Massivizing Distributed Systems = Making Systems Scalable, Reliable, High-Performance, etc., Yet Efficient

dr. ir. Alexandru Iosup
Parallel and Distributed Systems Group

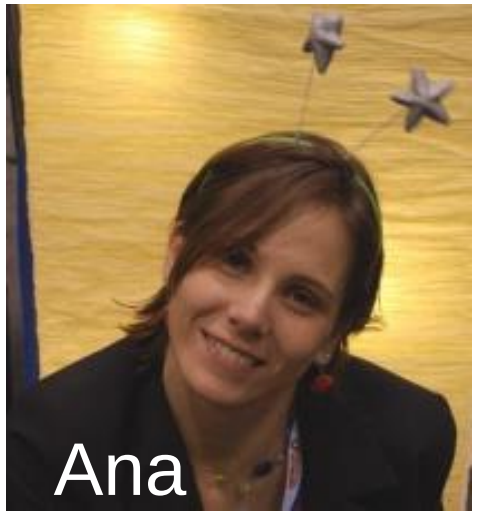
COMMIT/

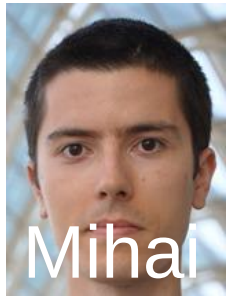
TU Delft

DELFT DATA SCIENCE

Important Things First:

Many, Many Thanks To Many Family, Friends,





Finding My Problem

(Being Ishmael, not Ahab)

In the Digital Economy, Data Services Are Expensive, Few Can Afford Being Successful!


Creator


Creator


Creator


Creator

Opportunity To Create

My Research

NETFLIX

Dropbox

CERN

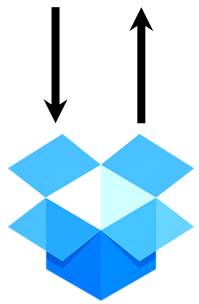
"ICT is vital for SMEs, SMEs are 60% GDP"
"15% ICT market is simple cloud services"
"Already 60+ bn.€/year"

Sources: Eurostat'15,EC Digital Agenda,IDC'14

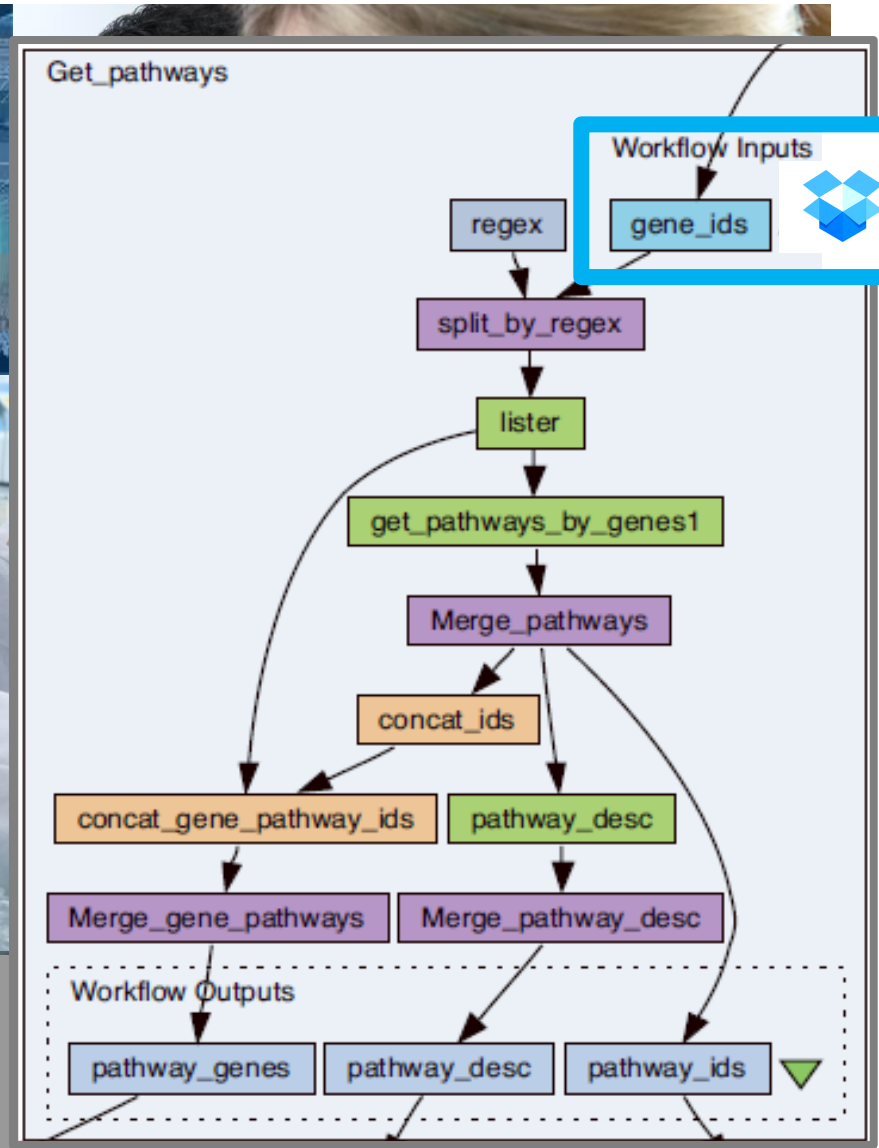
Simple Data Service vs. Complex Data Service

Dropbox

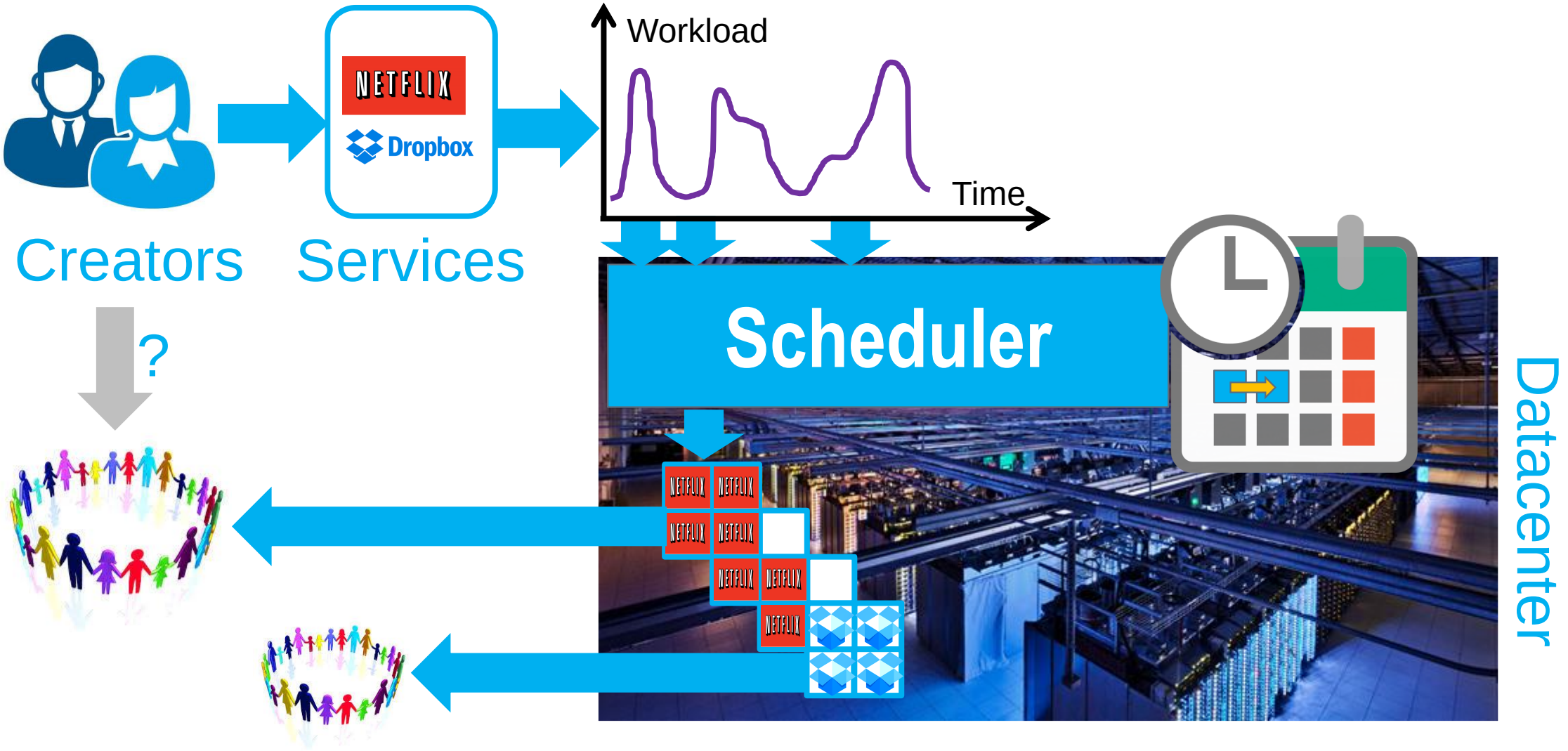
File Management



“too expensive, too inefficient”



Current Technology: Scheduler? Datacenter?



The Scheduling Challenge

“30—70% scheduler decisions incorrect in datacenters”

Source: IEEE Computer’15

“current schedulers not efficient for many users, diverse services”

Source: Dutch industry, CCGRID’15

→ Need Smarter Schedulers

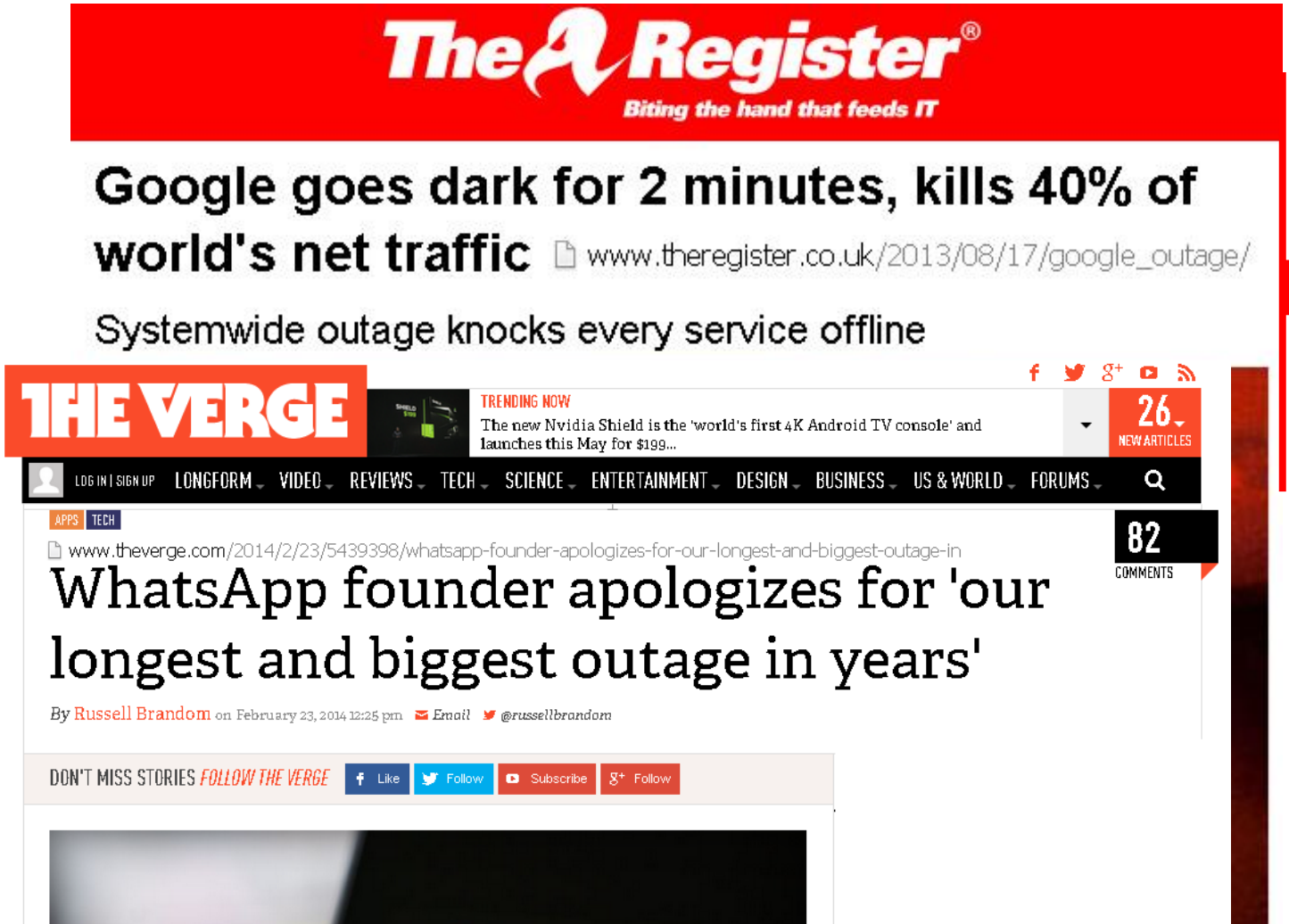
“new schedulers not used in datacenters, fear of failure”

Source: EuroPar’13,’14

→ Need to Select Schedulers

The Dependability* Challenge

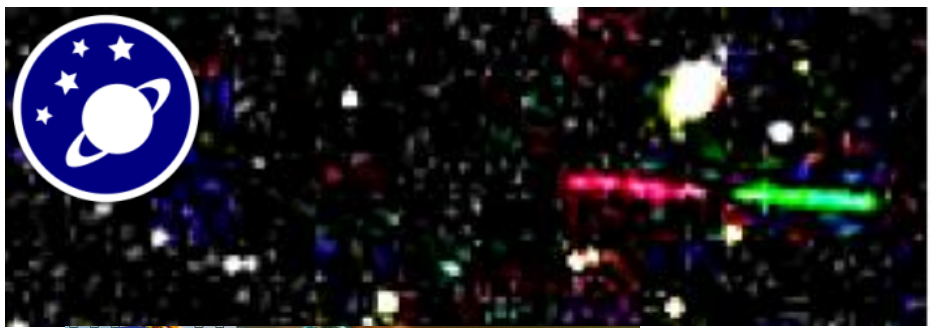
* Availability, Reliability, etc.

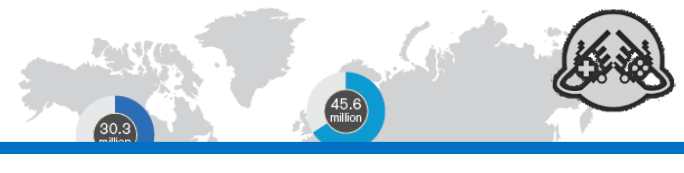


Need Dependable Systems

The New World Challenge

Cloud operator: new value-adding services, DevOps workloads





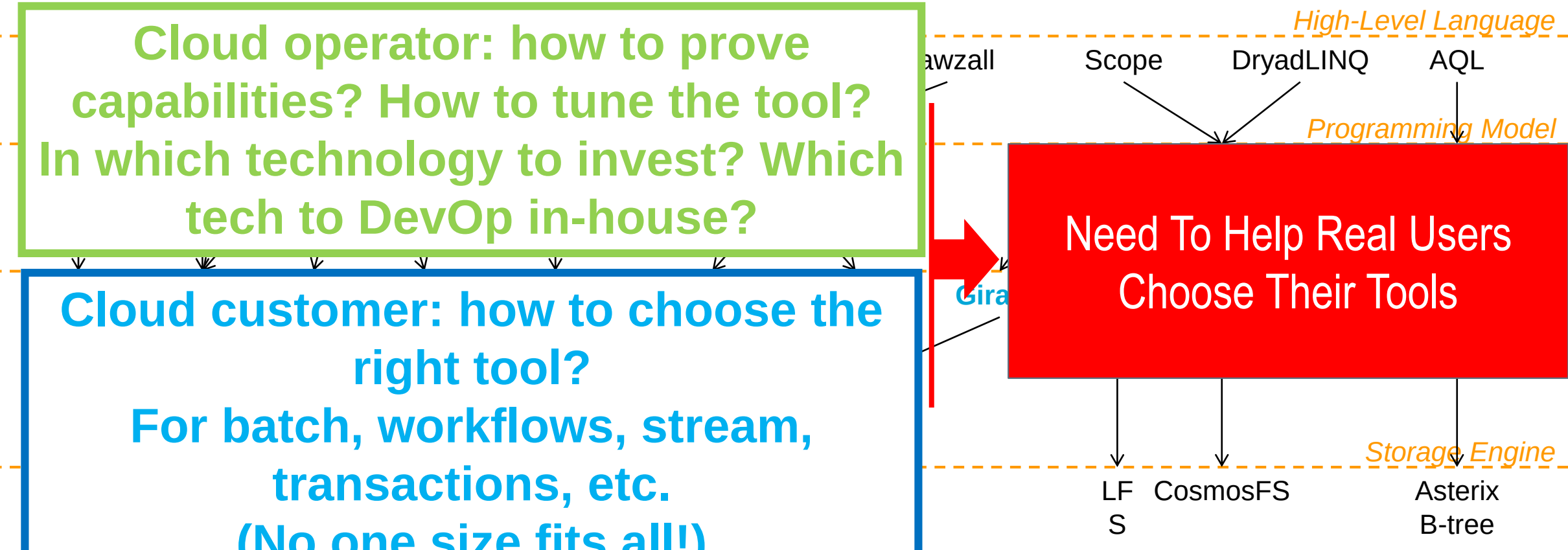
Need Operational Models

Cloud customer: new apps, new services, customers can become operators (value-chain)

The Ecosystem Navigation Challenge

Cloud operator: how to prove capabilities? How to tune the tool? In which technology to invest? Which tech to DevOp in-house?

Cloud customer: how to choose the right tool?
For batch, workflows, stream, transactions, etc.
(No one size fits all!)



Jevons Effect: More Efficient, Yet Less Capable

Over 500 YouTube videos have at least 100,000,000 viewers each.

If you want to help kill the planet:
https://www.youtube.com/playlist?list=PLirAqAtl_h2r5g8xGajEwdXd3x1s

Need To Be Much More Efficient, But Also To Educate Our Customers

PSY Gangnam consumed ~500GWh

= more than entire countries* in a year (*41 countries),
= over 50MW of 24/7/365 diesel, 135M liters of oil,
= 100,000 cars running for a year, ...

Source: Ian Bitterlin and Jon Summers, UoL, UK, Jul 2013.
Note: Psy has >3 billion views (Nov 2015).

The New "Jevons Effect": The "Data Deluge" Challenge

Data Deluge =

Need To Address The "Data Deluge"

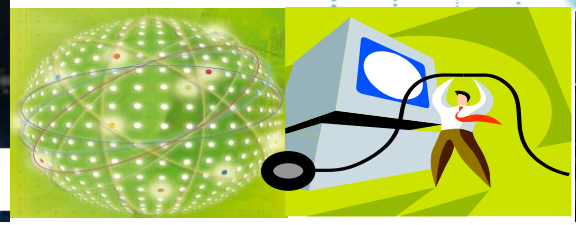
- Creating

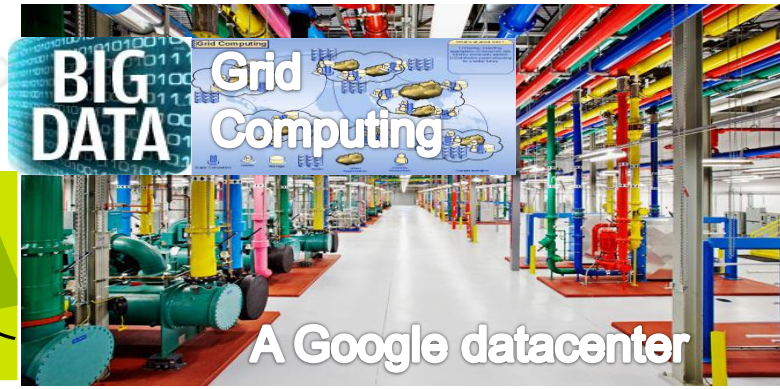
To be capable of processing Big Data, need to address Volume, Velocity, Variety of Big Data*

* Other Vs possible: ours is "vicissitude"

Sources: IDC, EMC.

Scientific and Technical Challenges





To build what we need as ICT infrastructure and platform

- Super-scalable, super-flexible
- End-to-end, large-scale automation
- Complex services forming dynamic workloads
- Evolving, heterogeneous hardware and software
- Under strict performance & cost & energy & reliability & … requirements
- … but efficient, and without assuming expertise from most customers

Need To Massivize

My Contribution To Computer Science, So Far

(thank you, Mary Shaw, for a great keynote!)

Massivizing Distributed Systems

Scheduling
Bags-Of-Tasks
Workflow
Mixed-Workload
Portfolio

Dependability
Failure Analysis*
Space-/Time-Correlation
Availability-On-Demand

New World
Workload Modeling
Interaction Graphs
Business-Critical
Online Gaming

Ecosystem Navigation
Performance Variability
Grid*, Cloud, Big Data
Benchmarking
Longitudinal Studies

Scalability/Elasticity
Delegated Matchmaking*
POGGI*
Area-Of-Simulation
BTWorld*
Auto-Scalers

Socially Aware Techniques
Collaborative Downloads*
Groups in Online Gaming
Toxicity Detection*

Software Artifacts
Graphalytics, etc.

Data Artifacts
A Distributed Systems Memex*

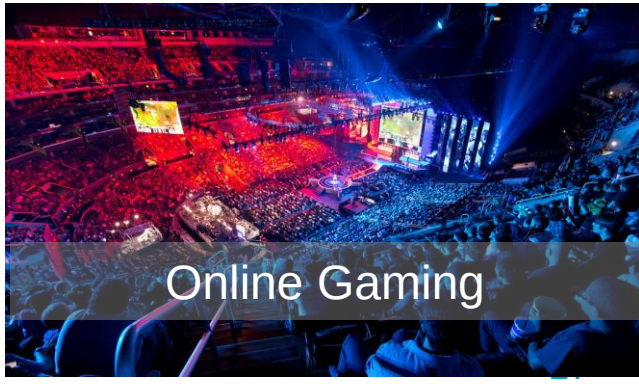
Fundamental Problems
My Contribution So Far (* Award-winning)

Knowledge Utilization and Impact

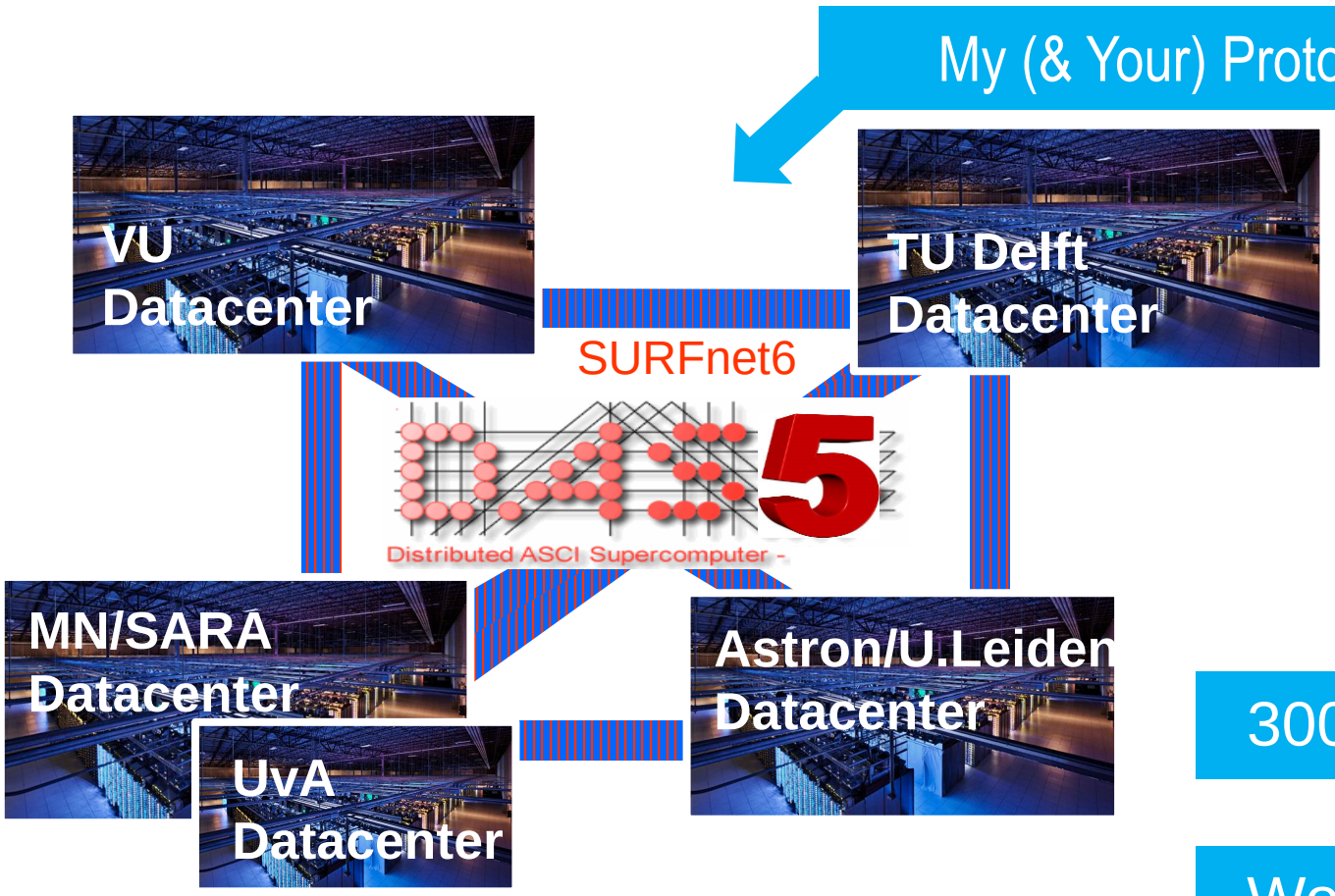
Typical Advisors Collaborators

+

Application Domains



Experimental Research Methodology
My Main Scientific Instrument: DAS-5



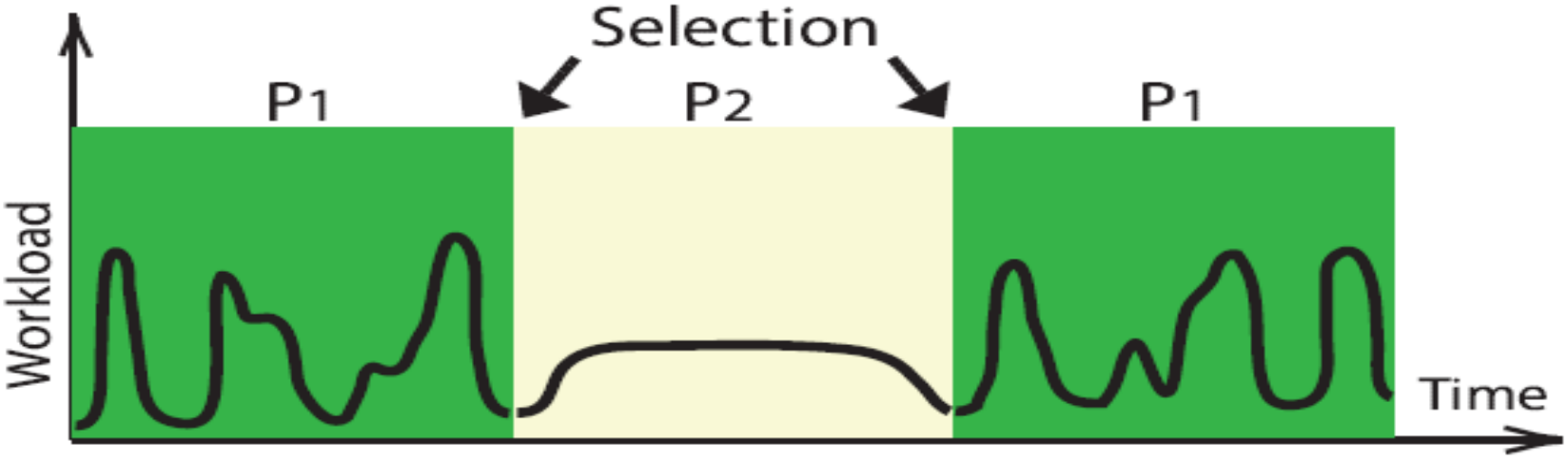
300+ scientists as users

Won IEEE Scale Challenge 2014

An Example:
Portfolio Scheduling for Datacenters
(what's in a name)

Portfolio Scheduling, In A Nutshell

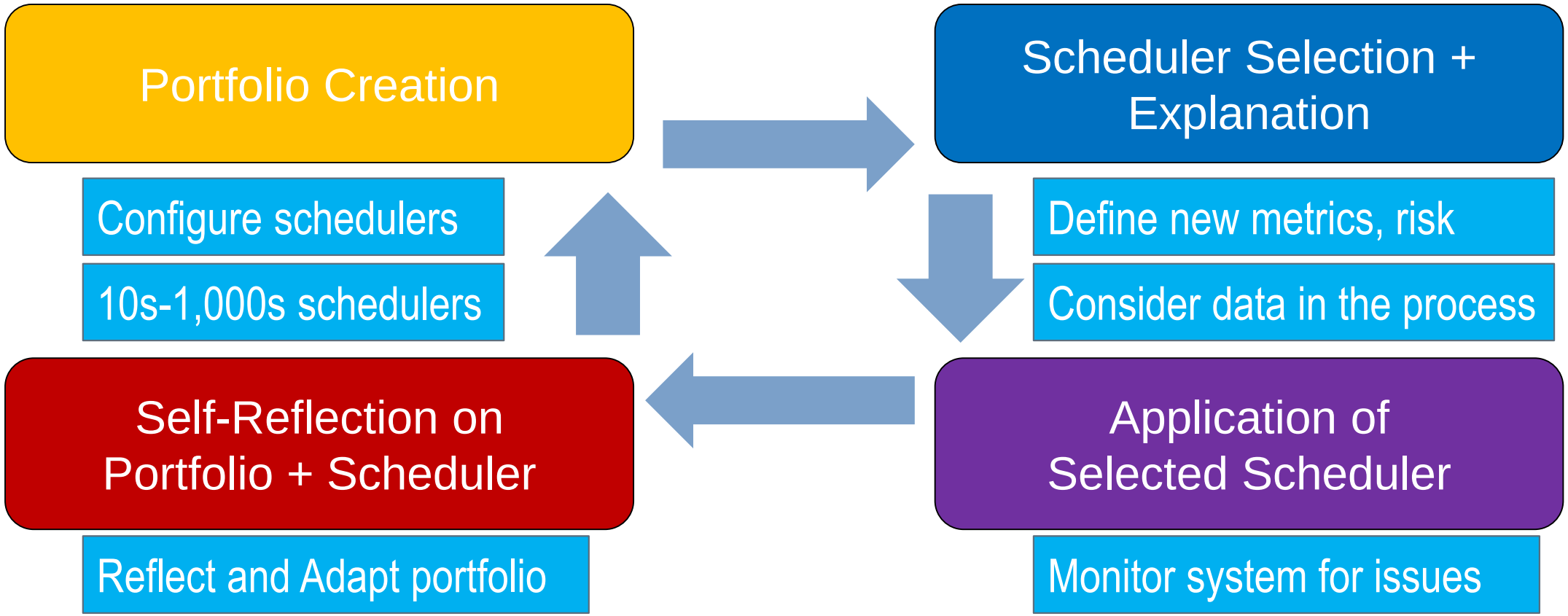
- Datacenters cannot work without one or even several schedulers
- Instead of ephemeral, risky schedulers, I propose to



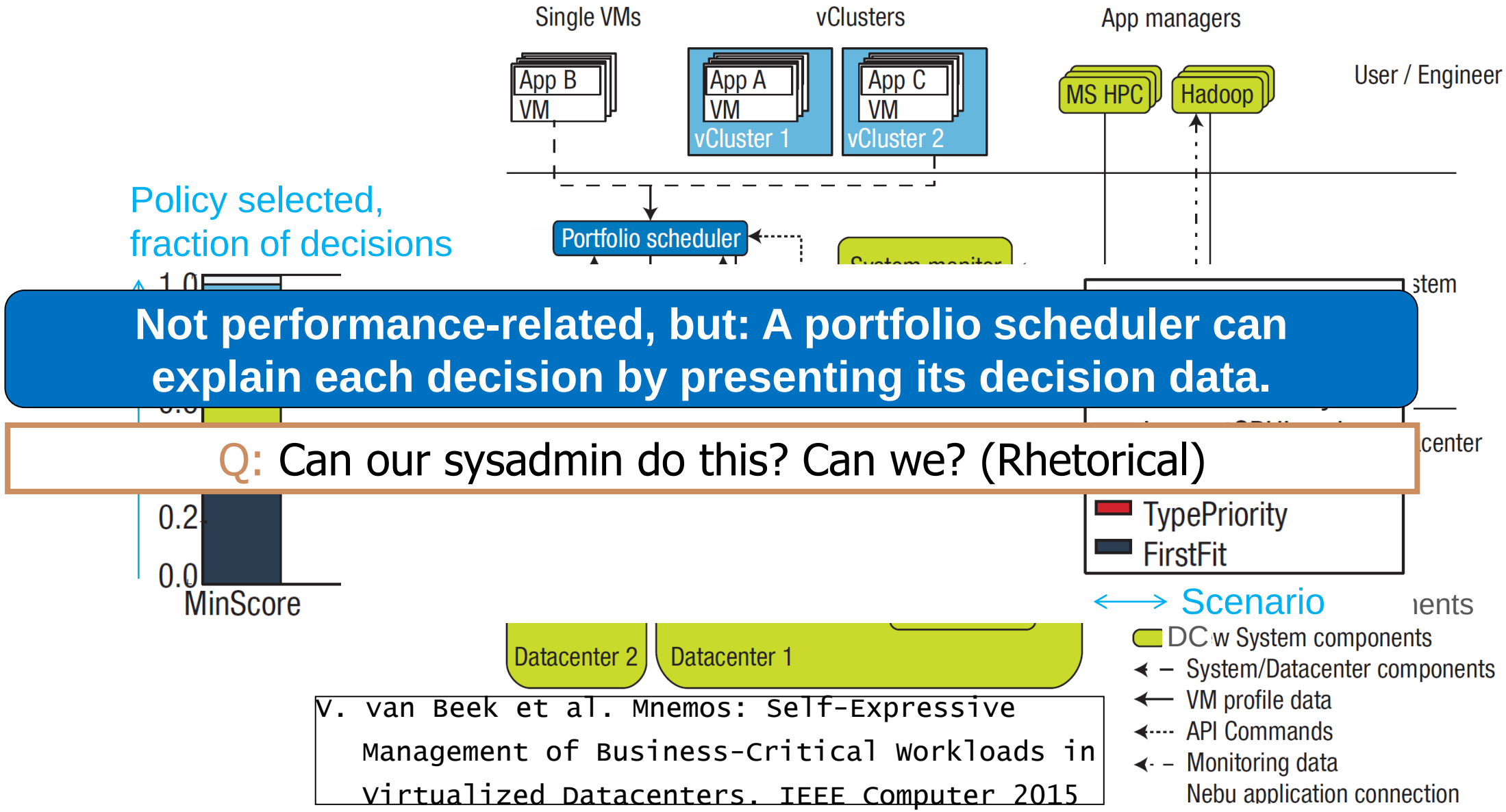
1. Create a set of schedulers
   - Resource provisioning and allocation policies for datacenters
2. Online selection of the active scheduler, for the next period

Portfolio Scheduling

Portfolio Scheduling for Datacenters running Complex Data Services



Portfolio Scheduling in Practice



Not performance-related, but: A portfolio scheduler can explain each decision by presenting its decision data.

Q: Can our sysadmin do this? Can we? (Rhetorical)

V. van Beek et al. Mnemos: Self-Expressive Management of Business-Critical Workloads in Virtualized Datacenters. IEEE Computer 2015

Ongoing Research

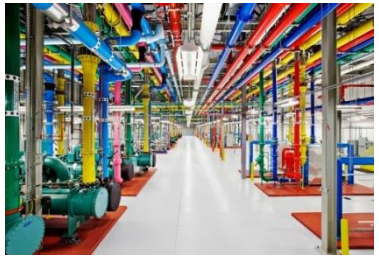
(for an incurable optimist,
the future is always better)

MagnaData: A Needed Step to Enable All Data Services for All People

Creator

MagnaData Scheduler

Efficiency



Portfolio Scheduling

+

Social Awareness

+

Per-Task Support

Thank you!

Massivizing Distributed Systems

Now everyone can be creative!

Innovation

Social Awareness

Portfolio Scheduling

Per-Task Support

Impact

salesforce

netherlands eScience center

spec Research

DELFT DATA SCIENCE

bitbrains IT SERVICES

Nerdalize

IBM

ORACLE

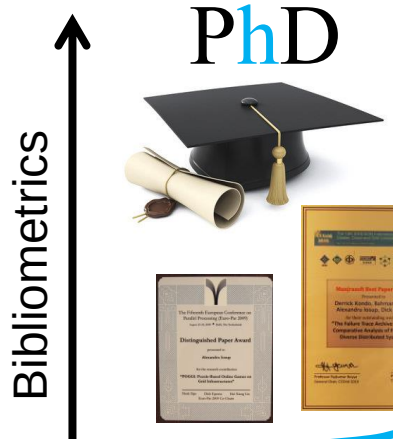
intel Google

SAP

Toward All Data Services to All People!

Thank you!
Extra Slides Follow.

MagnaData Is …
Also a Giant Leap for My Career



Icons