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)

USENIX ATC'11 (June 2011, Portland, OR)

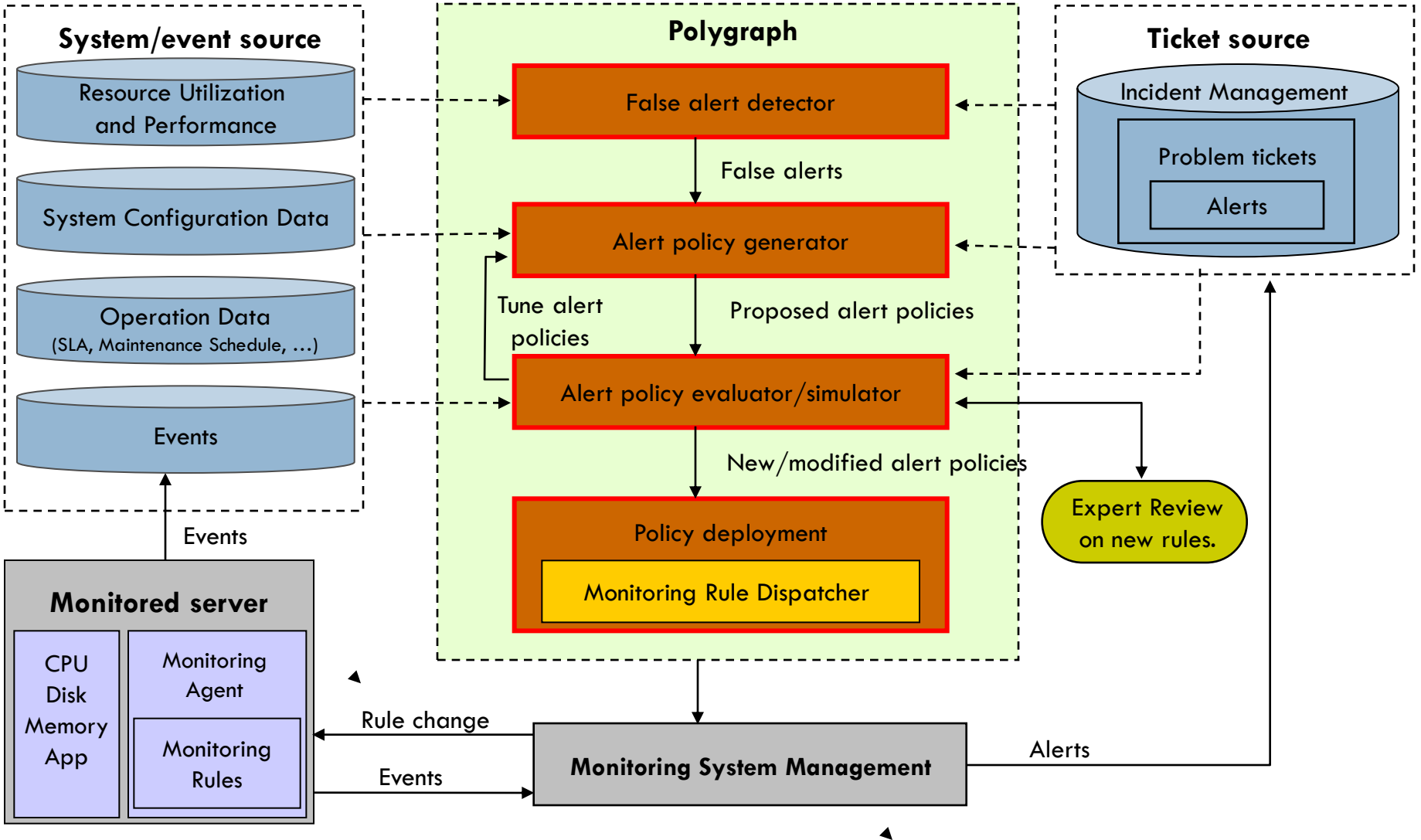
Background

- 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

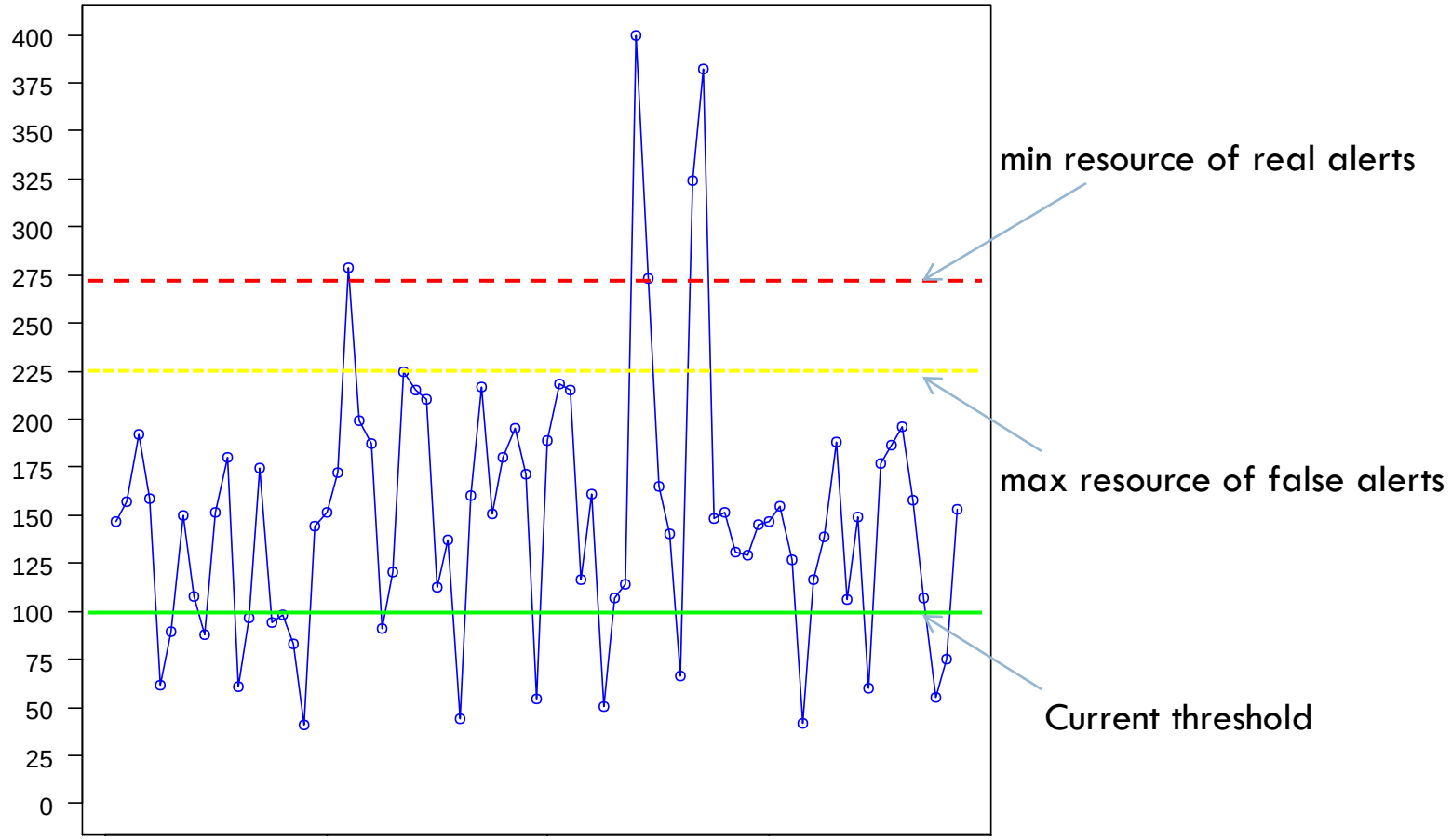
Basic Alert Policy Types

Type	Example
IF A;	IF (System.Virtual_Memory_Percent_Used > 90)
IF A AND B;	IF (NTPhysical_Disk.Disk_Time > 80) AND (NT_Physical_Disk.Disk_Time ≤ 90)
IF A OR B;	IF (SMP_CPU.CPU_Status = 'off-line') OR (SMP_CPU.Avg_CPU_Busy_15 > 95)

Polygraph System Architecture

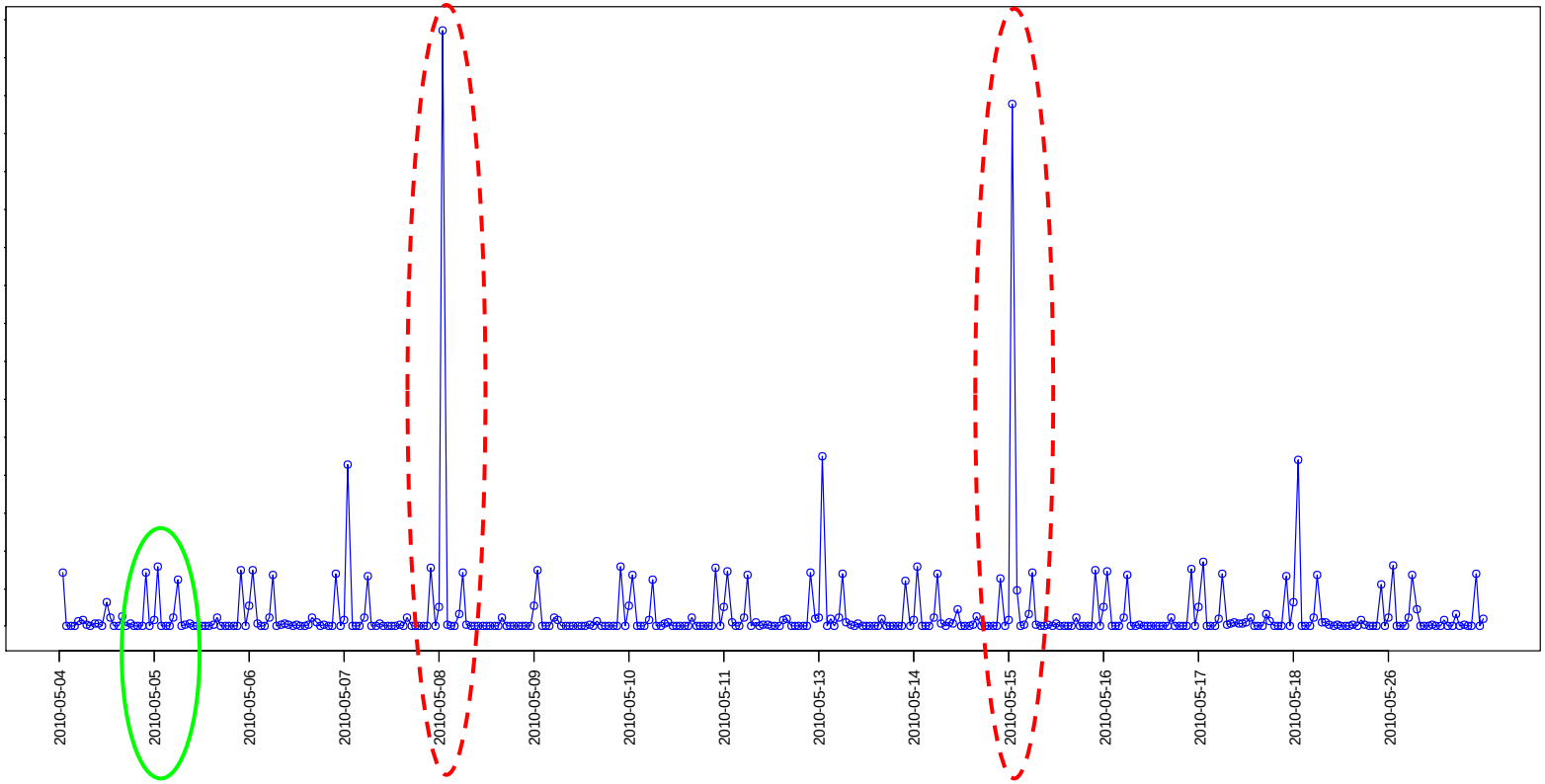


Host-based Alert Policy Threshold Adjustment



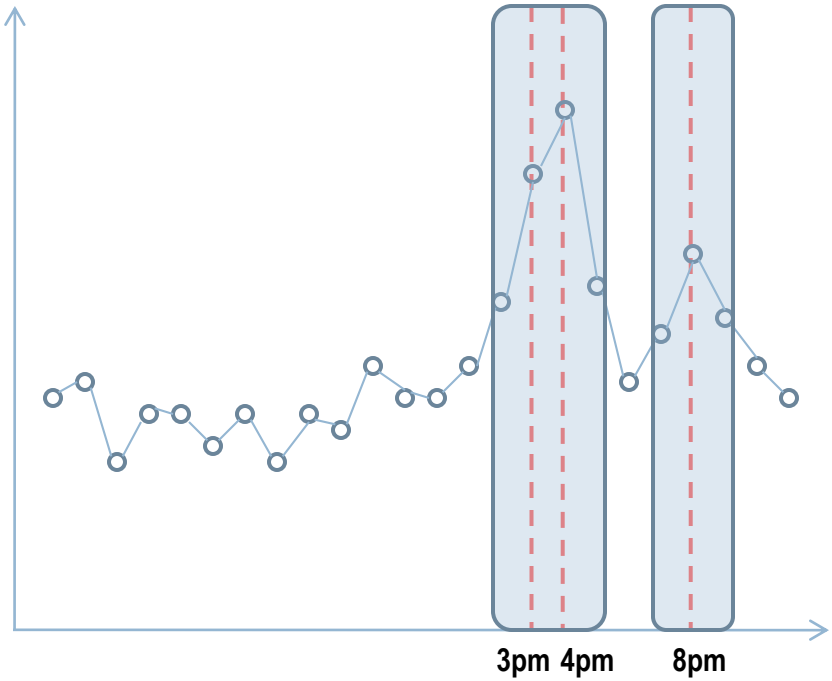
Time-based Alert Policy Threshold Adjustment (I)

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



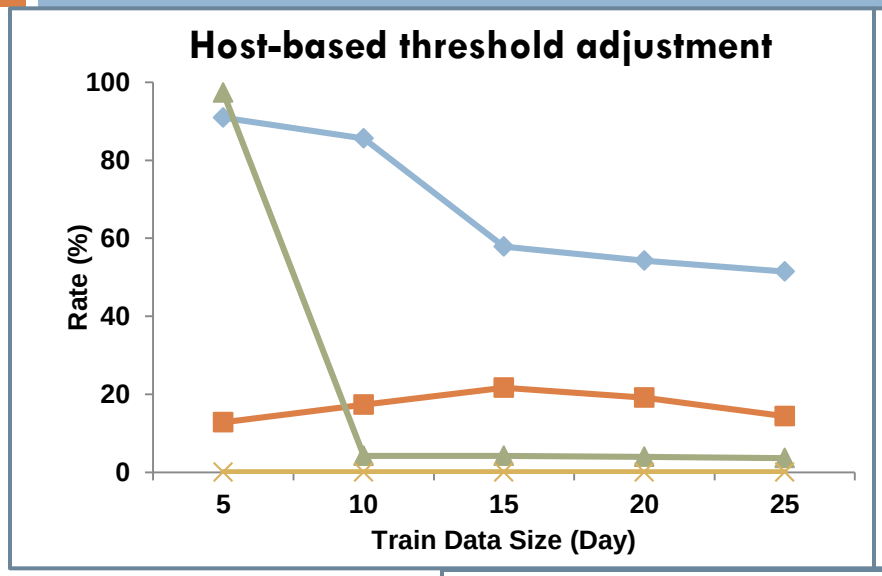
Time-based Alert Policy Threshold Adjustment (II)

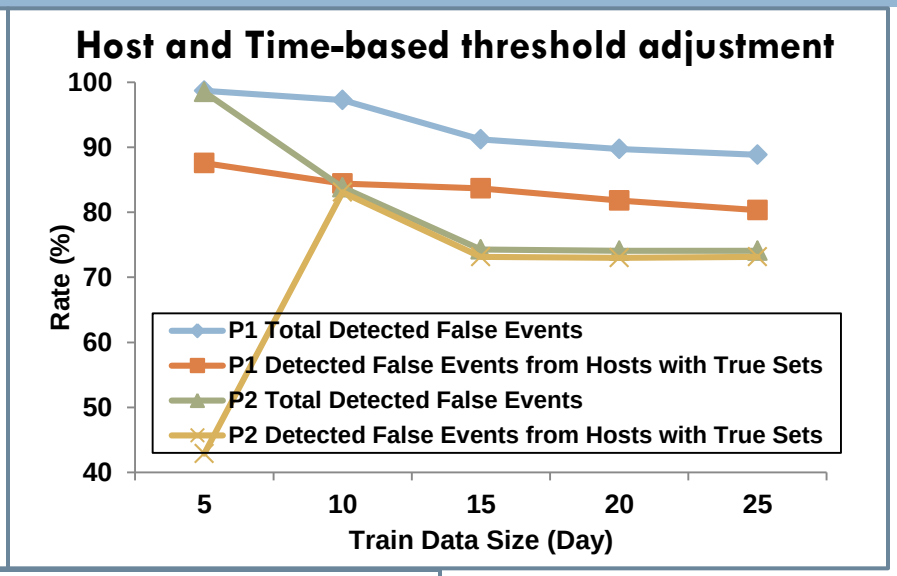
- Finding patterns for true alerts
  - Mine true ranges
    - User-specified threshold given to decide the width of true range

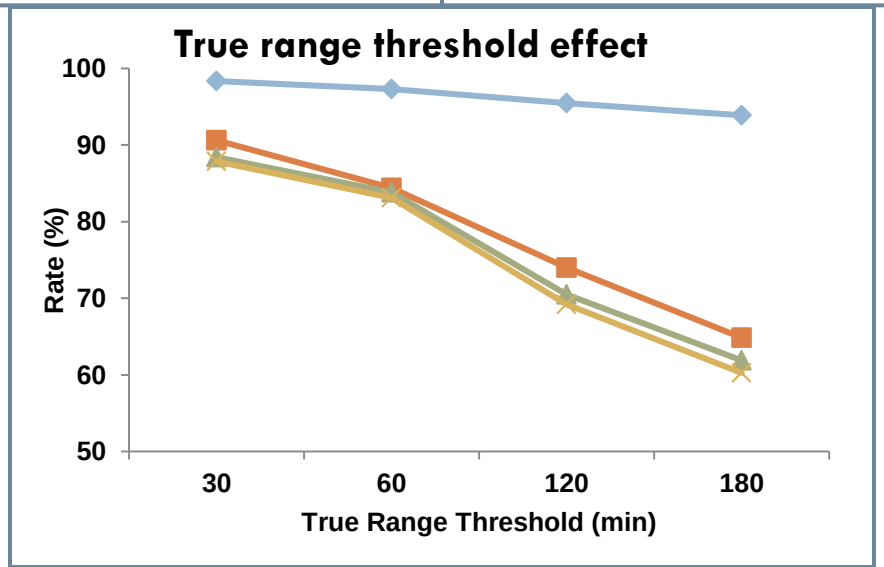


*True range threshold: 1 hour
True ranges: (2-5pm), (7-9pm)

Experiments







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 ?