Microsoft Research

Sensing Foundations for Environmental Science

Jie Liu

Senior Researcher

Microsoft Research

Redmond, WA 98052

liuj@microsoft.com

Microsoft-FAPESP Environmental Science Workshop, Nov. 11th, 2010

Networked Embedded Computing Group @ MSR

- Jie Liu
- Suman Nath
- Bodhi Priyantha
- Michel Goraczko
- Aman Kansal
- Dimitrios Lymberopoulos

Microsoft Research

Special Thanks to our LATAM Interns

- Andre Santanche
  Universidade Estadual de Campinas
- Börje Felipe Fernandes Karlsson
  Pontifícia Universidade Católica do Rio de Janeiro
- Leonardo Barbosa Oliveira
  Universidade Estadual de Campinas
- Heitor Soares Ramos Filho
  Universidade Federal de Minas Gerais
- Miguel Palomera
  Universidad Nacional Autónoma de México





Microsoft Research

Sensing Made Easy -- As It Appears

- Sensing, computing, and networking become increasing accessible
  - Low cost sensors, radios, and devices
  - Low barrier of entry for accessing cloud services
  - Ubiquitous cell and WiFi coverage
- Ride the consumer device wave
  - Sales of sensor-rich smartphones: 270 million in 2010. (IDC)
  - Potential of smart grid: 70 million single family houses in US.
- A round of academic investments around the world
  - US: Cyber-physical systems
  - EU: CONET
  - China: Internet of things

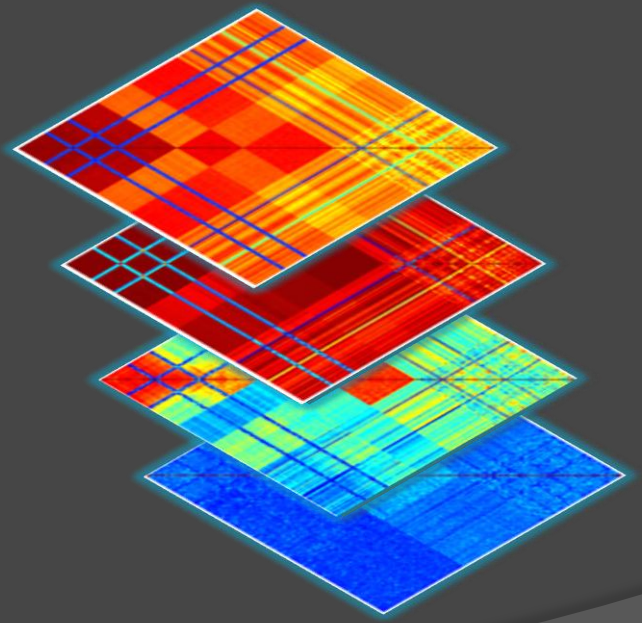
Microsoft Research

Sensing as a Foundation for Science

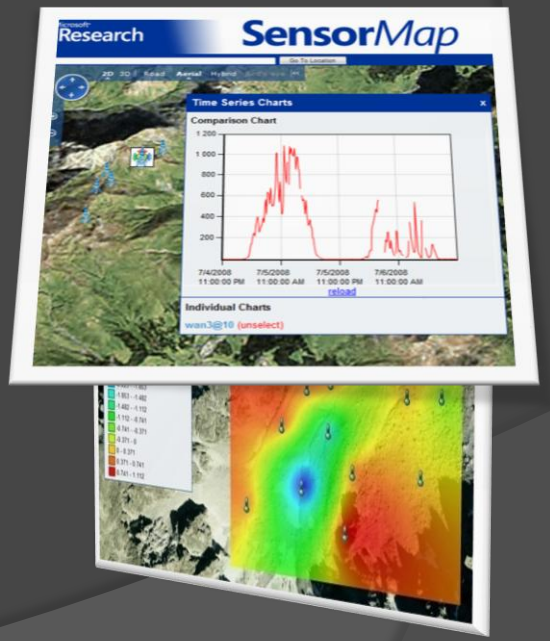
At Microsoft Research, we tackle platform and systems issues on sensing the physical world at scale.

Collection

Extraction



Collaboration



Microsoft Research

Data Centers

Data Center Genome

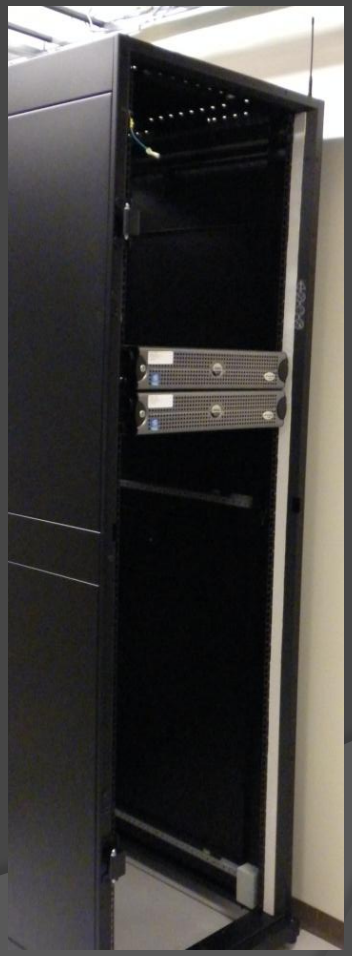
- A data center can have hundreds of thousand servers and other equipment and consumes tens of mega watts.
- The complex interactions among utilization, computing, and cooling make improve their efficiency challenging.

Microsoft Research

MeshID:

Low cost scalable asset location and tracking

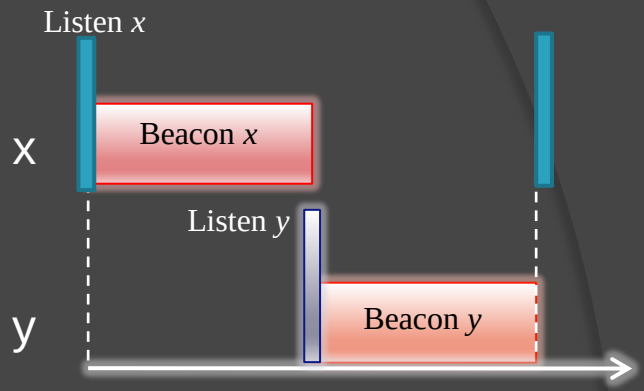
- Low cost RFIDs on each server
- In-rack RFID reader arrays
- Transit units to convert RFIDs to wireless sensors
- Mesh ad hoc network among transit units.

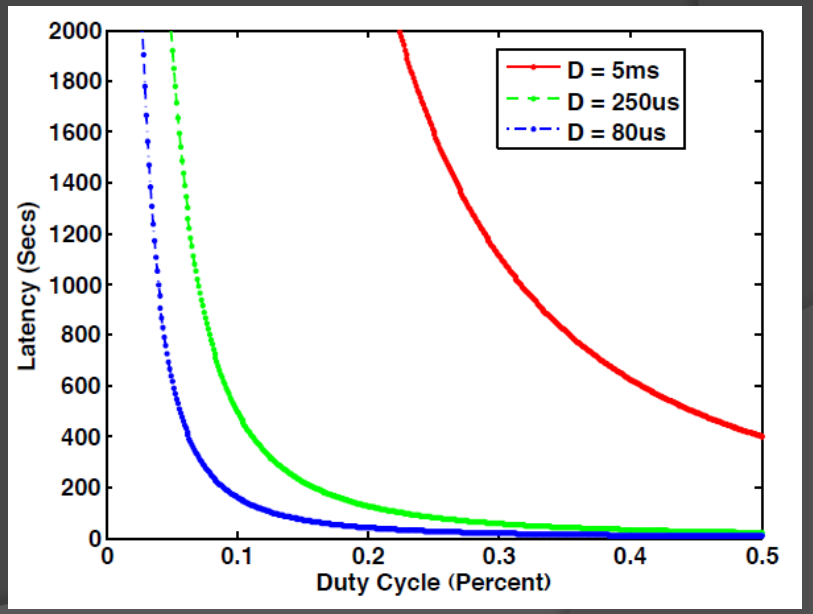


Microsoft Research

WiFlock: Ultra-Low Duty Cycle Discovery

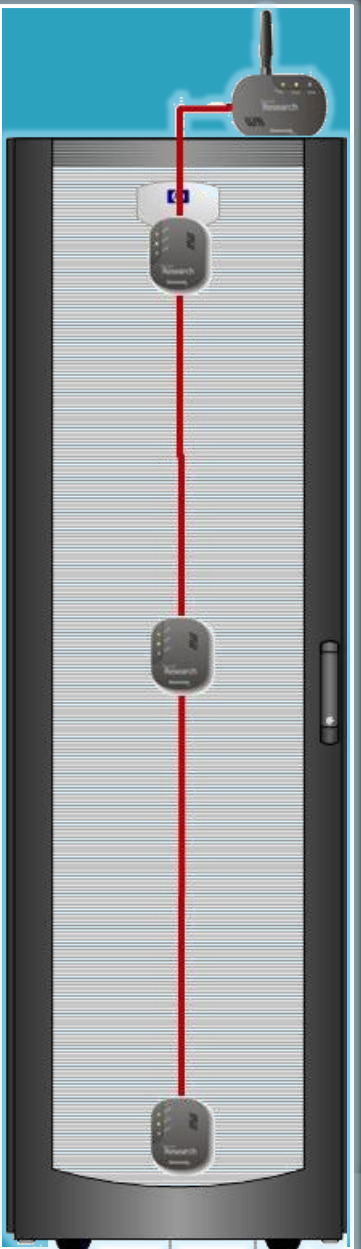
- Automatic discovery and group formation when nodes come to communication range.
- Lower duty cycle means more energy efficiency, but longer discovery latency.
- The faster a node can wake up the lower duty cycle it can achieve. We can do 80μs with CC2500, or equivalently 0.2% duty cycle.
- Expected life time: 5~7 years.



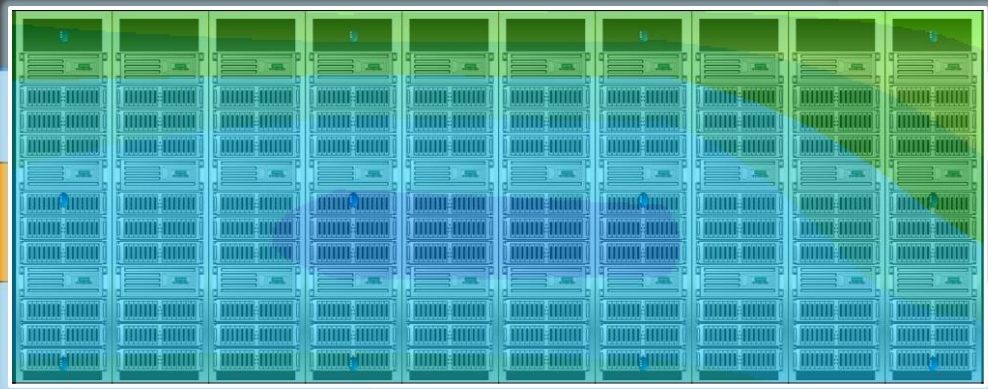


Microsoft Research

Genomotes

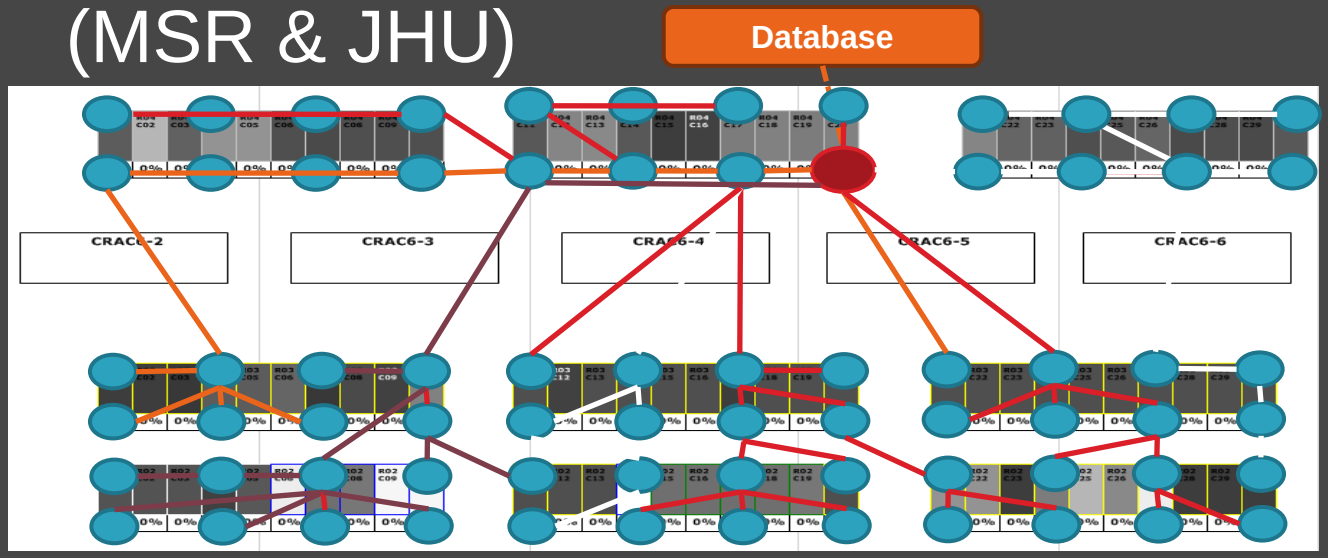


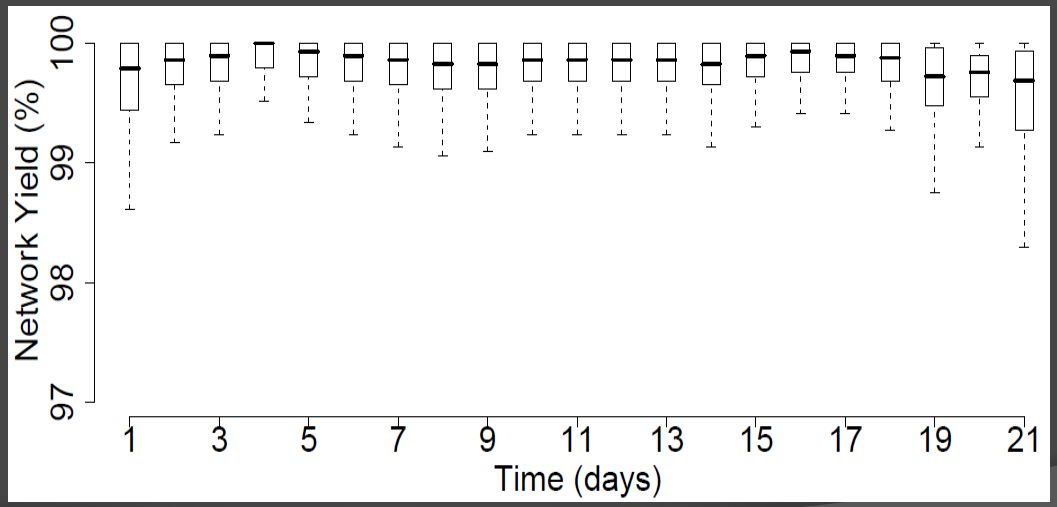
- Each server have a dozen fans.
- Together with AC fans, the temperature distribution is complex.
- Wireless sensors are ideal for non-intrusive continuous monitoring.



RACNet: Reliable ACquisition Network

(MSR & JHU)





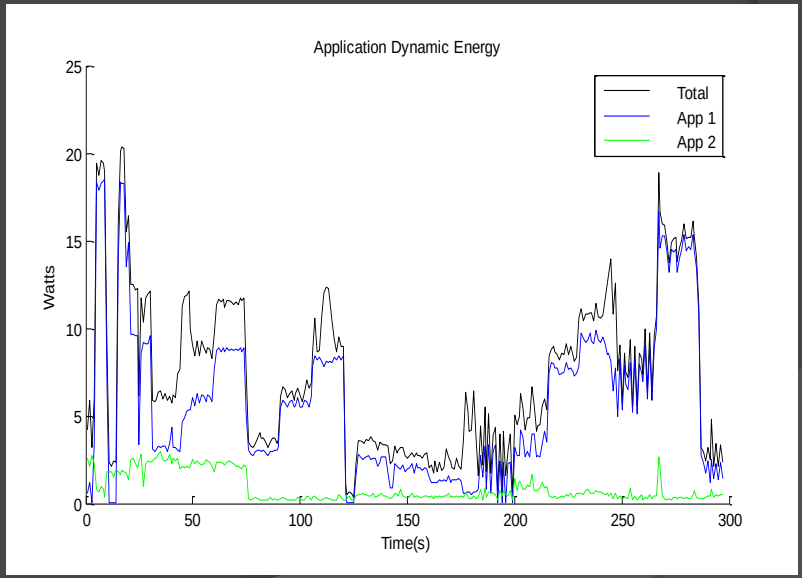
- Technology highlights
  - Local data caches
  - Multi-channel, multi-hop collection tree
  - Token passing data retrieval over collection trees
  - End-to-end reliability.

Microsoft Research

Virtual and Soft Sensors

- Not all variables of interest are measurable (e.g. per virtual machine power consumption.)
- Virtual and soft sensors compute the values indirectly.

Example:

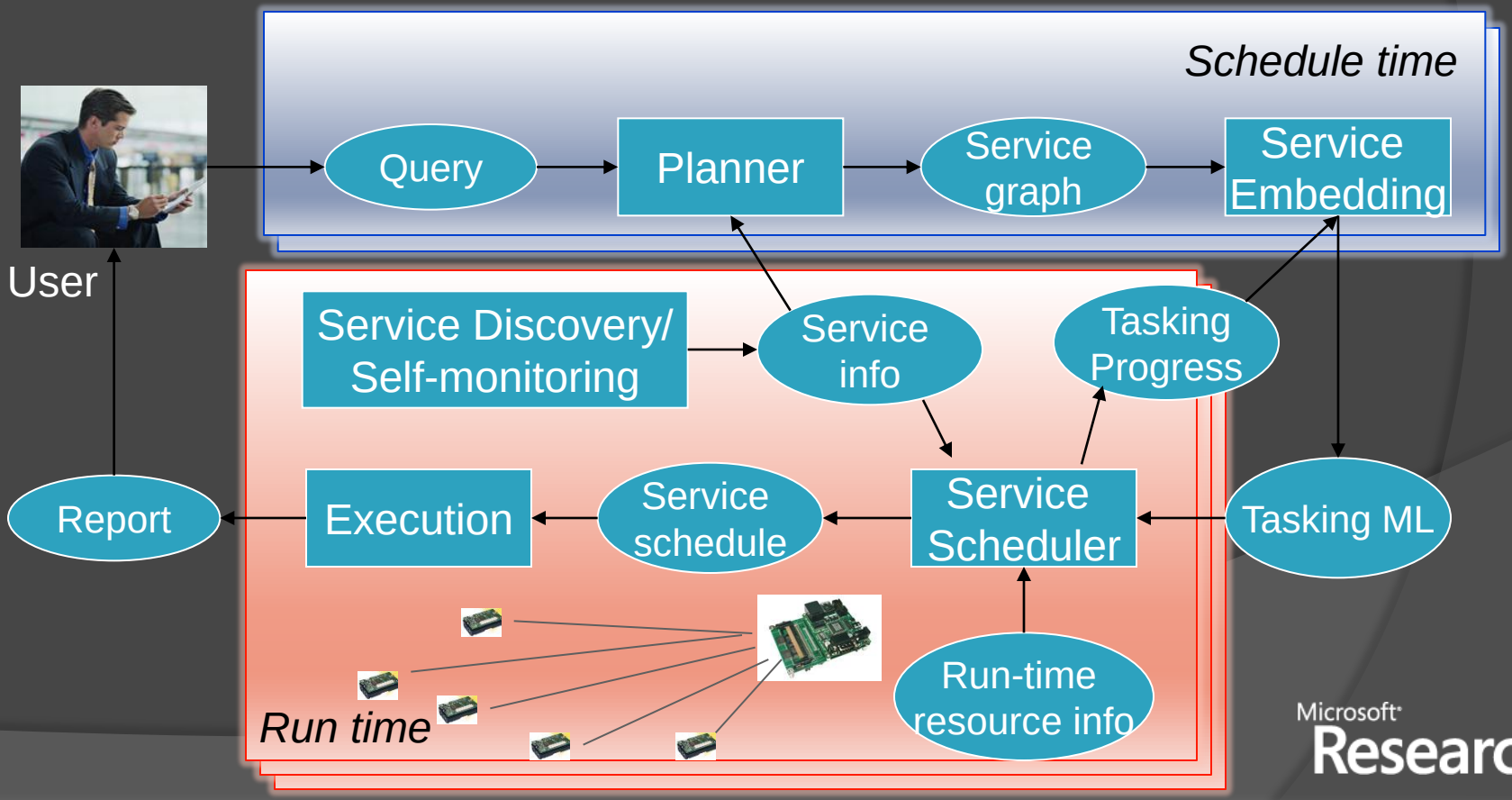


Microsoft Research

Software and Virtual Sensors

Sensor net as service providers

- Services encapsulate data and computation
- Input/output are label with clear semantics
- Services can be discovered, composited, and executed on demand



Microsoft Research

Scalable Data Management

- Unprecedented data collection capability creates massive amount of data
- Microsoft data centers collects ~1TB/day

- Data management challenges
  - Information extraction
  - Archiving
  - Analysis

Microsoft Research

Cypress Approach

— Semantic-based compression to reduce storage and I/O cost

— Achieve 100X compression with bounded maximum error (L_∞)

— Answer useful queries directly from compressed data

Column-oriented compression:

Precision reduction

[selection]

Trickles:

Single-stream spectrum analysis

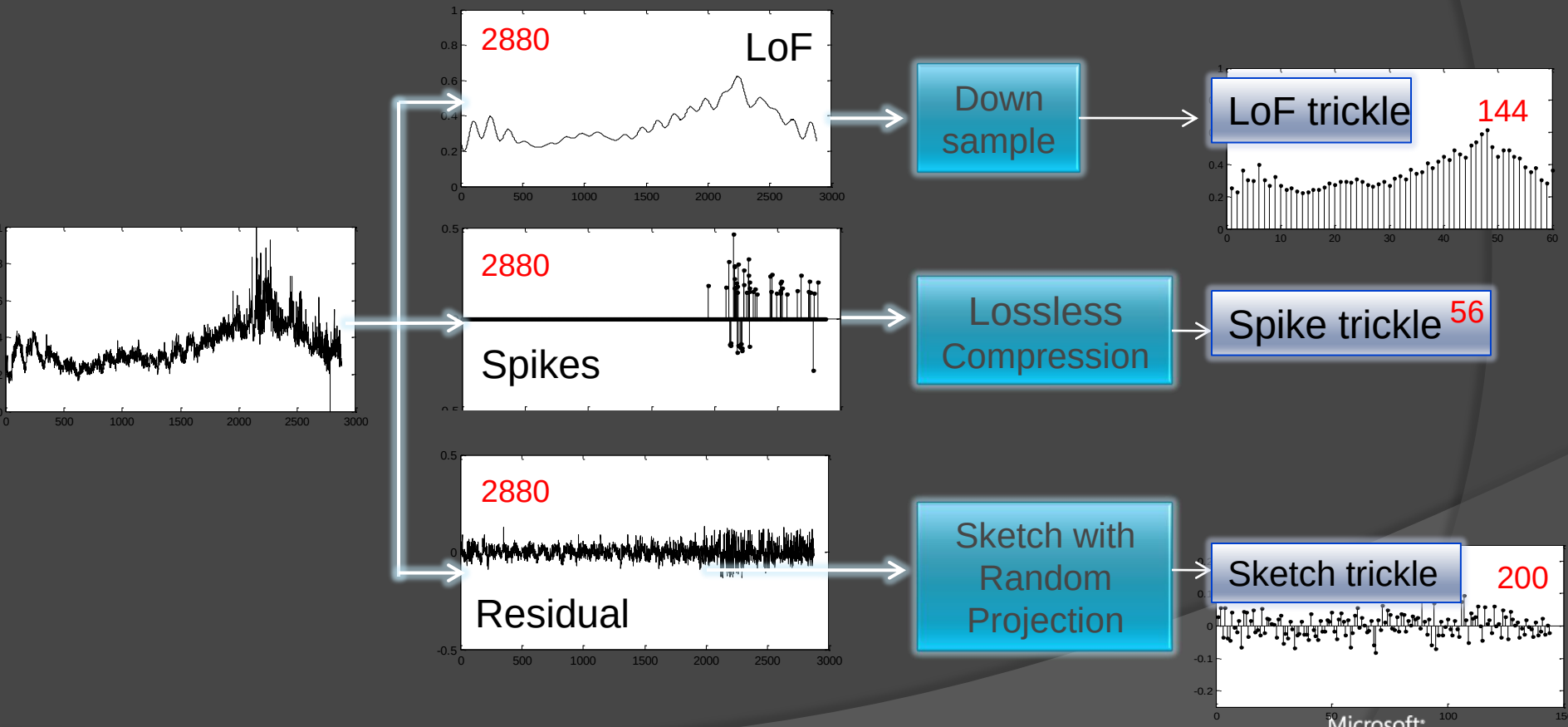
[correlation, histogram]

GAMPS:

Cross-stream compression

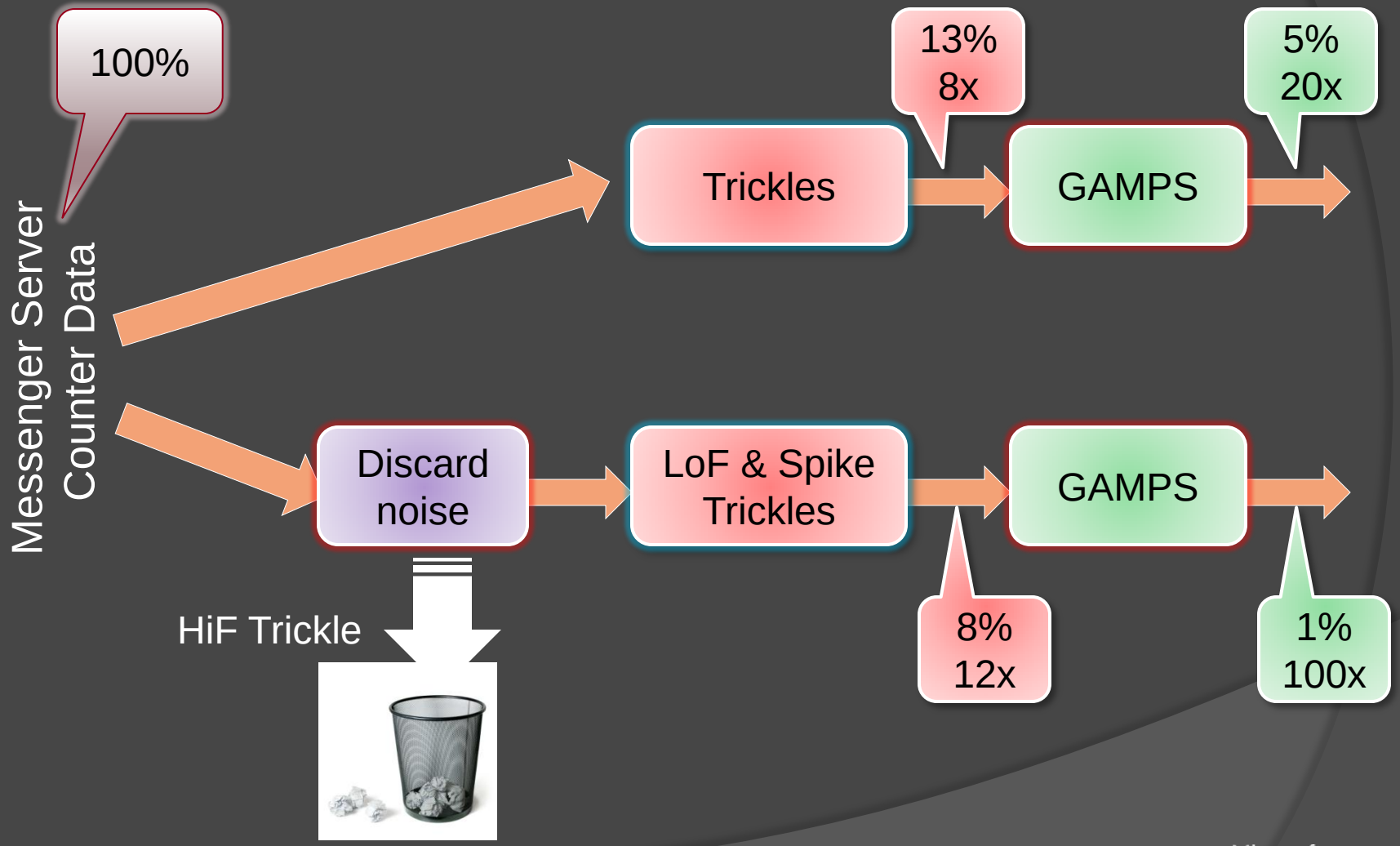
[selection, similarity]

Compress components into “trickles”



Microsoft Research

Trickles + GAMPS

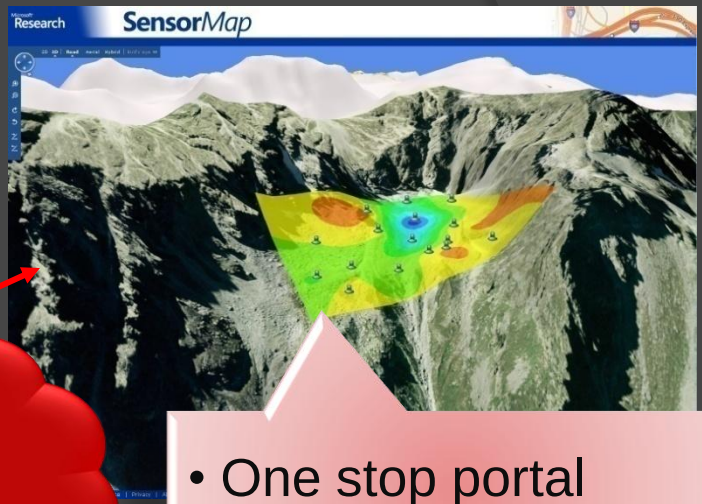


Microsoft Research

SenseWeb: Wikipedia of Sensors

- Sensor registration
- Metadata mgmt
- Store and query data
- Share data with others

SenseWeb



- One stop portal
- Discover data source
- Spatio-temporal viz.

SwissEx

USGS

Seamonster

...

http://atom.research.microsoft.com/sensewebv3/sensormap/

Microsoft Research

Thoughts

- Every system is different
  - Power options & power consumption
  - Infrastructure vs infrastructure-less
  - Stationary vs mobile
  - Engineered vs opportunistic vs participatory
- Extracting information from data is the challenge
  - Modeling, validation, learning, hypothesis testing
  - Archiving vs streaming
  - Network effects
  - Crowd sourcing/citizen scientists
  - Trust
- Can we design sensing systems at scale?
  - Computer scientists and *real* scientists need to collaborate
  - Fast prototype and feedback

Microsoft Research

ACM SenSys 2011 in Seattle

- Conference dates: Nov. 1-4, 2010
- Submission deadline: April 1, 2010
- Workshop proposals are welcome!

Microsoft Research