-gpu — Medusa: Simplified Graph Processing on GPUs · mapgraph Beta — Massively Parallel Graph processing on GPUs · TOTEM · Gunrock — High-performance Graph Primitives on GPU · VertexAPI2

Y. Guo, A. L. Varbanescu, D. Epema, and A. Iosup, "Design and Experimental Evaluation of Distributed Heterogeneous Graph-Processing Systems," Submission to *CCGrid, 2016.*

# Our approach: 3 Families of Distributed *and* Heterogeneous (CPU+GPU) Graph-Processing Systems



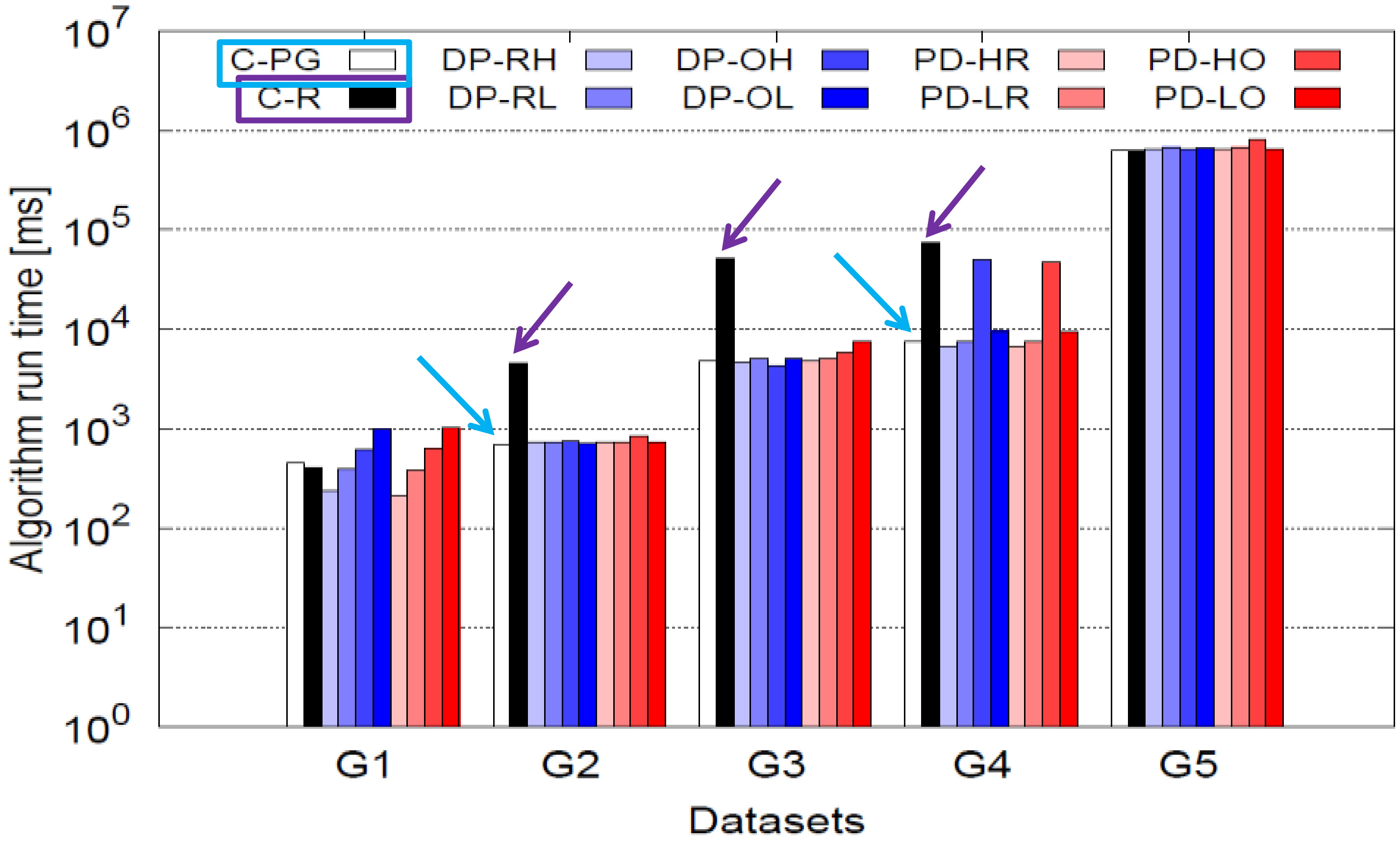

Distributed-then-Parallel (DP) Systems

Parallel-then-Distributed (PD) Systems

(Combined Par.-and-Distributed (C) Systems

Y. Guo, A. L. Varbanescu, D. Epema, and A. Iosup, "Design and Experimental Evaluation of Distributed Heterogeneous Graph-Processing Systems," Submission to *CCGrid, 2016.*

# 3 Families Explored: 2 Lessons Learned




- PageRank, 4 machines
- Also tried BFS and WCC

1. There is no overall winner, but C-R is in general the worst.
2. Our new PG policy for Combined systems shows good performance.

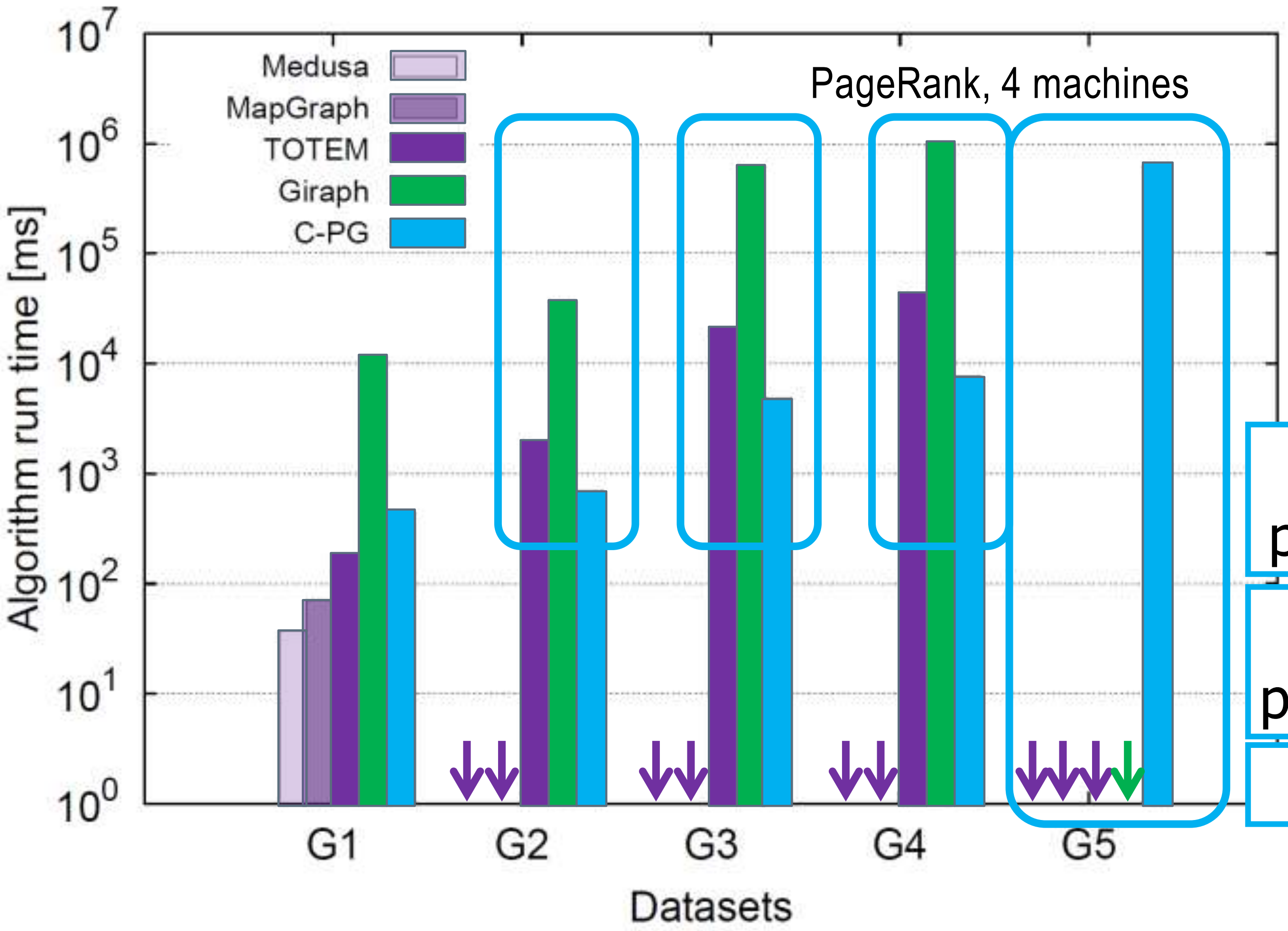
Y. Guo, A. L. Varbanescu, D. Epema, and A. Iosup, "Design and Experimental Evaluation of Distributed Heterogeneous Graph-Processing Systems," Submission to *CCGrid, 2016.*



TU Delft

DELFT DATA SCIENCE

# Promising Results for Distributed *and* Heterogeneous Graph-Processing Systems




Distributed *and* Heterogeneous

Distributed only

Heterogeneous only

Missing bar = ➔ = system failure

C-PG system has good performance iff. large data

TOTEM has good performance iff. in-memory

C-PG system scales

TU Delft

Y. Guo, A. L. Varbanescu, D. Epema, and A. Iosup, "Design and Experimental Evaluation of Distributed Heterogeneous Graph-Processing Systems," Submission to *CCGrid, 2016.*

DELFT DATA SCIENCE

# Massivizing Graph-Processing Systems

Interactive

5' — Pitch on Massivizing Graph-Processing Systems

5' — Two Exemplary Steps Forward

- Benchmarking distributed *or* heterogeneous graph-processing systems
- Designing distributed *and* heterogeneous graph-processing systems

5' — Towards a Taskforce on Data Science as a Service

# Take-Home Message

## (Cloud computing +) Big Data

## Graph processing as example

## Important New Challenges

1. Benchmarking
2. Distributed Heterogeneous Systems
3. …

TU Delft

DELFT DATA SCIENCE

# Next? A Taskforce on Data Science as a Service

## Identify industry needs in the Netherlands

- Stakeholders: datacenter operators, ICT designers, ICT analysts, ICT researchers, governance, ICT media

## Establish a joint research agenda, between fundamental and applied research

- Groundbreaking ideas for important challenges
- Prototypes and Proof-of-Concepts, not only ideas

## Build a solid, pragmatic collaboration

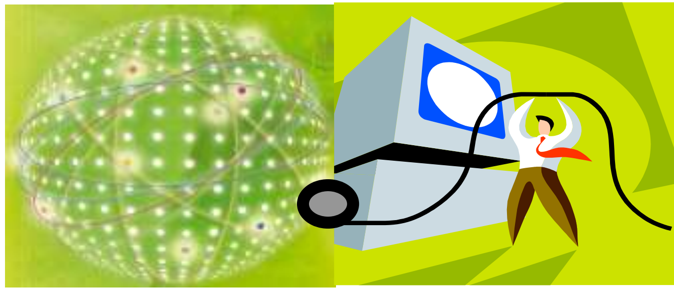
- Relevant recommendations for relevant problems
- Embedding of human resources, joint networking, etc.

Staff members

# Contact Our Team

Collaboration or discussion about:

- Leveraging open-source / open-access cloud computing and big data systems
- Distributed *and* heterogeneous graph-processing




*A.Iosup@tudelft.nl*

*+31-15-2784433*

*@AIosup*

*http://pds.twi.tudelft.nl/~iosup/*

*https://www.linkedin.com/in/aiosup*

PDS Group, Faculty EEMCS, TU Delft

Room HB07.050, Mekelweg 4, 2628CD Delft


TUDelft

KIVI Engineering Society

DELFT DATA SCIENCE

# Recommended Reading

## Elastic Big Data and Computing

- B. Ghit, N. Yigitbasi (Intel Research Labs, Portland), A. Iosup, and D. Epema. Balanced Resource Allocations Across Multiple Dynamic MapReduce Clusters. SIGMETRICS 2014
- L. Fei, B. Ghit, A. Iosup, D. H. J. Epema: KOALA-C: A task allocator for integrated multicluster and multicloud environments. CLUSTER 2014: 57-65
- K. Deng, J. Song, K. Ren, A. Iosup: Exploring portfolio scheduling for long-term execution of scientific workloads in IaaS clouds. SC 2013: 55

## Time-Based Analytics

- B. Ghit, M. Capota, T. Hegeman, J. Hidders, D. Epema, and A. Iosup. V for Vicissitude: The Challenge of Scaling Complex Big Data Workflows. Winners IEEE Scale Challenge 2014

## Graph Processing / Benchmarking

- M. Capota, T. Hegeman, A. Iosup, A. Prat-Pérez, O. Erling, P. A. Boncz: Graphalytics: A Big Data Benchmark for Graph-Processing Platforms. GRADES@SIGMOD/PODS 2015: 7:1-7:6
- A. L. Varbanescu, M. Verstraaten, C. de Laat, A. Penders, A. Iosup, H. J. Sips: Can Portability Improve Performance?: An Empirical Study of Parallel Graph Analytics. ICPE 2015: 277-287
- Y. Guo, A. L. Varbanescu, A. Iosup, and D. Epema. An Empirical Performance Evaluation of GPU-Enabled Graph-Processing Systems, IEEE/ACM CCGRID 2015: 423-432
- Y. Guo, M. Biczak, A. L. Varbanescu, A. Iosup, C. Martella, T. L. Willke: How Well Do Graph-Processing Platforms Perform? An Empirical Performance Evaluation and Analysis. IPDPS 2014: 395-404

TUDelft

DELFT DATA SCIENCE

# Disclaimer: images used in this presentation obtained via Google Images.

- Images used in this lecture courtesy to many anonymous contributors to Google Images, and to Google Image Search.
- Many thanks!

TUDelft

DELFT DATA SCIENCE