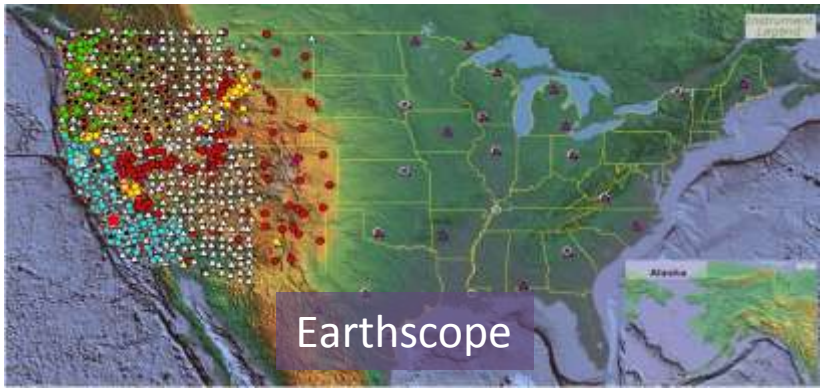
SenseWeb: Shared Macro-scopes for Scientific Exploration

Aman Kansal*, Suman Nath, Feng Zhao
Networked Embedded Computing
Microsoft Research

Instrumentation Is Hard

1. Share data via central archives
   - Swivel, Sloan sky survey, Fluxdata.org, BWC Data Server
2. Build macro-scopes: NEON, Earthscope
   - Can only address a few domains



3. **Share all instrumentation: SenseWeb**

Key Idea: Wikipedia of Sensors

Everyone deploys their own sensor network

Share all sensors using SenseWeb

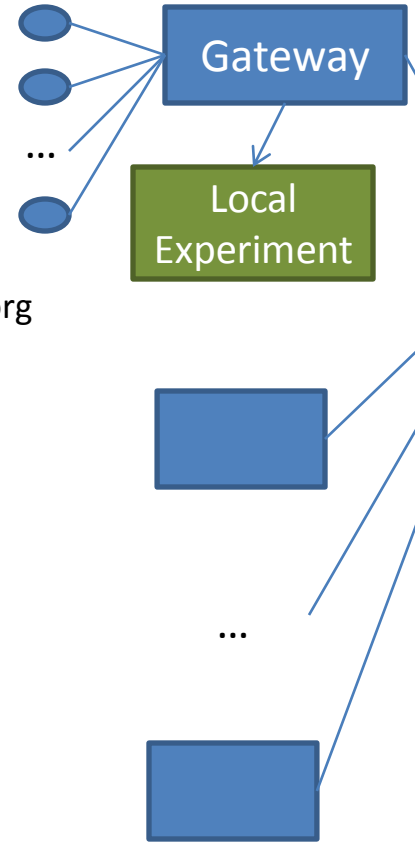
Everyone can run more experiments!

Soil Ecologists

Eg. LifeUnderYourFeet.org

USGS sensors

Other labs...



Outline

Case Study

- SeaMonster: Glacier, hydrology, and oceanographic exploration
- SensorMap Demo

SenseWeb Architecture

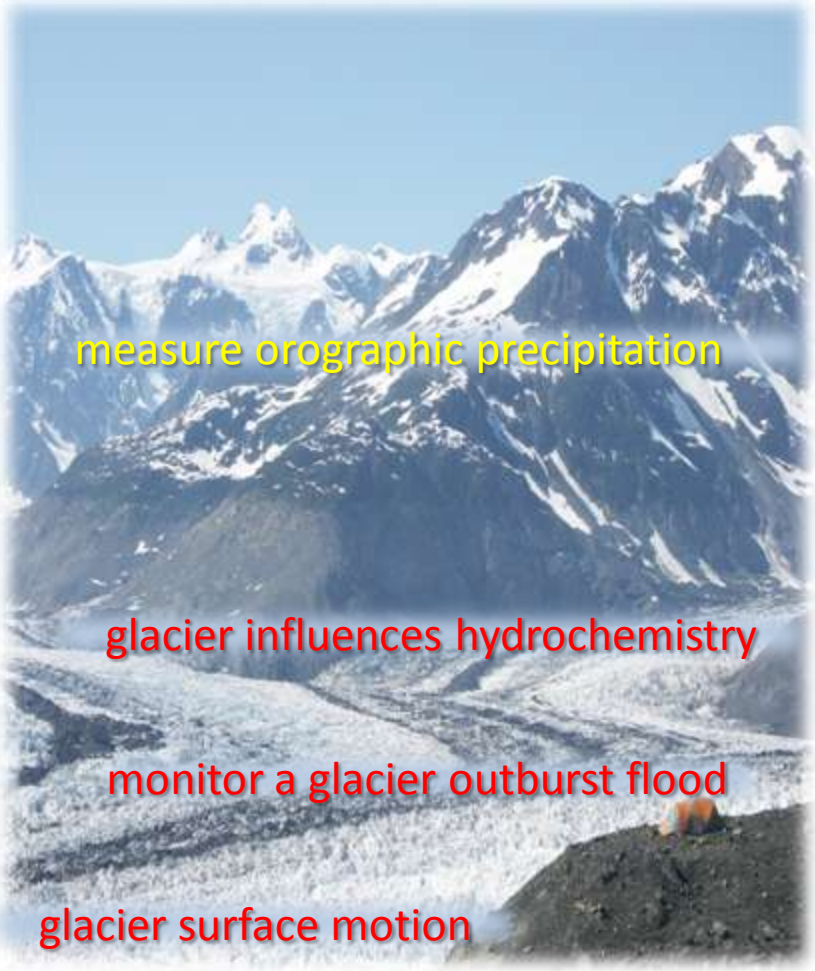
- Global or selective sensor stream sharing

Usage Examples

- Projects using SenseWeb

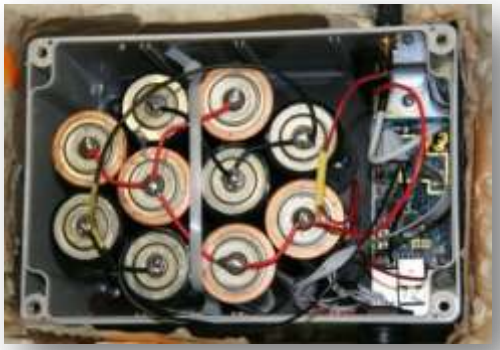
A Case Study: SeaMonster

- **S**outh **E**ast **A**laska **MO**nitoring **N**etwork for **S**cience, **T**elecommunications, **E**ducation, and **R**esearch
  - Collaborative environmental science with large volumes of environmental data
  - NASA, NOAA, Univ. of Alaska, Vexcel-Microsoft



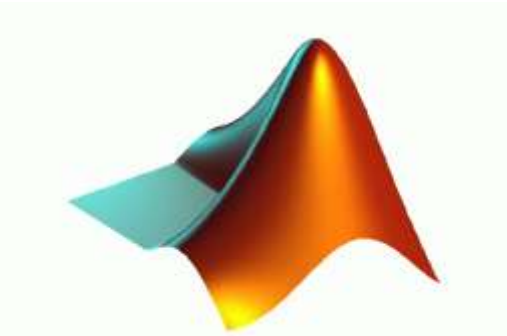
SeaMonster: Generation 1

- Deploy sensors with local storage



- Physically visit for data collection

- Process archived data offline

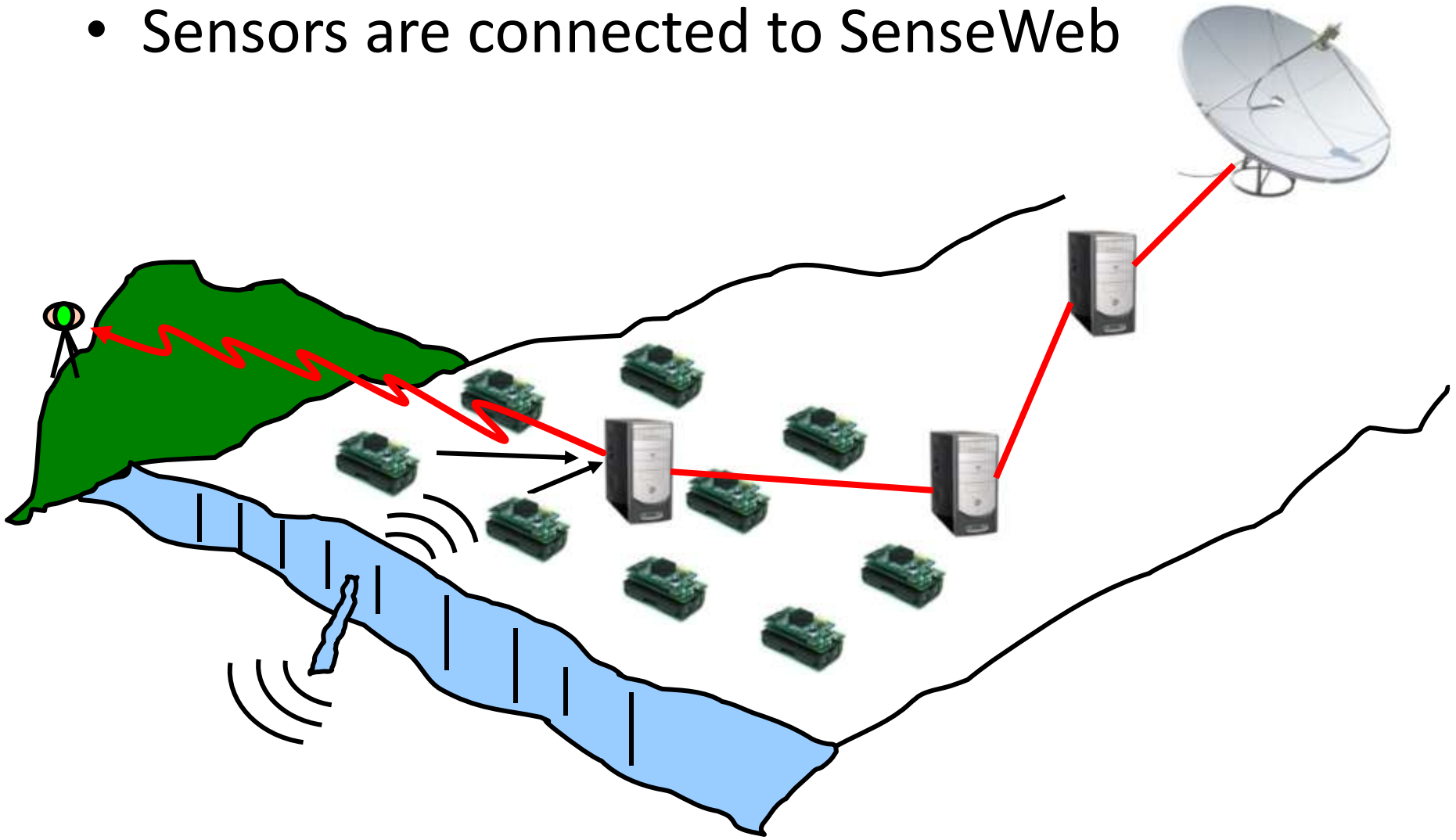


Generation 1

- Problems:
  - No real time feedback
  - No data if the device fails
  - No interactivity

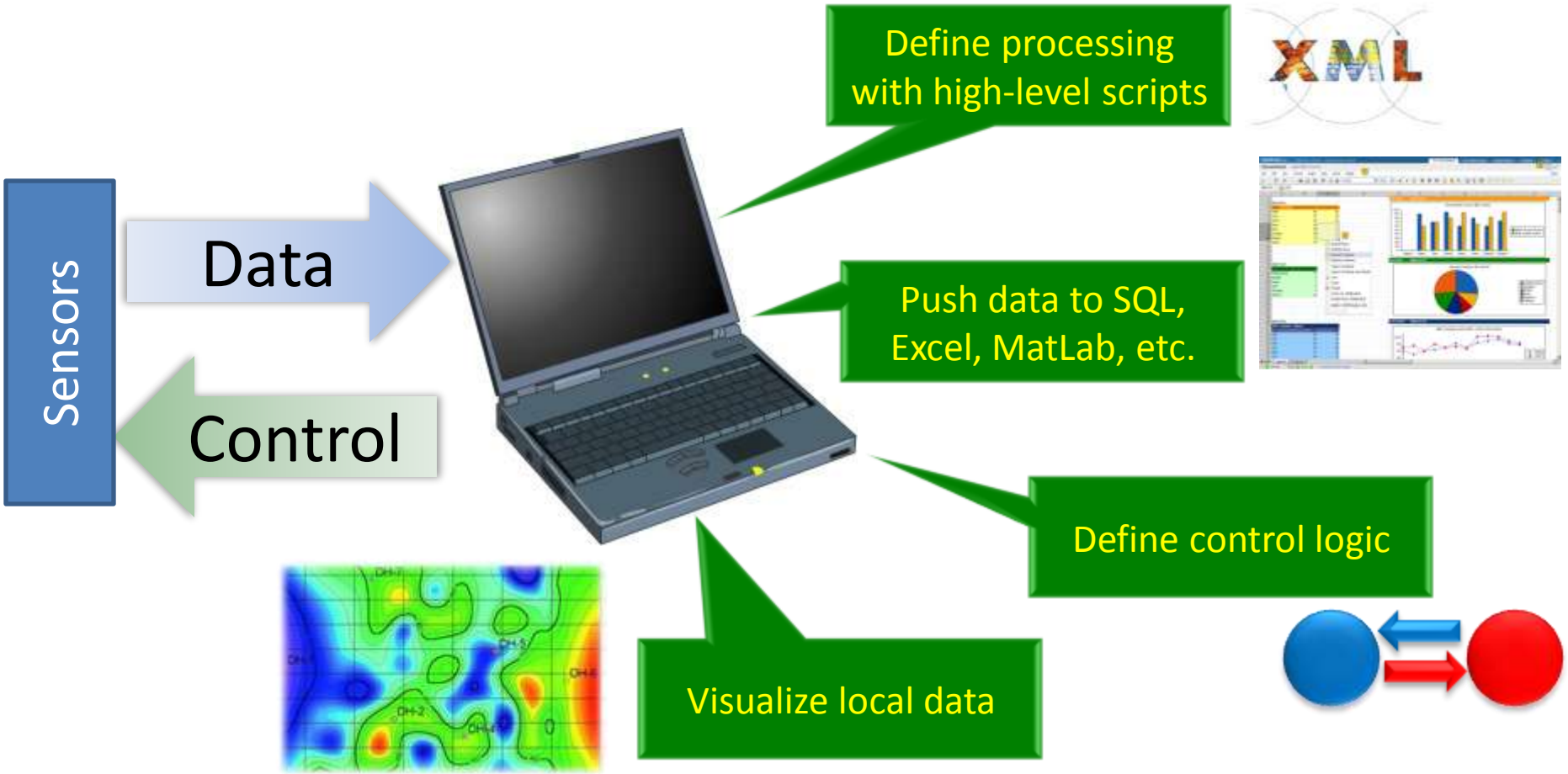
SeaMonster: Generation 2

- Sensors are connected to SenseWeb



MSRSense

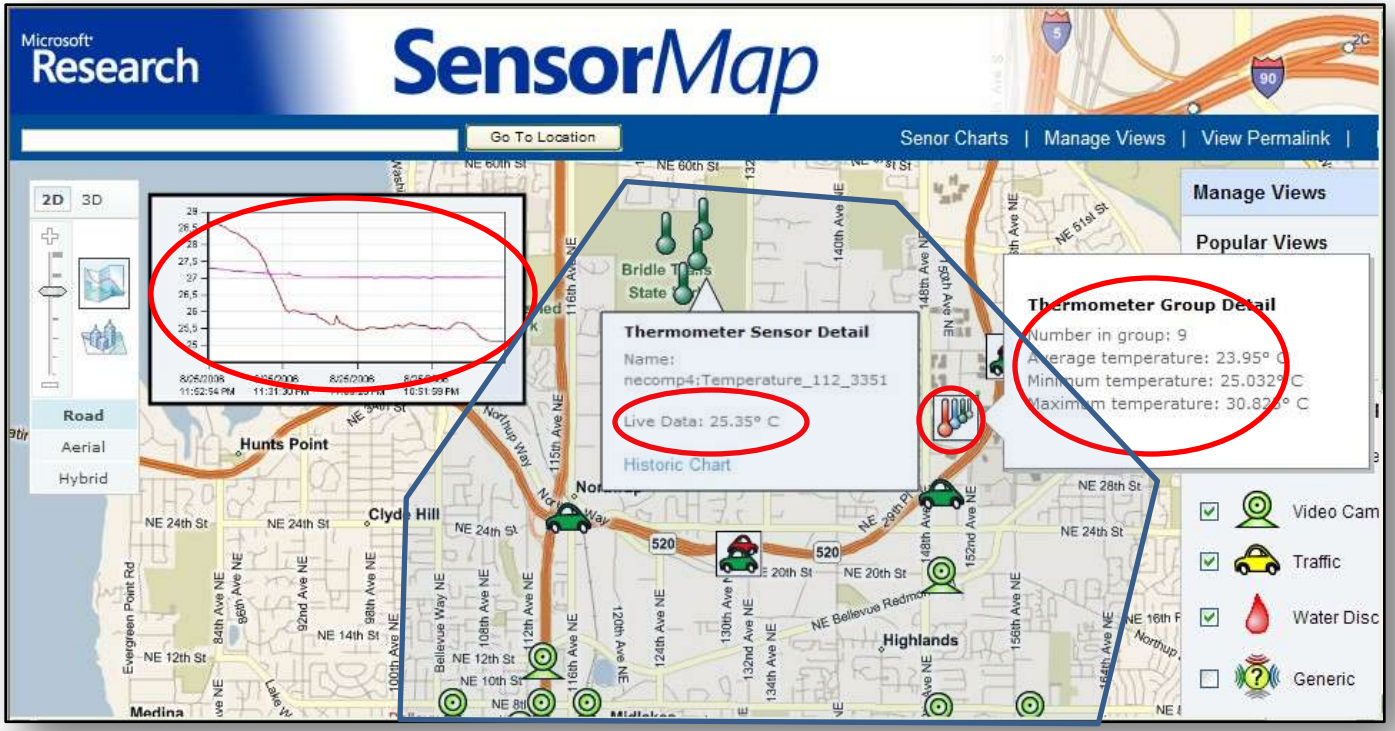
- Real-time data streaming and processing



SensorMap

Portal for finding sensors, eye-balling sensor data, and manage sensors

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

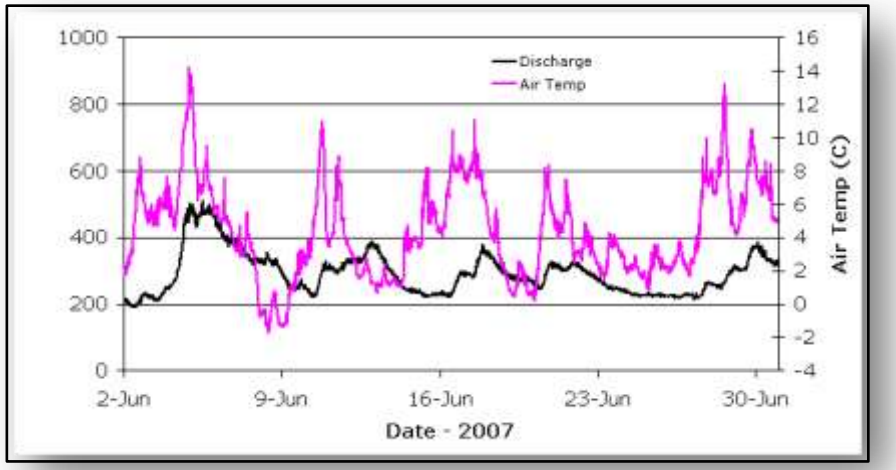


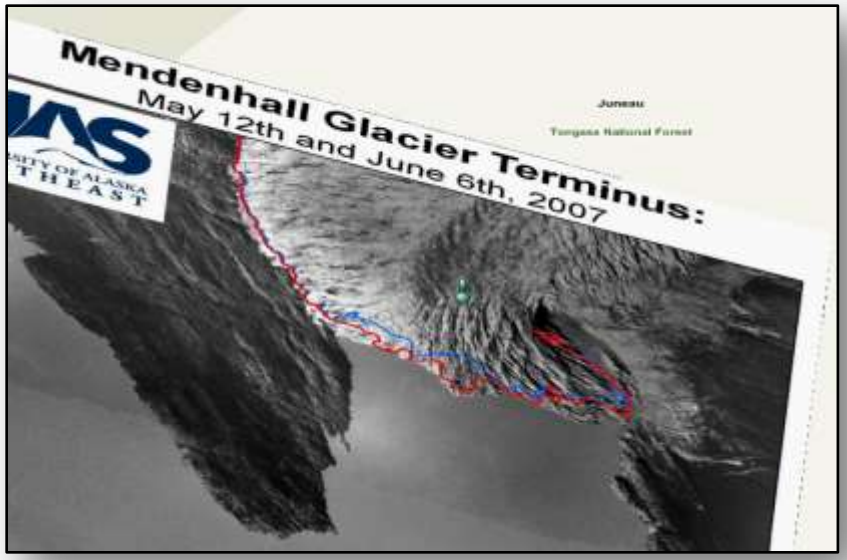
Sensors as Icons
Show real-time and archived data

Search sensors based on geography, type, keywords

Aggregate live data at different zoom levels

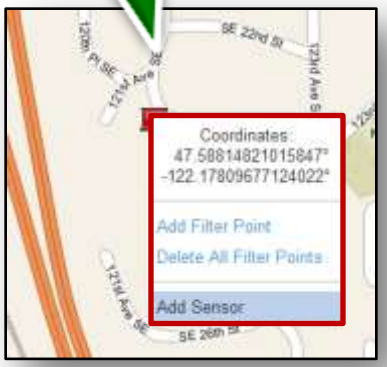
3D and Custom Visualization





Manage Sensors on SensorMap

Click directly on map to add sensors



Fill out form with sensor metadata

Publish Sensor

Sensor Name

Sensor Type*
Video Camera

File Type

Latitude 47.589682244791305

Longitude -122.17809677124022

Real Time URI

Keywords (optional)

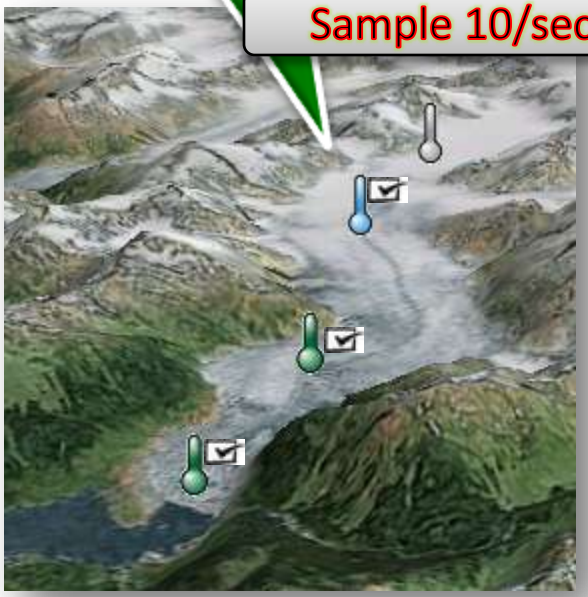
Sensor Description (optional)

Publish Cancel

* SensorMap only allows publishing of video and generic sensors. More sensor types may be supported soon.

Select multiple sensors and send command

Sample 10/sec



Outline

Case Study

- SeaMonster: Glacier, hydrology, and oceanographic exploration
- SensorMap Demo

SenseWeb Architecture

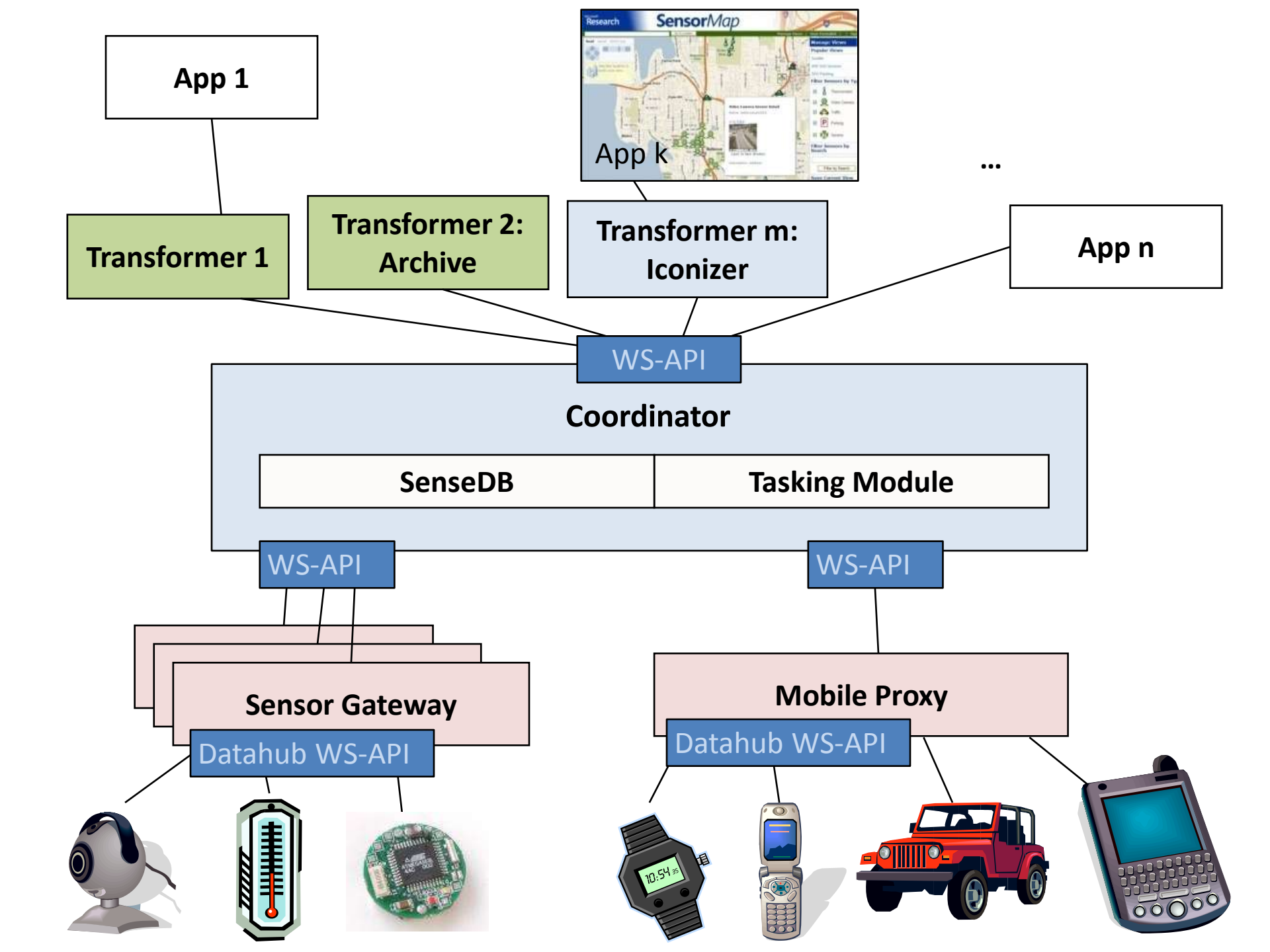
- Global or selective sensor stream sharing

Usage Examples

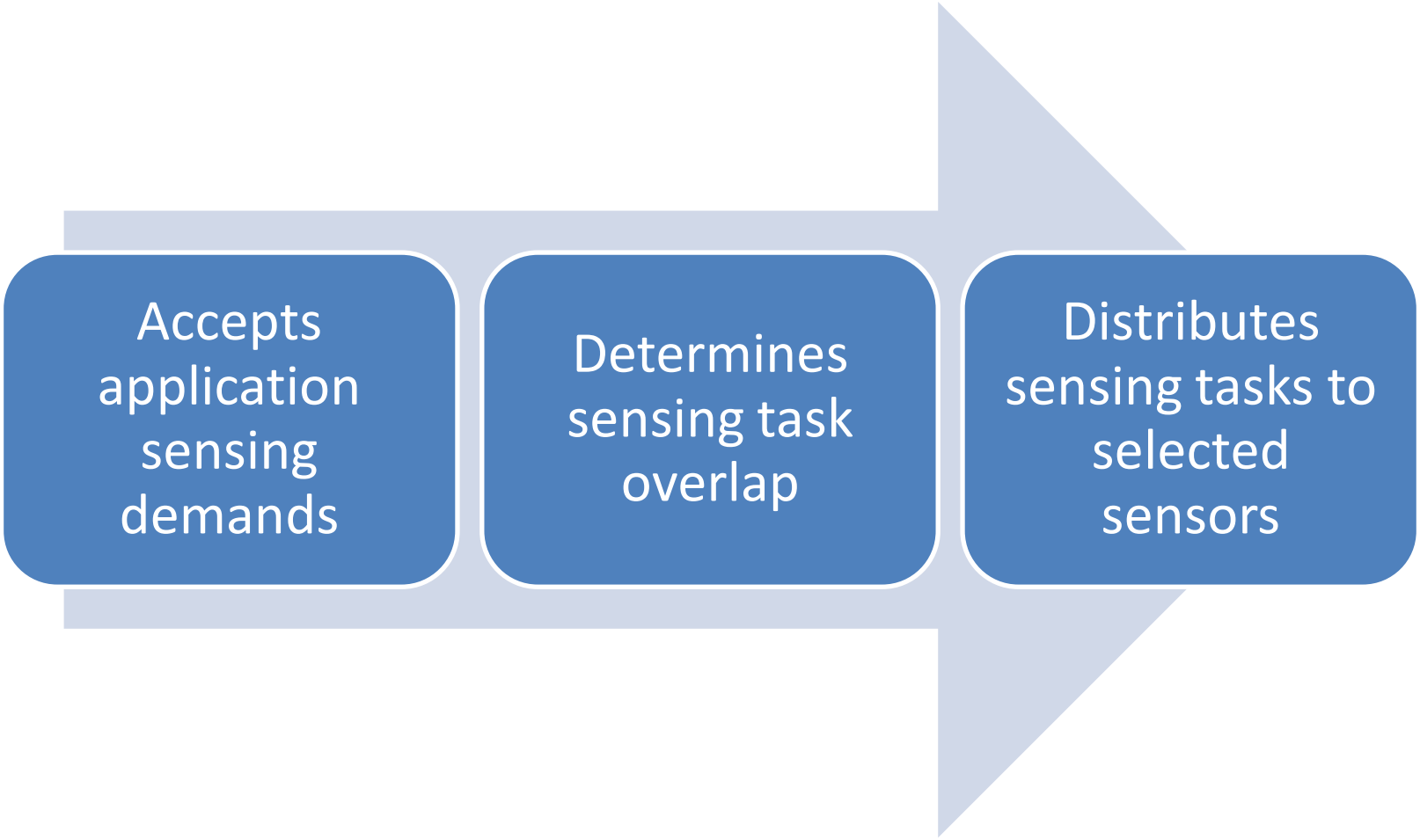
- Projects using SenseWeb

Architecture Design Challenges

- Heterogeneity
  - Resource capability: bandwidth, power, computation
  - Willingness to share
  - Measurement accuracy
- Scalability
  - Streaming all raw data from all sensors to all applications not feasible
- Security and Privacy
- Data Verifiability, Trust


App 1
Research
SensorMap
App k
...
Transformer 1
Transformer 2:
Archive
Transformer m:
Iconizer
App n
WS-API
Coordinator
SenseDB
Tasking Module
WS-API
WS-API
Sensor Gateway
Datahub WS-API
Mobile Proxy
Datahub WS-API

Coordinator



Data Re-use

- Many applications may need similar data
  - Within a tolerable latency of each other
  - From overlapping region
- Can cache data and aggregates to reduce load on sensors and network
  - Overlap may be partial: computed aggregates may need partial new data

Query Model

SELECT Count(*)

FROM Sensor

WHERE sensor.location in Polygon(A) AND
sensor.time BETWEEN now()-10 and now()+10

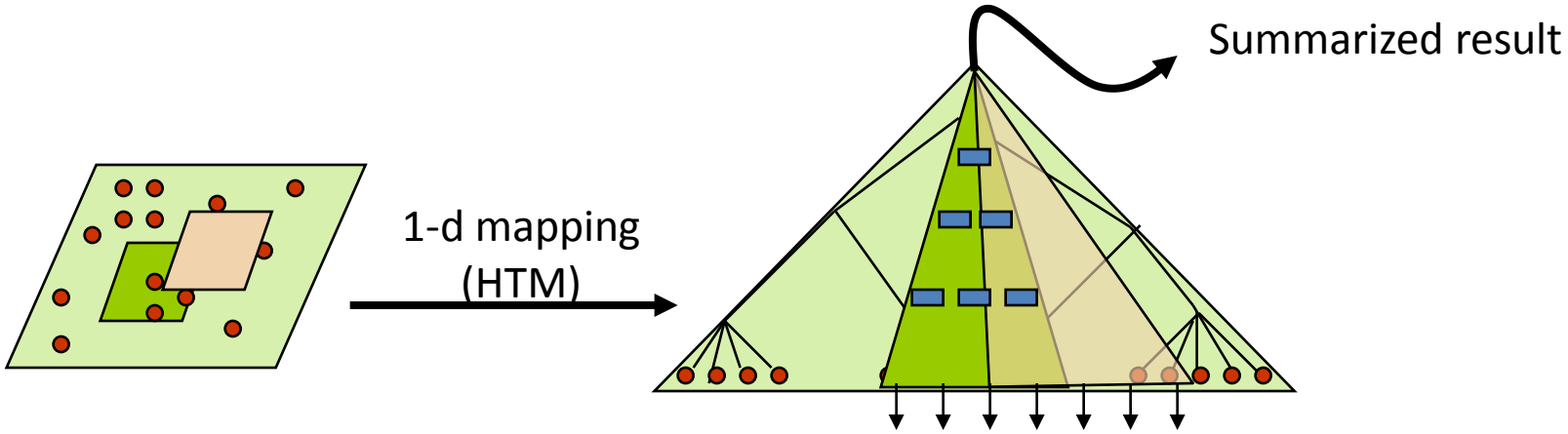
REPORTRATE 10 min

SAMPLESIZE 50

EVENT EventSpec(T>25)

COLR-Tree (COLlection R-Tree)

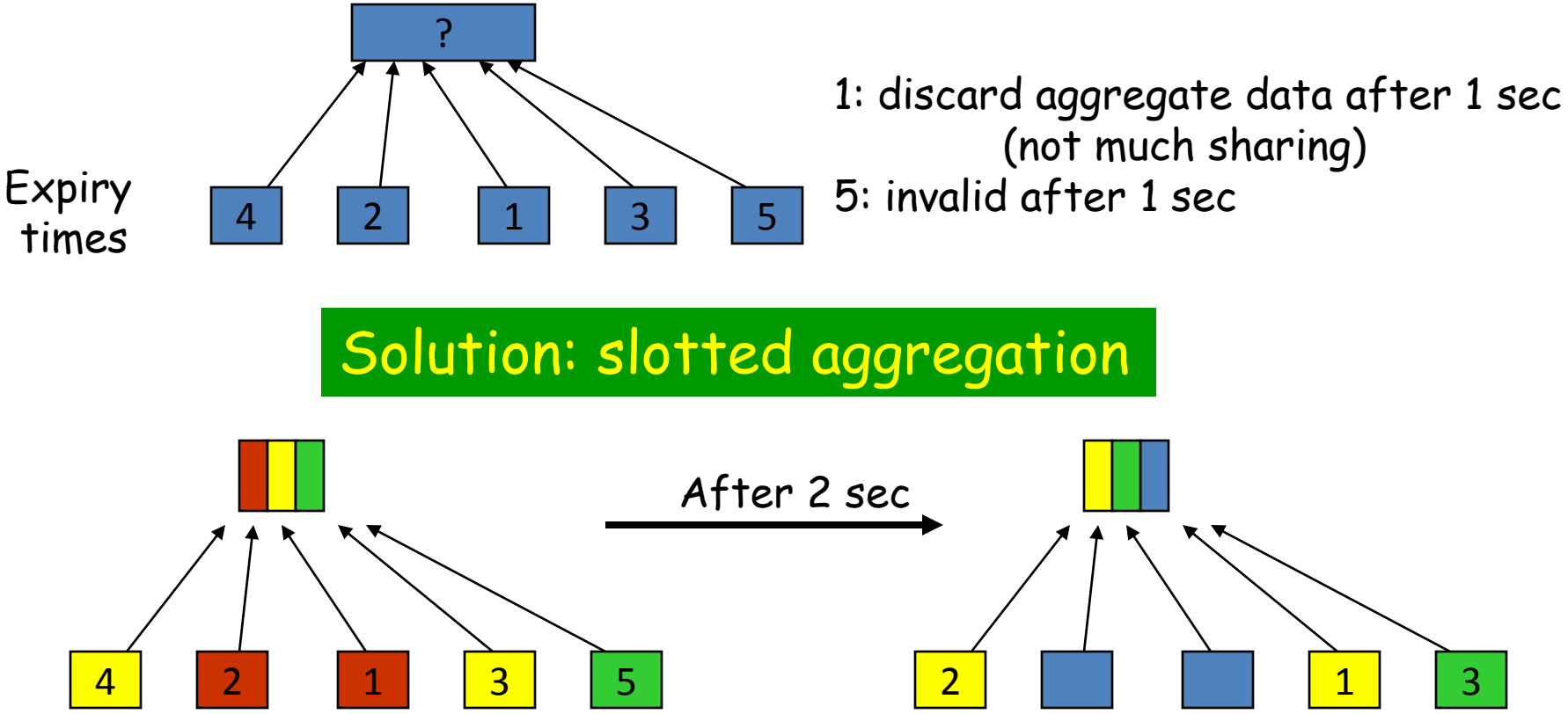
Index 2-D data with aggregates



- Minimizing sensor access
  - Cached data may have skewed distribution
  - Sample more from non-cached region
- Implemented on MS-SQL Server: usable with all SQL server capabilities

COLR-Tree: Aggregates

- Challenge: temporal aggregation



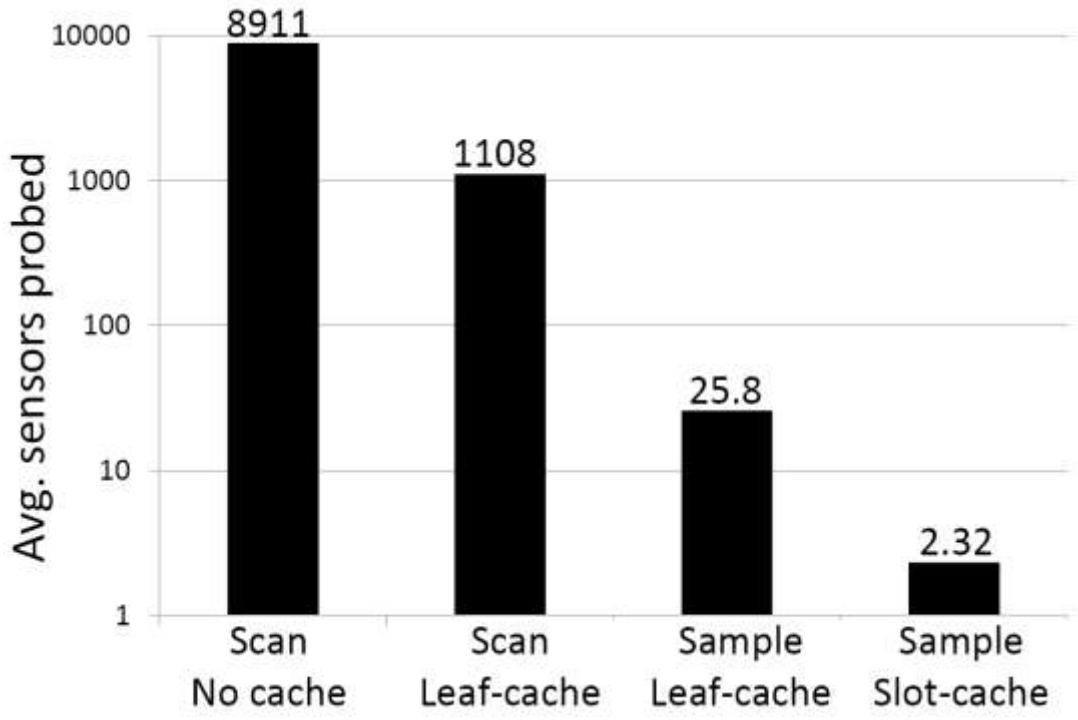
Spatial Sub-sampling

- Suppose sample size of R needed
- Layered sub-sampling along COLR-tree levels
- Partition R to achieve spatially uniform sample
  - BB(i): area covered by i-th child, c(i): data cached for i-th child, w(i): sensors under i-th child, q: query region
  - For each child I at next level:

$$R(i) = R * \frac{w(i) - c(i)}{\sum_i w(i) * |overlap(BB(i), q)|}$$

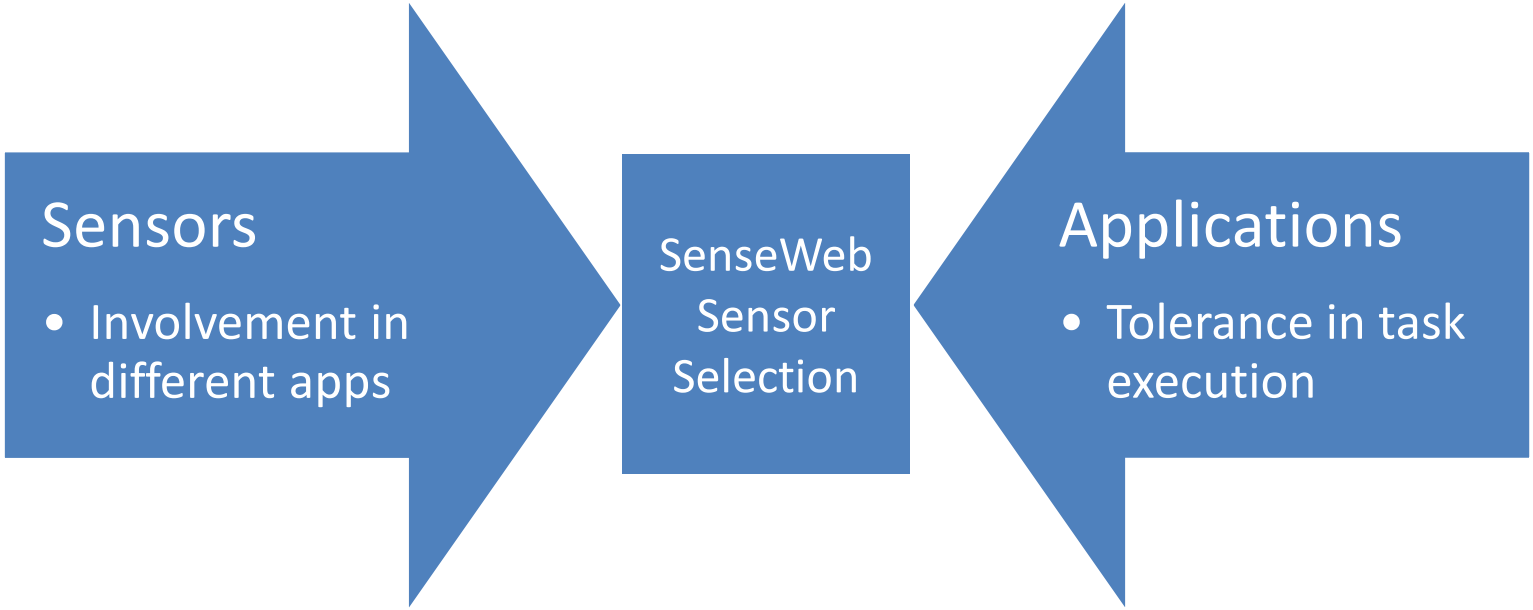
COLR-Tree Evaluation

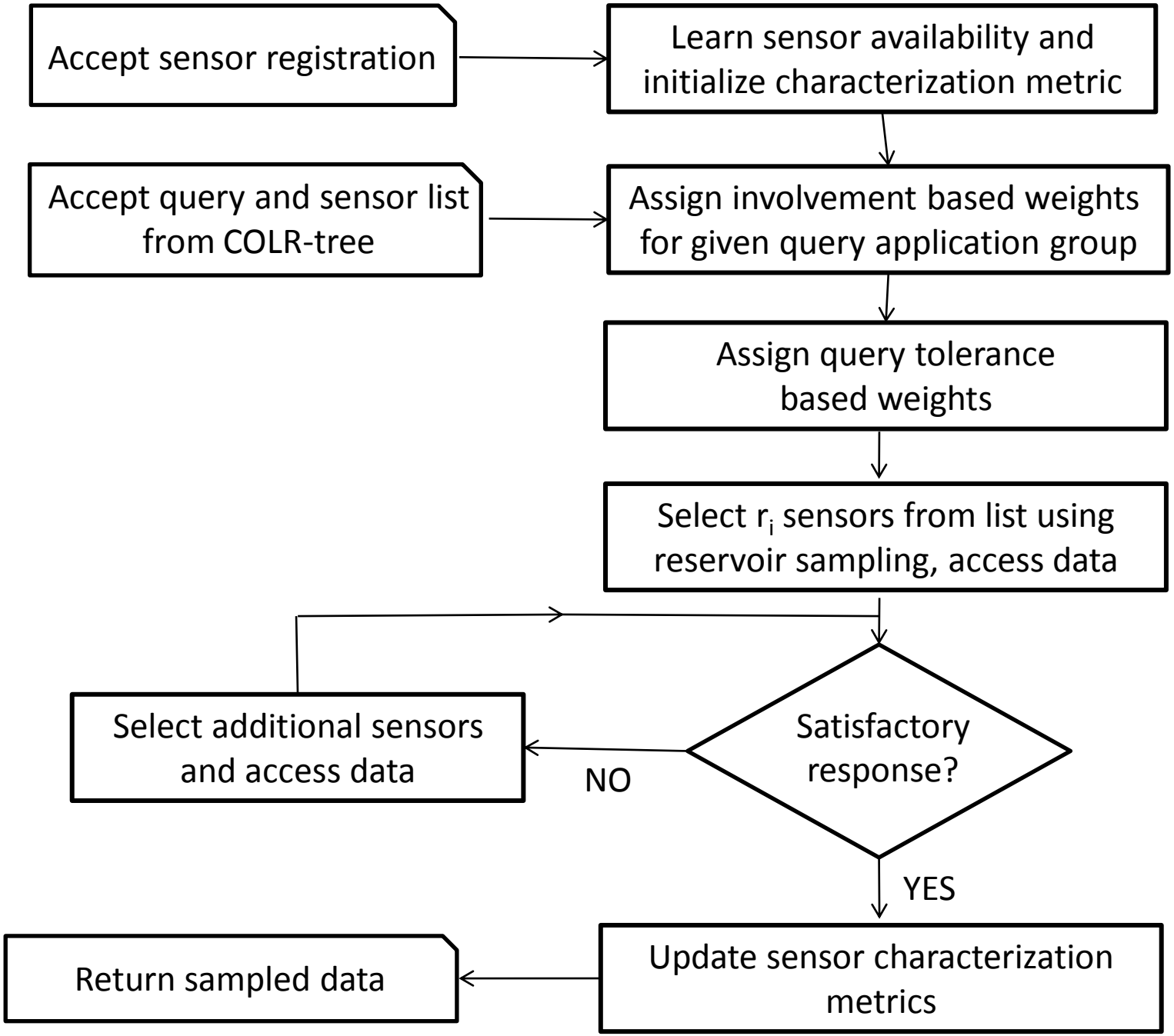
- Test data
  - 400K points from VE Yellow Pages
  - Regions queried: Virtual Earth usage trace



Tasking Heterogeneous Sensors

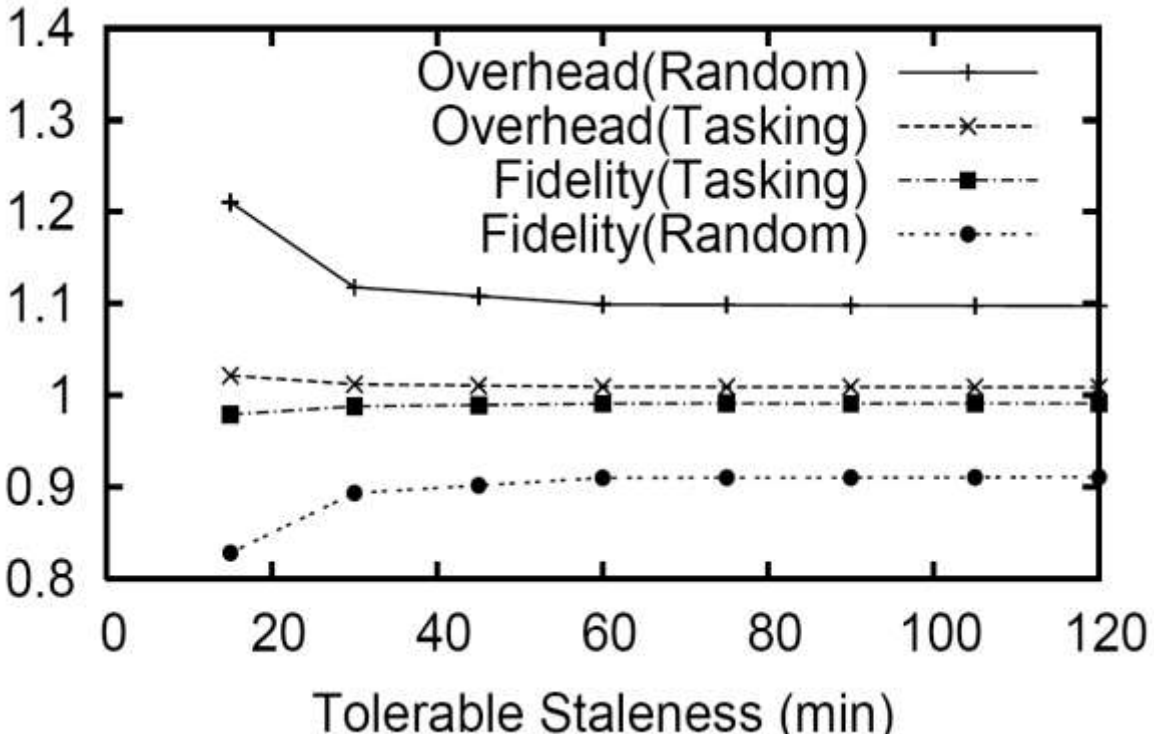
- Select uniformly rather than overloading the best sensors
- Leverage lower capability sensors when usable for a query
- Learn and adapt to sensor characteristics: availability, bandwidth
- Weighted reservoir sampling
  - Weighted random selection, with desired number of sensors




Accept sensor registration
Learn sensor availability and initialize characterization metric
Accept query and sensor list from COLR-tree
Assign involvement based weights for given query application group
Assign query tolerance based weights
Select r_i sensors from list using reservoir sampling, access data
Satisfactory response?
NO
Select additional sensors and access data
YES
Update sensor characterization metrics
Return sampled data

Tasking Algorithm Performance

- Test on USGS stream water sensors
  - Random selection vs. Weighted reservoir sampling

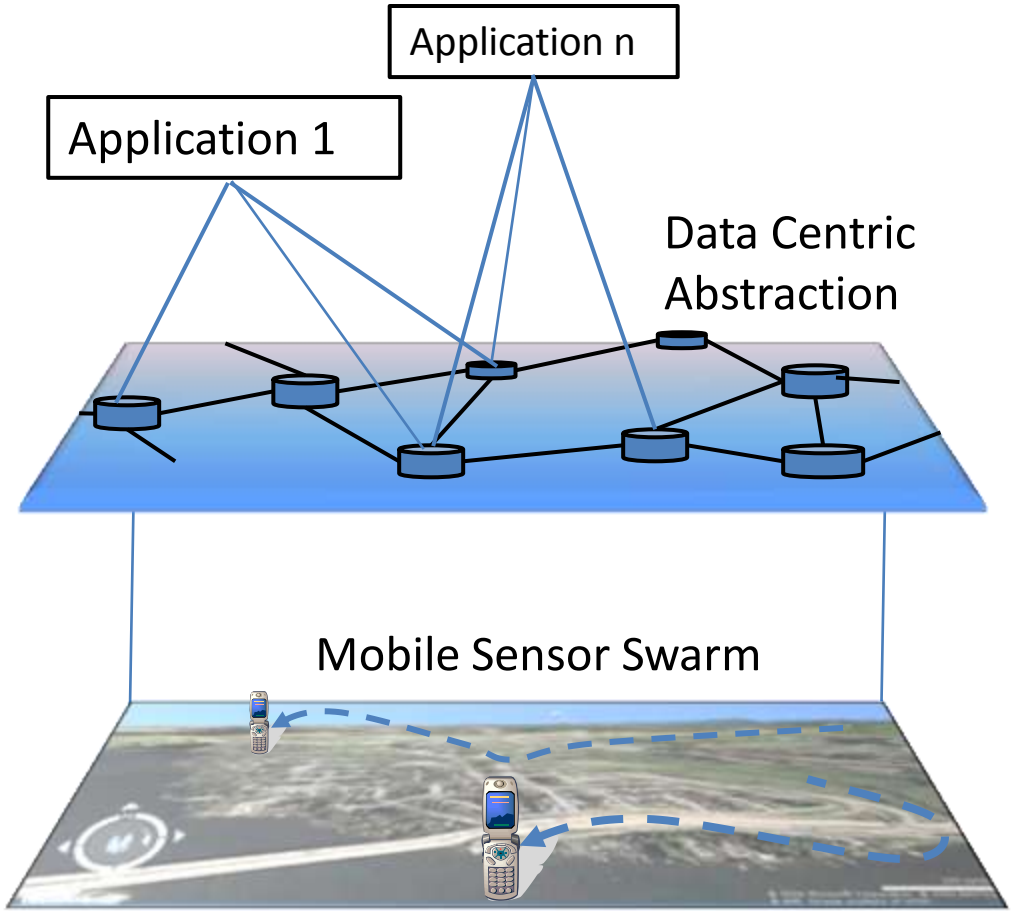


Mobile Sensors in SenseWeb

More coverage but

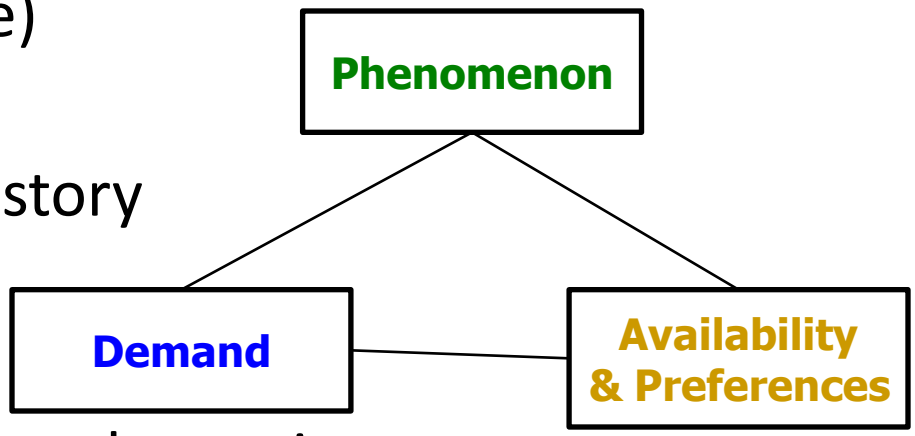
Hard for application to track relevant devices

- **Solution:** data centric abstraction
  - Location based indexing
    - using GPS, cell-tower triangulation, content based location



Community Sensing

- Leverage roving sensors to measure urban/social phenomenon
  - **Information value** (collapse uncertainty)
  - **Demand** (“utilitarian” usage)
- Sensor availability
  - **Predict location** based on history
- Preferences
  - **Abide by preferences**
  - E.g., Frequency / number of probes, min. inter-probe interval
  - Other constraints: e.g., “Not near my home!”



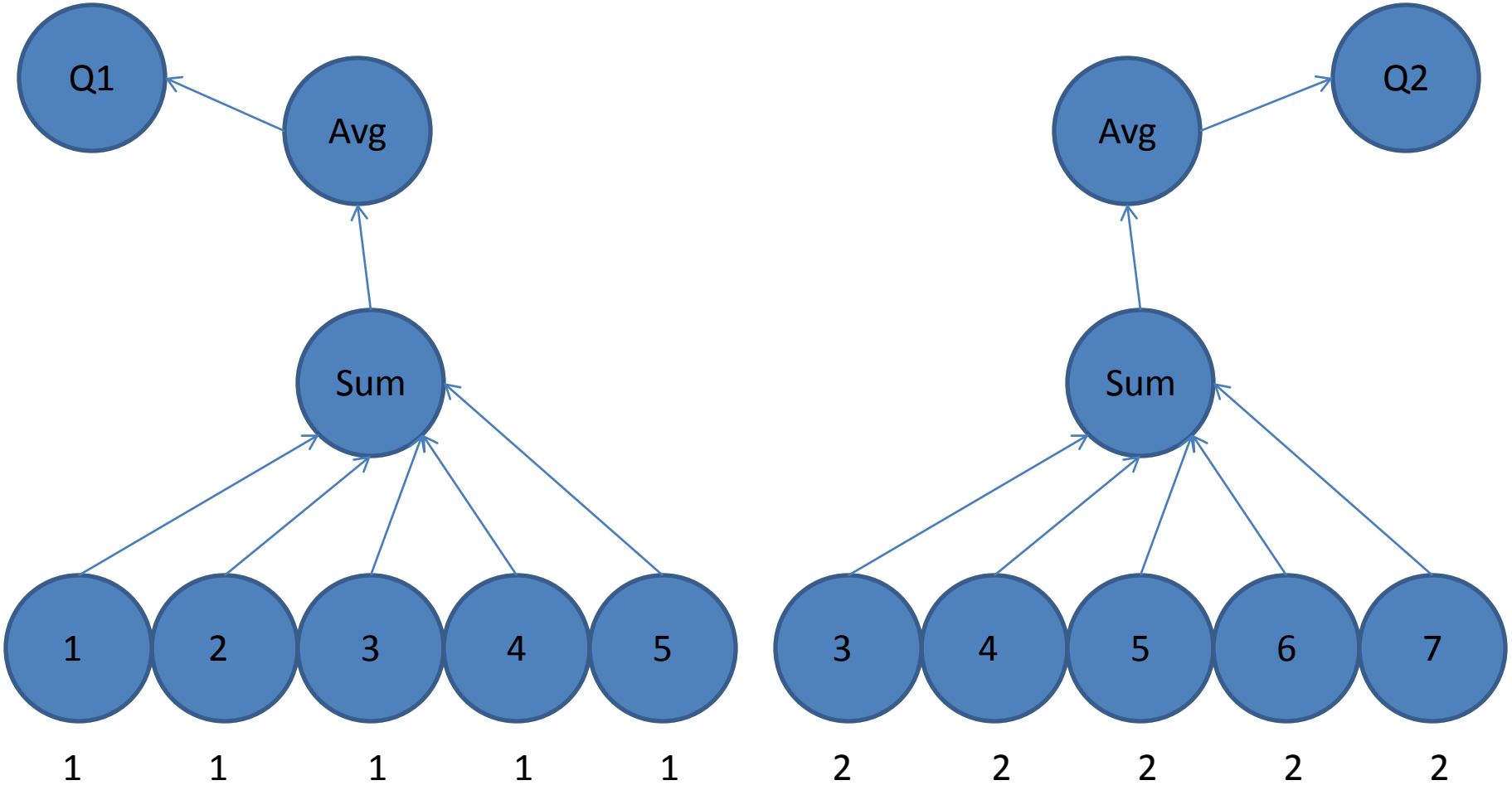
(With Andreas Krause, Eric Horvitz)

Shared Streaming

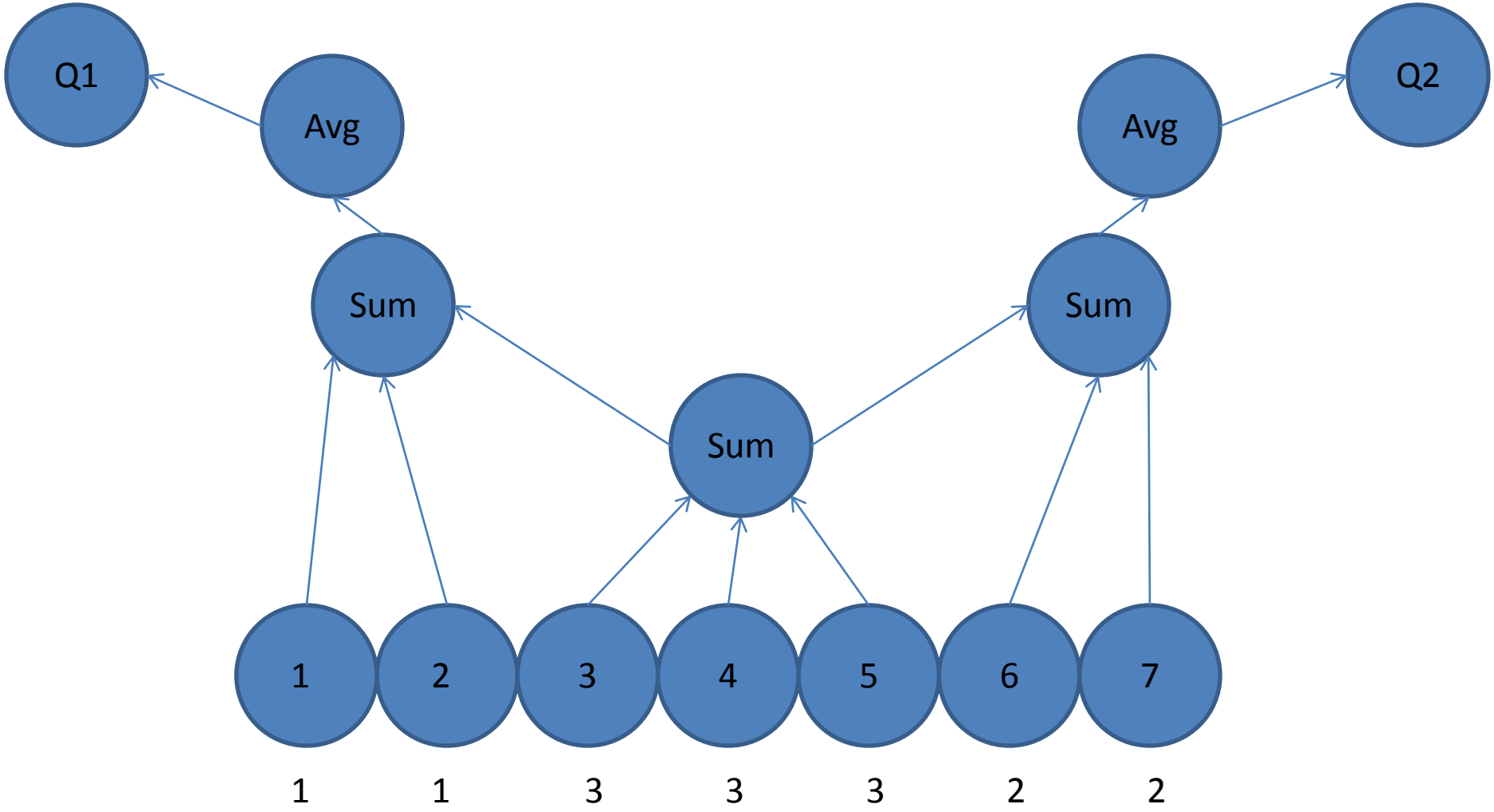
- Multiple apps. need data from similar sensors
- Problems
  - Sensor resources limited
    - Upload bandwidth, connectivity
    - Energy
  - Scalability of aggregation and streaming
- Solution
  - Cache data: identify relevant cache efficiently
  - Share aggregation and processing

(With Arsalan Tavakoli)

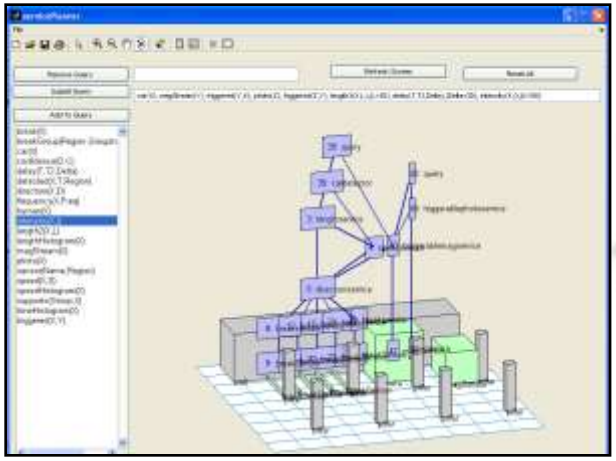
Query DAG's



Optimized Shared Query DAG



Tools for Sensor Contributors



For mote networks

- Automatic data collection and sharing
- Simplified processing and application composition

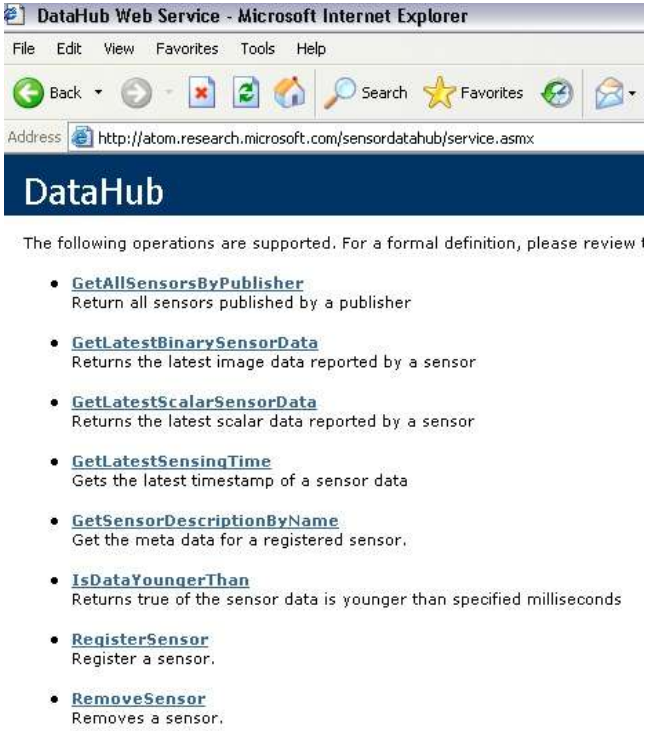
Client for cell-phones

- Allows users to take pictures
- Automatically uploads data to server
- Location stamps using inbuilt/Bluetooth GPS

Webcam data processing and sharing tool

Tools for Sensor Contributors

- Gateway for sensor contributors
  - Web service API: Datahub
  - Supports several sensor types via semantic hierarchy
  - Also archives sensor data
- Tools available for download
  - Tutorials available online



Outline

Case Study

- SeaMonster: Glacier, hydrology, and oceanographic exploration
- SensorMap Demo

SenseWeb Architecture

- Global or selective sensor stream sharing

Usage Examples

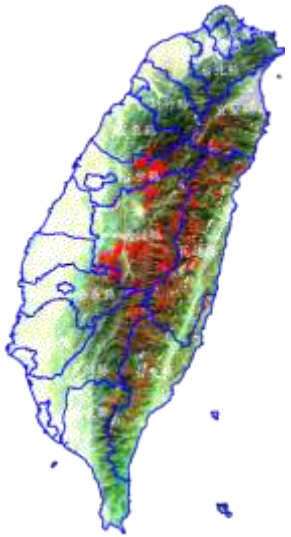
- Projects using SenseWeb

Current Projects

- Urban air quality
  - Vanderbilt, Harvard Univ

- Life Under Your Feet
  - John Hopkins Univ.

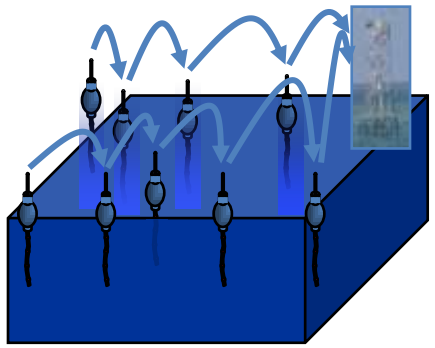
- Debris Flow
  - National Tsing Hua University, China



Current Projects

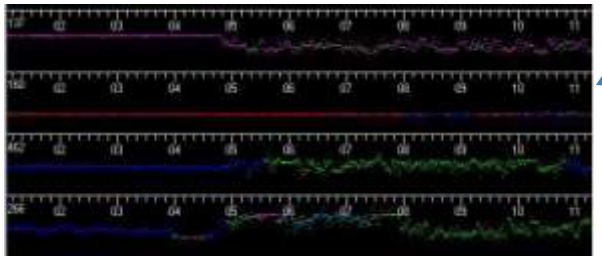
- National Weather
  - NTU Singapore

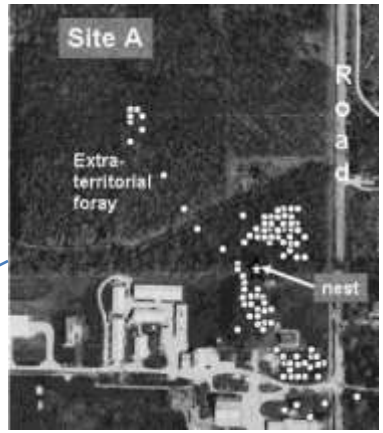
- Coral reef ecosystem in The Great Barrier Reef
  - U. Melbourne



Current Projects

- Bioscope: bird call streaming
  - UIUC





- Swiss-Experiment
  - EPFL, ETH, others

Applications Beyond Science

Community Fitness and Recreation

- Runners: Where are sidewalks broken? Construction finished on 24th St?
- Mountain Bikers: Average biker heart rate at Adams Pass on trail 320? [SlamXR]
- Surfer: What is the wave level and wind speed at Venice Beach now?

Real Time Information

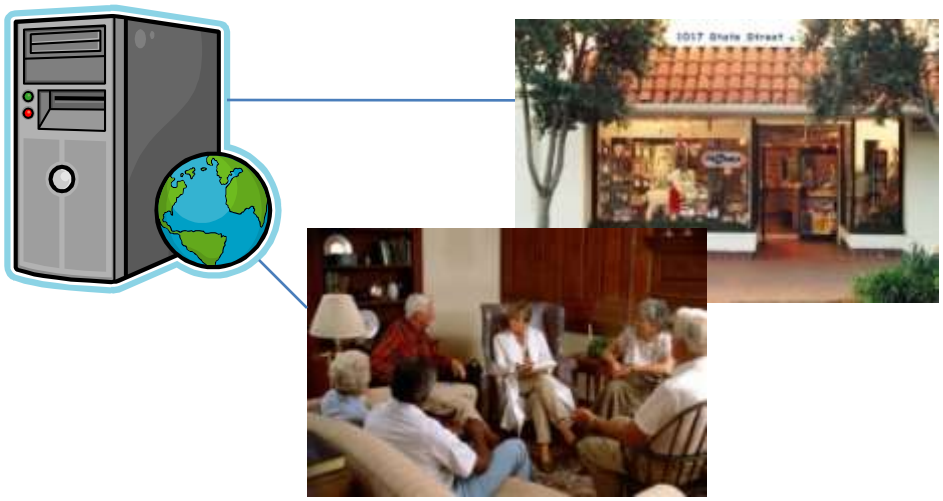
- Public initiated instant news coverage
- Road traffic monitoring from shared car GPS receivers

Business

- What are people doing tonight? Restaurant waiting times in downtown?
- Mall visitor activity and parking usage across franchise outlets worldwide
- Share pictures of suspected restaurant hygiene issues

Current Projects

- Urban-Net
  - Shopper interest
  - Assisted living
  - U. Virginia

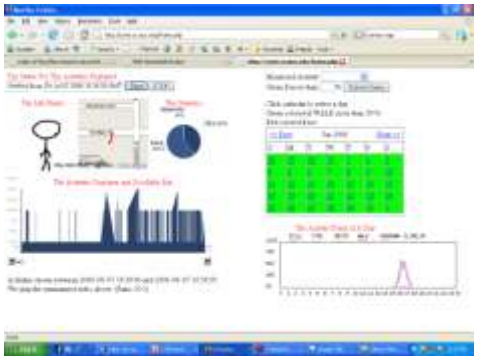


- Indoor events
  - U. Washington

Current Projects

- Large scale urban monitoring
  - Harvard

- Human Activity View
  - UIUC



Summary

- SenseWeb
  - Share sensor networks
  - Generic data and sensor management
- SensorMap
  - Interact with sensors in real time
  - Eye-ball sensor data
- MSRSense
  - Domain specific data analysis/mining
- Details: http://research.microsoft.com/nec/senseweb/