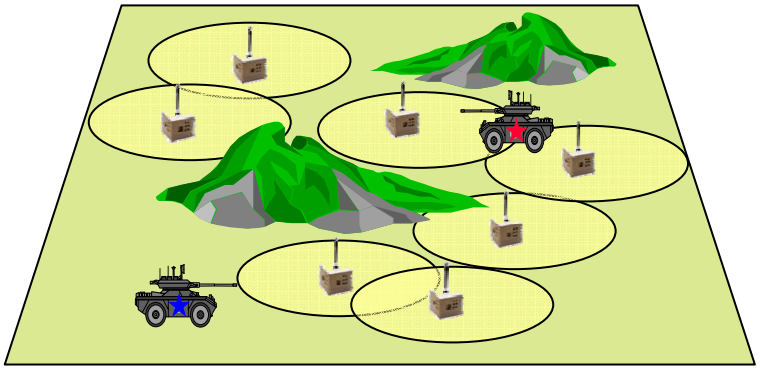
***PRESTO**: Feedback-driven Data Management in Sensor Networks*

(*__PRE__dictive __STO__rage)

Ming Li, Deepak Ganesan, and Prashant Shenoy
University of Massachusetts, Amherst

Emerging large-scale sensor networks

- *Wired/wireless network of spatially distributed sensing devices*
- *Closely and densely monitor physical phenomena*



Object Tracking

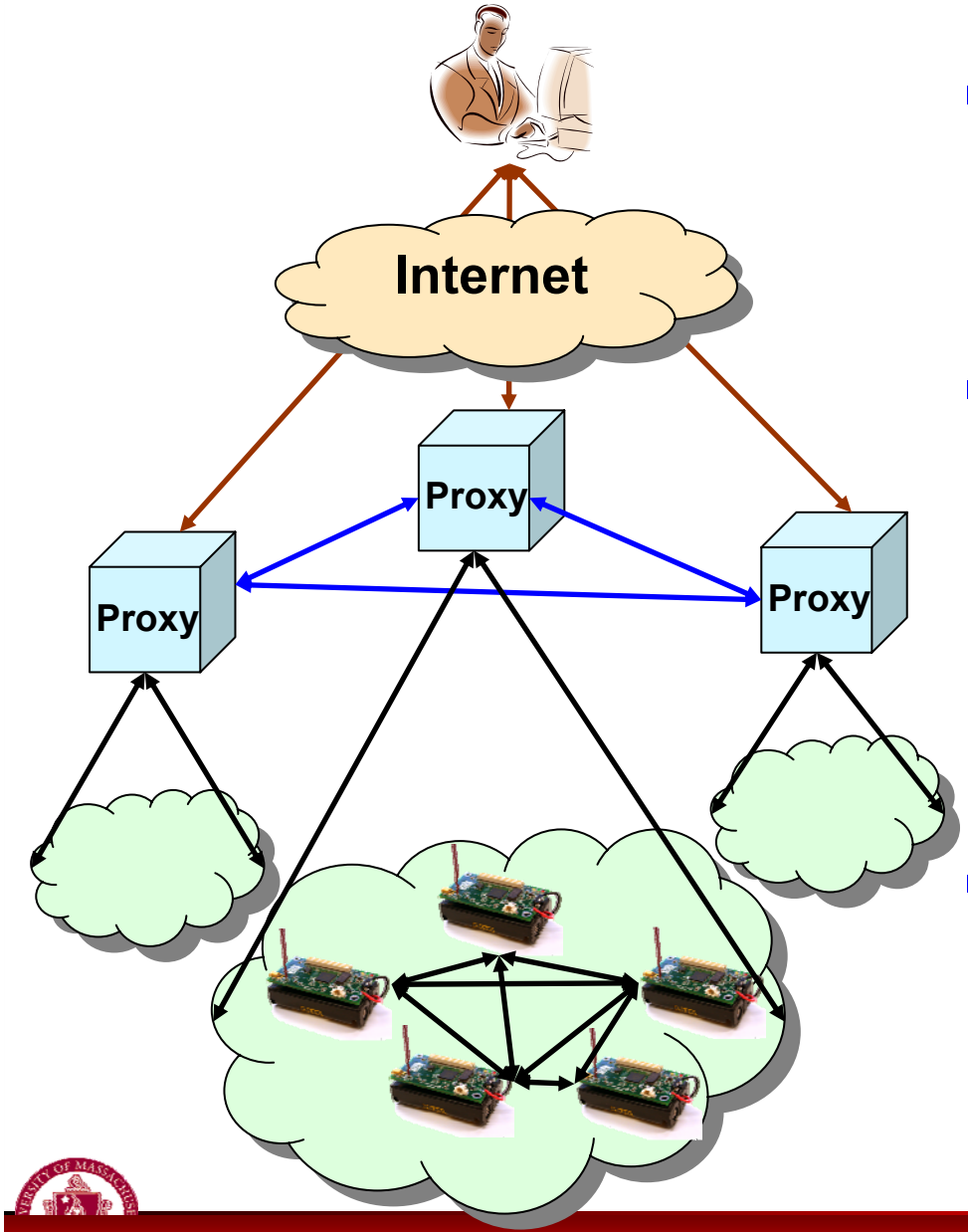
Ecosystem Monitoring

Habitat Monitoring

Resource Constraints

- Energy -> 1 AA battery for 1 year
- Computation < Storage << Communication

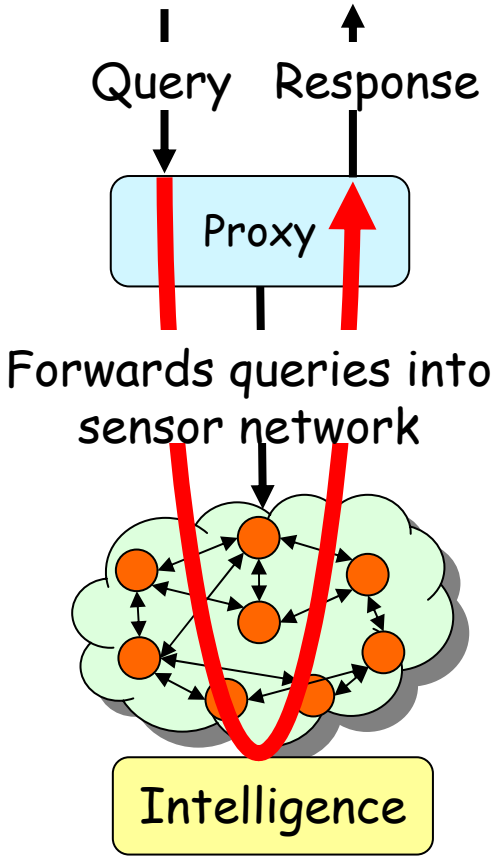
Hierarchical Sensor Network Architecture



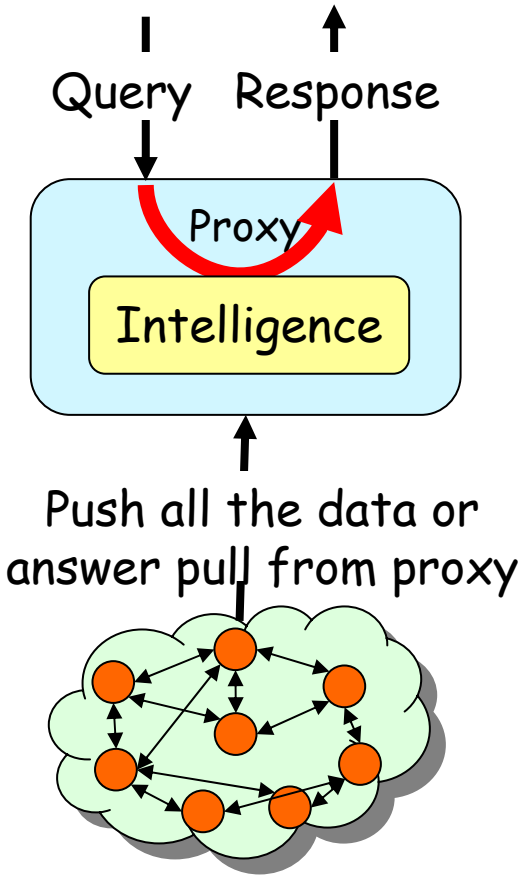
- User
  - Issues queries to the proxy.
- Proxy
  - Gateway between user and sensor
  - PDA class microcomputer
  - Tethered or solar-powered
- Sensor
  - Sensing physical world
  - Resource constrained

Approaches to Proxy-Sensor Interaction

Sensor-centric

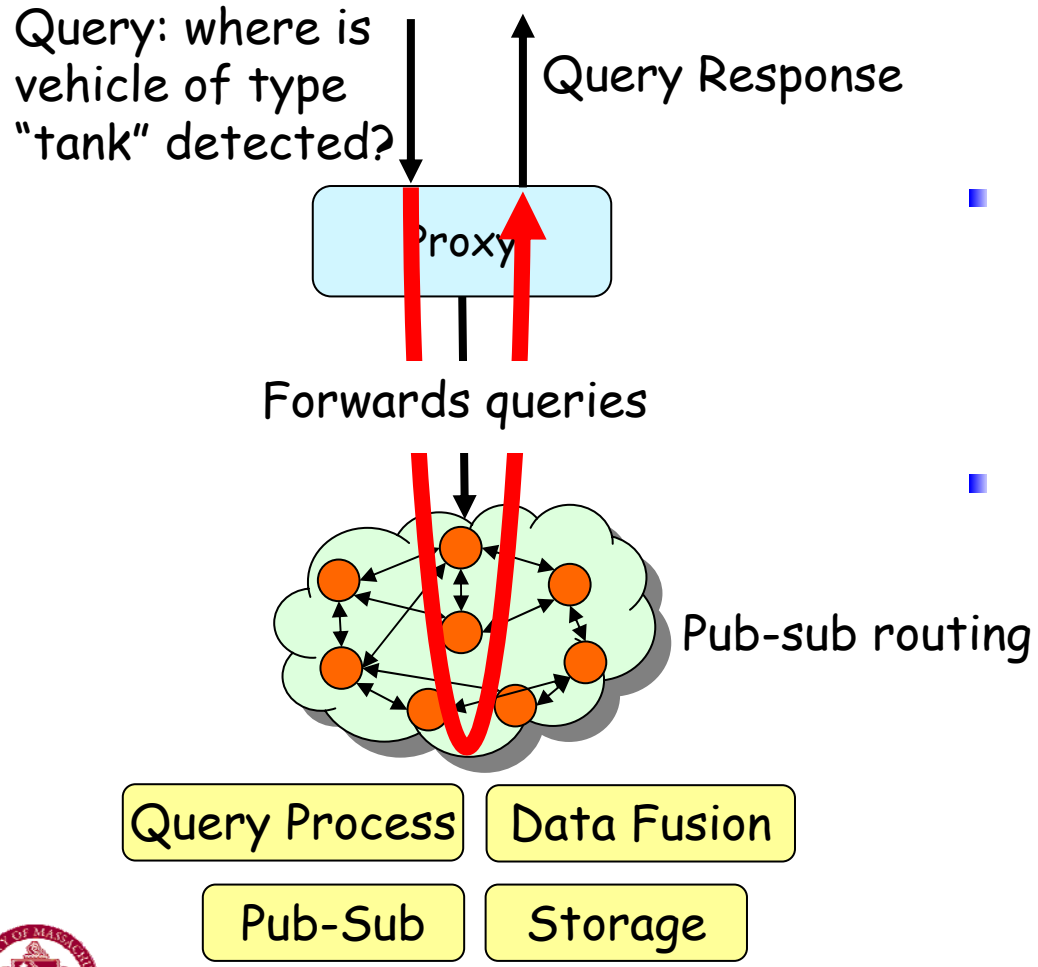


Proxy-centric



Sensor-Centric Architecture

Directed Diffusion [Heidemann01]



- Sensors perform data fusion, query processing and filtering within the network.
- Pros:
  - Greater query accuracy
  - Better energy-efficiency.
- Cons:
  - Greater sensor complexity.
  - Greater query latency.

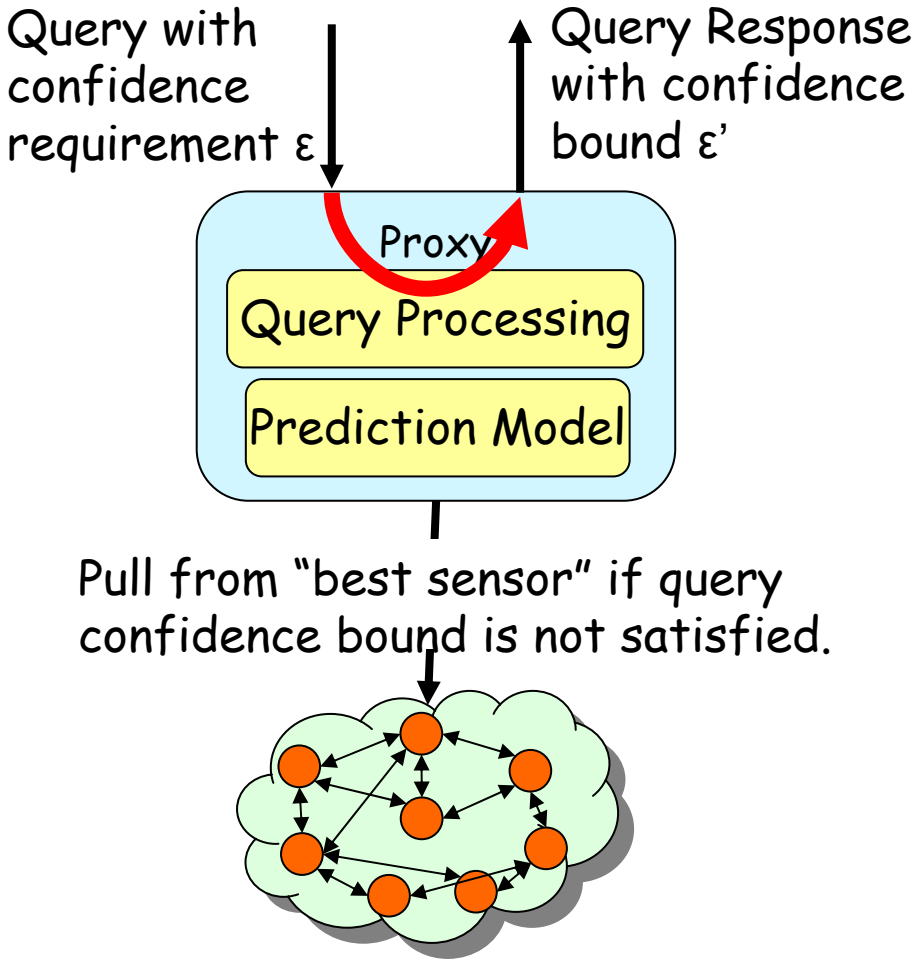
Proxy-Centric Architecture

- Proxy processes queries and uses complex spatial/temporal model to predict

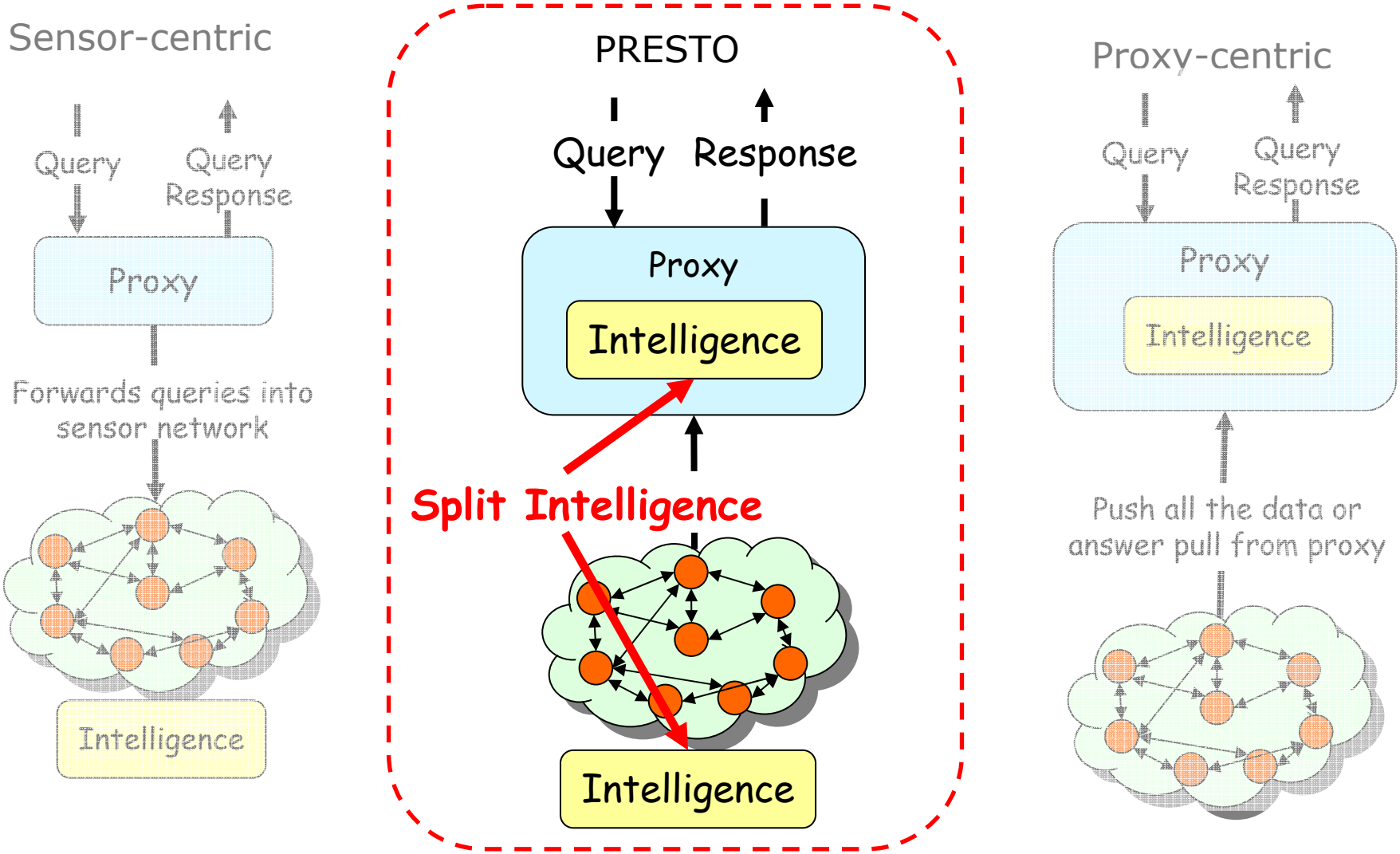
- Pros:
  - Intelligence placed where resources are available.
  - Lower query latency

- Cons:
  - Hard to detect anomalies.
  - Sacrifice energy-efficiency or query accuracy.

BBQ [Deshpande04]



PRESTO


Sensor-centric
Query
Query Response
Proxy
Forwards queries into sensor network
Intelligence
PRESTO
Query
Response
Proxy
Intelligence
Split Intelligence
Intelligence
Proxy-centric
Query
Query Response
Proxy
Intelligence
Push all the data or answer pull from proxy

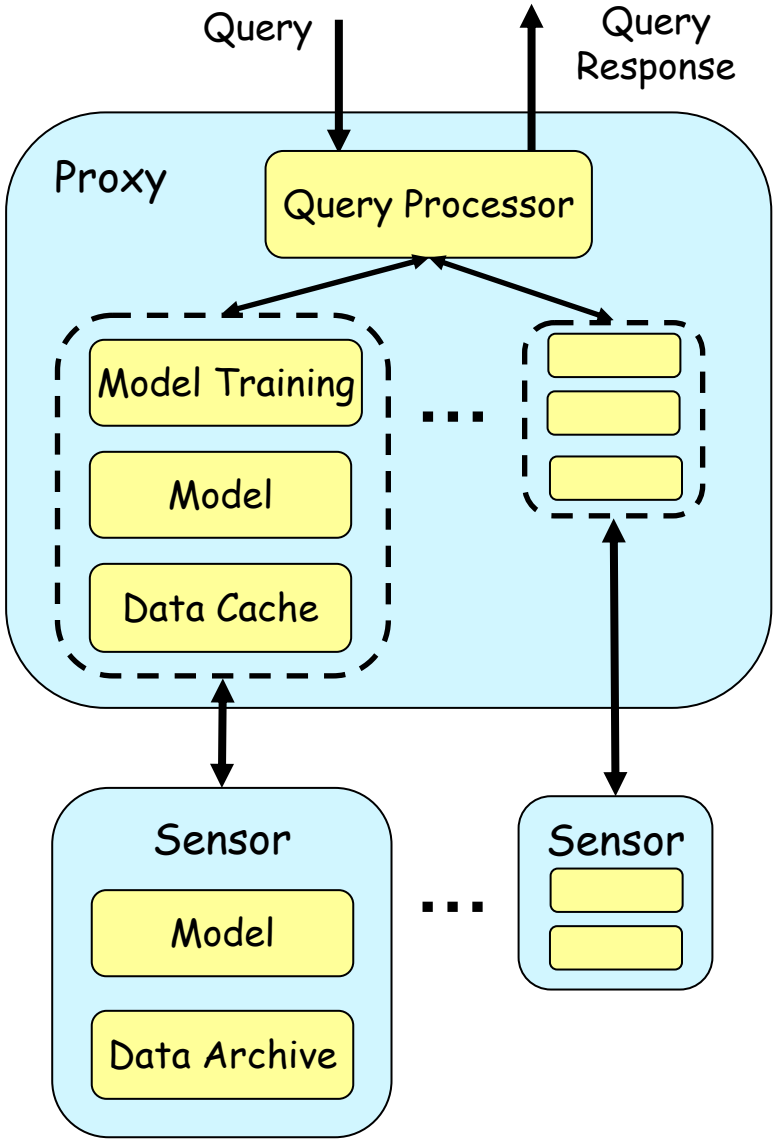
Key Ideas in PRESTO

- Split intelligence between the proxy which has greater resources, and the sensor which has ground truth.
- Exploit predictable structure in sensor data
- Use continual feedback from the proxy to sensors to adapt to data and query dynamics

Outline

- Motivation & Key Ideas
- **PRESTO Architecture**
- ARIMA Model
- Evaluation
- Summary & Future Work

PRESTO Architecture



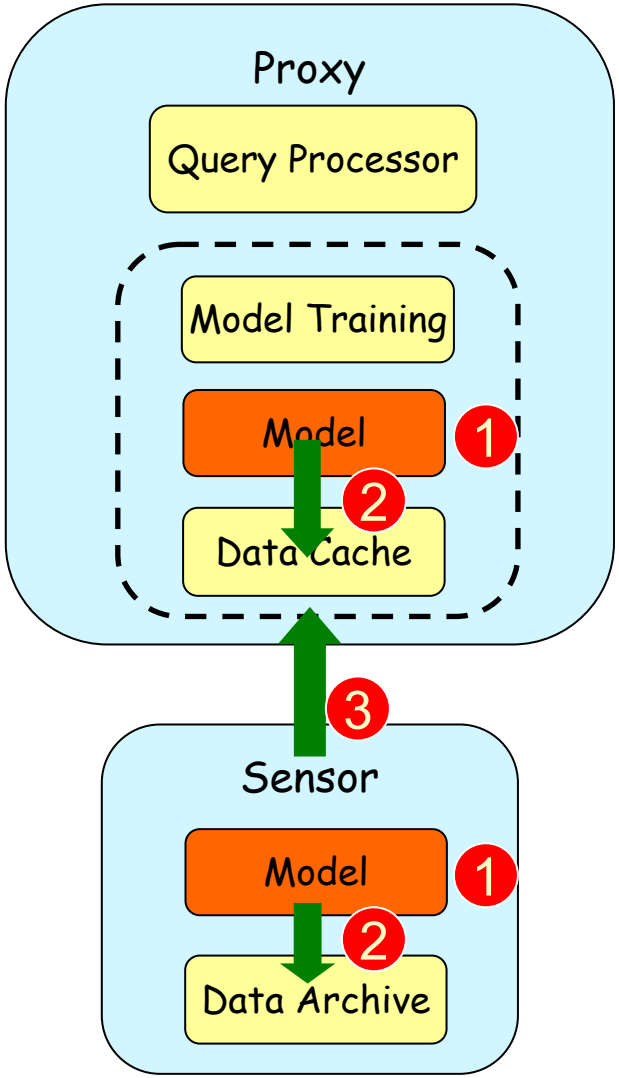
Query

Which sensor to query
Which time point to query
Confidence bound of the result

Key Components

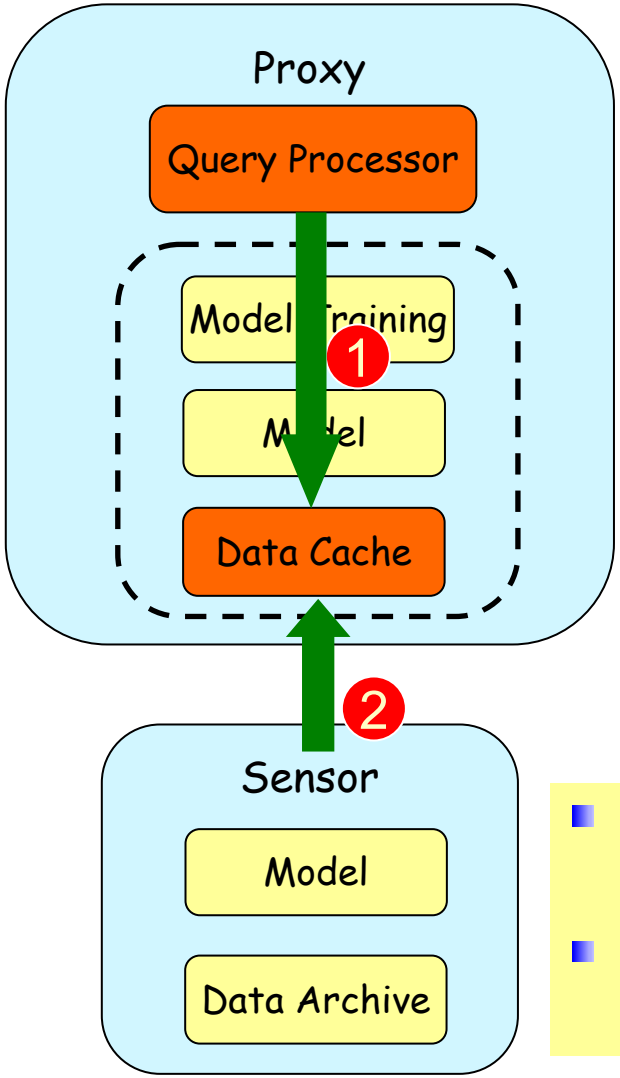
- Model-driven Push
- Model-driven Pull
- Cache Update
- Adaptive Feedback

Model-driven Push



- Capture the unpredictable data to guarantee the accuracy of the prediction.
- Procedure
  1. In each sampling period, sensor and proxy both predict current data
  2. Proxy and sensor cache the result. Proxy also caches the confidence of the prediction
  3. Sensor pushes data if difference between the prediction and the real data is higher than a threshold
- Only data anomalies are transmitted
- Bound the error of cache

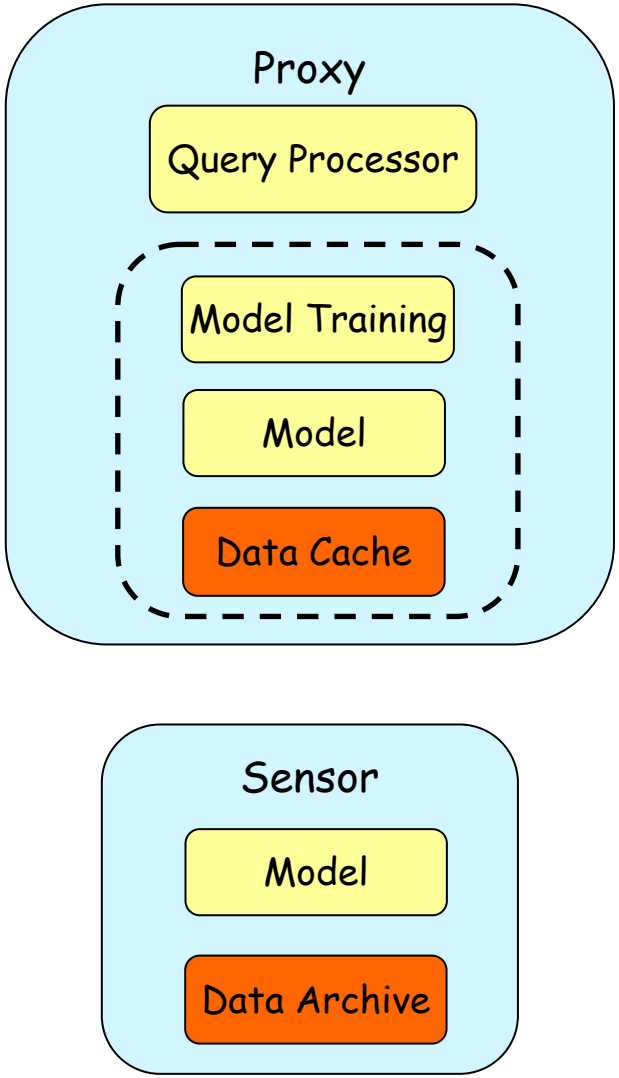
Model-driven Pull



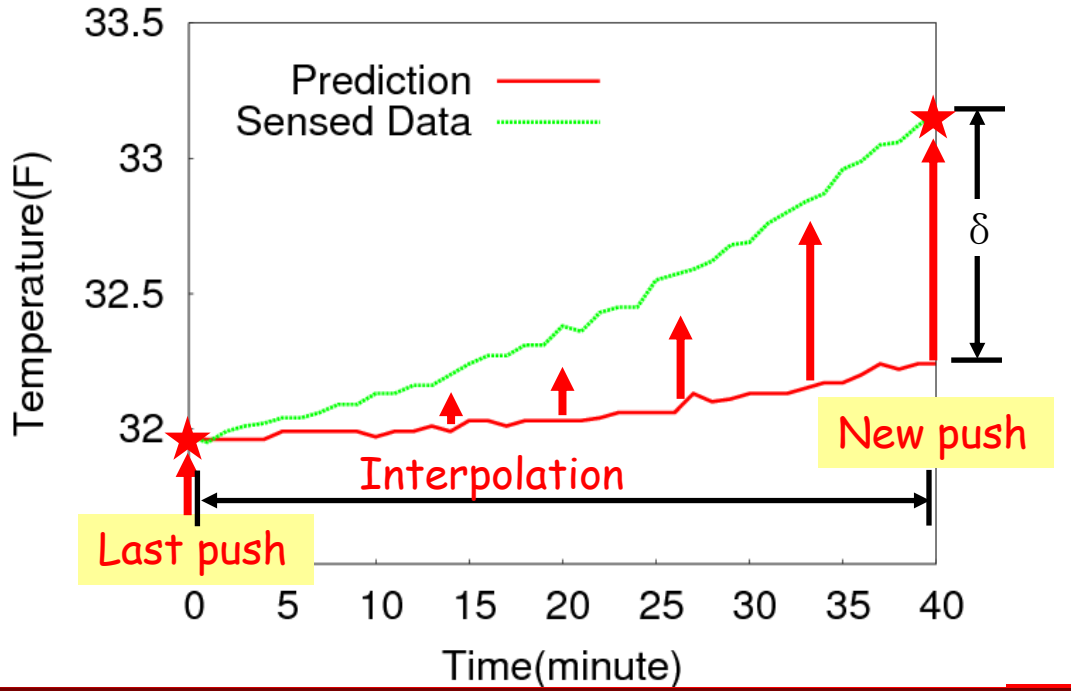
- Answer the queries which can not be satisfied by the cached data
- Procedure
  1. Proxy answers query using cached data if it is accurate enough
  2. Proxy pulls data from sensor if the cached data does not satisfy the query

- Ensure the accuracy of the query response
- Most of the queries can be answered using cached data.

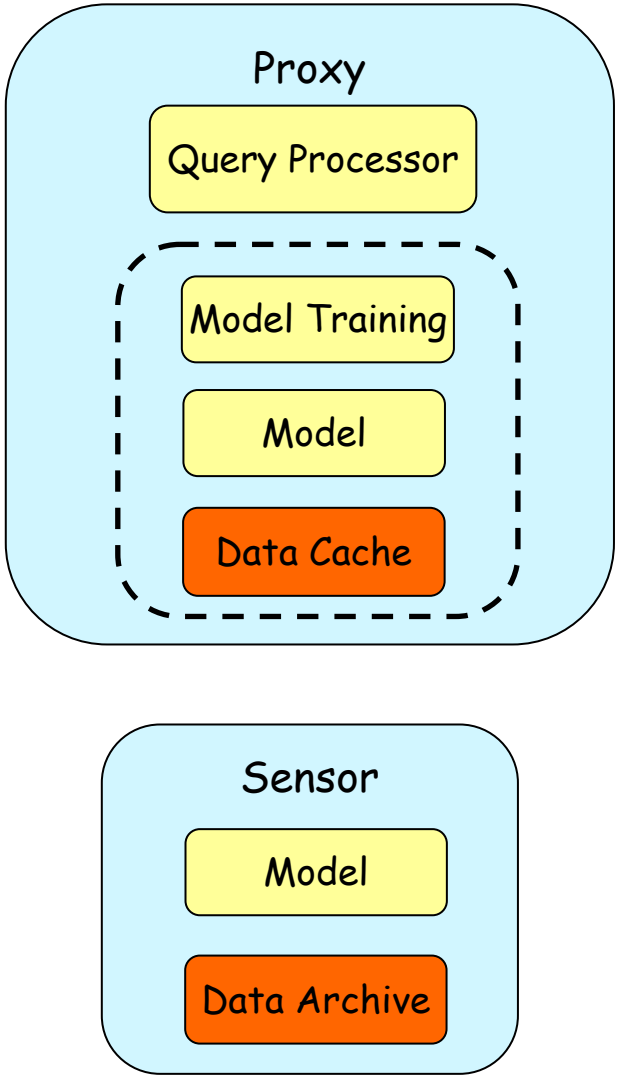
Cache Updating



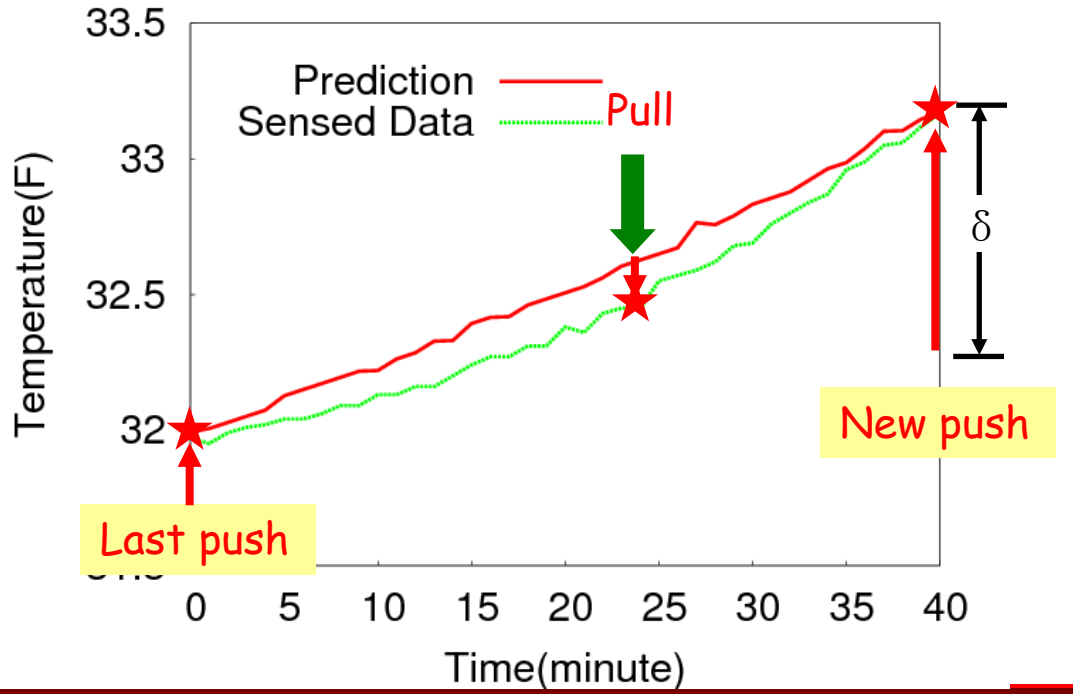
- Use any information from sensor to refine the accuracy of data cache
- Cache is refined after each push/pull
  - Interpolation
  - Re-prediction



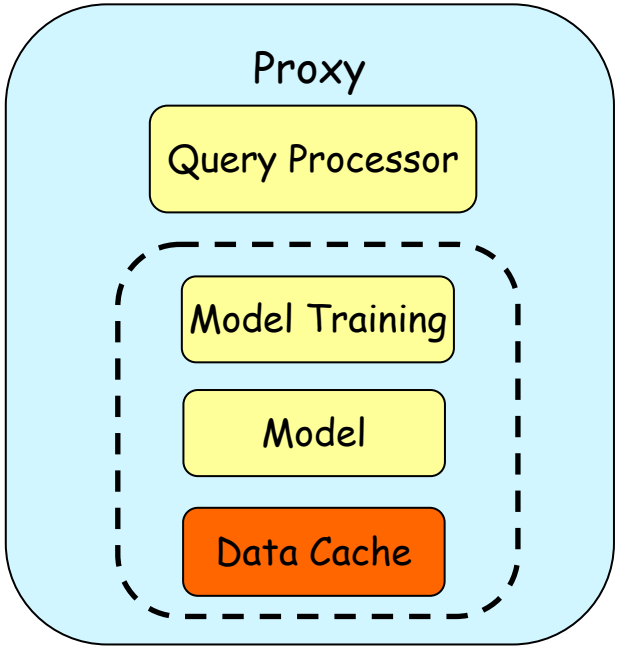
Cache Updating

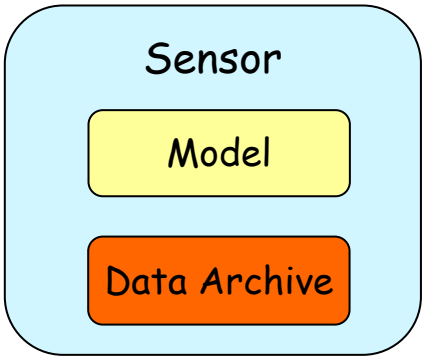


- Use any information from sensor to refine the accuracy of data cache
- Cache is refined after each push/pull
  - Interpolation
  - Re-prediction

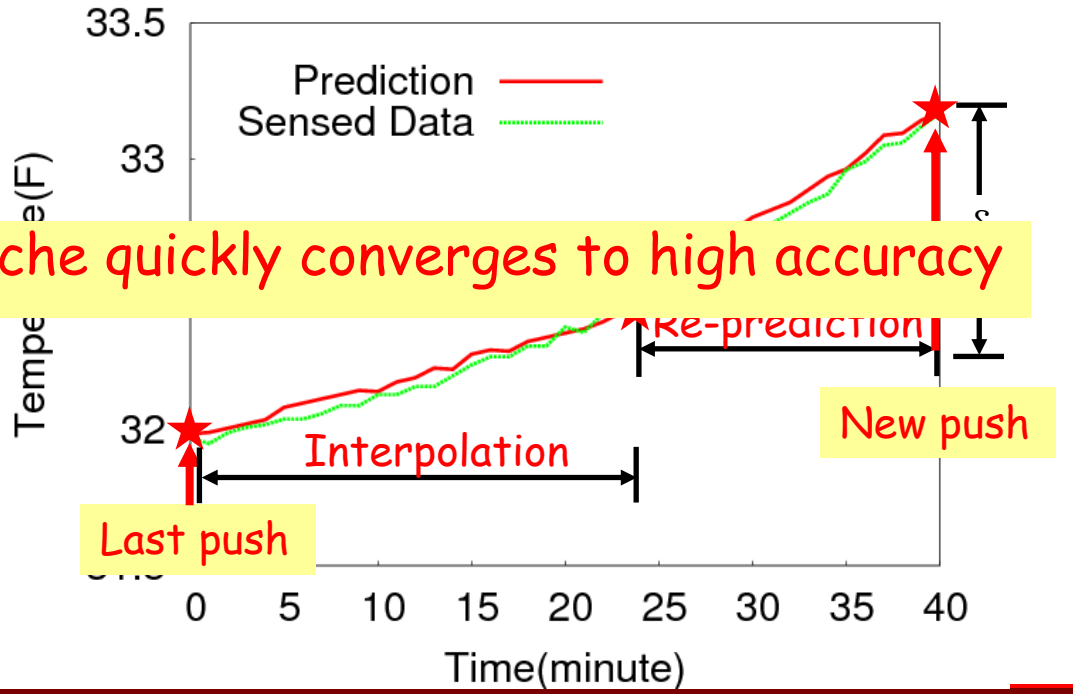


Cache Updating

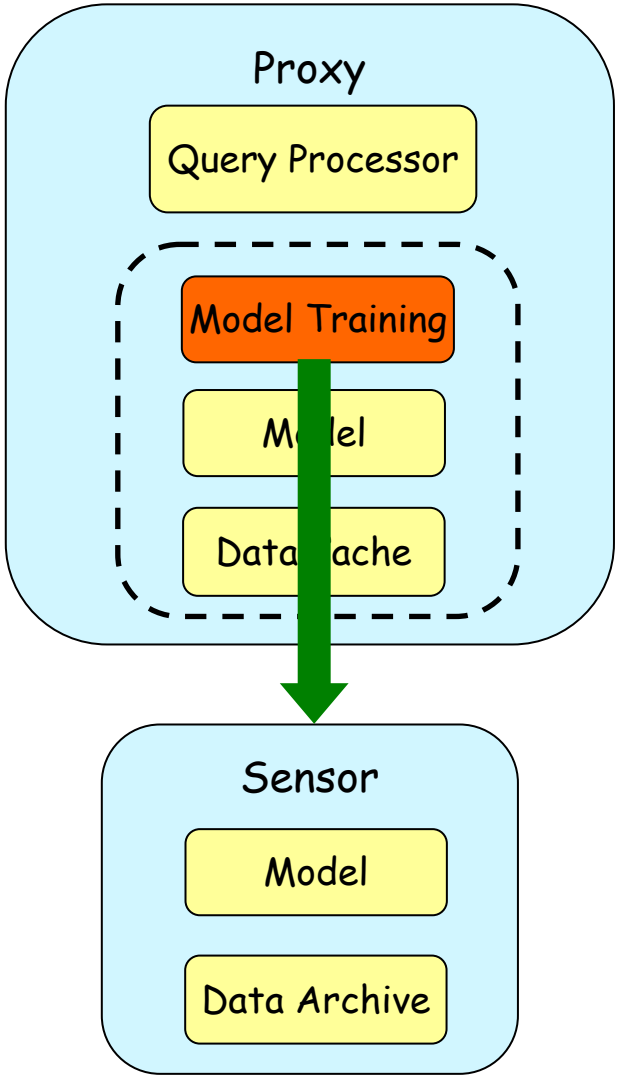




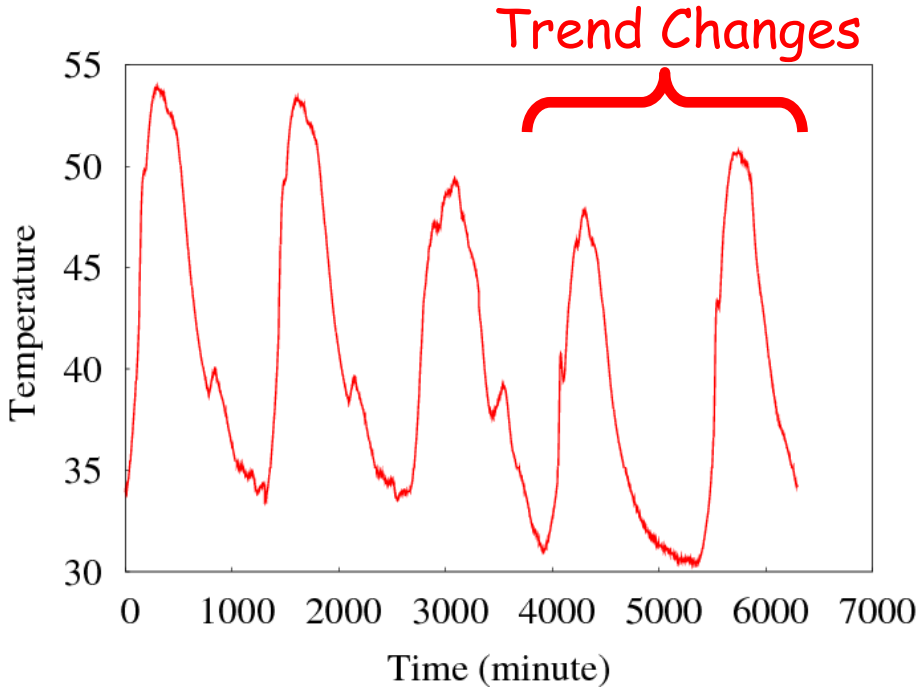
- Use any information from sensor to refine the accuracy of data cache
- Cache is refined after each push/pull
  - Interpolation
  - Re-prediction
- Cache quickly converges to high accuracy



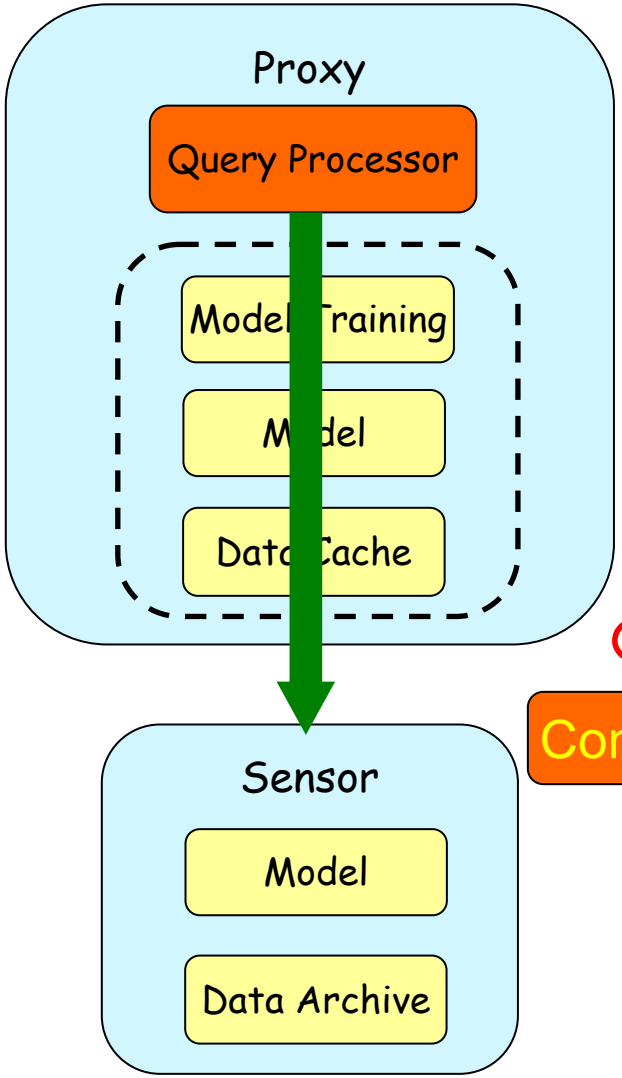
Adaptive Feedback



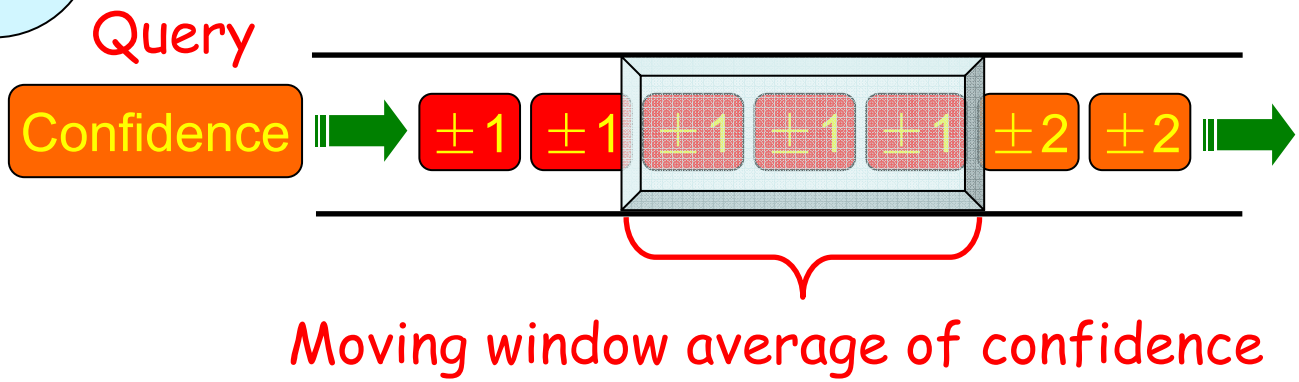
- Data dynamics
  - Periodically retrains and provides feedback of model parameters to sensor



Adaptive Feedback



- Data dynamics
  - Periodically retrains and provides feedback of model parameters to sensor
- Query dynamics
  - Computes and provides feedback of optimal δ to sensor

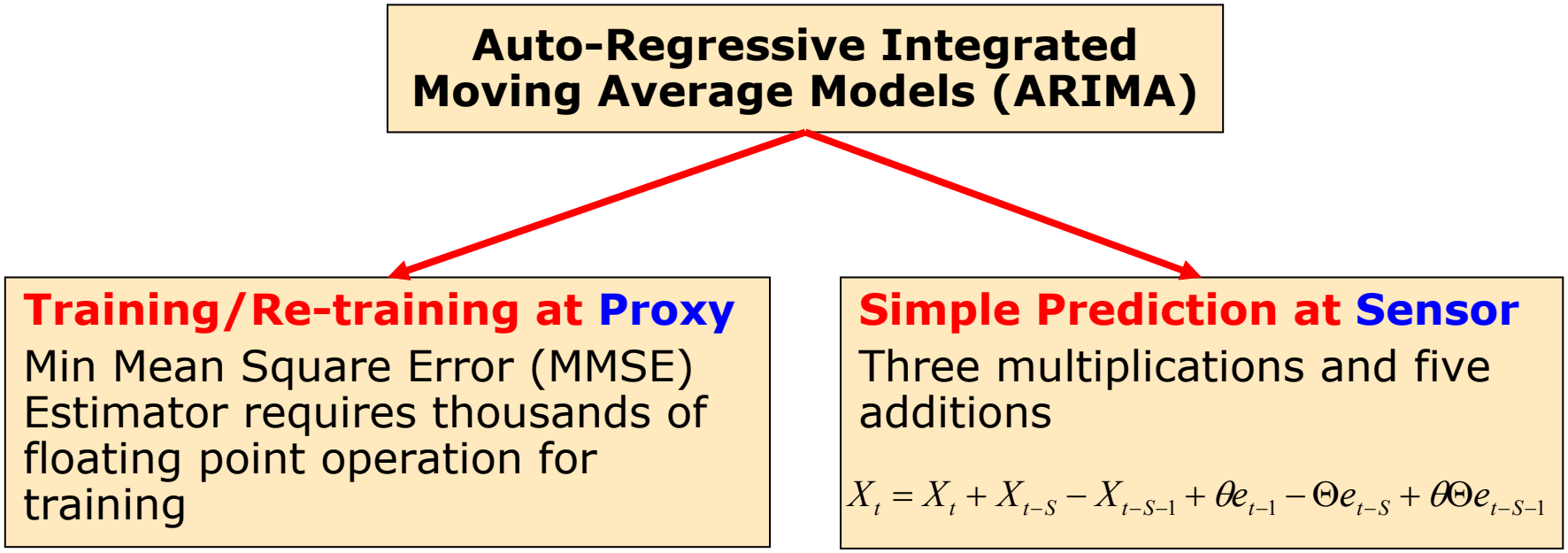


Outline

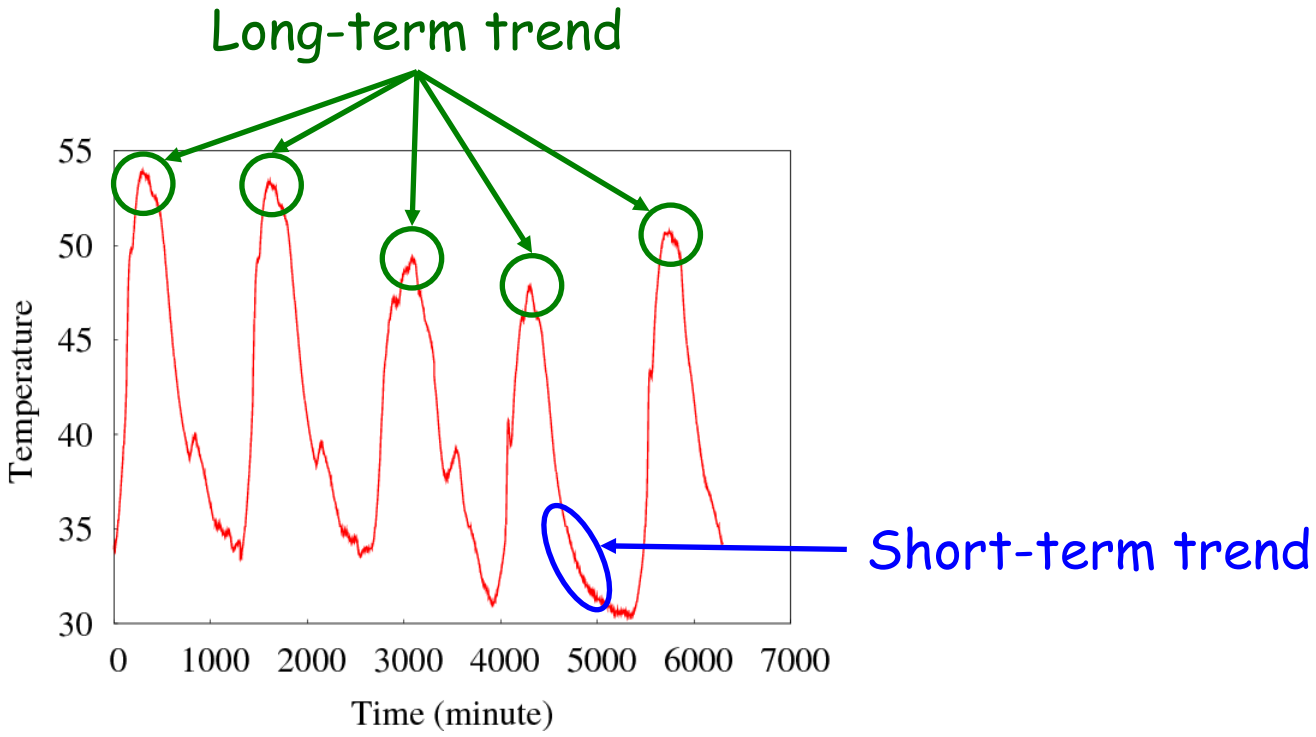
- Motivation & Key Ideas
- PRESTO Architecture
- **ARIMA Model**
- Evaluation
- Summary & Future Work

Split Computation Load

- Allocate computation between proxy and sensor according to their resources.



Capture Long-term and Short-term Trends



$$\Phi_P(B^S)\cdot\phi_p(B)\cdot(1-B^S)^D\cdot(1-B)^d\cdot X_t=\Theta_Q(B^S)\cdot\theta_q(B)\cdot e_t$$

Seasonal Parameters

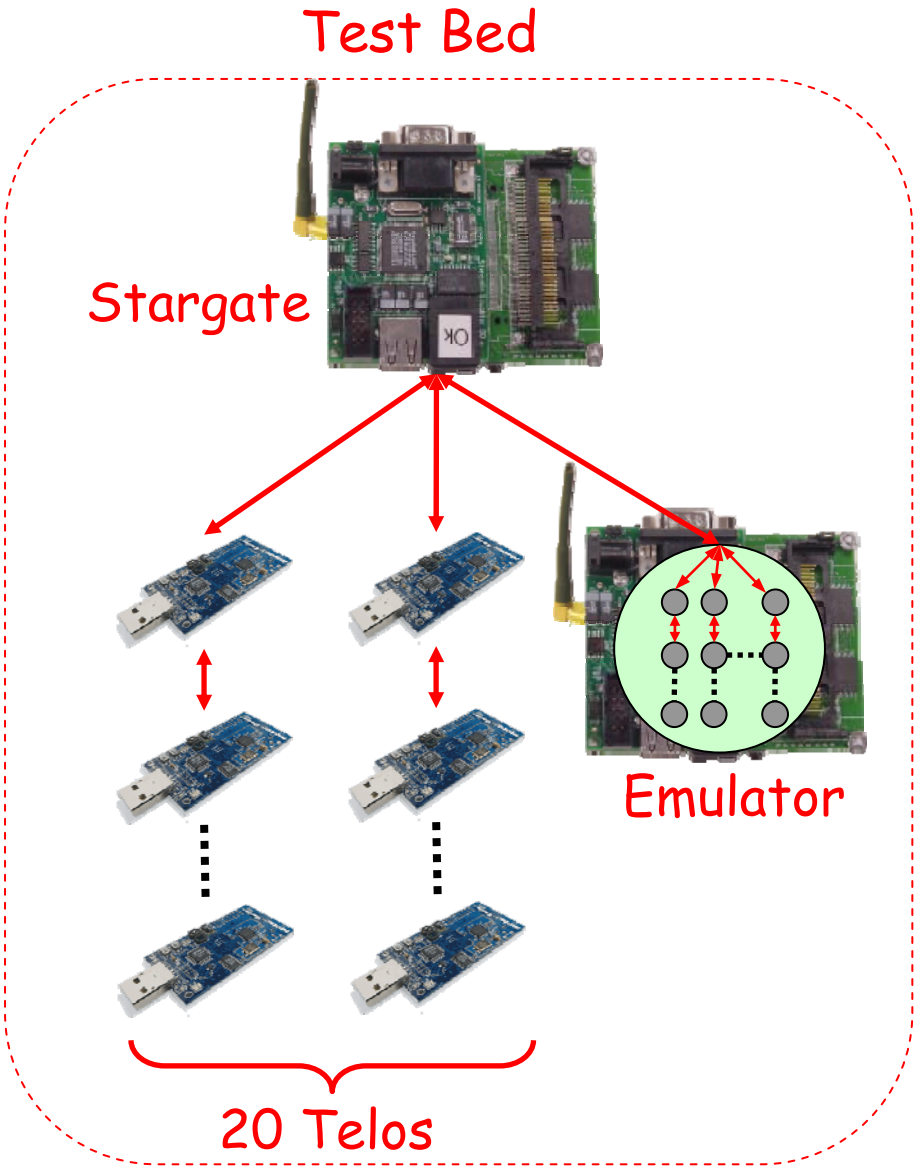
One Step Parameters

Outline

- Motivation & Key Ideas
- PRESTO Architecture
- ARIMA Model
- **Evaluation**
- Summary & Future Work

Evaluations

- Uses both numerical simulations and real deployments
- Datasets
  - One month of temperature readings from from James Reserve at CENS, UCLA.
  - Live temperature monitoring at Umass Amherst.
- Metrics
  - Number of messages
  - Query latency
  - Mean-square error



Sensor Micro-Benchmark

Breakdown of Energy Costs

Component	Operation	Energy (nJ)
Model	Predict 1 Sample	27
Data Archive	Read + Write + Erase 1 sample	21
Push/Pull	Transmit 1 sample + Receive 1 ACK	3300

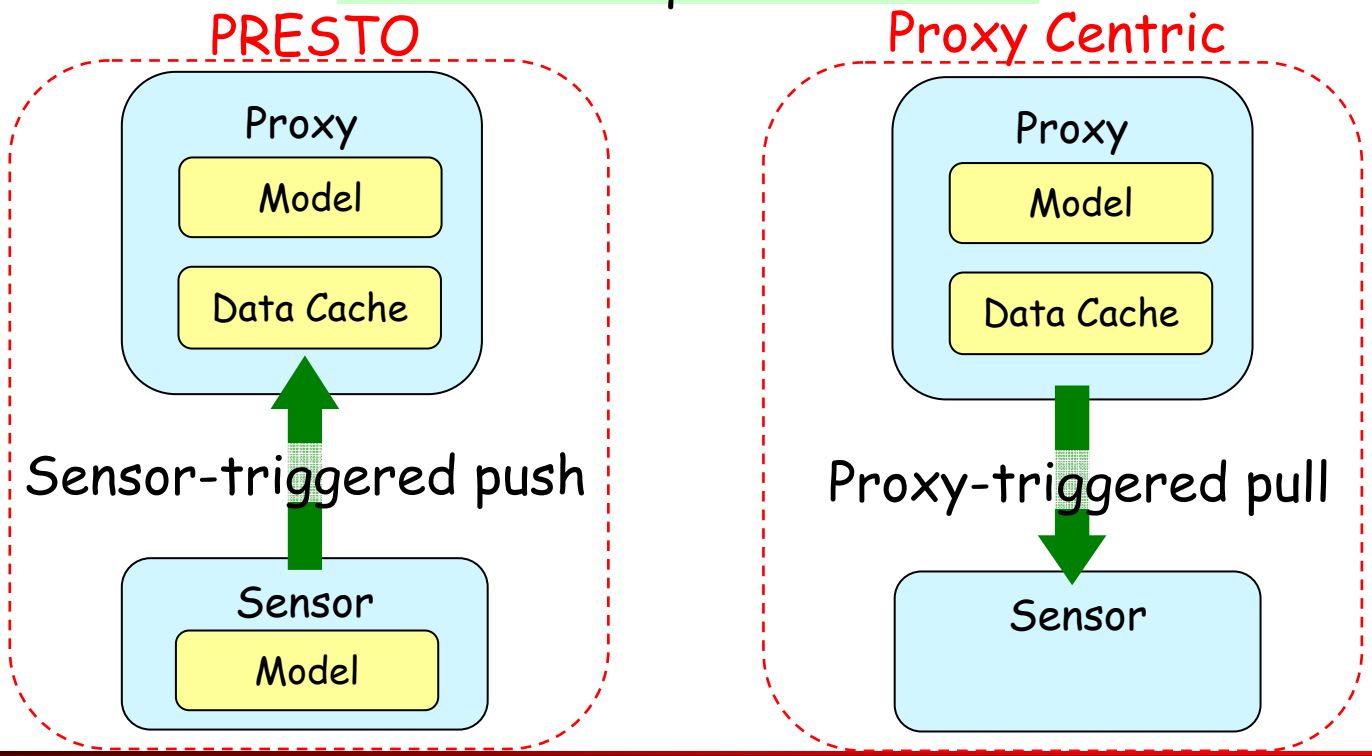
- PRESTO exploits modeling and archival to save the cost of pushes and pulls.
- Prediction operation at the sensor incurs low energy cost

PRESTO vs. Proxy Centric

- Compare energy efficiency and cache accuracy of PRESTO against proxy-centric approach

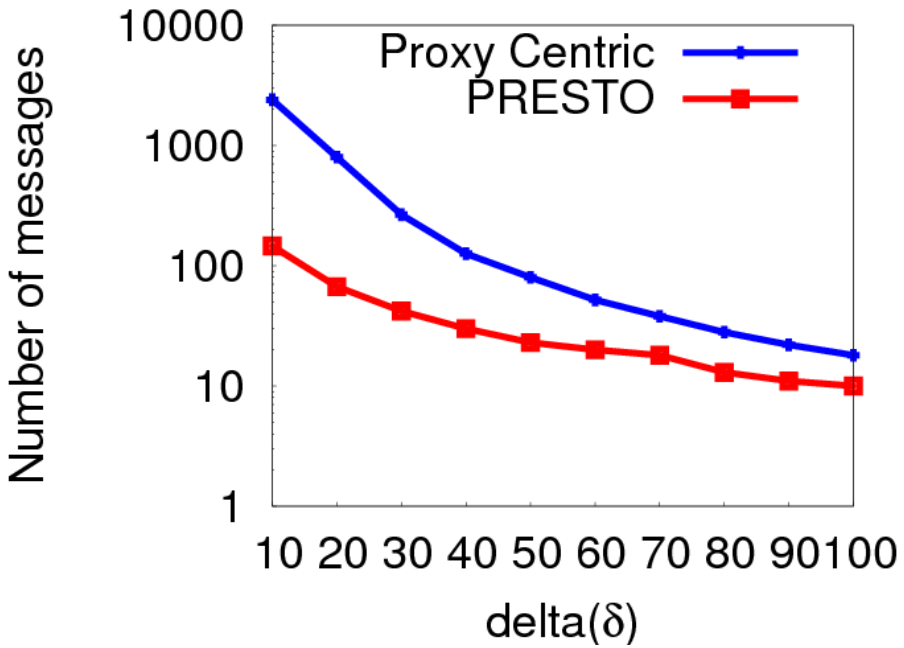
Setup

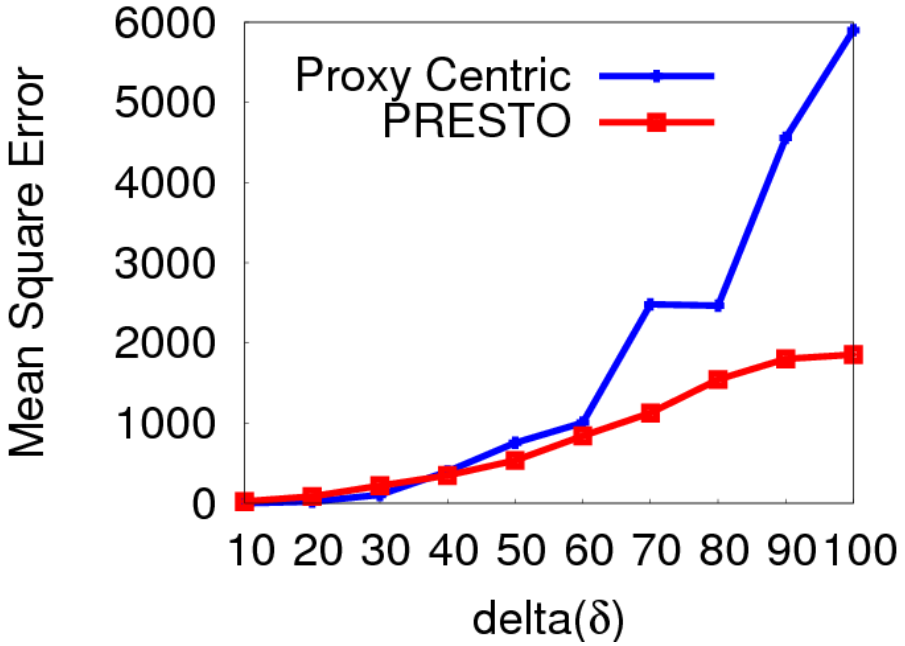
- Matlab simulations
- Real temperature trace
- Without queries



PRESTO vs. Proxy Centric

- Compare energy efficiency and cache accuracy of PRESTO against proxy-centric approach





- PRESTO is 2-20 times more energy efficient and up to 3 times more accurate

Scalability

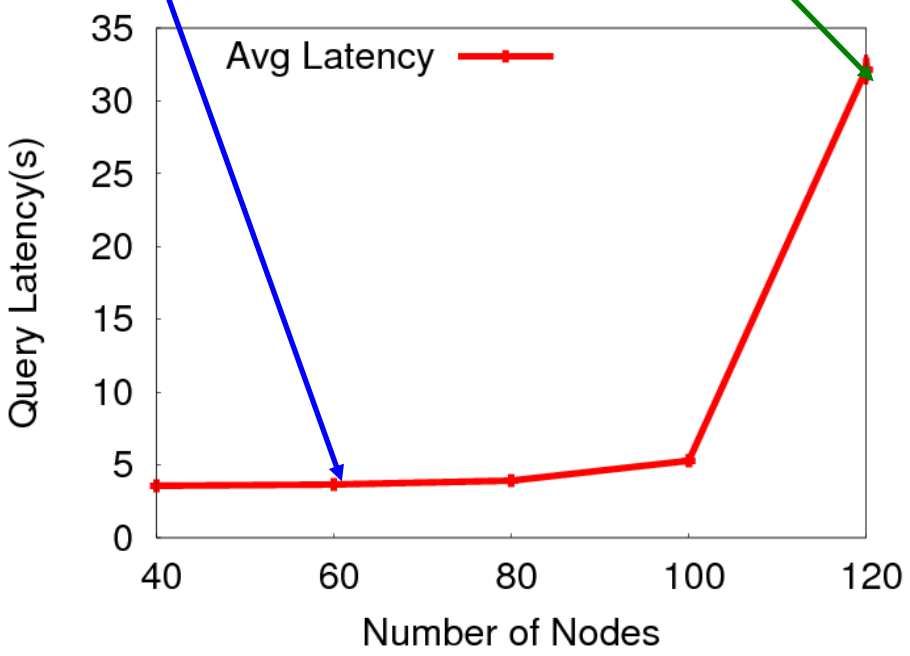
- How many sensors can a single proxy support at high query rate?

Congestion point of PRESTO

Congestion point of sensor-centric system

Setup

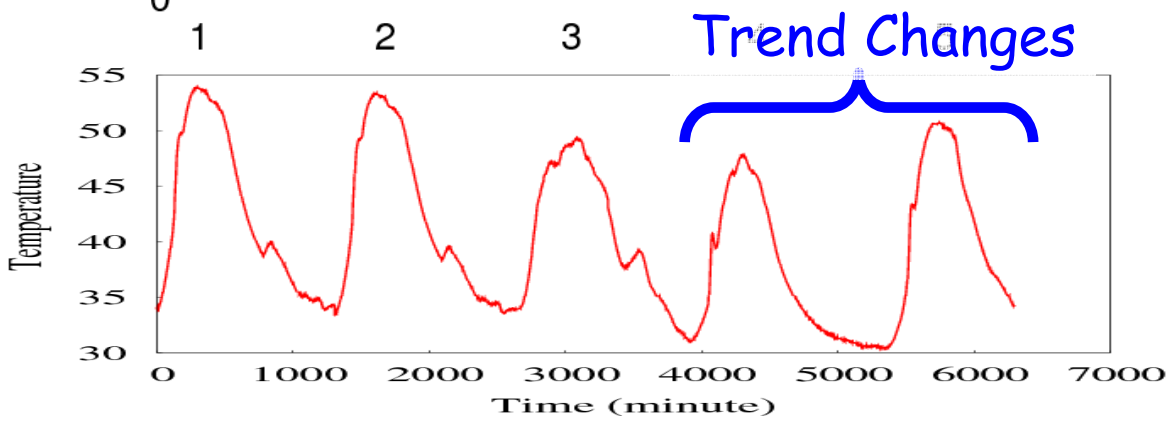
- Runs on test bed
- Real temperature trace
- Queries request data at each sensor as Poisson process at the rate of 1 query/node/min
- Network is duty-cycled and can support at most 1 query/node/sec



- One proxy supports 100% more sensors than a sensor-centric system

Adapt to Data Dynamics

- What the energy benefit of feedback-driven adaptation



Setup

- Runs on test bed
- Real temperature trace
- Feedback is sent in the end of each day

- Reduce communication by 30% compared to non-adaptive scheme

Summary

- PRESTO achieves low energy cost, low query latency and high accuracy by:
  - Splitting intelligence between the sensor and proxy using model-driven push and pull.
  - Using interpolation and prediction to maintain a precise proxy cache at low energy cost.
  - Feedback to update data and query model parameters.

Future Work

- Explore other models besides ARIMA that can be used to split complexity between the proxy and sensor.
- Extend to consider spatial correlation.
- Explore use of PRESTO in new sensor application domains.

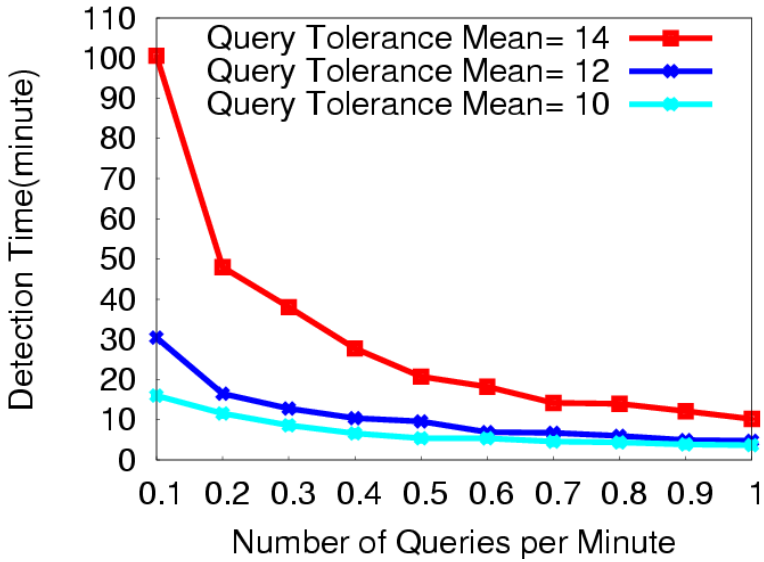
The End

Thank You!

Questions?

Failure Detection

- Detection latencies in different query tolerances and query frequencies
- Setup
  - Runs on one Tmote
  - Using James Reserve data set
  - Queries arrive as Poisson arrival



Longest detection latency less than 2 hours