Load Shedding in Network Monitoring Applications

Pere Barlet-Ros¹  Gianluca Iannaccone²
Josep Sanjuàs¹  Diego Amores¹  Josep Solé-Pareta¹

¹Technical University of Catalonia (UPC)
Barcelona, Spain

²Intel Research
Berkeley, CA

Intel Research Berkeley, July 26, 2007
Outline

1. Introduction
   - Motivation
   - Case Study: Intel CoMo

2. Load Shedding
   - Prediction Method
   - System Overview

3. Evaluation and Operational Results
   - Performance Results
   - Accuracy Results
Motivation

- Building robust network monitoring applications is hard
  - Unpredictable nature of network traffic
  - Anomalous traffic, extreme data mixes, highly variable data rates

- Processing requirements have greatly increased in recent years
  - E.g., intrusion and anomaly detection
Building robust network monitoring applications is hard
- Unpredictable nature of network traffic
- Anomalous traffic, extreme data mixes, highly variable data rates

Processing requirements have greatly increased in recent years
- E.g., intrusion and anomaly detection

The problem
- Efficiently handling extreme overload situations
- Over-provisioning is not possible
Case Study: Intel CoMo

CoMo (Continuous Monitoring)¹
  - Open-source passive monitoring system
  - Fast implementation and deployment of monitoring applications

Traffic queries are defined as *plug-in* modules written in C
  - Contain complex stateful computations

¹ http://como.sourceforge.net
Case Study: Intel CoMo

CoMo (Continuous Monitoring)¹
- Open-source passive monitoring system
- Fast implementation and deployment of monitoring applications

Traffic queries are defined as *plug-in* modules written in C
- Contain complex stateful computations

Traffic queries are **black boxes**
- Arbitrary computations and data structures
- Load shedding cannot use knowledge about the queries

¹http://como.sourceforge.net
Load Shedding Approach

Main idea

1. Find correlation between traffic features and CPU usage
   - Features are query agnostic with deterministic worst case cost
2. Leverage correlation to predict CPU load
3. Use prediction to guide the load shedding procedure
Load Shedding Approach

Main idea

1. Find correlation between traffic features and CPU usage
   - Features are query agnostic with deterministic worst case cost
2. Leverage correlation to predict CPU load
3. Use prediction to guide the load shedding procedure

Novelty: No *a priori* knowledge of the queries is needed

- Preserves high degree of flexibility
- Increases possible applications and network scenarios
Traffic Features vs CPU Usage

Figure: CPU usage compared to the number of packets, bytes and flows
System Overview

**Figure:** Prediction and Load Shedding Subsystem
Load Shedding Performance

Figure: Stacked CPU usage (Predictive Load Shedding)
Load Shedding Performance

![CDF of the CPU usage per batch](image)

**Figure:** CDF of the CPU usage per batch
Queries estimate their unsampled output by multiplying their results by the inverse of the sampling rate.

Errors in the query results ($mean \pm stdev$)

<table>
<thead>
<tr>
<th>Query</th>
<th>original</th>
<th>reactive</th>
<th>predictive</th>
</tr>
</thead>
<tbody>
<tr>
<td>application (pkts)</td>
<td>55.38% ±11.80</td>
<td>10.61% ±7.78</td>
<td>1.03% ±0.65</td>
</tr>
<tr>
<td>application (bytes)</td>
<td>55.39% ±11.80</td>
<td>11.90% ±8.22</td>
<td>1.17% ±0.76</td>
</tr>
<tr>
<td>flows</td>
<td>38.48% ±902.13</td>
<td>12.46% ±7.28</td>
<td>2.88% ±3.34</td>
</tr>
<tr>
<td>high-watermark</td>
<td>8.68% ±8.13</td>
<td>8.94% ±9.46</td>
<td>2.19% ±2.30</td>
</tr>
<tr>
<td>link-count (pkts)</td>
<td>55.03% ±11.45</td>
<td>9.71% ±8.41</td>
<td>0.54% ±0.50</td>
</tr>
<tr>
<td>link-count (bytes)</td>
<td>55.06% ±11.45</td>
<td>10.24% ±8.39</td>
<td>0.66% ±0.60</td>
</tr>
<tr>
<td>top destinations</td>
<td>21.63 ±31.94</td>
<td>41.86 ±44.64</td>
<td>1.41 ±3.32</td>
</tr>
</tbody>
</table>
Ongoing and Future Work

- **Ongoing Work**
  - Query utility functions
  - Custom load shedding
  - Fairness of service with non-cooperative users
  - Scheduling CPU access vs. packet stream

- **Future Work**
  - Distributed load shedding
  - Other system resources (memory, disk bandwidth, storage space)
Availability

- The source code of our load shedding prototype is publicly available at http://loadshedding.ccaba.upc.edu

- The CoMo monitoring system is available at http://como.sourceforge.net

Acknowledgments

- This work was funded by a University Research Grant awarded by the Intel Research Council and the Spanish Ministry of Education under contract TEC2005-08051-C03-01

- Authors would also like to thank the Supercomputing Center of Catalonia (CESCA) for giving them access to the Catalan RREN
Work Hypothesis

**Our thesis**
- Cost of maintaining data structures needed to execute a query can be modeled looking at a set of traffic features

**Empirical observation**
- Different overhead when performing basic operations on the state while processing incoming traffic
  - E.g., creating or updating entries, looking for a valid match, etc.
- Cost of a query is mostly dominated by the overhead of some of these operations
Work Hypothesis

Our thesis

- Cost of maintaining data structures needed to execute a query can be modeled looking at a set of traffic features

Empirical observation

- Different overhead when performing basic operations on the state while processing incoming traffic
  - E.g., creating or updating entries, looking for a valid match, etc.
- Cost of a query is mostly dominated by the overhead of some of these operations

Our method

Models queries’ cost by considering the right set of traffic features
Traffic Features vs CPU Usage

**Figure:** CPU usage versus the number of packets and flows
Multiple Linear Regression (MLR)

**Linear Regression Model**

\[ Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_p X_{pi} + \varepsilon_i, \quad i = 1, 2, \ldots, n. \]

- \( Y_i \) = \( n \) observations of the response variable (measured cycles)
- \( X_{ji} \) = \( n \) observations of the \( p \) predictors (traffic features)
- \( \beta_j \) = \( p \) regression coefficients (unknown parameters to estimate)
- \( \varepsilon_i \) = \( n \) residuals (OLS minimizes SSE)

**Feature Selection**

- Variant of the Fast Correlation-Based Filter\(^2\) (FCBF)
- Removes **irrelevant** and **redundant** predictors
- Reduces significantly the cost of the MLR

---

System Overview

Prediction and Load Shedding subsystem

1. Each 100ms of traffic is grouped into a batch of packets
2. The traffic features are efficiently extracted from the batch (multi-resolution bitmaps)
3. The most relevant features are selected (using FCBF) to be used by the MLR
4. MLR predicts the CPU cycles required by the query to run
5. Load shedding is performed to discard a portion of the batch
6. CPU usage is measured (using TSC) and fed back to the prediction system
Load Shedding

When to shed load
- When the prediction exceeds the available cycles
  \[ \text{avail}_\text{cycles} = (0.1 \times \text{CPU frequency}) - \text{overhead} \]
  - Corrected according to prediction error and buffer space
  - Overhead is measured using the time-stamp counter (TSC)

How and where to shed load
- Packet and Flow sampling (hash based)
- The same sampling rate is applied to all queries

How much load to shed
- Maximum sampling rate that keeps CPU usage < \text{avail}_\text{cycles}
  \[ \text{srate} = \frac{\text{avail}_\text{cycles}}{\text{pred}_\text{cycles}} \]
## Load Shedding

### When to shed load
- When the prediction exceeds the available cycles
  \[ \text{avail\_cycles} = (0.1 \times \text{CPU frequency}) - \text{overhead} \]
  - Corrected according to prediction error and buffer space
  - Overhead is measured using the time-stamp counter (TSC)

### How and where to shed load
- Packet and Flow sampling (hash based)
- The same sampling rate is applied to all queries

### How much load to shed
- Maximum sampling rate that keeps CPU usage < \text{avail\_cycles}
- \[ srate = \frac{\text{avail\_cycles}}{\text{pred\_cycles}} \]
Load Shedding

When to shed load
- When the prediction exceeds the available cycles
- \( avail\_cycles = (0.1 \times CPU\ frequency) - overhead \)
  - Corrected according to prediction error and buffer space
  - Overhead is measured using the time-stamp counter (TSC)

How and where to shed load
- Packet and Flow sampling (hash based)
- The same sampling rate is applied to all queries

How much load to shed
- Maximum sampling rate that keeps CPU usage \(< avail\_cycles\)
- \( srate = \frac{avail\_cycles}{pred\_cycles} \)
## Load Shedding Algorithm

Load shedding algorithm (simplified version)

```
pred_cycles = 0;
foreach q_i in Q do
    f_i = feature_extraction(b_i);
    s_i = feature_selection(f_i, h_i);
    pred_cycles += mlr(f_i, s_i, h_i);

if avail_cycles < pred_cycles × (1 + \hat{\text{error}}) then
    foreach q_i in Q do
        b_i = sampling(b_i, q_i, srate);
        f_i = feature_extraction(b_i);

foreach q_i in Q do
    query_cycles_i = run_query(b_i, q_i, srate);
    h_i = update_mlr_history(h_i, f_i, query_cycles_i);
```
Testbed Scenario

- **Equipment and network scenario**
  - 2 × Intel® Pentium™ 4 running at 3 GHz
  - 2 × Endace® DAG 4.3GE cards
  - 1 × Gbps link connecting Catalan RREN to Spanish NREN

- **Executions**

<table>
<thead>
<tr>
<th>Execution</th>
<th>Date</th>
<th>Time</th>
<th>Link load (Mbps) mean/max/min</th>
</tr>
</thead>
<tbody>
<tr>
<td>predictive</td>
<td>24/Oct/06</td>
<td>9am:5pm</td>
<td>750.4/973.6/129.0</td>
</tr>
<tr>
<td>original</td>
<td>25/Oct/06</td>
<td>9am:5pm</td>
<td>719.9/967.5/218.0</td>
</tr>
<tr>
<td>reactive</td>
<td>05/Dec/06</td>
<td>9am:5pm</td>
<td>403.3/771.6/131.0</td>
</tr>
</tbody>
</table>

- **Queries (from the standard distribution of CoMo)**

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>application</td>
<td>Port-based application classification</td>
</tr>
<tr>
<td>counter</td>
<td>Traffic load in packets and bytes</td>
</tr>
<tr>
<td>flows</td>
<td>Per-flow counters</td>
</tr>
<tr>
<td>high-watermark</td>
<td>High watermark of link utilization</td>
</tr>
<tr>
<td>pattern search</td>
<td>Finds sequences of bytes in the payload</td>
</tr>
<tr>
<td>top destinations</td>
<td>List of the top-10 destination IPs</td>
</tr>
<tr>
<td>trace</td>
<td>Full-payload collection</td>
</tr>
</tbody>
</table>
Packet Loss

Figure: Link load and packet drops
Related Work

Network Monitoring Systems
- Only consider a pre-defined set of metrics
- Filtering, aggregation, sampling, etc.

Data Stream Management Systems
- Define a declarative query language (small set of operators)
- Operators’ resource usage is assumed to be known
- Selectively discard tuples, compute summaries, etc.
Related Work

<table>
<thead>
<tr>
<th>Network Monitoring Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Only consider a pre-defined set of metrics</td>
</tr>
<tr>
<td>• Filtering, aggregation, sampling, etc.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data Stream Management Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Define a declarative query language (small set of operators)</td>
</tr>
<tr>
<td>• Operators’ resource usage is assumed to be known</td>
</tr>
<tr>
<td>• Selectively discard tuples, compute summaries, etc.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Restrict the type of metrics and possible uses</td>
</tr>
<tr>
<td>• Assume explicit knowledge of operators’ cost and selectivity</td>
</tr>
</tbody>
</table>
Conclusions and Future Work

- Effective load shedding methods are now a basic requirement
  - Rapidly increasing data rates, number of users and complexity of analysis methods

- Load shedding operates without knowledge of the traffic queries
  - Quickly adapts to overload situations by gracefully degrading accuracy via packet and flow sampling

- Operational results in a research ISP network show that:
  - The system is robust to severe overload
  - The impact on the accuracy of the results is minimized

- Limitations and Future work
  - Load shedding methods for queries non robust against sampling
  - Load shedding strategies to maximize the overall system utility
  - Other system resources (memory, disk bandwidth, storage space)