POLYGRAPH: SYSTEM FOR DYNAMIC REDUCTION OF FALSE ALERTS IN LARGE-SCALE IT SERVICE DELIVERY ENVIRONMENTS

SANGKYUM KIM (UIUC)
WINNIE CHENG, SHANG GUO, LAURA LUAN, DANIELA ROSU (IBM RESEARCH)
ABHIJIT BOSE (GOOGLE)
Background

- Large-scale IT service delivery systems
  - No longer confined to racks within a single data center
  - Increasing adoption of virtualization and cloud computing

- Our focus
  - Monitoring alerts
  - Significant portion of alerts are false

- Polygraph
  - Mine historical alerts to dynamically adjust monitoring policies
## Basic Alert Policy Types

<table>
<thead>
<tr>
<th>Type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>IF A;</td>
<td>IF (System.Virtual_Memory_Percent_Used &gt; 90)</td>
</tr>
<tr>
<td>IF A AND B;</td>
<td>IF (NTPhysical_Disk.Disk_Time &gt; 80) AND (NT_Physical_Disk.Disk_Time ≤ 90)</td>
</tr>
<tr>
<td>IF A OR B;</td>
<td>IF (SMP_CPU.CPU_Status = ‘off-line’) OR (SMP_CPU.Avg_CPU_Busy_15 &gt; 95)</td>
</tr>
</tbody>
</table>
Polygraph System Architecture

System/event source
- Resource Utilization and Performance
- System Configuration Data
- Operation Data (SLA, Maintenance Schedule, ...)
- Events

Ticket source
- Incident Management
  - Problem tickets
  - Alerts

Polygraph
- False alert detector
  - False alerts
- Alert policy generator
  - Tune alert policies
  - Proposed alert policies
- Alert policy evaluator/simulator
  - New/modified alert policies
- Policy deployment
  - Monitoring Rule Dispatcher

Monitoring System Management
- Monitoring Agent
- Monitoring Rules

Monitored server
- CPU
- Disk
- Memory
- App

Events
- Operation Data (SLA, Maintenance Schedule, ...)
- Rule change
- Alerts

Proposed alert policies
- Problem tickets
- New/modified alert policies

Tune alert policies
Host-based Alert Policy Threshold Adjustment

- min resource of real alerts
- max resource of false alerts
- Current threshold
Finding patterns for false alerts

- Example: periodic patterns
- They might include true alerts
Finding patterns for true alerts

Mine true ranges

- User-specified threshold given to decide the width of true range

*True range threshold: 1 hour
True ranges: (2-5pm), (7-9pm)
Experiments

**Host-based threshold adjustment**

**Host and Time-based threshold adjustment**

**True range threshold effect**
Discussion

- Leverage operational data for alert policy tuning
  - Anti virus (20% of a customer’s alerts)
- Weighted scheme
  - Put emphasis on recent input
- Impact of change operations
  - Integration of service management data is necessary
- Leverage server similarity
  - Grouping similar servers provides a better training dataset
Conclusion

- How to reduce false alerts
  - Polygraph tunes alert policies based on historical data
    - To improve recall, we utilized
      - Localized feature: Host
        - High recall, barely miss true events
      - Time-dependent behavior
        - Higher recall, reasonable precision
Questions ?