Distributed Real-time Monitoring with Accuracy Objectives

Alberto Gonzalez Prieto

Laboratory for Communication Networks
KTH Royal Institute of Technology

Stockholm, Sweden

Outline

- Continuous Monitoring
- Accuracy in Monitoring
- A-GAP: a Distributed Solution
- Evaluation through Simulation
- Discussion

Continuous Monitoring

- Provide **continuous estimates** (in time) of management variables (aggregates)
- Aggregates
  - Network-wide variables computed from device variables accross the network
  - **SUM**, AVERAGE, MAX, MIN, etc
  - Examples
    - Number of VoIP flows in the domain
    - Most loaded link(s)

Accuracy

- Monitoring an aggregate without error
  - Requires negligable network delays and processing delays
  - Pull approach
    - Poll at Nyquist frequency
  - Push approach
    - Report local variables changes as they occur
  - **Unfeasible** in **large networks**
    - Large management overhead
      - Processing load
      - Traffic
    - Non-scalable
- Fundamental Trade-off: accuracy vs overhead

System Functionality

- Problem statement
  - Minimize the maximum load over all nodes
    for a configurable average (absolute) error of the estimation of the global aggregate

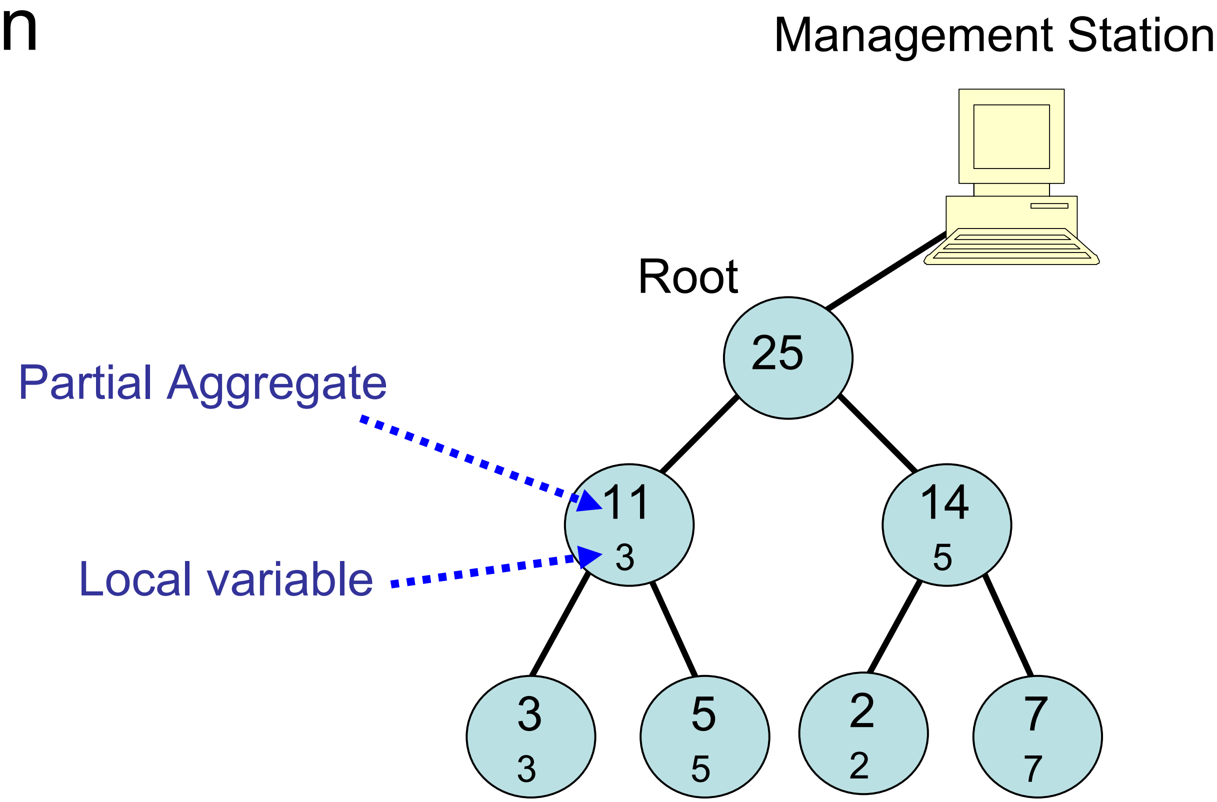
- Examples
  - Estimate number of ongoing voice conversations in the domain with an average error below 5 %
  - Estimate P2P traffic volume in the domain with an average error below 10 %

Related Work

- Olston et al.:
  - First to discuss the accuracy vs overhead **trade-off** (ACM Sigmod 2001)
  - **Centralized** approach (Sigmod 2001 & 2003, IEEE Data Engineering Bulletin 2005)
- Deligiannakis et al.:
  - **Distributed** approach: makes use of an aggregation tree (EDBT 2004)
- In both works
  - Accuracy expressed in terms of **maximum error**
  - Overhead metric: **overall management traffic**
  - Proposals are based on **filters**
  - Use globally **synchronized** rounds of operation

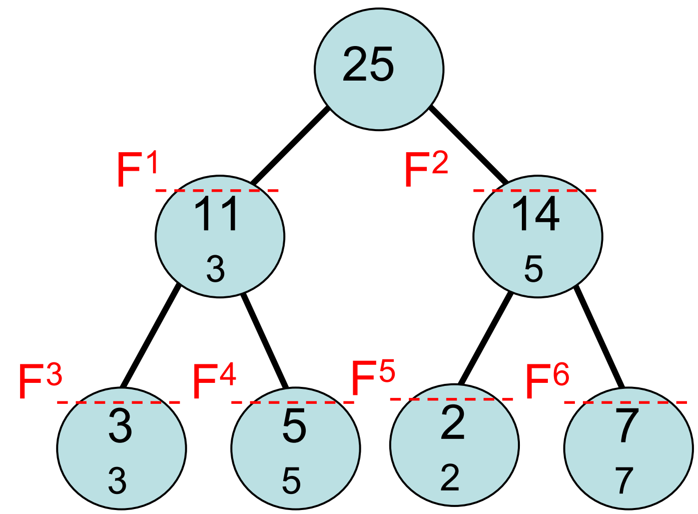
GAP: Generic Aggregation Protocol

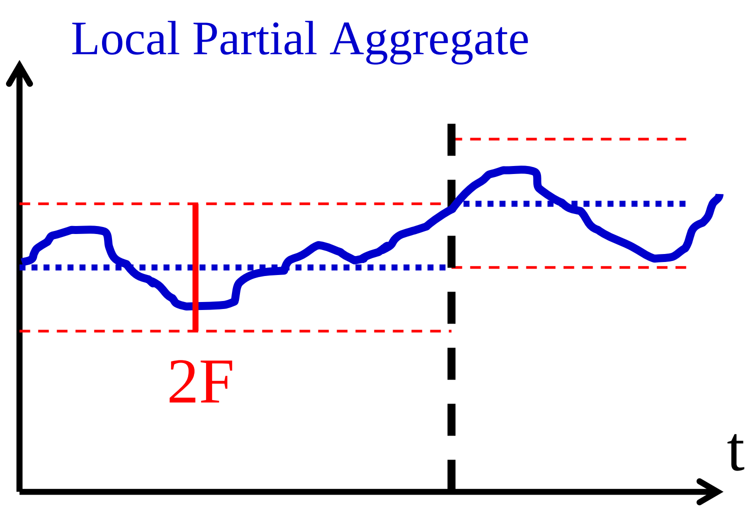
- Creates and maintains and **aggregation tree**
  - Spanning tree
- Event-driven



Filter-based Approaches

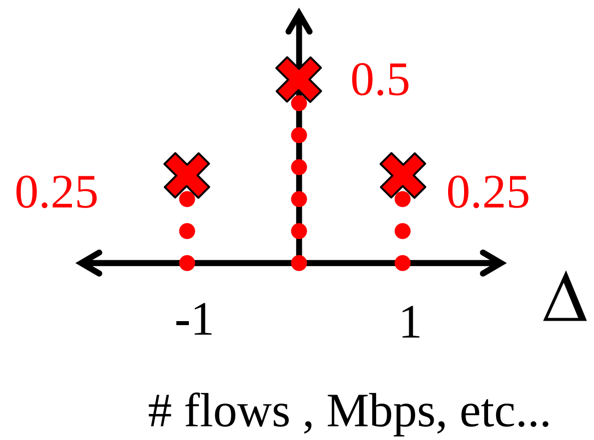
- Each node (n) in the aggregation tree has a configurable filter
- Filter functionality:
  - Filter out updates
  - Filter width: F^n
  - On arrival of an update from a child
    - Calculate new partial aggregated value
    - If variation > F^n
      - Send update to parent in the tree
  - F^n determines:
    - Accuracy
    - Generated overhead
- F^n : decision variables





A-GAP: Stochastic Model

- Modeling the Dynamics of Local Variables
  - **discretized** in time and value
  - one-dimensional **random walk** model
    - variation over time: sequence of steps of different sizes: sspdf(Δ).
    - Steps are independent
    - Applied in similar contexts before
    - Facilitates an algorithmic solution



A-GAP: Stochastic Model (2)

- Modeling Aggregation Trees
  - Model to relate
    - Filter widths
    - Management overhead (load)
    - Estimation error
  - Assumes statistical independence among local variables
- Lack of a closed form (for now)
  - However, we can evaluate candidate solutions

A-GAP: Stochastic Model (3)
Modeling a Node

(1) $$T_{ij}^{n} = \begin{cases} sspdf_{in}^{n}(j-i) & \text{for } j\neq 0 \\ sspdf_{in}^{n}(-i) + \sum_{k=-\infty}^{-(F^{n}+i+1)} sspdf_{in}^{n}(k) + \sum_{k=F^{n}+i+1}^{\infty} sspdf_{in}^{n}(k) & \text{for } j=0 \end{cases}$$

(2) $$sspdf_{out}^{n}(\Delta) = \begin{cases} sspdf_{in}^{n}(\Delta) * \pi^{n}(\Delta) & \text{for } |\Delta|>F \\ \sum_{\Delta=-F^{n}}^{F^{n}} sspdf_{in}^{n}(\Delta) * \pi^{n}(\Delta) & \text{for } \Delta=0 \\ 0 & \text{for any other } \Delta \end{cases}$$

(3) $epdf^{n}_{out}(\Delta) = epdf^{n}_{in}(\Delta) * \pi^{n}(\Delta)$

(4) $\lambda^{n} = nc^{n}\,(1 - sspdf^{n}_{out}(0))$

* stands for convolution

A-GAP: Stochastic Model (4)
Modeling the Aggregation Tree

(5) $epfd^{n}_{in}(\Delta) = epdf^{c1}_{out}(\Delta) * epdf^{c2}_{out}(\Delta) * epdf^{c3}_{out}(\Delta) * \ldots$

(6) $sspdf^{n}_{in}(\Delta) = \sum_{c \in children} sspdf^{c}_{out}(\Delta) nc^{c} \Big/ \sum_{c \in children} nc^{c}$

A-GAP: Design Principles

- Decentralized and Asynchronous
  - Scalability and robustness
- Based on **GAP**
  - Spaning tree
  - Incremental aggregation
  - All nodes execute same code
- **Heuristic**
  - Global problem mapped into local problems
    - Minimize processing load for a given accuracy regarding its partial aggregate
    - Mechanism: periodically re-compute filters of children

A-GAP: Pseudo-code

F1:

```
every τ seconds
  if backoff flag is clear
        request statistics from children (sspdf_in, epdf_in, nc)
        select η children in a round-robin fashion
        compute new filters for selected children
        compute new accuracy constraints for children (obj_i=α·obj_n·e_ratio)
        send new filters and constraints to children
        compute local statistics (sspdf_in, epdf_in, nc)
  if backoff flag checked
        send backoff message to children
```

A-GAP: Pseudo-code

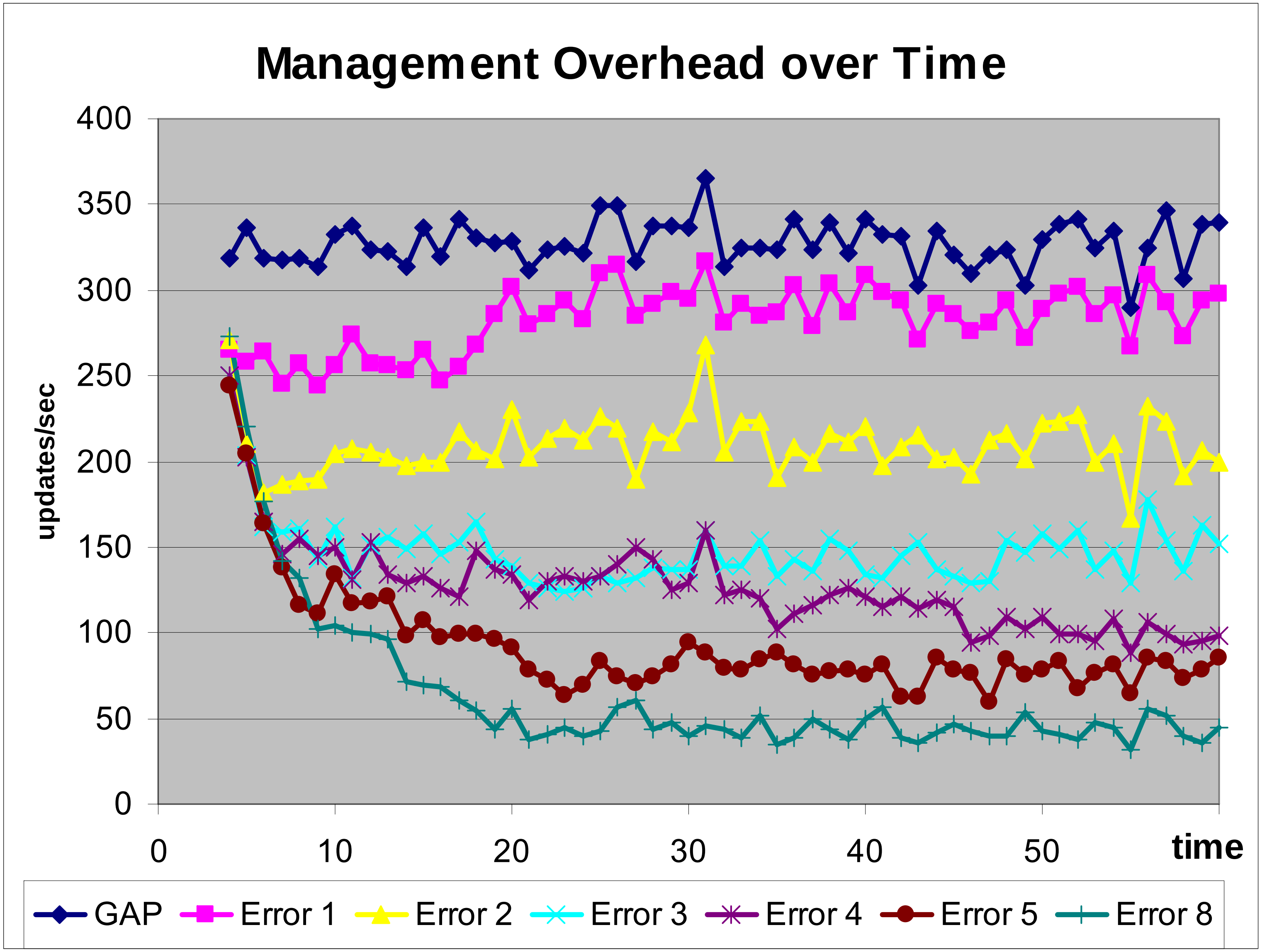
F2:

```
on message received from parent
if message type is new filter and accuracy constraint
    set new filter and accuracy constraint
    if new filter width < current filter width
        set backoff flag
    else
        clear backoff flag
if message type is backoff
    restore filter from the last cycle
    set backoff flag
```

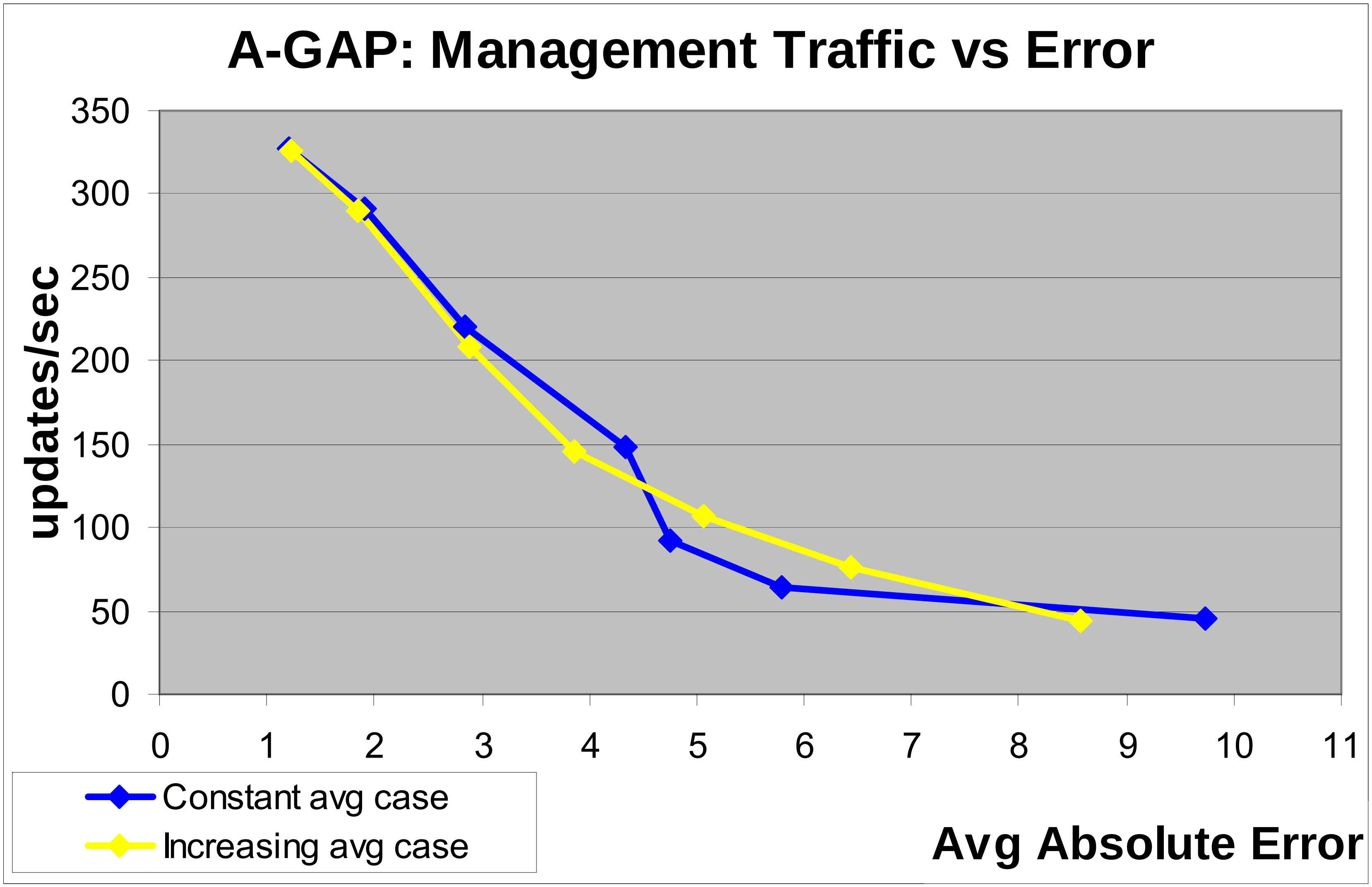
Evaluation through Simulation

- Scenario
  - Abovenet
    - 654 nodes
    - 1332 links
  - Filters re-computed once per second asynchronously
  - Scenario 1: Constant Average Aggregate
    - 50% sources have a step size of 1
    - 50% sources have a step size of 2
  - Scenario 2: Increasing Average Aggregate
    - Value of the sum increases linearly with time
      - ~60/ sec

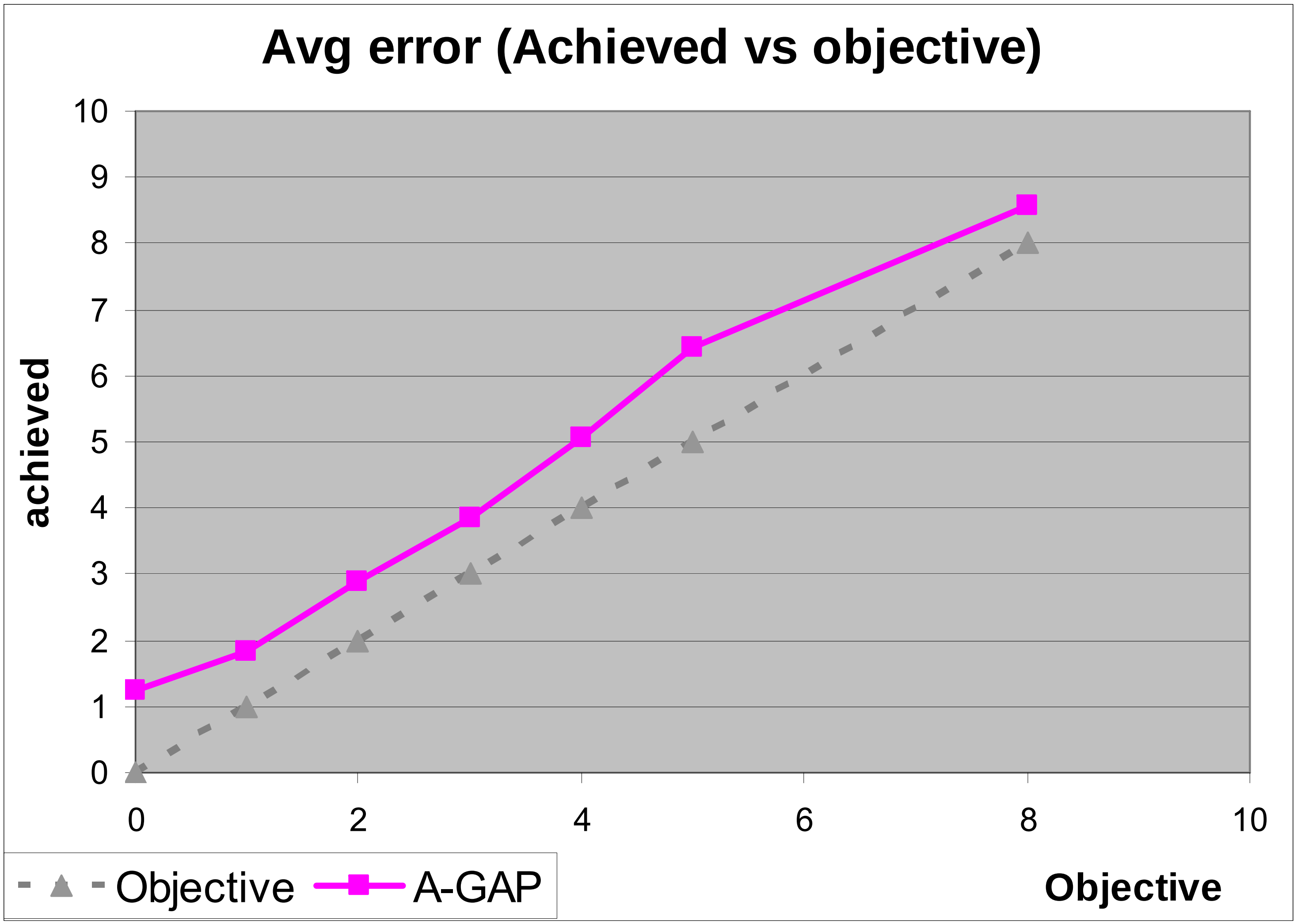
Evolution over Time



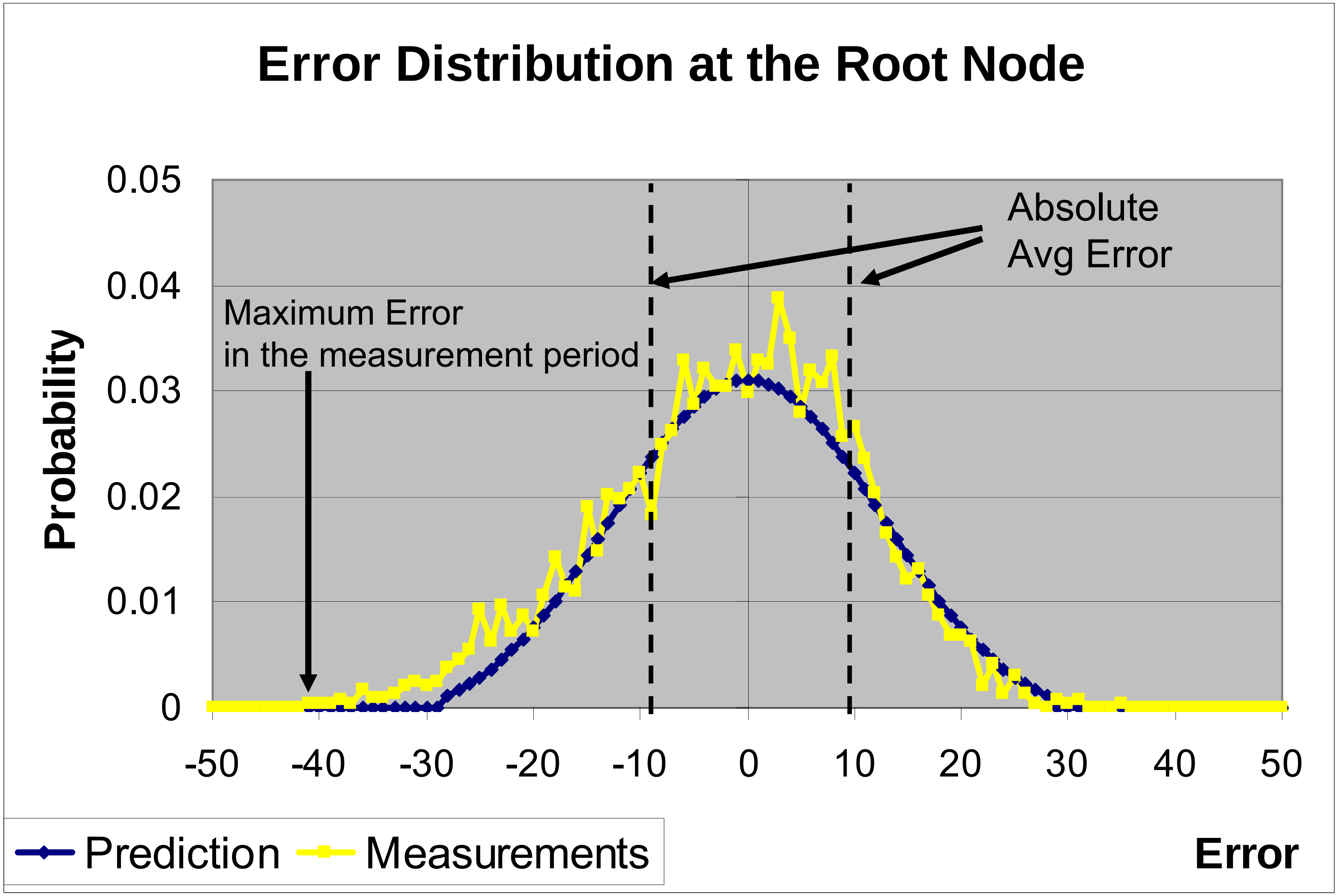
Trade-off


A-GAP: Management Traffic vs Error
updates/sec
350
300
250
200
150
100
50
0
0
1
2
3
4
5
6
7
8
9
10
11
Constant avg case
Increasing avg case
Avg Absolute Error

Achieving the Objective

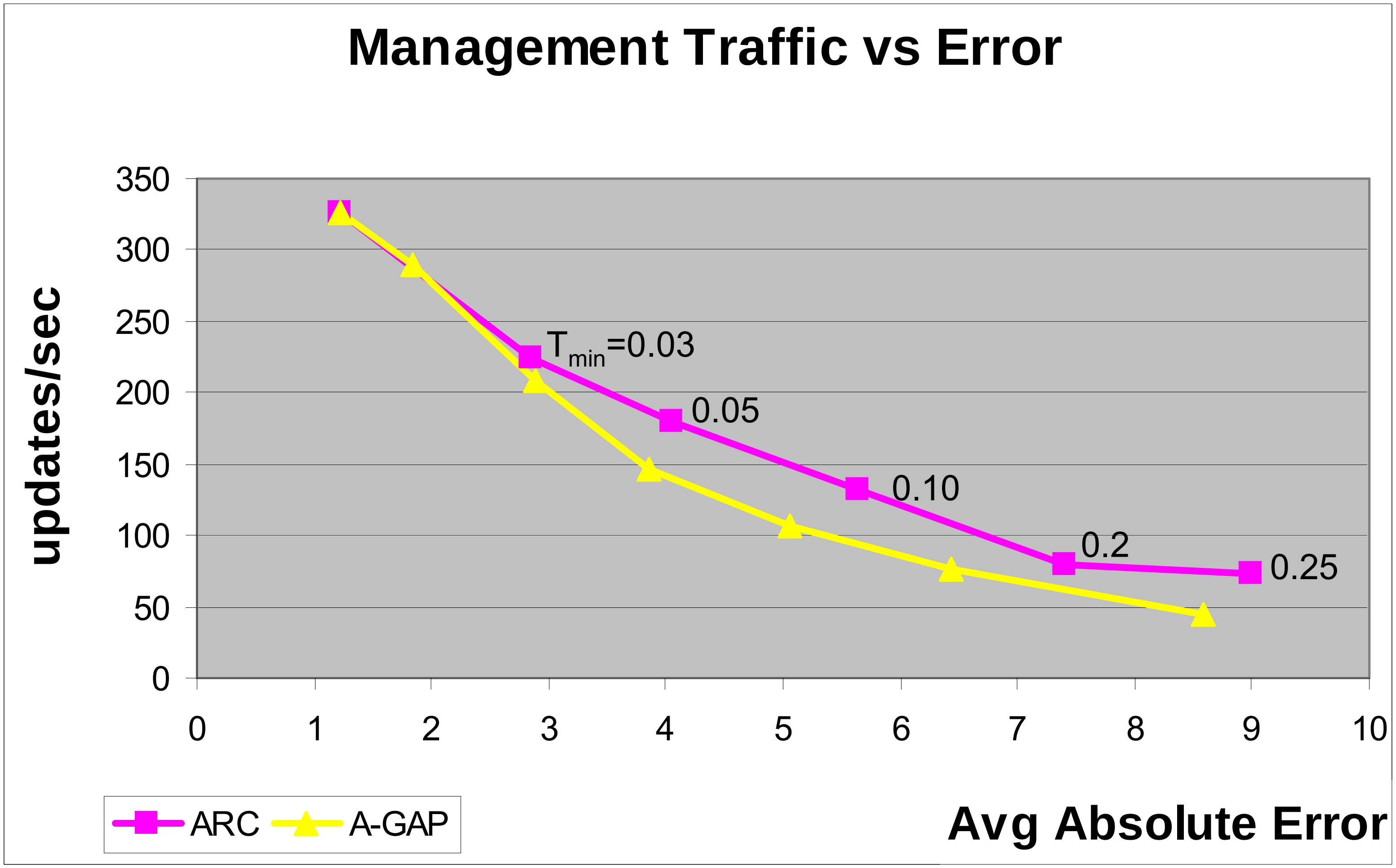


Distribution of the Estimation Error



- Very long tail: maximum error is 180

Benchmarking against a Rate-control Scheme



- Determination of the operating point is difficult for a rate-control scheme

Computational Cost of A-GAP

- Can be qualitatively controlled
  - Recomputation interval
  - Number of filters re-computed per interval

Discussion

- Goal of A-GAP
  - Accuracy objectives in terms of average absolute error (vs max)
- Characteristics of A-GAP
  - Decentralized, asynchronous
  - Heuristic: global problem mapped into local problems
- Trade-off (achievements)
  - Can be controlled
  - Significant overhead reductions: 85% for objective of 8 units
- A-GAP enables Perfomance Prediction at the MS
  - Computation of trade-off curve at run-time (overload avoidance)
  - Achieve estimation of error distribution

Future Work

- Real-time estimation of random walk model parameters
- Implementation of A-GAP on testbed