The Edge, the Cloud and the Supercomputer: Welcome to the Age of the Digital Continuum!

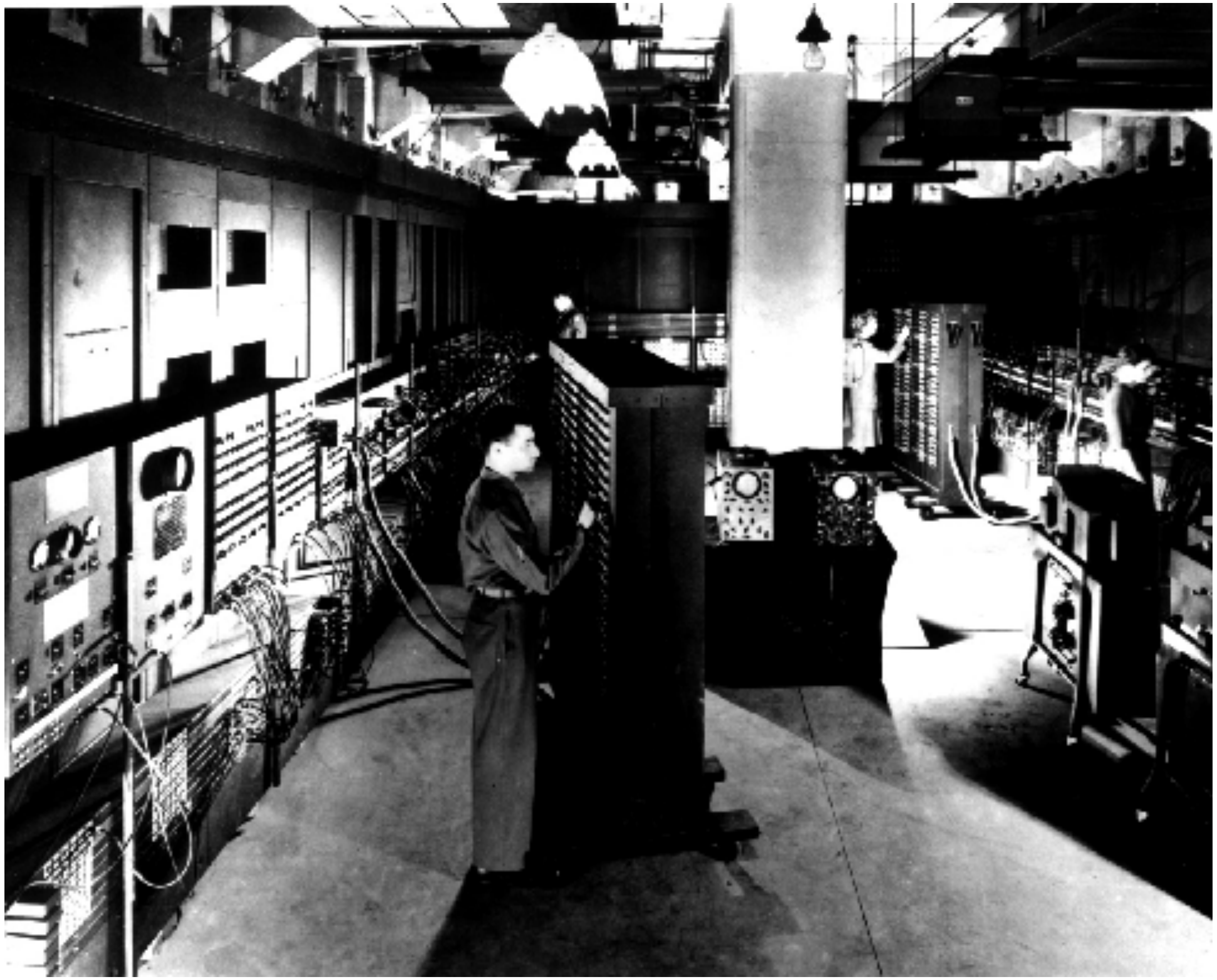
Gabriel Antoniu
Inria, Rennes

16 September 2020

Inria
UMR IRISA

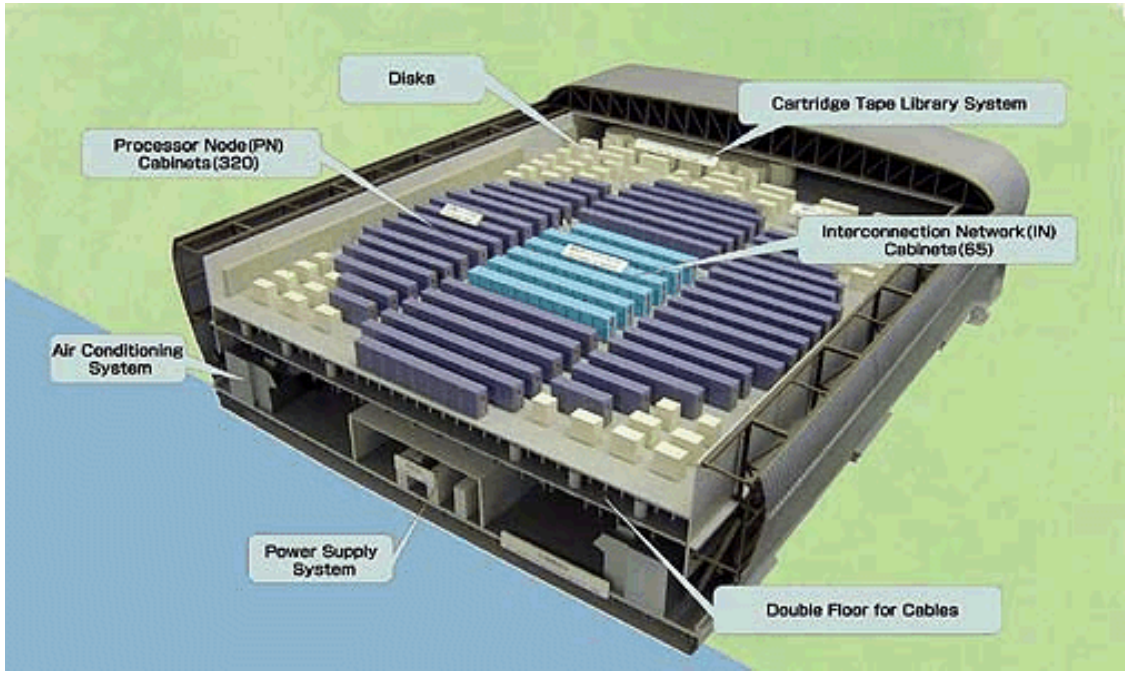
In the forties … the central computer

- Feb., 14 1946
- ENIAC
- 18.000 tubes, 30 tons, 170 m²
- 2.000 tubes replaced every months by 6 technicians



Earth simulator (2002)

- First computer to reach the Teraflops (10^{12} flops)
- Target Application: CFD-Weather, Climate
- 640 NEC SX/6 (mod)
  - 5120 Vector CPUs
- 40TeraFlops (peak)
- $400 million
  - $20-$30/y maintenance
- Size of a large concert hall
- Homogeneous, Centralized, Proprietary, Expensive!



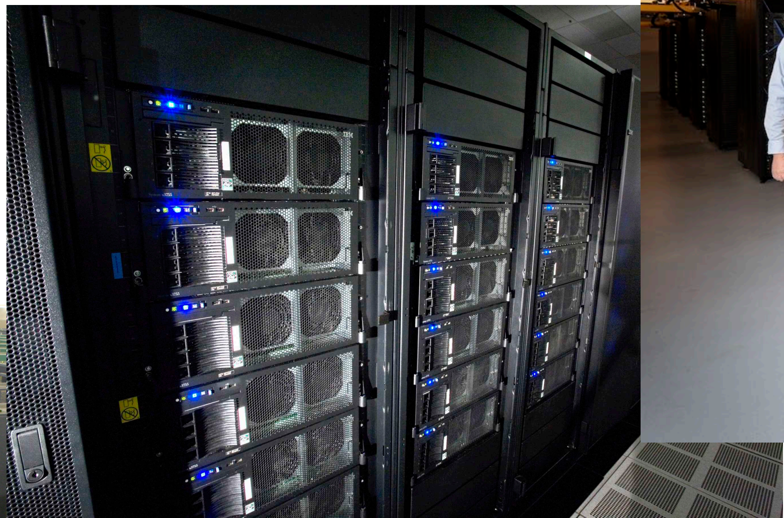
(Earth Simulator Picture from JAERI web page)

IBM Roadrunner (2008)

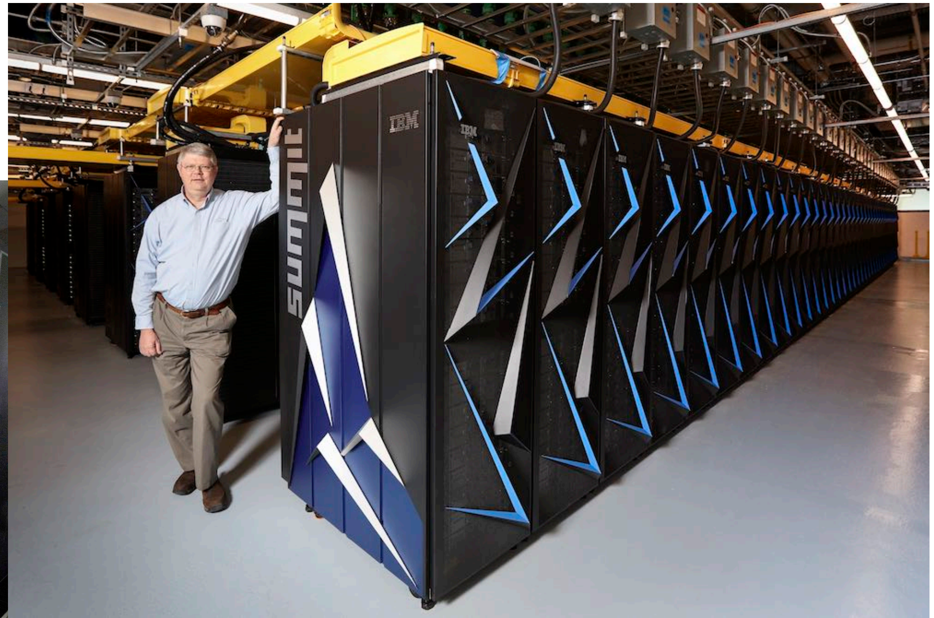
- First computer to reach the Petaflops (10^{15} flops)
- Roadrunner runs on
  - 6,948 dual-core AMD Opteron chips on IBM Model LS21 blade servers,
  - 12,960 Cell engines (same as PS3) on IBM Model QS22 blade servers.
- With 80 terabytes of memory, the Roadrunner system and is housed in 288 IBM BladeCentre racks occupying 6,000 square feet.
- 10,000 connections, both infiniband and gigabit Ethernet, with 57 miles of fibre-optic cable.

The Supercomputer and the Race for Performance

2002 - Earth Simulator
1 Teraflops (10^{12} flops)
NEC, Japan

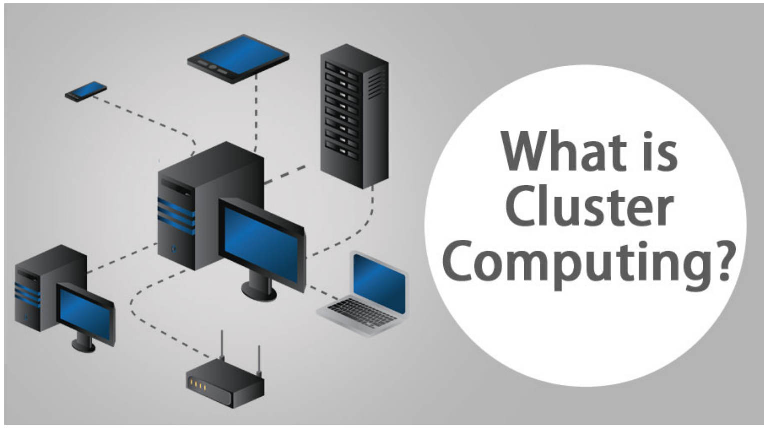


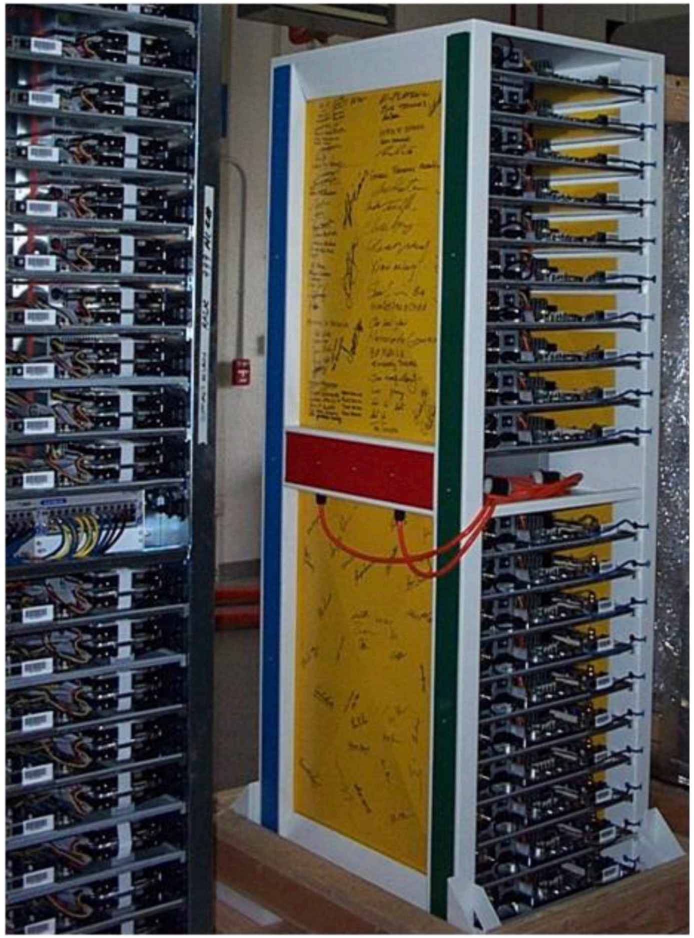
2008 - RoadRunner
1 Petaflops (10^{15} flops)
IBM, USA



Today - Summit
200 Petaflops
IBM, USA

Democratizing Performance: Cluster Computing

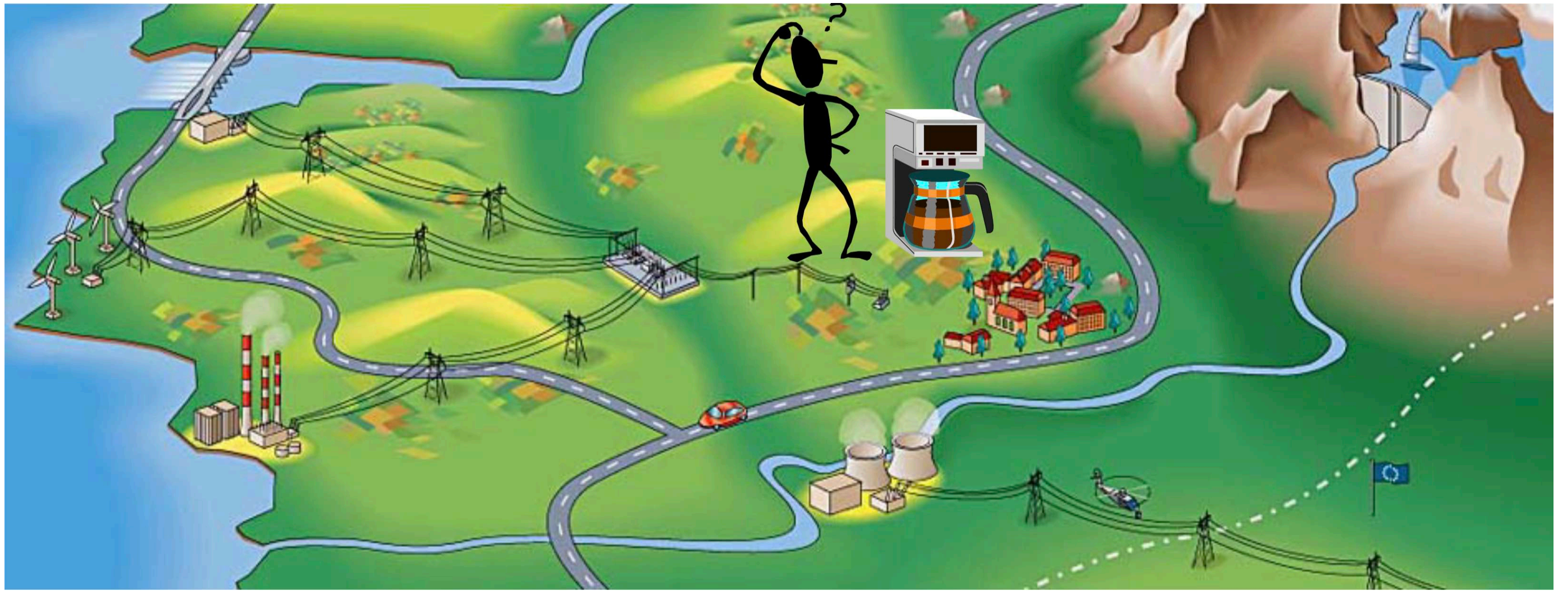




Google Cluster in 2011

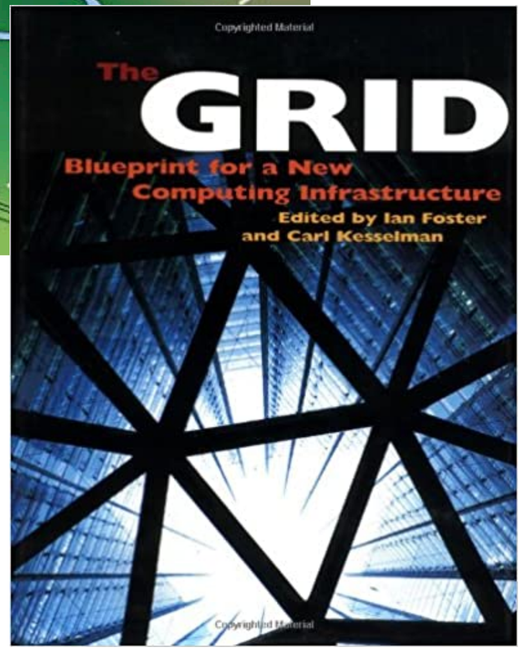
Google Cluster in 1997

The Power Grid Concept



Distributing electrical power

The Grid Computing Concept



Distributing computing power

Electric Power Grids and Computational Grids

- **Term coined by Ian Foster in the early 90s**
- **Power Grid analogy**
  – Power producers: machines, software, networks, storage systems
  – Power consumers: user applications
- **Applications draw power from the Grid the way appliances draw electricity from the power utility**
  – Seamless, High-performance, Ubiquitous, Dependable
- **Why the Computational Grid is like the Electric Power Grid**
  – Electric power is ubiquitous
  – Don't need to know the source of the power (transformer, generator) or the power company that serves it
- **Why the Computational Grid is different from the Electric Power Grid**
  – Wider spectrum of performance
  – Wider spectrum of services
  – Access governed by more complicated issues: Security, Performance

Credits: D. Petcu

Some books and a lot of hype !

Why Grids ?

- First need: supercomputing at a national or international scale
- Large size problems (grand challenge) need a collaboration between several codes/supercomputing centers
- Always a need of more computing power, memory capacity, and disk storage
- The power of any single resource is always small compared to the aggregation of several resources
- Network connectivity increased fast !

- Many available resources
  - Many clusters
  - Supercomputers
  - Millions of PC and workstations connected
  - Sharing or renting resources

- Increasing complexity of applications
  - Multi-scale
  - Multi-disciplinary
  - Huge data set produced
  - Heterogeneity

GRID Computing

- Allowing communities (virtual organizations) to share distributed resources to face common goals without
  - Centralized control
  - Global knowledge
  - Trust relations

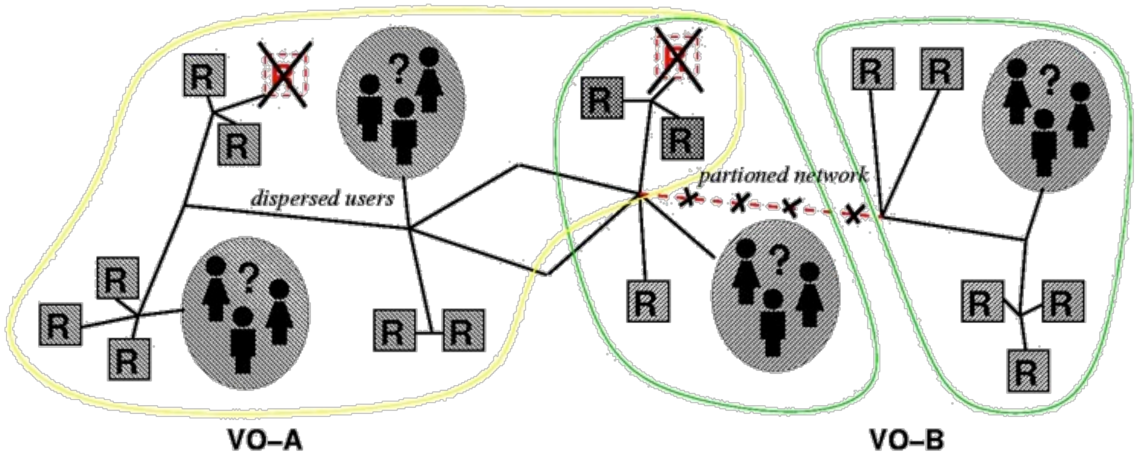
- Resources (supercomputers, visualization systems, sensors, instruments, people) are integrated through middleware
  - Goal : facilitate the use of these resources (or at least try to hide their complexity)

Elements of the Problem

- Sharing resources
  - Computers, storage, sensors, network, …
  - Their share is always conditional: trust problems, access policies, negotiation, payment, …

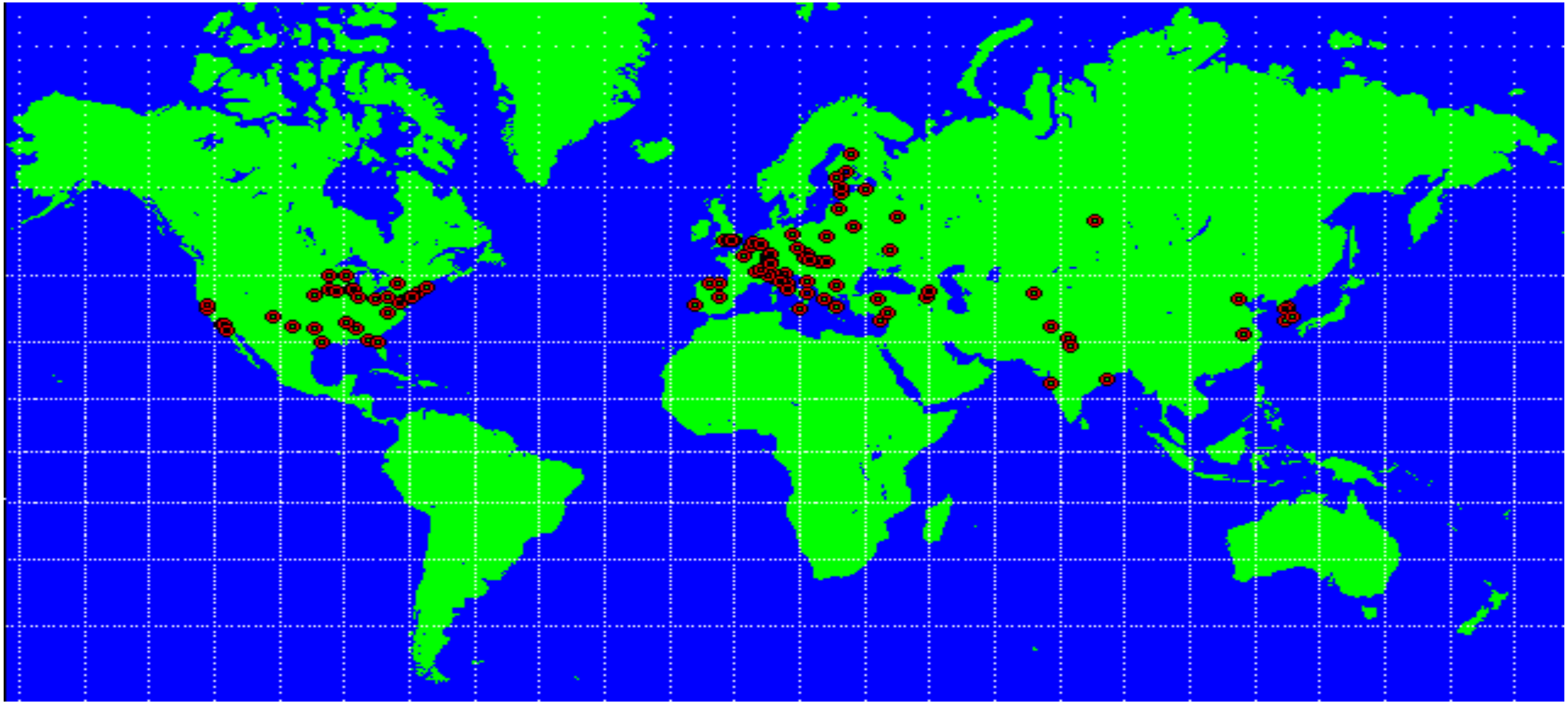
- Coordinated problem solving
  - More than “classical” client-server approach: data analysis, computation, distributed collaborations

- Dynamic and multi-institutional Virtual Organizations (VOs)
  - Overlap over classical organization structures
  - Small or large, static or dynamic



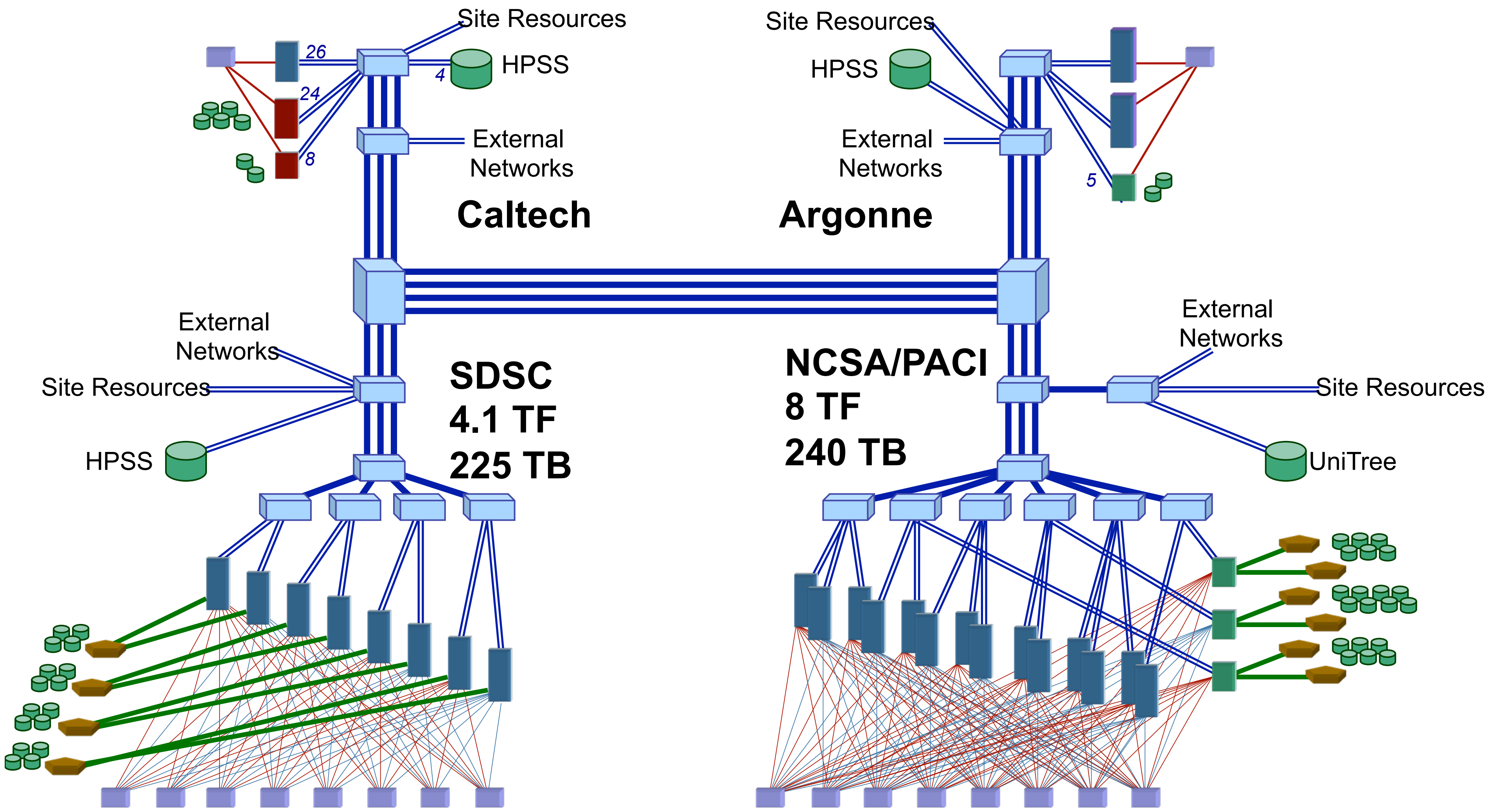
One example of a large VO: *CERN's Large Hadron Collider*

1800 physicists, 150 institutes, 32 countries



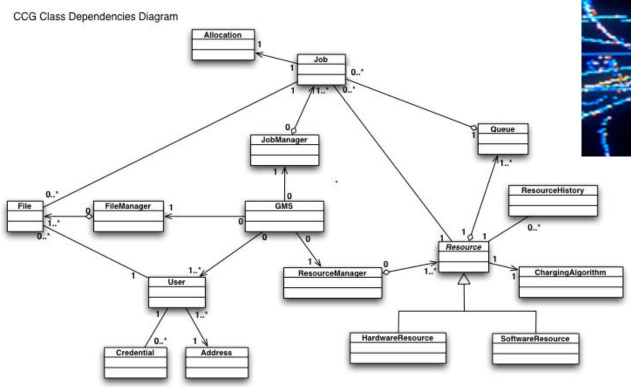
100 PB of data around 2010; 50,000 CPUs?

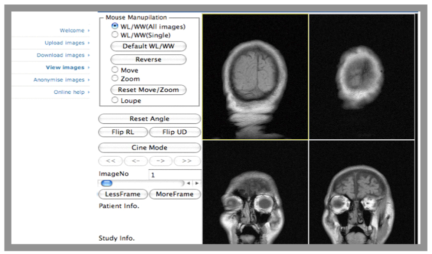
The 13.6 TF TeraGrid: Computing at 40 Gb/s

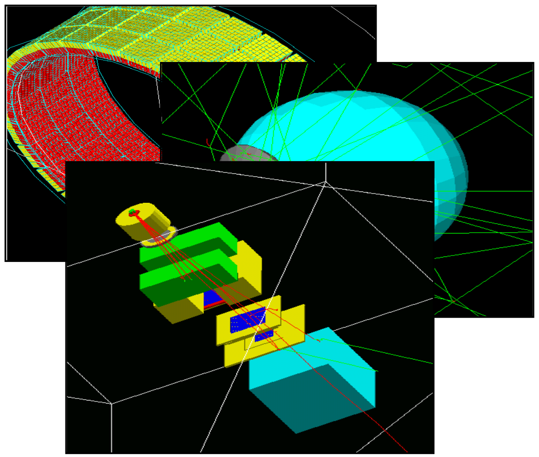


TeraGrid/DTF: NCSA, SDSC, Caltech, Argonne www.teragrid.org

Huge Applications Needs



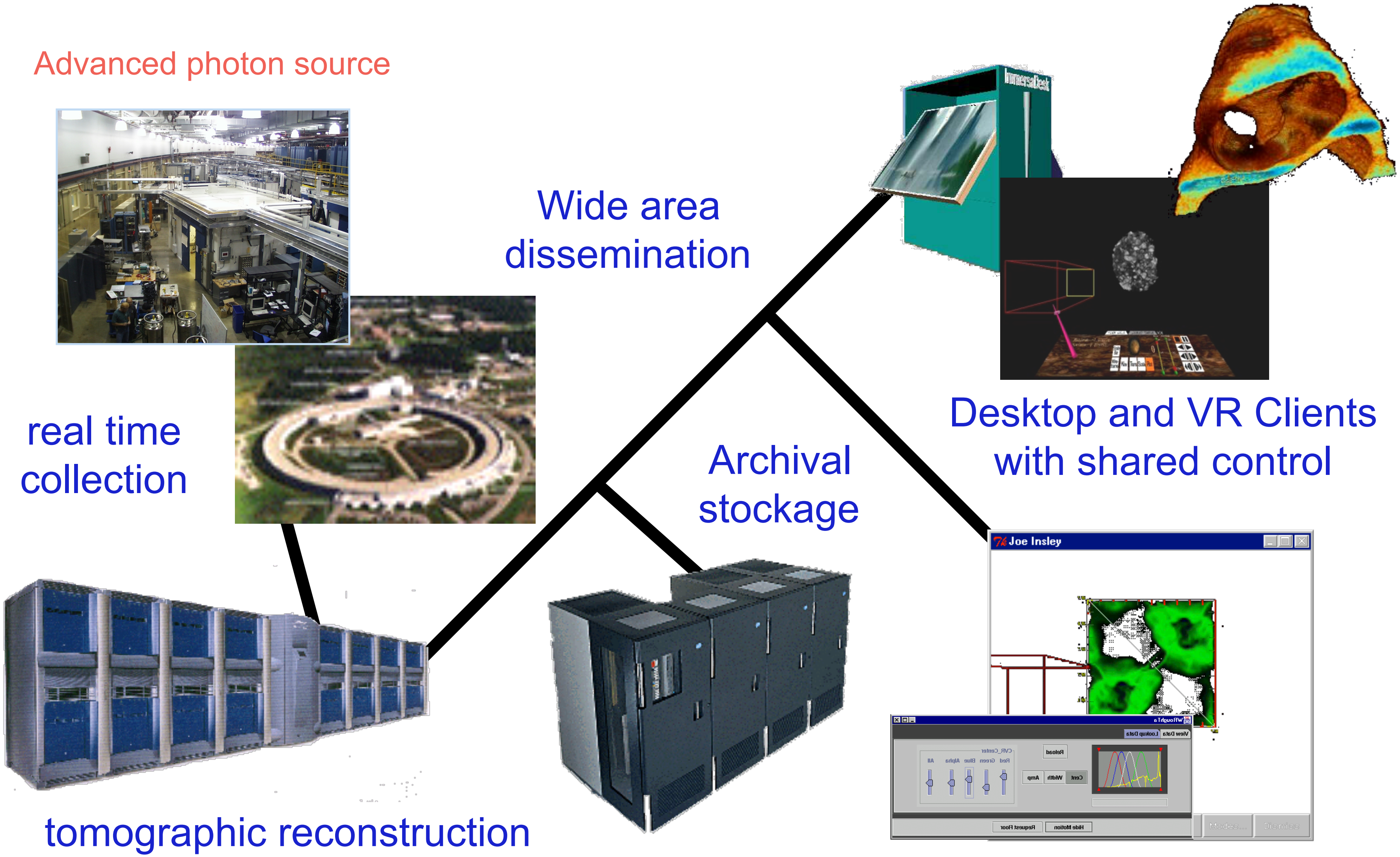




Different problems to be solved on Grids

- Computer-centric problems
  - Need teraflops, lots of them!
  - The Grid combines large computational resources
- Data-centric problems
  - The Grid is used to collect, store and analyze data maintained in geographically distributed repositories, digital libraries, and databases
- Community-centric problems
  - Collaborative applications --enable and enhance human-to- human interactions.
  - Provide a “virtual shared space”

Online Instruments

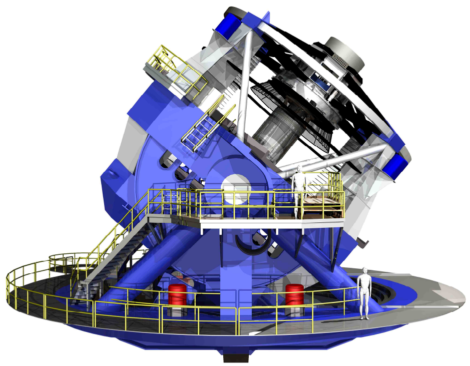


DOE X-Ray grand challenge: ANL, USC/ISI, U. Chicago

Large Synoptic Survey Telescope (LSST)

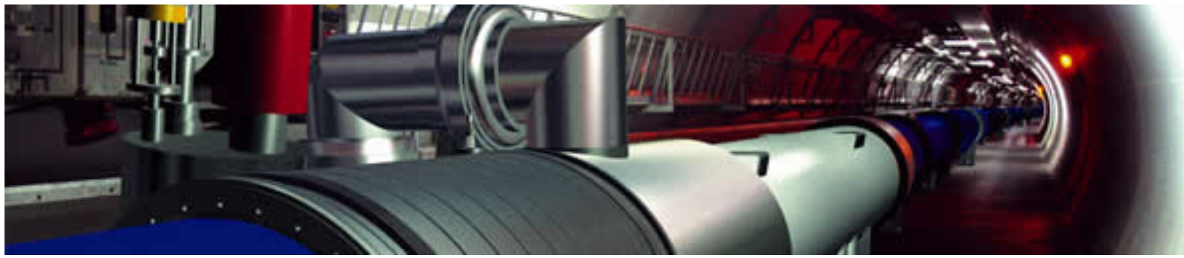
- Ground-based 8.4-meter, 10 square-degree-field telescope
- will provide digital imaging of faint astronomical objects across the entire sky, night after night
- 3 gigapixels detector for wide field imaging
- LSST will cover the available sky every three nights, opening a movie-like window on objects that change or move on rapid timescales:
  - exploding supernovae, potentially hazardous near-Earth asteroids, and distant Kuiper Belt Objects.
  - images used to trace billions of remote galaxies and measure the distortions in their shapes produced by lumps of Dark Matter, providing multiple tests of the mysterious Dark Energy.
- Data
  - >30 TB of data/night !

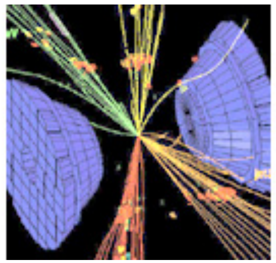
Large Synoptic Survey Telescope



Large Hadron Collider (LHC)

- Higher energy collisions are the key to further discoveries of more massive particles ($E=mc^2$)
- One particle predicted by theorists remains elusive: the Higgs boson
- The LHC is the most powerful instrument ever built to investigate elementary particles
- Beams of protons collision at an energy of 14 TeV with a 27 km circumference instrument.
- Data
  - 40 million collisions per second
  - After filtering, 100 collisions of interest per second
  - A Megabyte of data digitised for each collision = recording rate of 0.1 Gigabytes/sec
  - 1010 collisions recorded each year
  - = 10 Petabytes/year of data





Google

- Single search query touches 700 to up to 1k machines in less than 0.25sec
- There are more than 200 Google File System clusters
- The largest BigTable instance manages about 6 petabytes of data spread across thousands of machines

- MapReduce is increasingly used within Google
  - 29,000 jobs in August 2004 and 2.2 million in September 2007
  - Average time to complete a job has dropped from 634 seconds to 395 seconds
  - Output of MapReduce tasks has risen from 193 terabytes to 14,018 terabytes
  - Typical day will run about 100,000 MapReduce jobs each occupies about 400 servers
  - takes about 5 to 10 minutes to finish

Distributed Entertainment

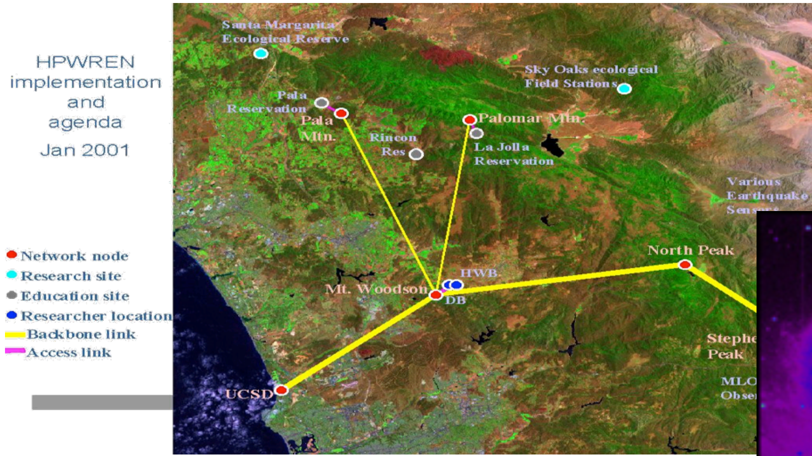
Everquest

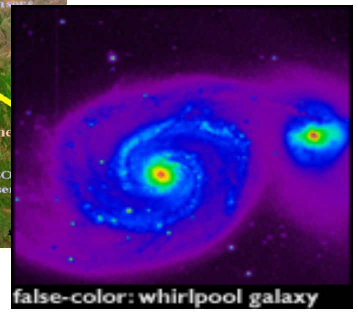
- 45 communal “world servers” (26 high-end PCs per server) supporting 430,000 players
- Real-time interaction, individualized database management, back channel communication between players
- Data management adapted to span both client PC and server to mitigate communication delays
- Game masters interact with players for real-time game management

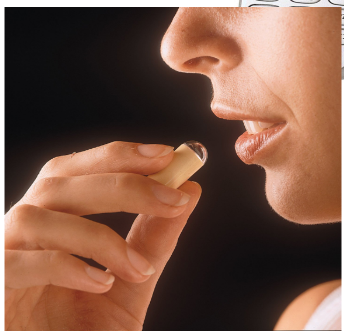
Next generation Grids will include new technologies

New devices

– PDAs, sensors, cars, clothes, smart dust, smart bandaids, …







Wired and Wireless

- HPWREN, Roadnet (Hans-Werner Braun, Frank Vernon et al.)
  – 45 Mbps between Mount Laguna telescope and SDSU, wireless access to Pala, Rincon, La Jolla Indian Reservations, etc.
  – Roadnet expanding to waterways, etc.

What about production platforms used for experimental purpose?

No reproducible experimental conditions:

- Scientific studies require reproducible experimental conditions

Not designed for experiments:

- Many researchers run short length, highly parallel & distributed algos
- Preparation and execution of experiments are highly interactive

Not optimized for experiments:

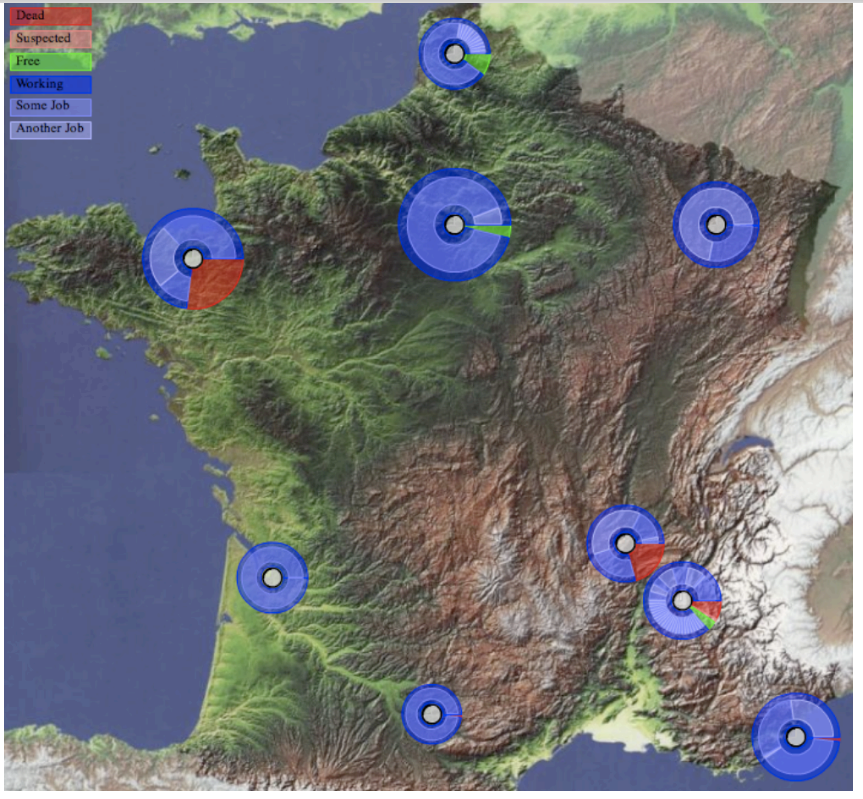
- Experimental platforms should exhibit a low utilization rate to allow researchers executing large collocated experiments

Not reconfigurable:

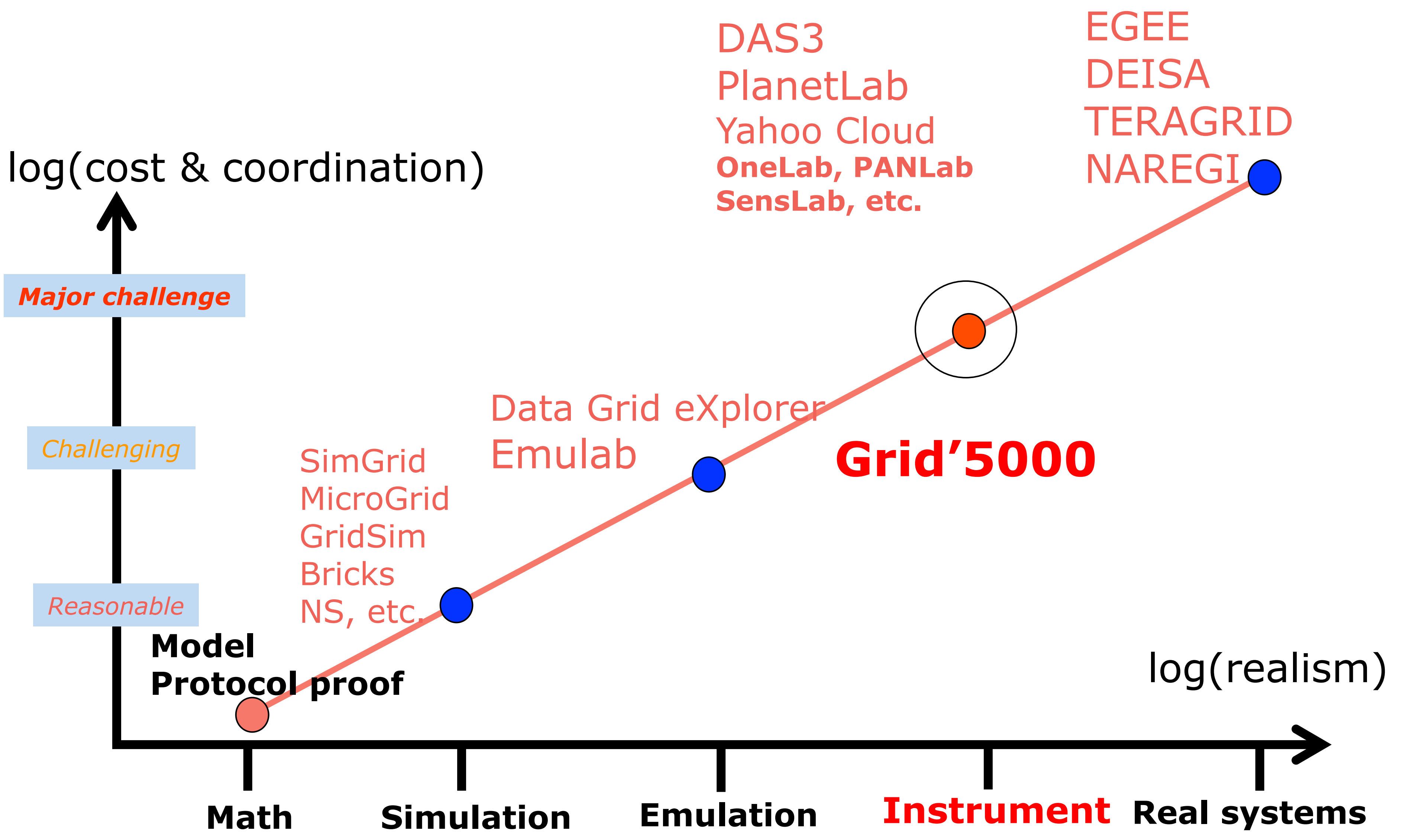
- Many projects require experiments on OS and networks,
- Some projects require the installation of specific hardware

→Nowhere to test networking/OS/middleware ideas, to measure real application performance

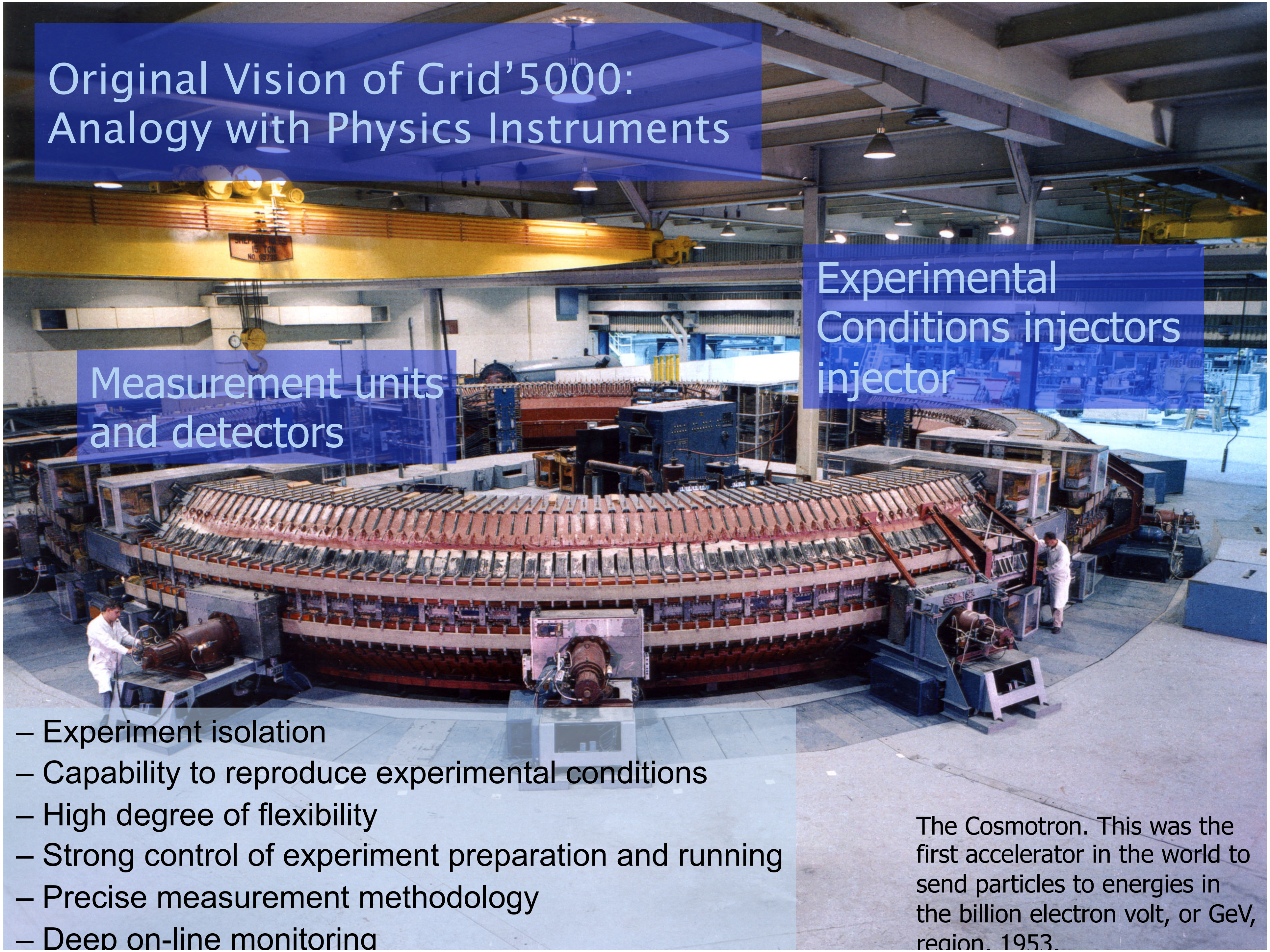
Grid'5000



Exploration tools



Original Vision of Grid’5000: Analogy with Physics Instruments



Measurement units and detectors

Experimental Conditions injectors injector

– Experiment isolation
– Capability to reproduce experimental conditions
– High degree of flexibility
– Strong control of experiment preparation and running
– Precise measurement methodology
– Deep on-line monitoring

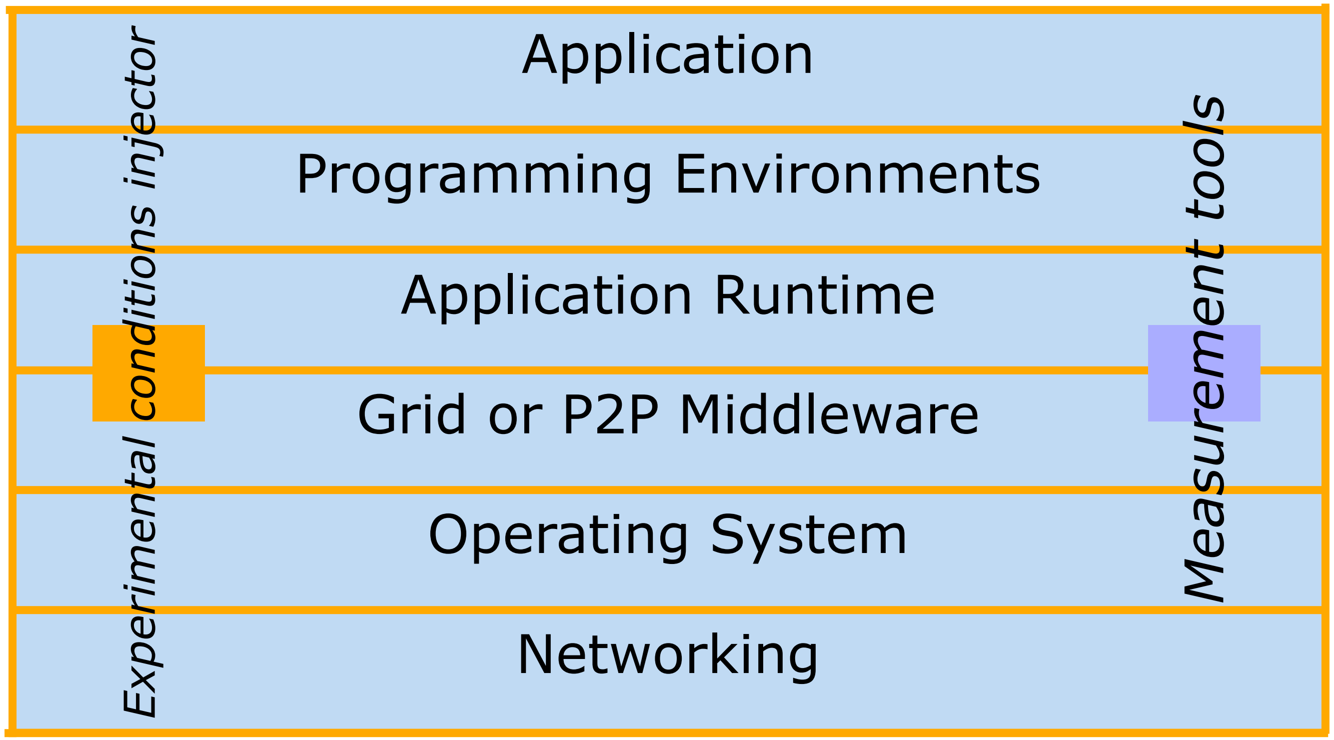
The Cosmotron. This was the first accelerator in the world to send particles to energies in the billion electron volt, or GeV, region. 1953.

Grid'5000 experimental model

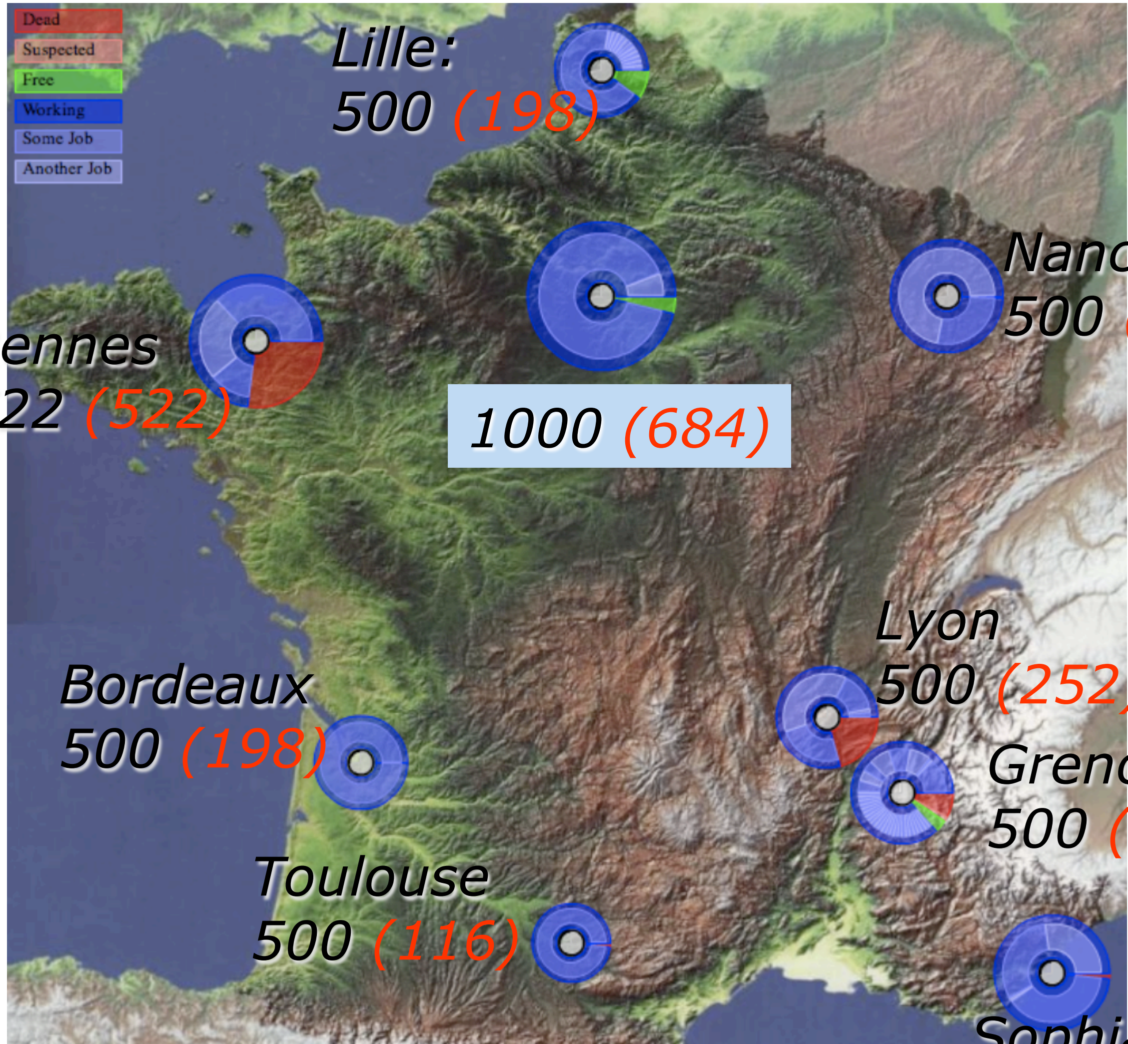
A highly controllable and reconfigurable experimental platform

Let users:

1) reserve experimental resources
2) create their software stack including injectors and probes
3) deploy their software environment on the reserved experimental resources and run their experiments

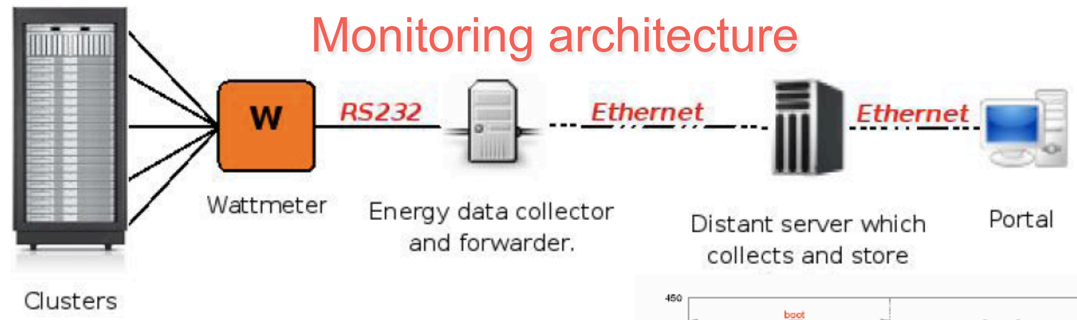


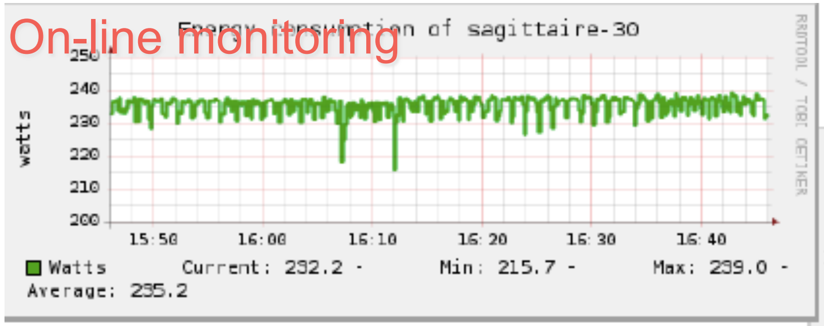
Grid'5000 map

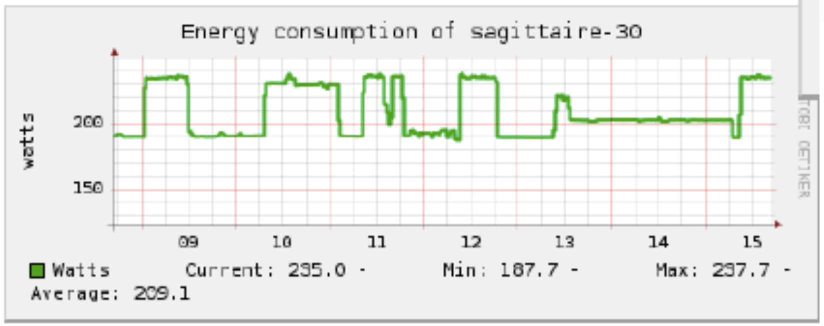


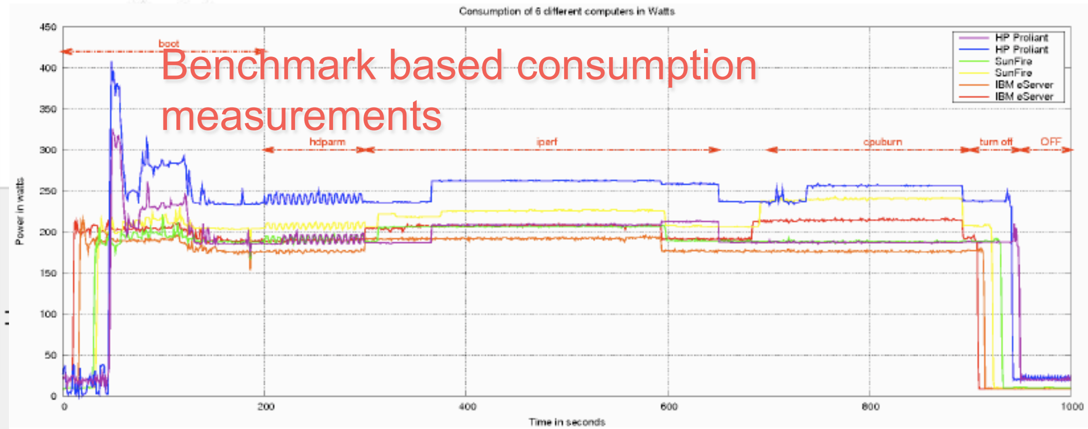
Dynamic Energy Consumption Monitoring

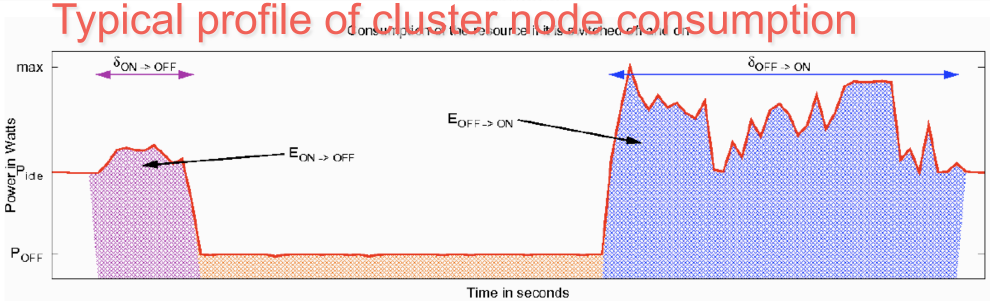
- Separate system monitoring Grid’5000
- Controllable and external energy sensors
  - --> on-line energy consumption measurements
  - --> required for developing energy saving algorithms











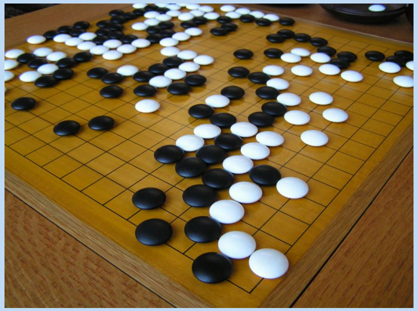
Example : Accelerating Research

Solving large instances of combinatorial optimization problems attracts the attention of researcher because of its algorithmic challenge. It always requires a huge number of computational resources.

Many success stories in combinatorial optimizations:

-solve the n-Queens problem for n=25, in 2005

-one of the **most promising one**, in 2008:
→Grid'5000 was used to design and improve the algorithm (MOGO) used in the **first computer victory against a professional Go player** (5 Dan) on a 9x9 plate in the last Paris tournament! (it's close to the Dan!)



•First

•

•

•

→Nev
comb
grid c

→ Pr
wit

Involv ... and
Toulo
The optimal solution required a wall-clock time of 25 days.

Challenges !

One View of Requirements

Identity & authentication
Authorization & policy
Resource discovery
Resource characterization
Resource allocation
(Co-)reservation, workflow
Distributed algorithms
Remote data access
High-speed data transfer
Performance guarantees
Monitoring

Adaptation
Intrusion detection
Resource management
Accounting & payment
Fault management
System evolution
Etc.
Etc.
…

Credits: Globus team

Another View: "Three Obstacles to Making Grid Computing Routine"

- New approaches to problem solving
  - Data Grids, distributed computing, peer-to-peer, collaboration grids, …
- Structuring and writing programs
  - Abstractions, tools

Programming Problem

- Enabling resource sharing across distinct institutions
  - Resource discovery, access, reservation, allocation; authentication, authorization, policy; communication; fault detection and notification; …

System Problem

Credits: Globus team

Programming & Systems Problems

- The programming problem
  - Facilitate development of sophisticated apps
  - Facilitate code sharing
  - Requires programming environments
    - APIs, SDKs, tools
- The system problem
  - Facilitate coordinated use of diverse resources
  - Facilitate infrastructure sharing
    - e.g., certificate authorities, information services
  - Requires systems
    - protocols, services

Credits: Globus team

What do we need (user point of view) ?

- **Single sign-on**
  - authentication to any resource of the Grid gives access to all resources
- **Single compute space**
  - one scheduler for all Grid resources
- **Single data space**
  - can address files and data from any Grid resources
- **Single development environment**
  - Grid tools and libraries that work on all grid resources

Credits: Globus team

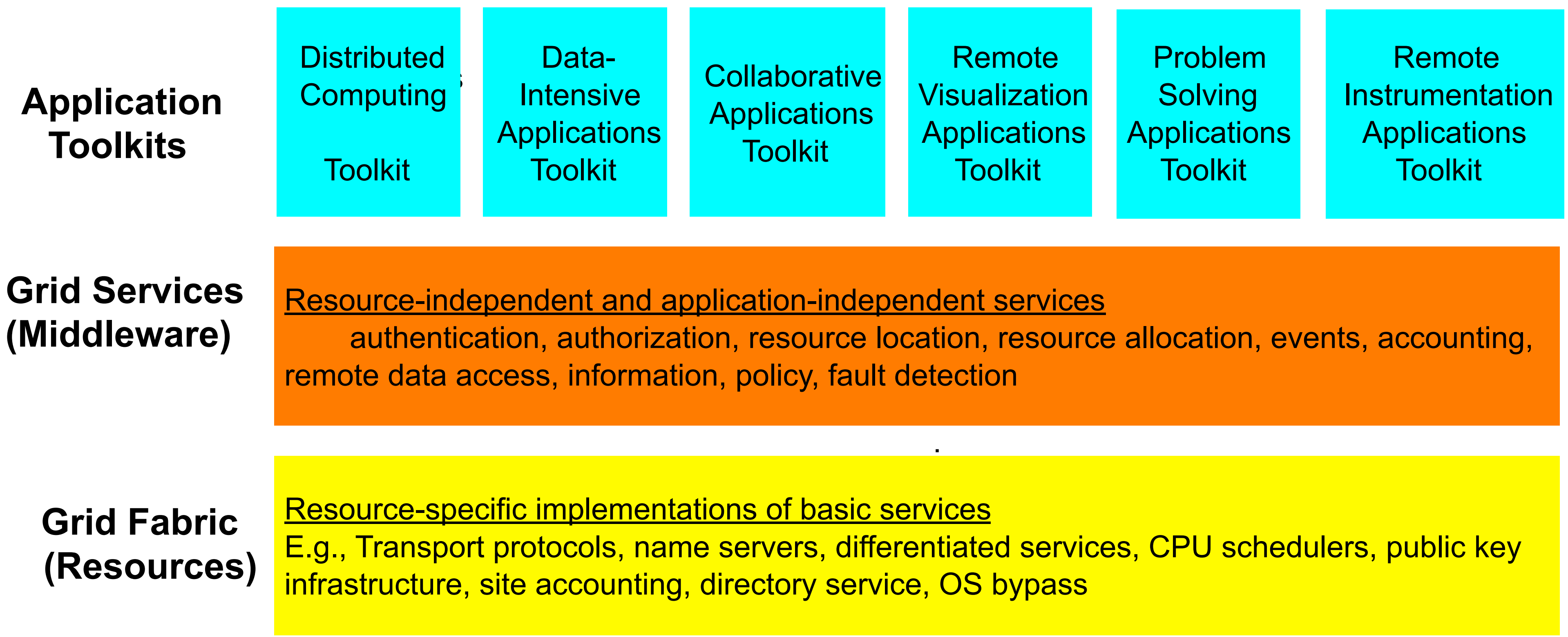
The System Problem: Resource Sharing Mechanisms That

- Address security and policy concerns of resource owners and users
- Are flexible enough to deal with many resource types and sharing modalities
- Scale to large number of resources, many participants, many program components
- Operate efficiently when dealing with large amounts of data & computation

Credits: Globus team

Which approach for grid programming

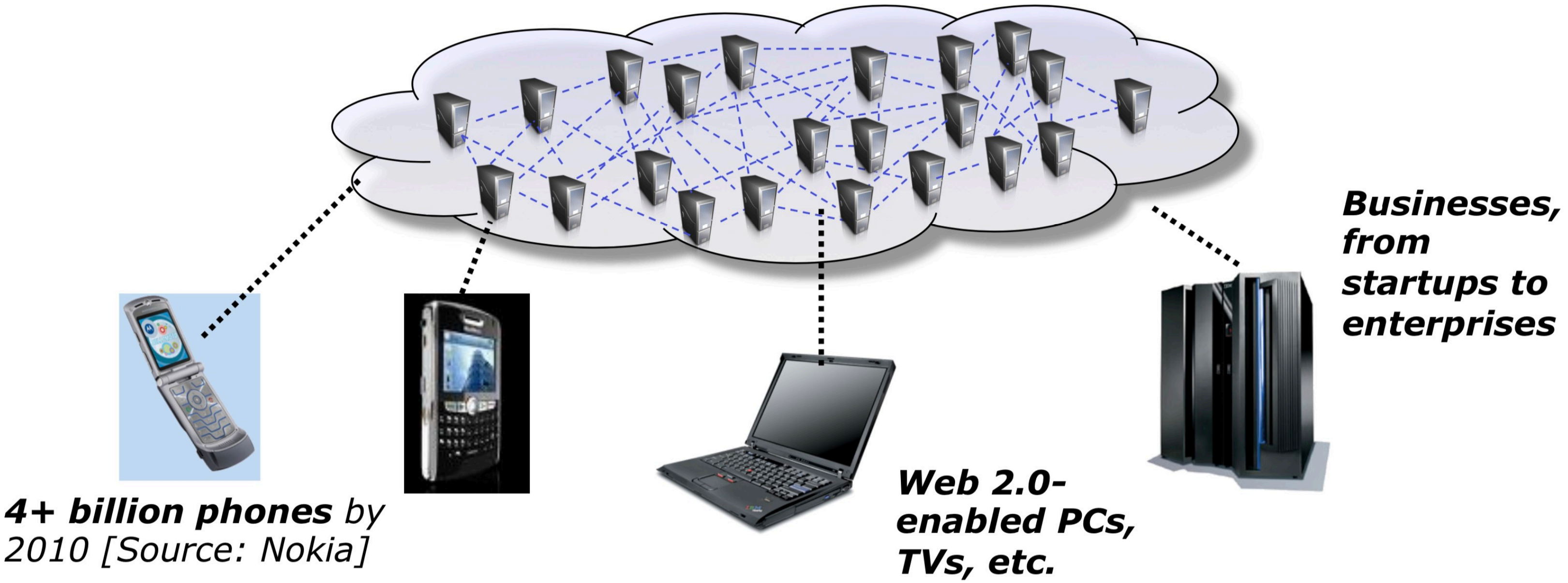
- MPICH-G2: classical message passing
- CoG kits, GridProt: web portals
- Condor-G, Pegasus, Taverna, …: workflows management
- Legion, Proactive: object model for the grid
- NetSolve/Ninf/DIET: Network Enabled Solver systems (GridRPC)
- Components approach
- Web services



2007: the Birth of Cloud Computing

What is Cloud Computing?

An emerging computing paradigm where data and services reside in massively scalable data centers and can be ubiquitously accessed from any connected devices over the internet.



Credit: IBM Corp.

Cloud Computing

- Two key concepts
  - Processing 1000x more data does not have to be 1000x harder
  - Cycles and bytes, not hardware are the new commodity
- Cloud computing is
  - Providing services on virtual machines allocated on top of a large physical machine pool
  - A method to address scalability and availability concerns for large scale applications
  - Democratized distributed computing
- Functionnalities
  - SaaS: Software as a Service
  - HaaS: Hardware as a Service
  - DaaS: Data as a Service
  - PaaS: Platform as a Service
  - IaaS: Infrastructure as a Service

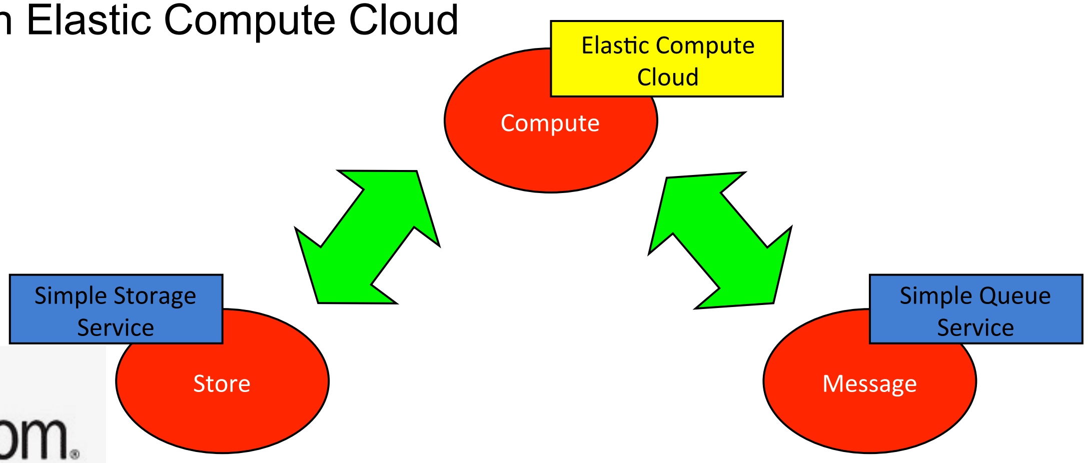
Cloud Computing Vendors

- http://www.johnmwillis.com/mysql/cloud-vendors-a-to-z/
- 3Tera, Adobe Air, Akamai, Amazon EC2, Amazon S3, Amazon SimpleDB, Apache CouchDB, Apache Hadoop, Areti Internet, Box-Net, Cassatt Corporation, Citrix (XenSource), CohesiveFT, Dell DCS, Elastra, EMC Mozy, Enki, Enomoly, EnomolyElasticDrive, EnterpriseDB, Flexiscale, Fortress ITX, Google Apps, HP AIaaS, IBM Blue Cloud, iCloud, Joyent, JungleDisk, Layered Technology, LongJump, Microsoft SSDS, MorphExchange, Mosso, Rackspace, Rightscale, Salesforce.com, Sun Caroline, Sun MySQL, Terremark, VMWare

Amazon Elastic Compute Cloud

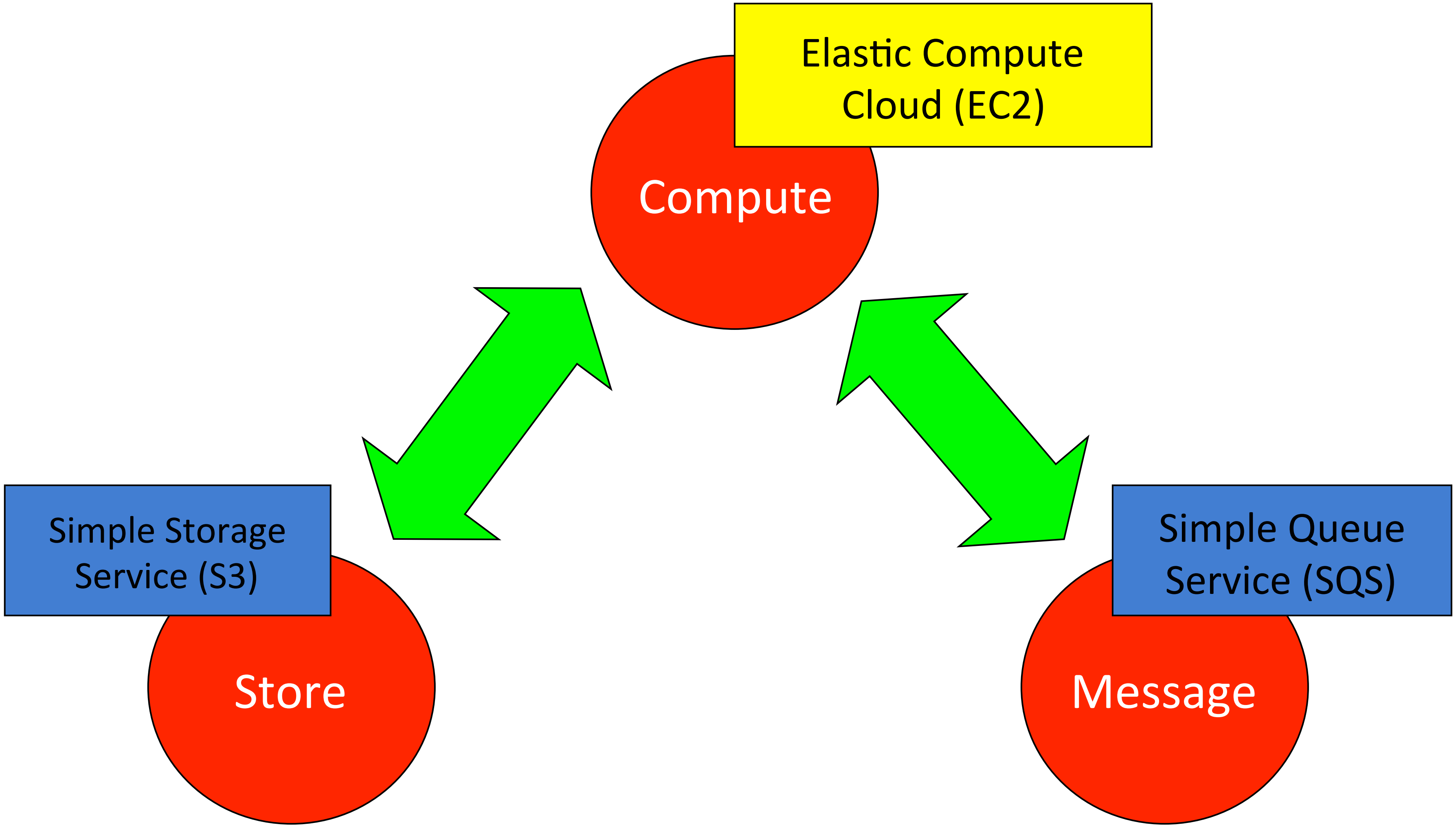
A set of APIs and business models which give developer-level access to Amazon's infrastructure and content:

- **Data As A Service**
  - Amazon E-Commerce Service
  - Amazon Historical Pricing
- **Search As A Service**
  - Alexa Web Information Service
  - Alexa Top Sites
  - Alexa Site Thumbnail
  - Alexa Web Search Platform
- **Infrastructure As A Service**
  - Amazon Simple Queue Service
  - Amazon Simple Storage Service
  - Amazon Elastic Compute Cloud
- **People As A Service**
  - Amazon Mechanical Turk



Credits: Jeff Barr, Amazon

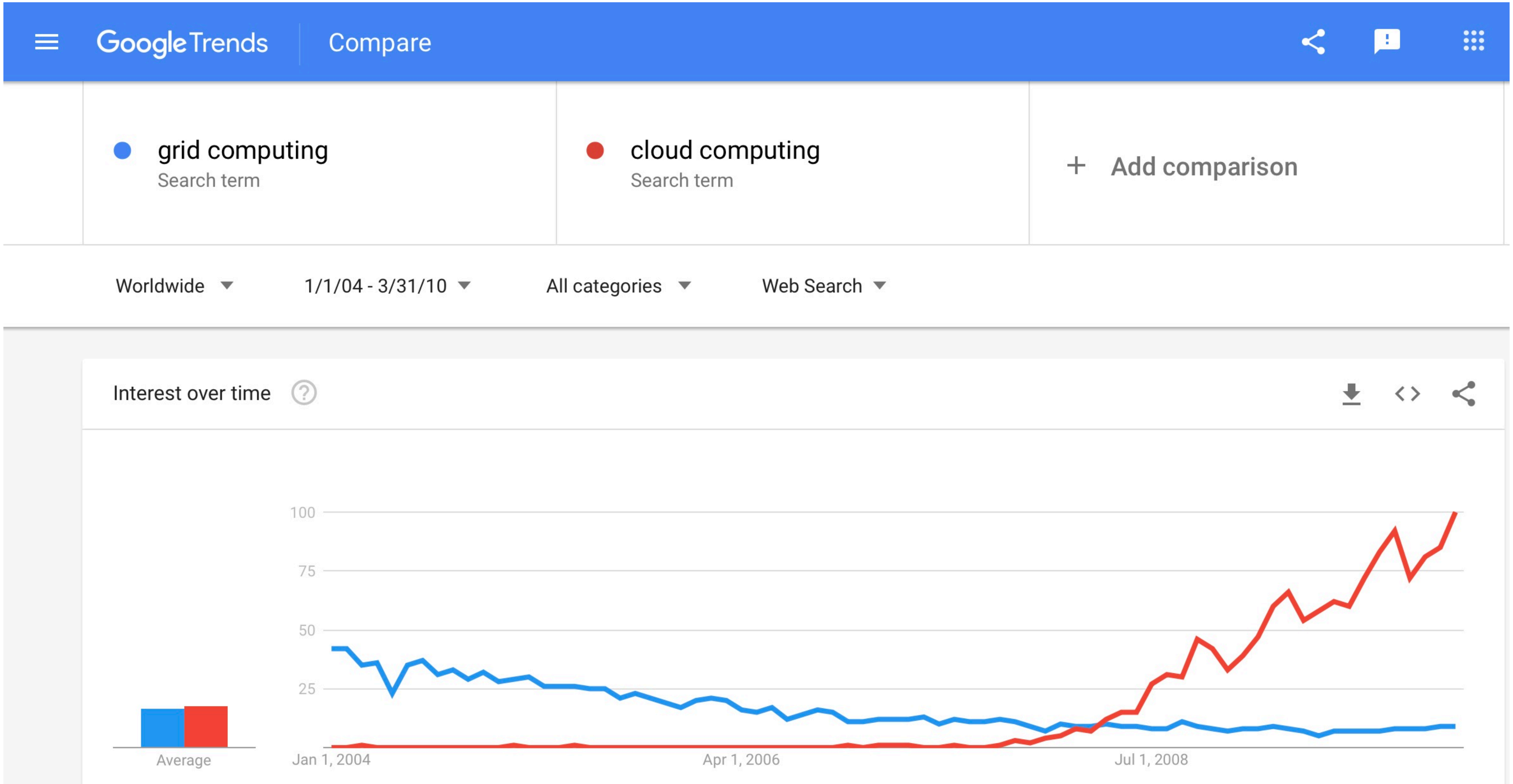
Amazon Web Services



Credits: Jeff Barr, Amazon

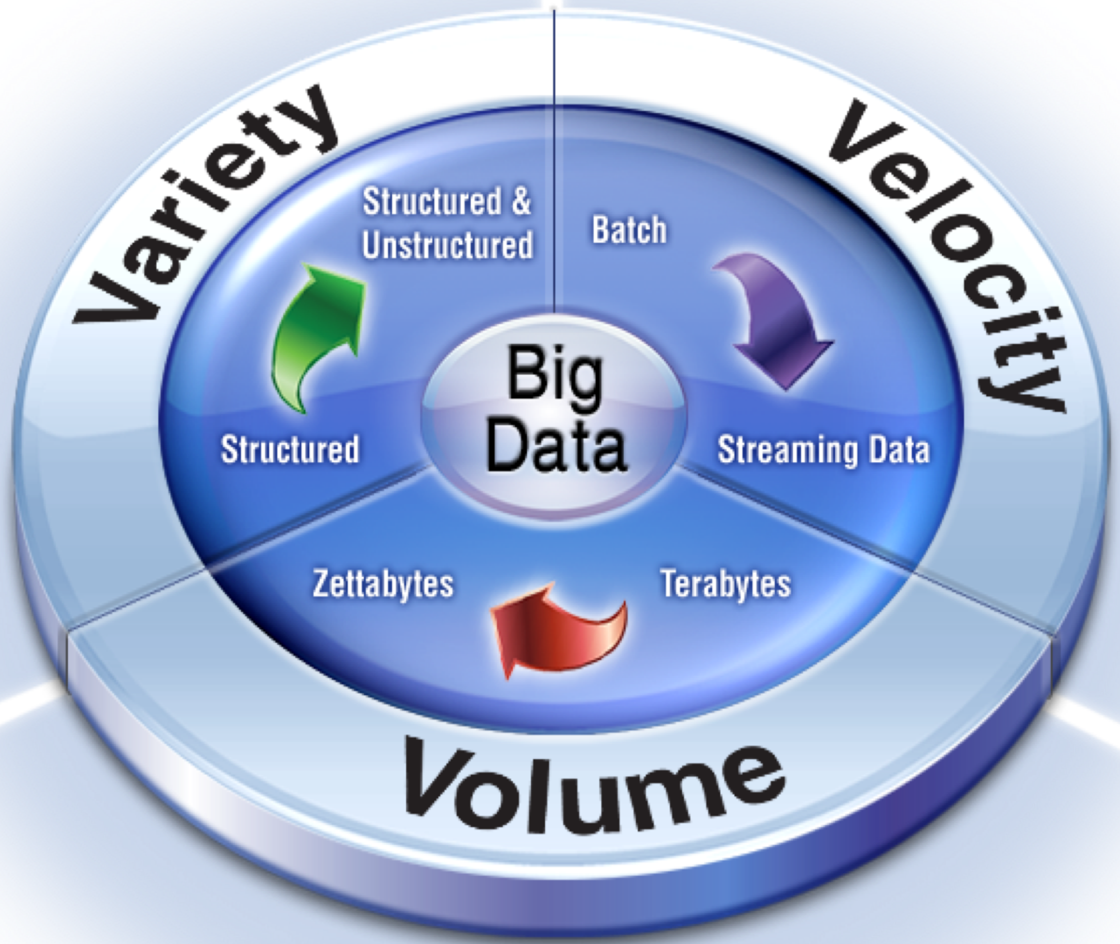
Cloud Datacenters

Beware the Cloud Hype!

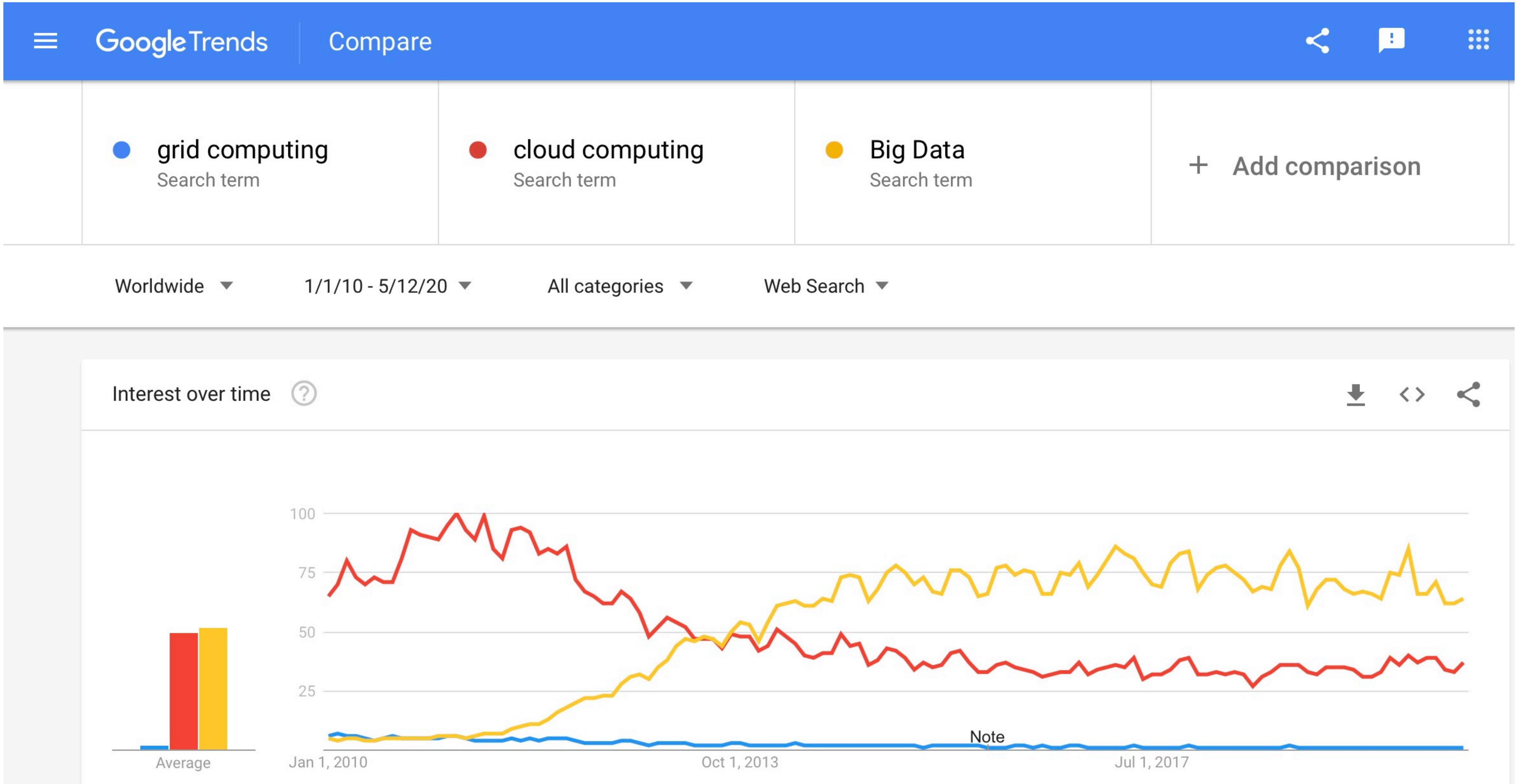


The Way to the Cloud is Long…

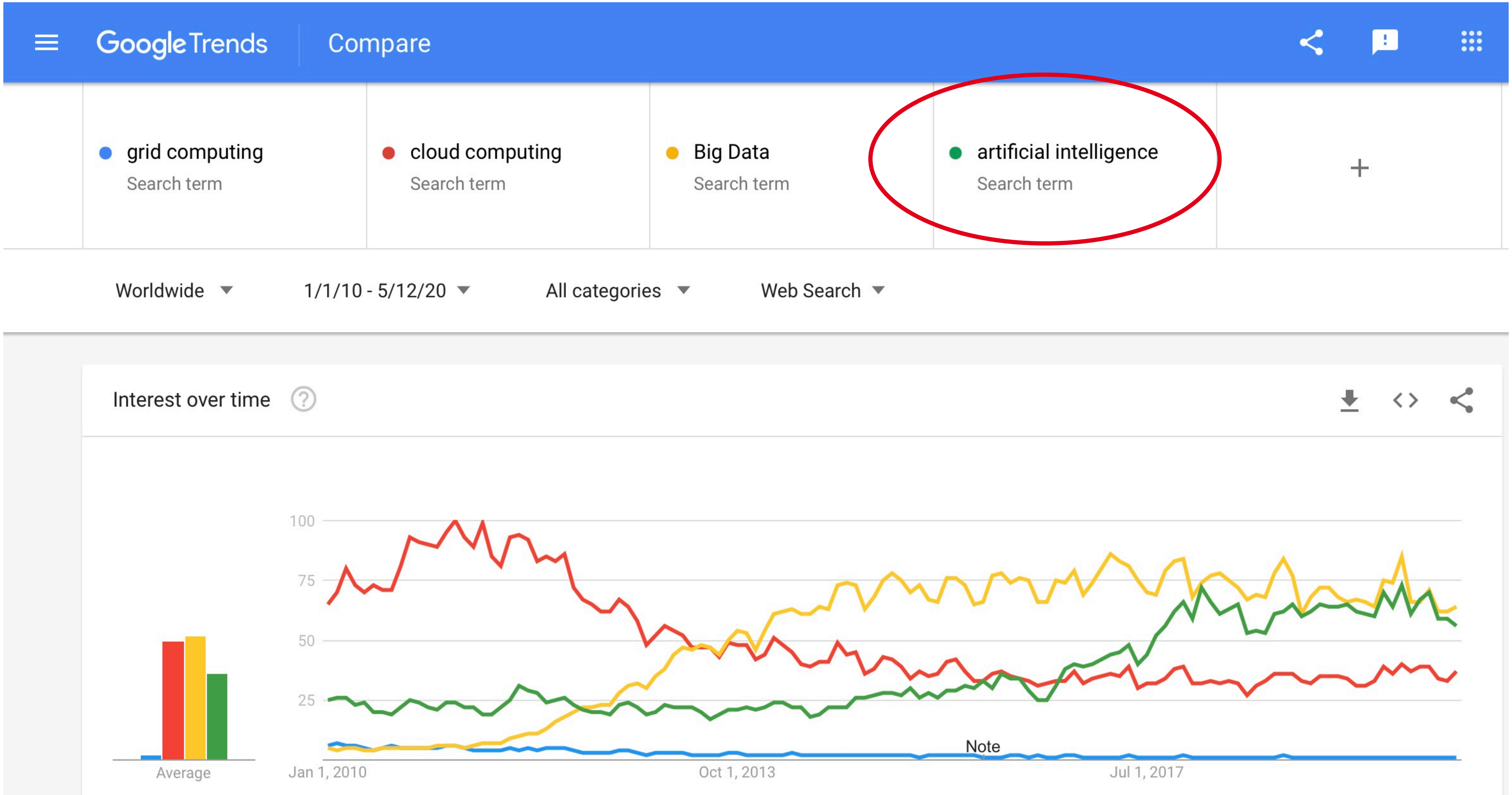
Big Data Challenges: The Three Vs



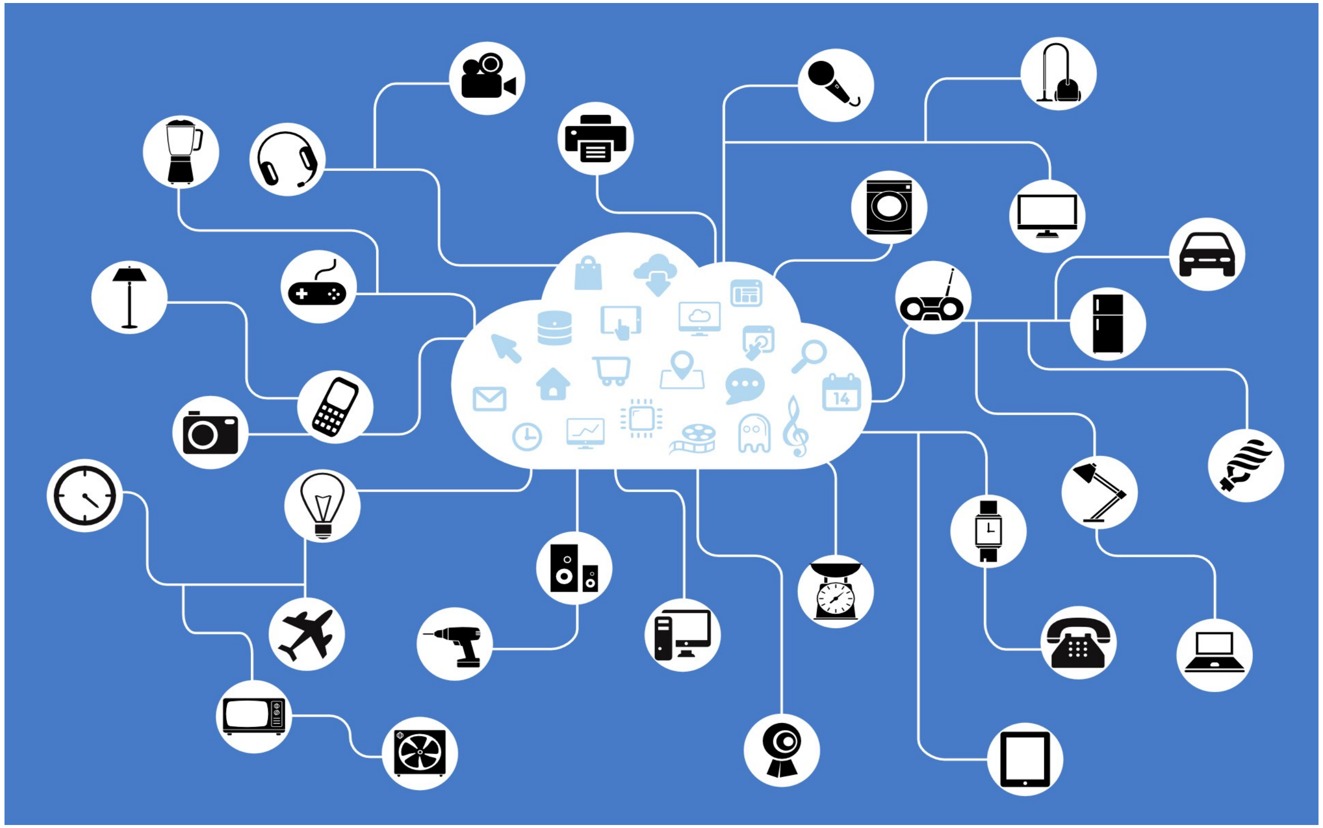
Beware the Big Data Hype!



What is the Next Hype?



Data Shifts to the Edge



By 2022 Gartner predicts that 75% of enterprise-generated data will be created and processed outside of the data center and cloud infrastructures compared with 10% today.

Source: Smarter with Gartner, What Edge Computing Means for Infrastructure and Operations, October 3, 2018
Extract from: BullSequana Edge positioning paper (Atos)

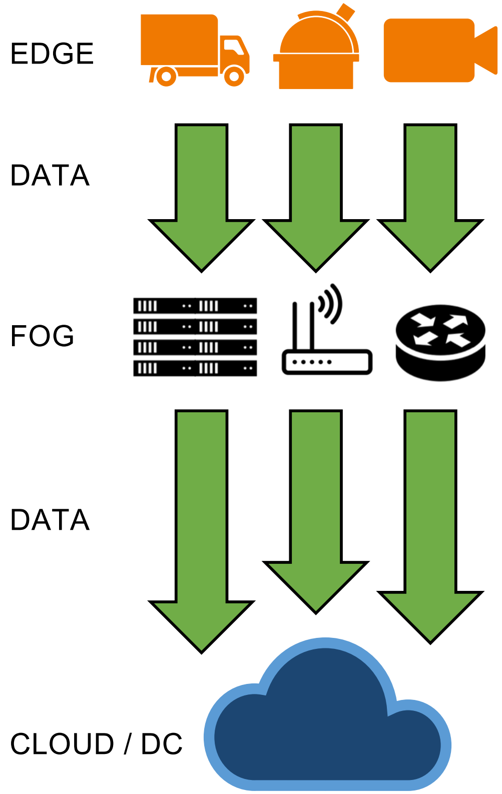
Sensors Everywhere:
Towards Autonomous Driving

Sensors Everywhere: Precision Agriculture

Welcome to Edge Computing!

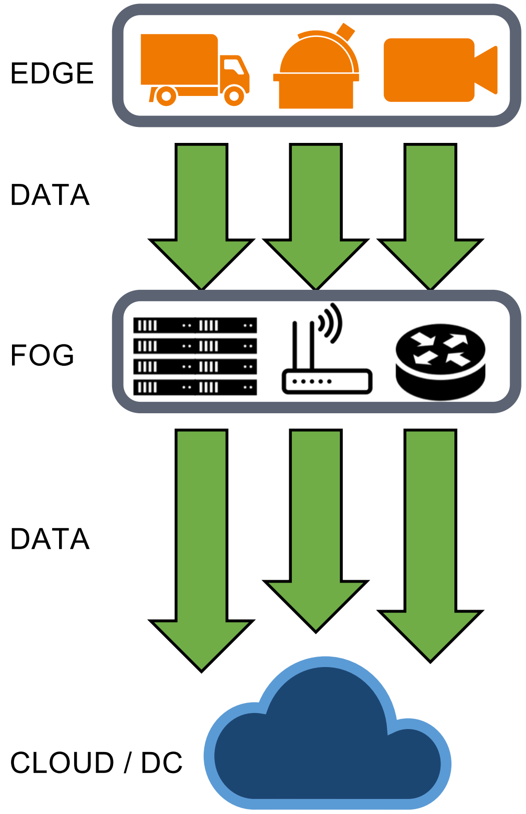
Advantages

- Easier access to data
- Bandwidth saving
- Privacy
- High potential parallelism



A Plethora of Edge Processing Tools

- Custom software
- Generic frameworks
  - Apache Edgent
  - Amazon Greengrass
  - Azure Stream Analytics
  - IBM Watson IoT
  - Intel IoT
  - Oracle Edge Analytics



Edge Processing Tools Are Great! ☺

Respond to local events in near real-time

AWS IoT Greengrass devices can act locally on the data they generate so they can respond quickly to local events,

Simplified device programming with AWS Lambda

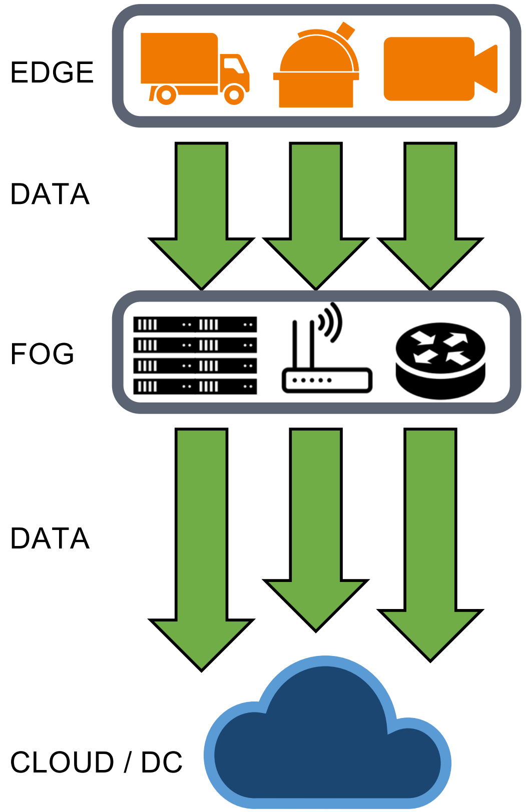
You can develop code in the cloud and then deploy it seamlessly to your devices with AWS Lambda. AWS IoT

Capture data in real time

Process IoT data instantly to help identify valuable insights related to device behavior and operations in the field. Spot

Instantly analyze data from all your IoT devices and gateways

As more and more data is generated from a variety of connected devices and sensors, transforming this data into actionable insights and predictions in near real-time is now an operational necessity.



How Great?

Respond to local events in near real-time

AWS IoT Greengrass devices can act locally on the data they generate so they ca

Simplified device pro
AWS Lambda

You can develop code in the clo
seamlessly to your devices with

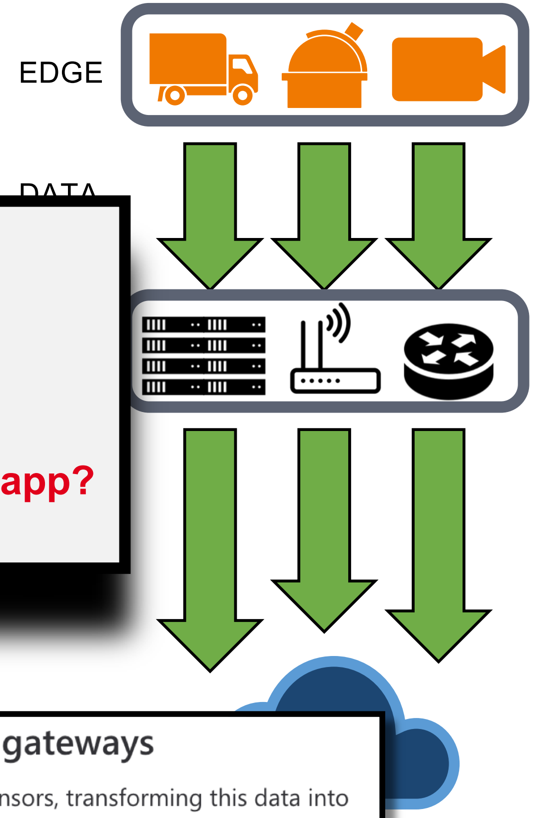
Capture

Process IoT data instantly
related to device behavior and operations in the field. Spot

What is their performance?

Under which conditions?

Do they integrate well with my app?

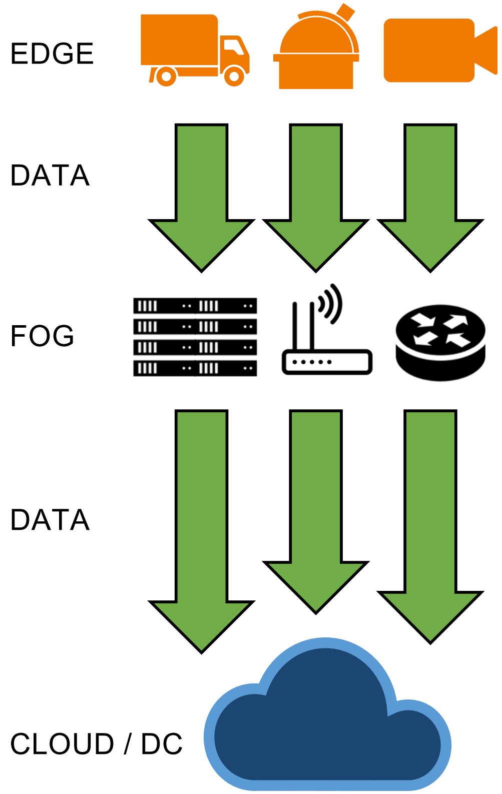


Instantly analyze data from all your IoT devices and gateways

As more and more data is generated from a variety of connected devices and sensors, transforming this data into actionable insights and predictions in near real-time is now an operational necessity.

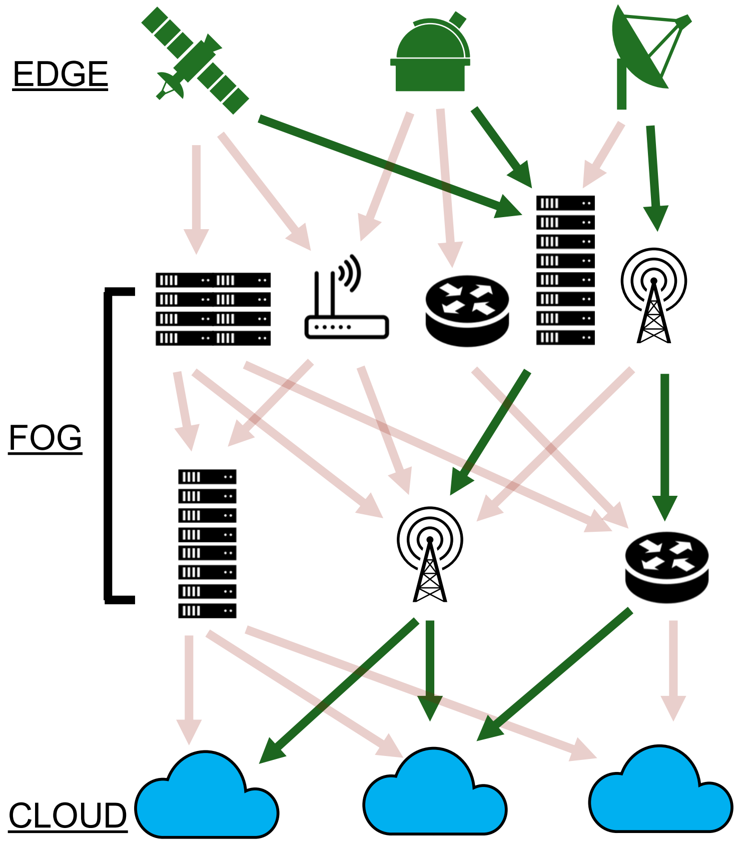
We Need Benchmarking!

Goal: **Understand** performance



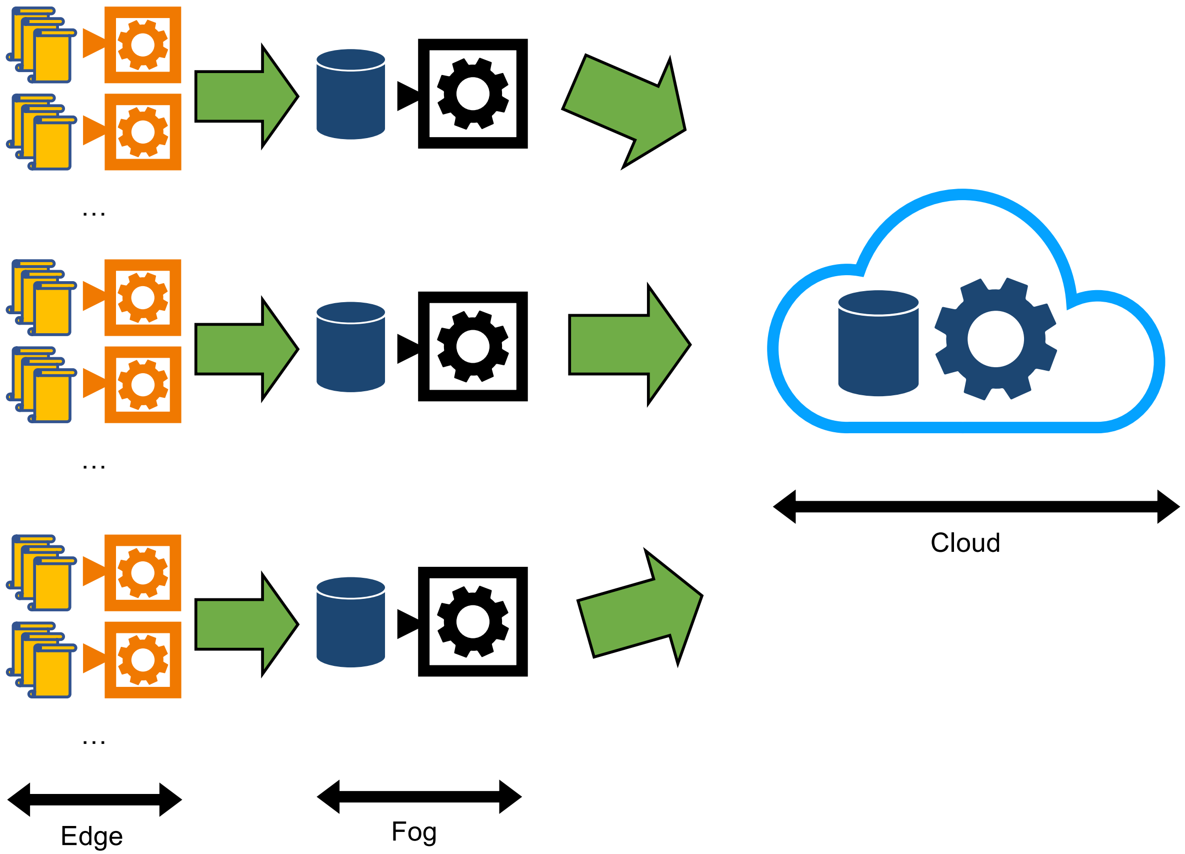
Benchmarking in a Hybrid World: Questions

- ❑ Are the **cost models** precise?
- ❑ What is the impact of **networking** on the performance?
- ❑ How do my algorithms react to **real-time** scenarios?
- ❑ How does my hybrid approach **compare** to a fully centralized solution?
- ❑ **Viability** on different infrastructure configurations
- ❑ What methodology for **reproducible** experiments?

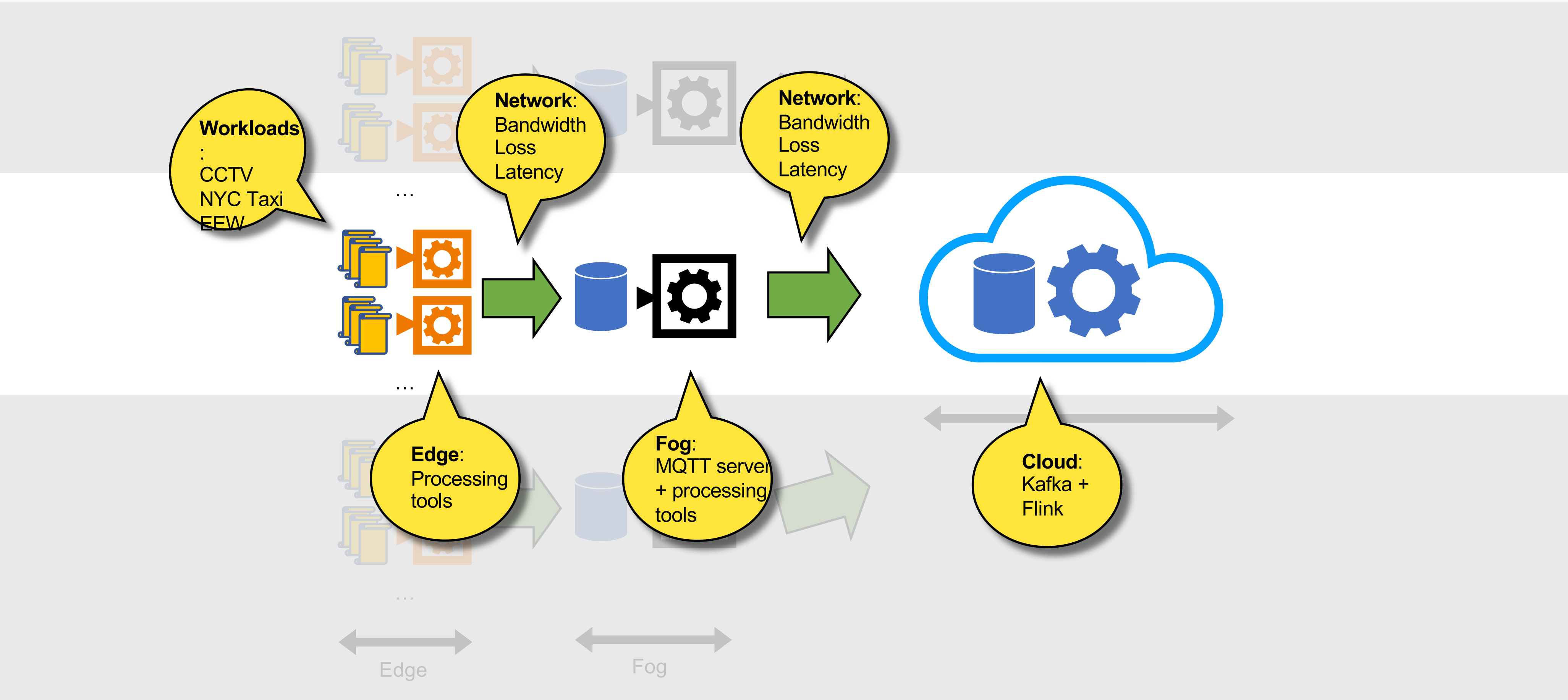


❑ SILVA, P., COSTAN A. and ANTONIU, G., Towards a Methodology for Benchmarking Edge Processing Frameworks. 1st Workshop on Parallel AI and Systems for the Edge (PAISE workshop collocated with IPDPS 2019).

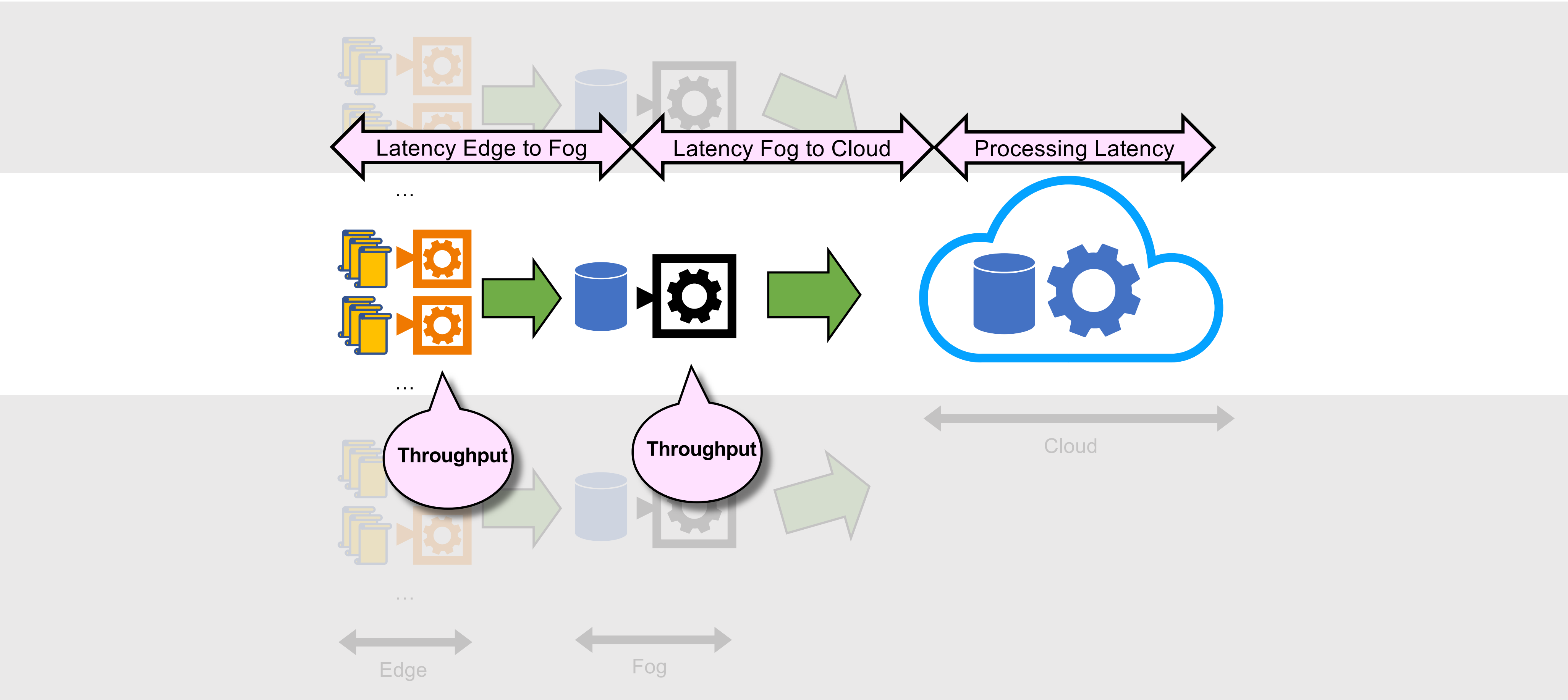
Benchmarking Hybrid Processing: Zoom



Benchmarking Hybrid Processing: Parameters

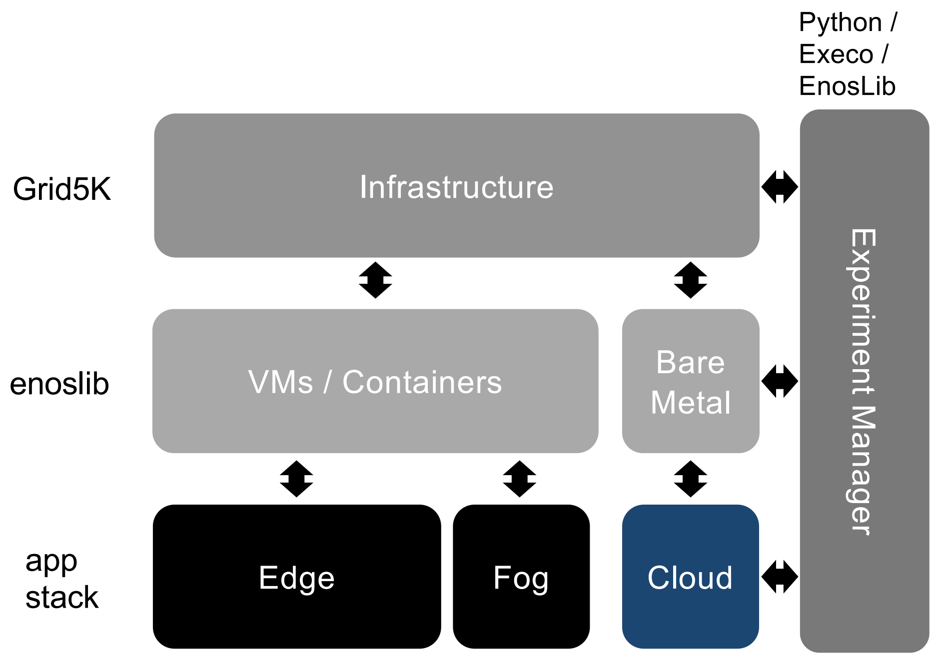


Benchmarking Hybrid Processing: Metrics



Benchmarking Platform: Implementation

- **Experiment manager**
  - Configures the infrastructure
  - Deploys frameworks/tools
  - Deploys applications and manages their executions
  - Monitors resource usage
  - Gathers metrics and logs



- **Edge+Fog+Cloud processing management**
  - Metric generation, configuration, connection

An Example of Application: Earthquake Early Warning Systems

Earthquakes

Large ground motions which cause substantial loss and damage to the environment across areas

spanning hundreds of kilometers from their origins (tsunamis, fires, landslides)

Earthquake Early Warning (EEW)

Objectives

> Automatically detect and characterize earthquakes as they happen

> Deliver alerts before the ground motion actually reaches sensitive areas

FAUVEL, K. ; BALOUEK-THOMERT, D. ; MELGAR, D. ; SILVA, P., SIMONET, A. ; ANTONIU G. ; COSTAN, A ; MASSON, V ; PARASHAR, M. ; RODERO, I. ; TERMIER, A. A Distributed Multi-Sensor Machine Learning Approach to Earthquake Early Warning. AAAI 2020, February, New York. **Outstanding Paper Award for Social Impact.**

EEW - Current Challenges

Earthquake Category vs Sensors

Two categories of earthquakes with damaging potential

Large Earthquakes
(6 ≤ magnitude, Richter scale)

- ✖ Seismometers: saturation issue
- ✔ GPS Stations

Medium Earthquakes
(5 ≤ magnitude < 6, Richter scale)

- ✔ Seismometers
- ✖ GPS Stations: noise

Response Time and Robustness

✖ Traditional centralized approach impractical

> **Large number of sensors** with **high frequency** data streams

> Sensors are **geographically distributed**

EEW - Our Positioning

Data

More and more data **available** from sensors (e.g. seismometers, GPS stations)

Solutions

EEW solutions exist (Yoon et al. 2015; Li et al. 2018; Perol, Gharbi, and Denolle 2018)

> No approach combining seismometers and GPS stations

> No distributed solution

POSITIONING

Integrating two **complementary** types of **sensors** in real-time based on a **distributed hybrid infrastructure** to cover the **whole spectrum** of potentially damaging earthquakes

Supervised Learning Problem

Seismometer

3 Time Series
East-West, North-South
and Up-Down

GPS Stations

3 Time Series
East-West, North-South
and Up-Down

Multivariate Time Series (MTS) Classifier

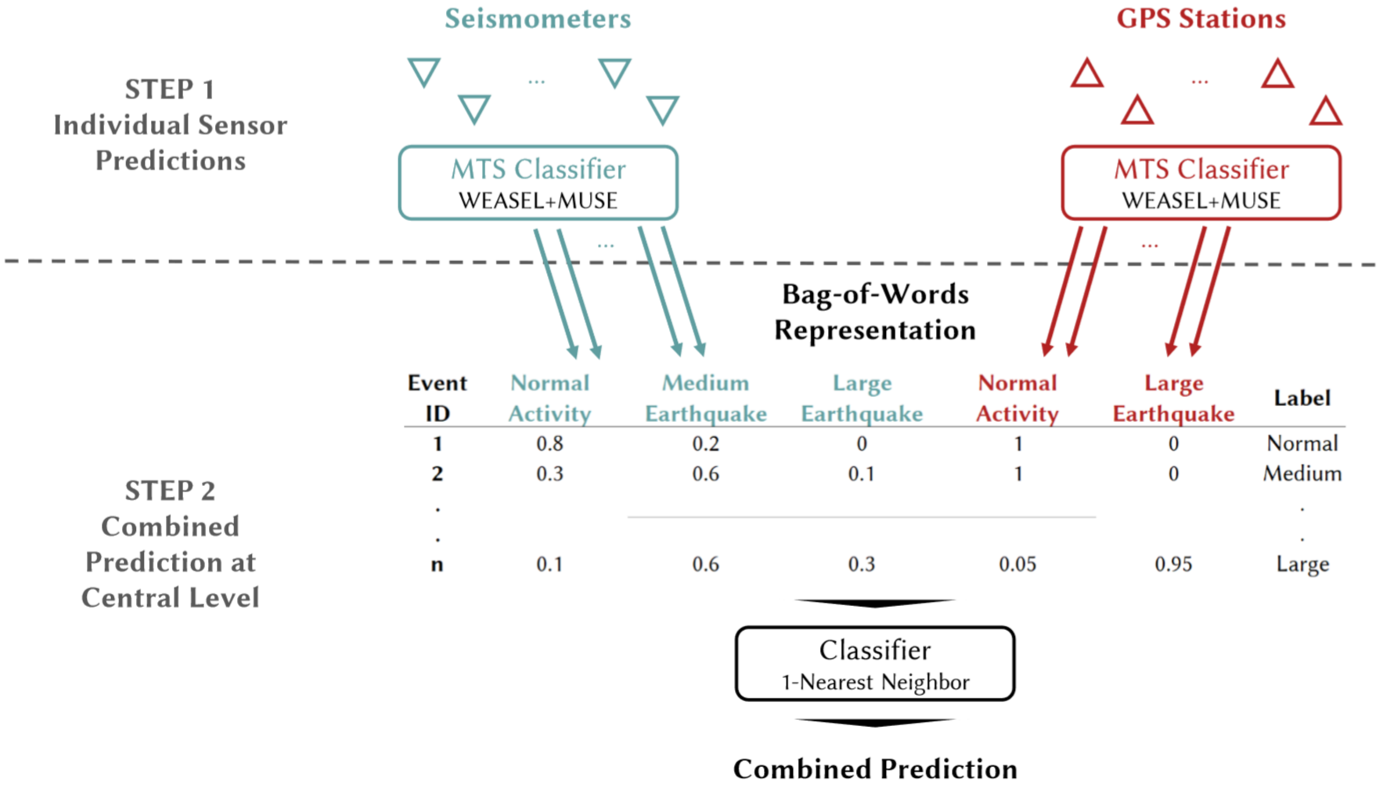
Earthquake Category?

Ground Truth:
P-waves follow a propagation model IASP91 (Kennet 1991)

Our Machine Learning Approach

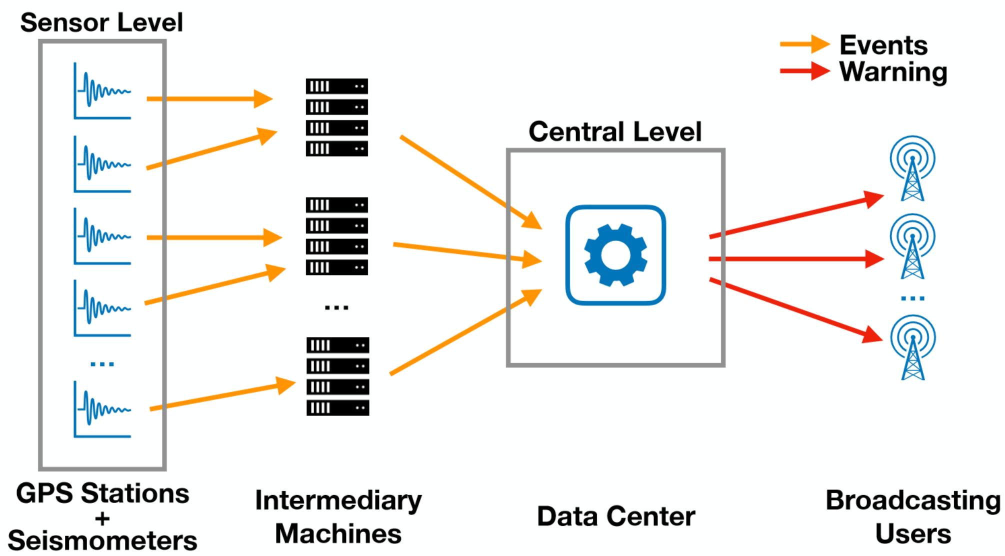
Distributed Multi-Sensor Earthquake Early Warning (DMSEEW)

A New Stacking Ensemble Method



Event ID	Normal Activity	Medium Earthquake	Large Earthquake	Normal Activity	Large Earthquake	Label
1	0.8	0.2	0	1	0	Normal
2	0.3	0.6	0.1	1	0	Medium
.						.
.						.
n	0.1	0.6	0.3	0.05	0.95	Large

Distributed, Hybrid Infrastructure

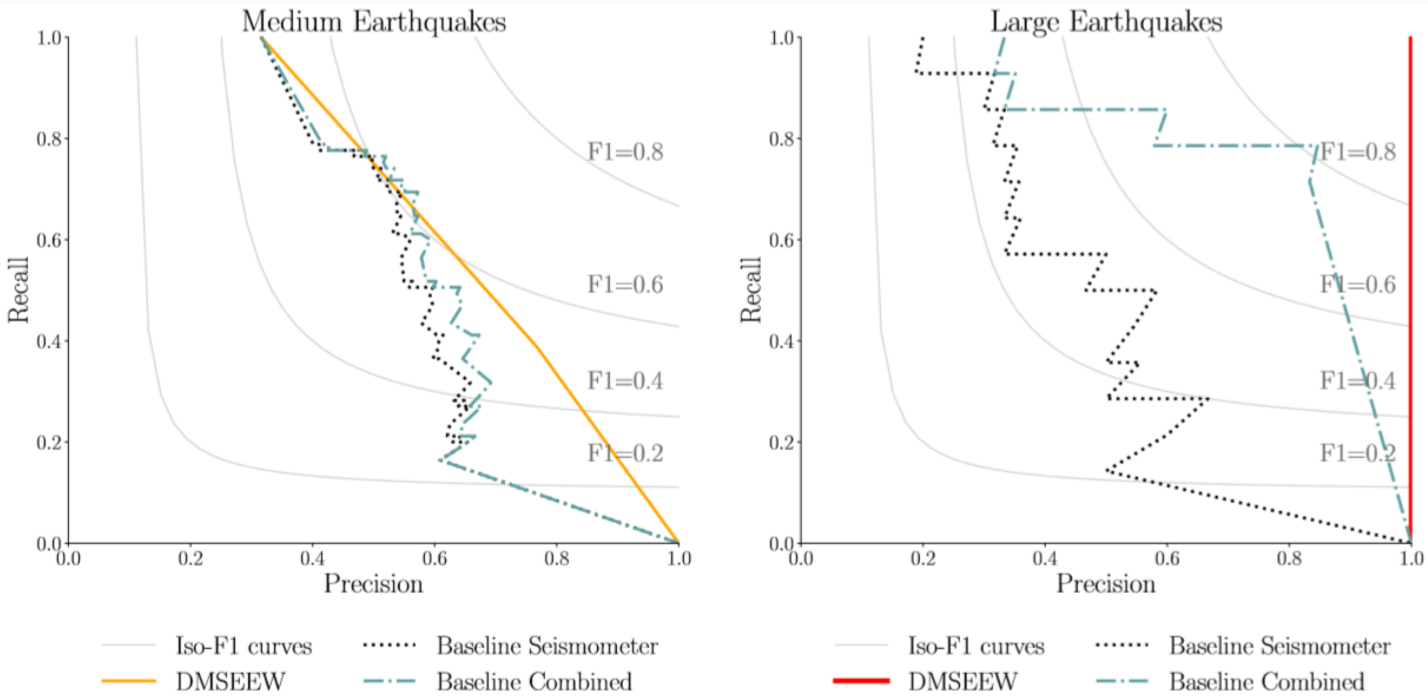


Sensor devices with computing capabilities

- Reduce the amount of data over the network
- Robustness in case of partial infrastructure failures

Edge + Cloud Computing

DMSEEW Evaluation



✓ DMSEEW **outperforms both** seismometer only and combined **baselines**

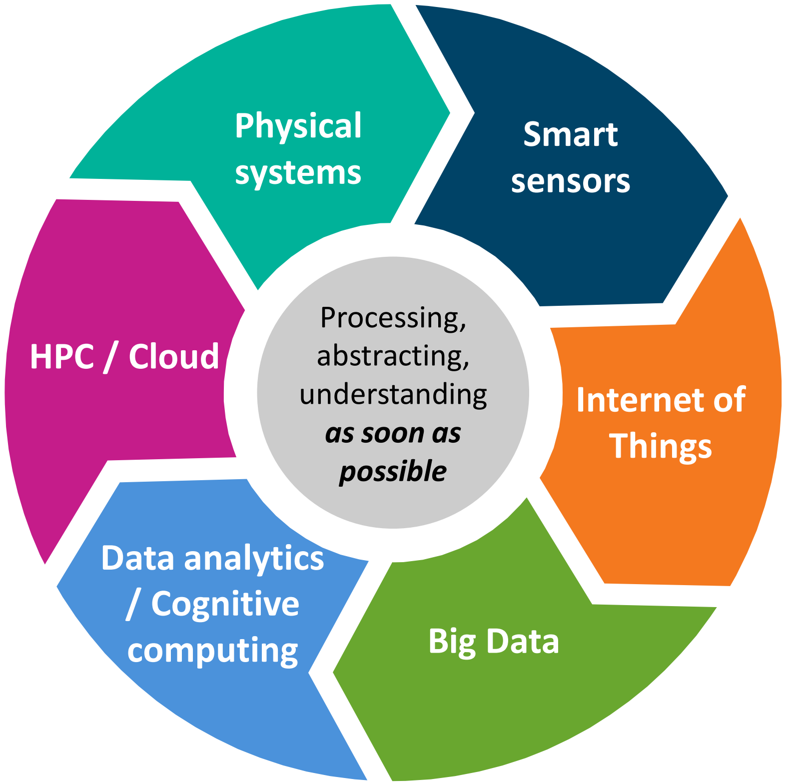
✓ DMSEEW detects **all large earthquakes without false alert**

> Geographically distributed infrastructure
- **Reduce the volume of data** transmitted in the network
- **Increase the robustness** of the EEW system

Supercomputer → Cloud → Edge

Is this a directional evolution?

Welcome to the Age of the Digital Continuum!



Enabling Intelligent data processing at the edge:

- Fog computing
- Edge computing
- Stream analytics

Transforming data into information as soon as possible

Collaboration between edge devices and the HPC/cloud ensuring:

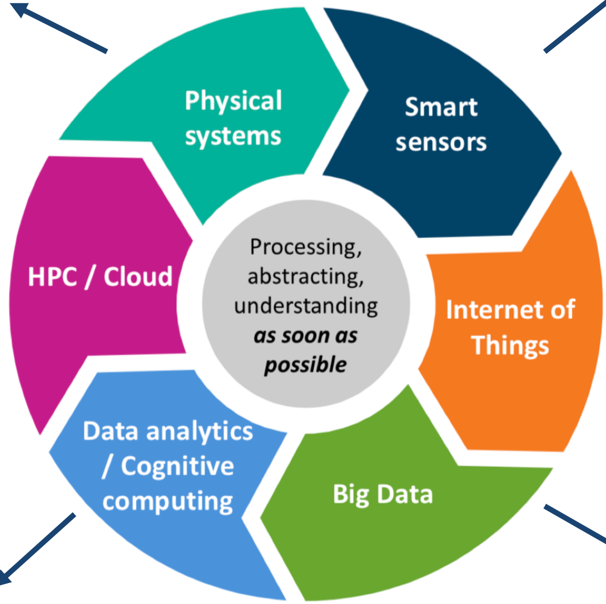
- Data security and Privacy
- Lower bandwidth
- Better use of HPC/Cloud

→ creating a continuous flow

Mapping “Extremes’ prediction in the Digital Continuum” to the loop

Cyber-physical entanglement of real-time, digital twin of Earth system, optimal information of user-specific information on demand and physical systems in impact areas: energy, water, health, law

Smart sensors in mobile phones, cars, wind farms, agriculture farms, cities etc. for weather and air quality



Extreme-scale computing to generate digital twin

IoT for accessing relevant data & IoT for disseminating user specific information

Machine learning for information extraction, product tailoring, surrogate modelling

Extreme-scale data collection-processing-dissemination on tight schedules

Novel
Emerging
Established

Thank you!

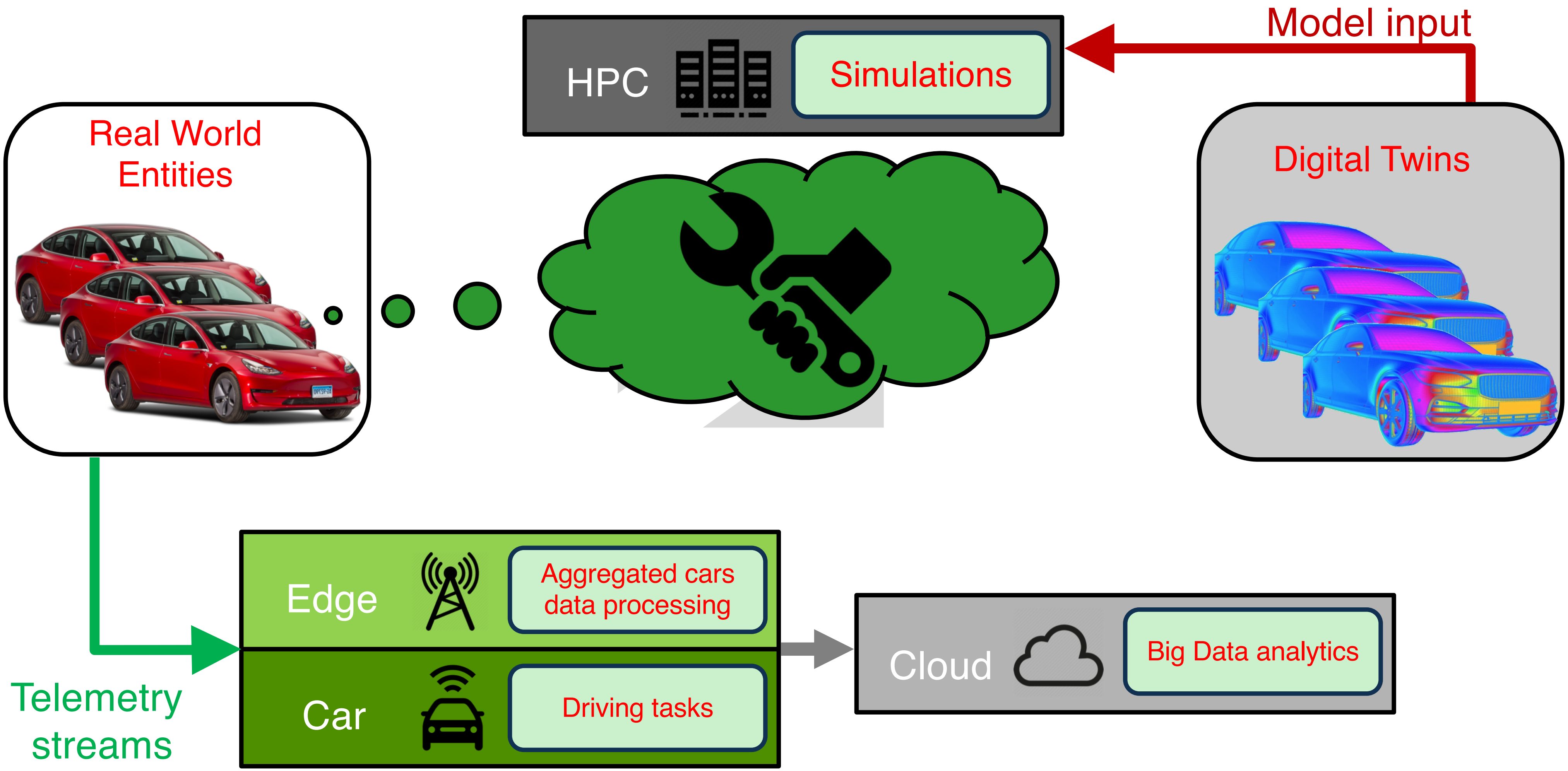
Contact: gabriel.antoniu@inria.fr

Inria

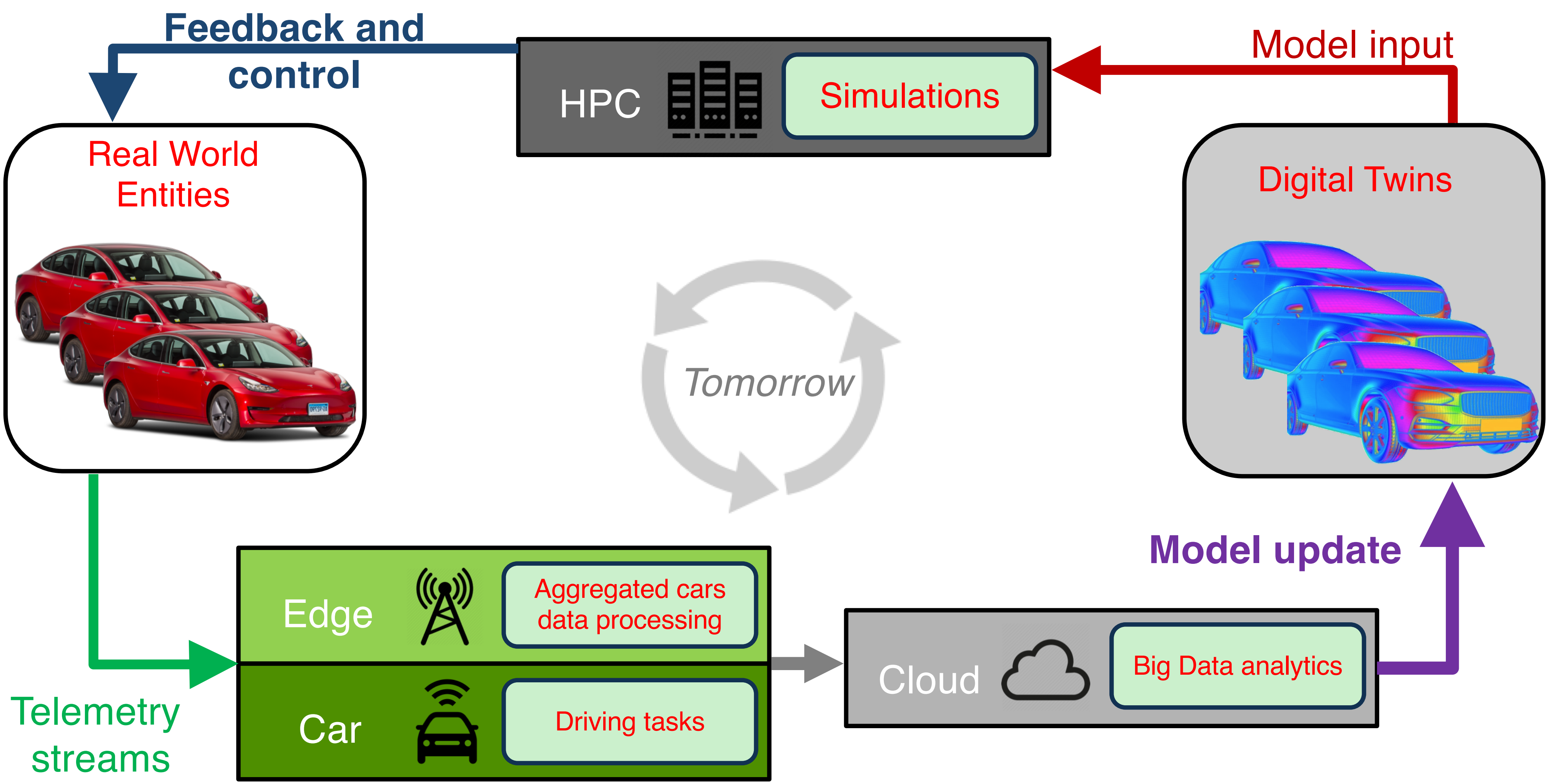
BACKUP SLIDES

INVENTEURS DU MONDE NUMÉRIQUE

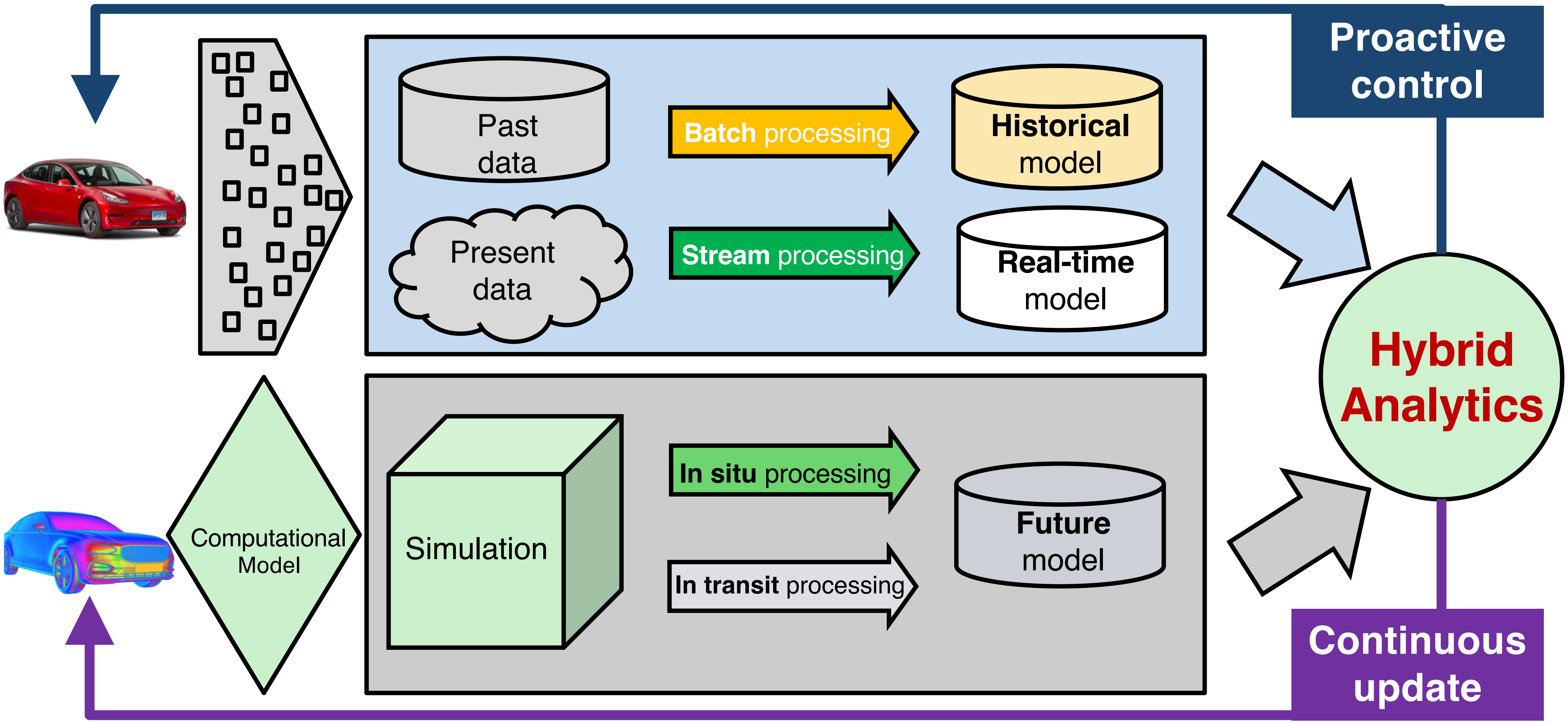
Scenario: Digital Twins TODAY



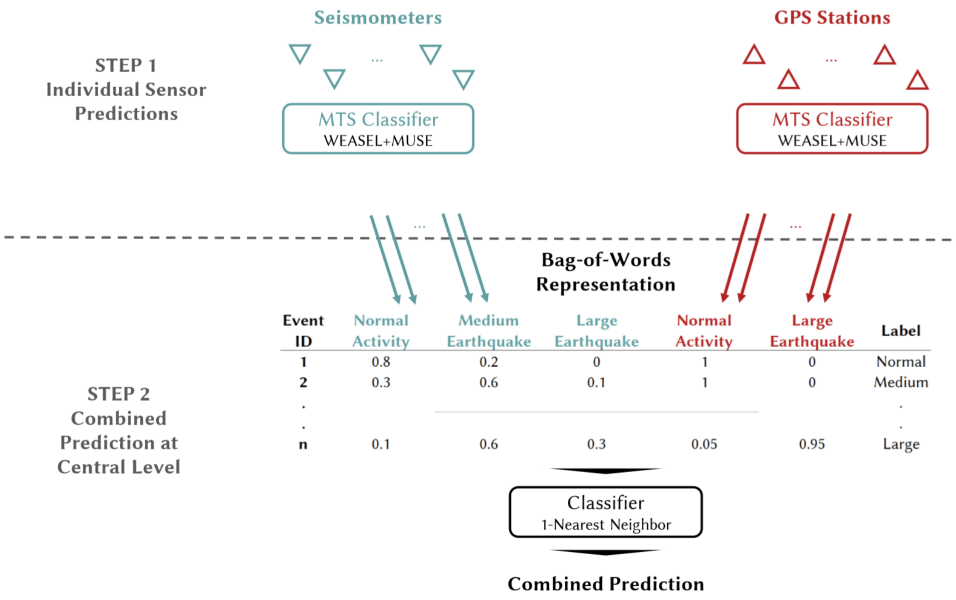
Scenario: Digital Twins TOMORROW



Our vision: Enabling Hybrid Analytics Through Unified Data Processing



Experiments



> **Metric:** Accuracy

> **Performance:** 3-Fold Cross-Validation

> **Step 1 - MTS Classifiers:** DTW_D , DTW_I , MLSTM-FCN, WEASEL+MUSE

> **Step 2 - Classifiers:** k-NN, Elastic Net, SVM, Random Forest, Extreme Gradient Boosting

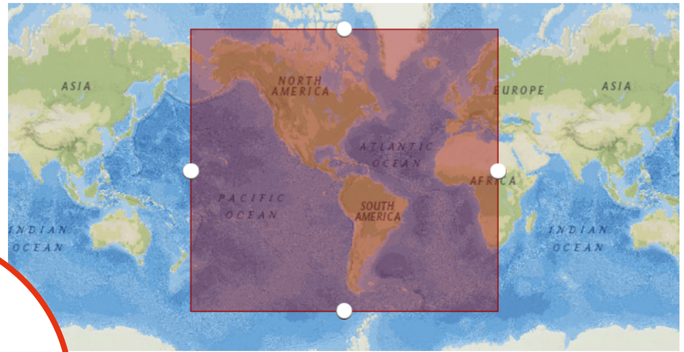
> **Benchmark:**

- Traditional Seismometer Only
- Combined Sensors with the Rule of Relative Strength

Experiments

269 Events (2001-2018)

(170 normal activity/85 medium/14 large)



Dataset

Time Series

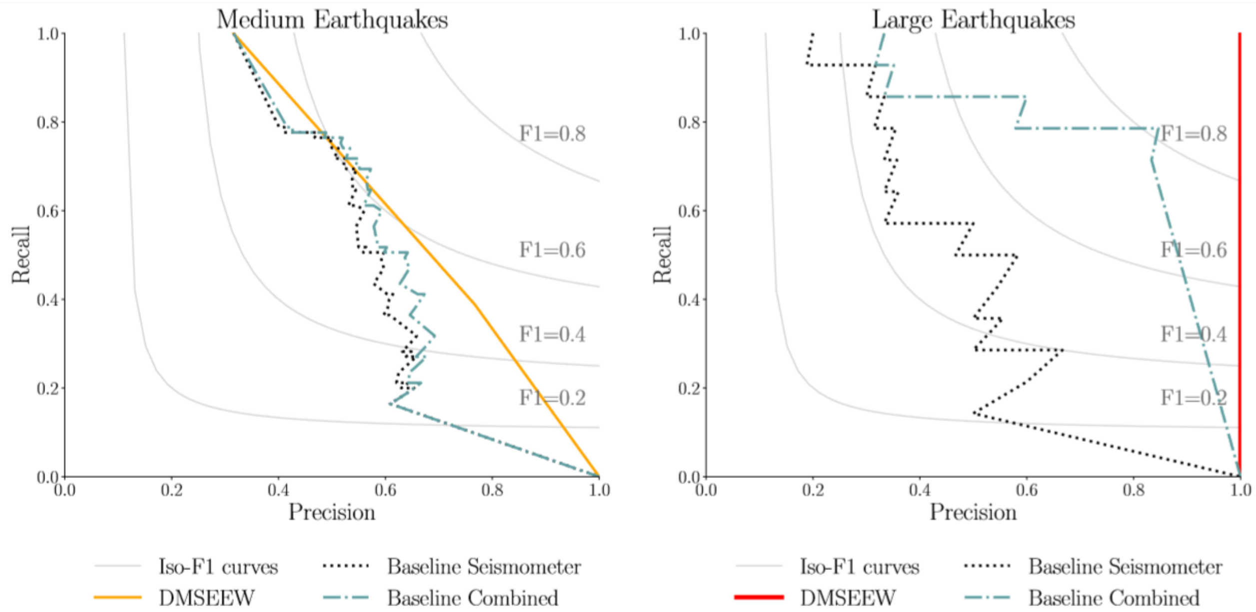
> MTS length: 60 seconds
> 1 timestamp / second

Preprocessing

> Digital signal to physical signal
> Standardization

[1] *https://figshare.com/articles/Earthquake_Early_Warning_Dataset/9758555*

Results



✔ DMSEEW **outperforms both** seismometer only and combined **baselines**

✔ DMSEEW detects **all large earthquakes without false alert**

Conclusion

> **Performance improvement** of DMSEEW on the detection of **earthquakes with damaging potential** / traditional seismometer only and rule based combined approaches

> Detection of **all large earthquakes without false alerts**

> Geographically distributed infrastructure
  - **Reduce the volume of data** transmitted in the network
  - **Increase the robustness** of the EEW system

> **Perspectives**:
  - Simulate different scenarios in an existing EEW execution platform to evaluate DMSEEW response time and robustness
  - Extend the evaluation of DMSEEW on the other main seismic network (Japanese NIED)

Hybrid Processing Architecture