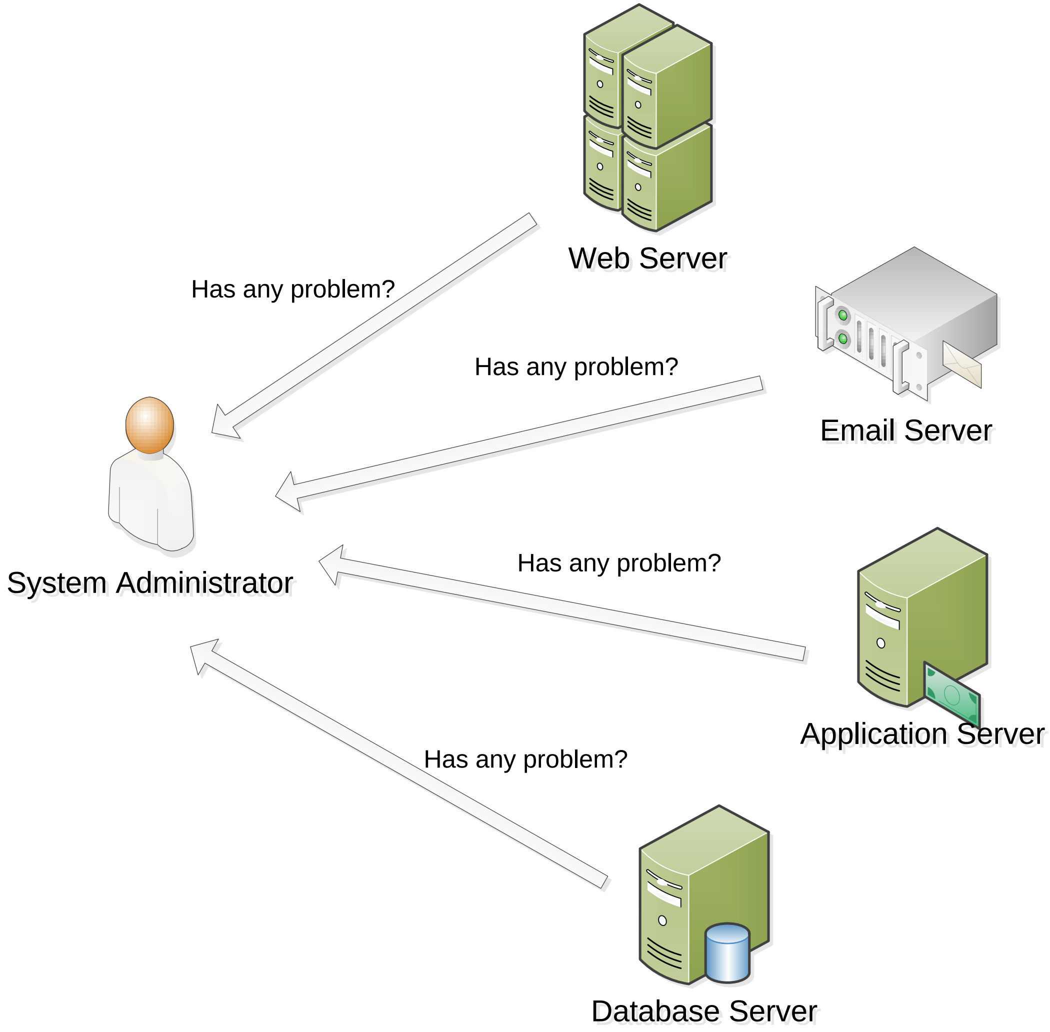
An Integrated Framework for Optimizing Automatic Monitoring Systems in Large IT Infrastructures

Liang Tang, Tao Li
{ltang002,taoli}@cs.fiu.edu

Larisa Shwartz, Florian Pinel
{lshwart,pinel}@us.ibm.com

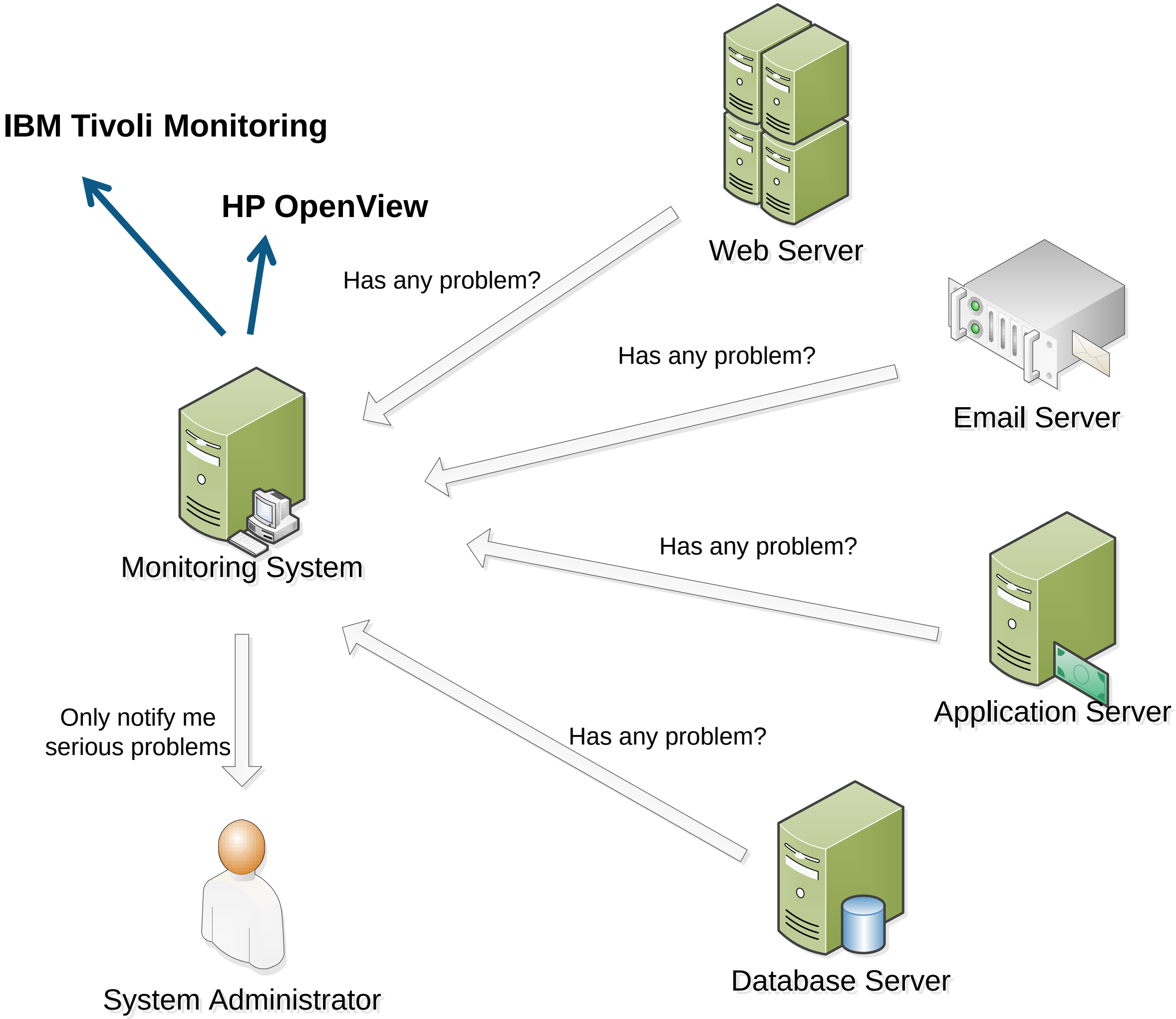
Genady Ya. Grabarnik
grabarng@stjohns.edu

Manual System Monitoring



In large IT infrastructures, the system admin cannot manually monitor so many machines.

Automatic System Monitoring



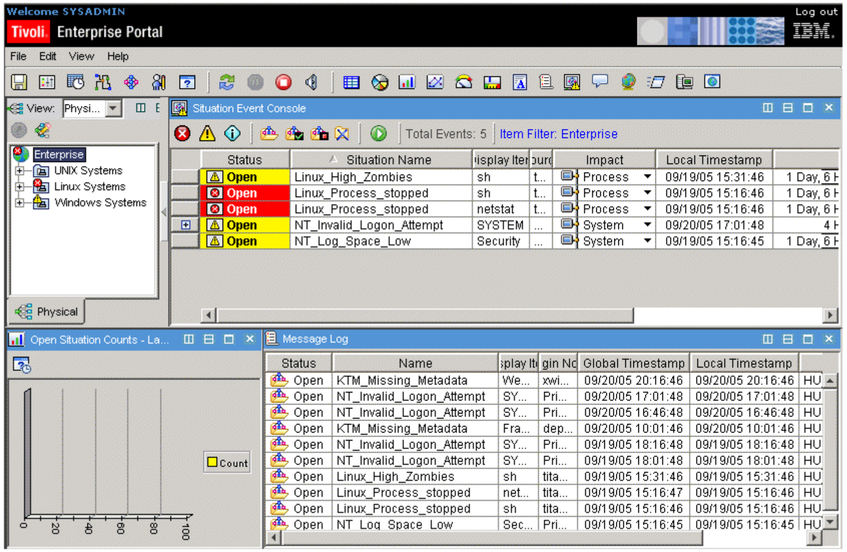
Monitoring system: monitor those servers, notify the system admin only when a problem happens.

Configurations of Monitoring Systems are Complicated

- In Large IT infrastructures, there are different machines, different software products…

- IBM Tivoli monitoring defines a lot of monitoring situations for monitoring different alerts
  - High CPU utilization
  - Low disk space
  - Process offline
  - …

IBM Tivoli Monitoring

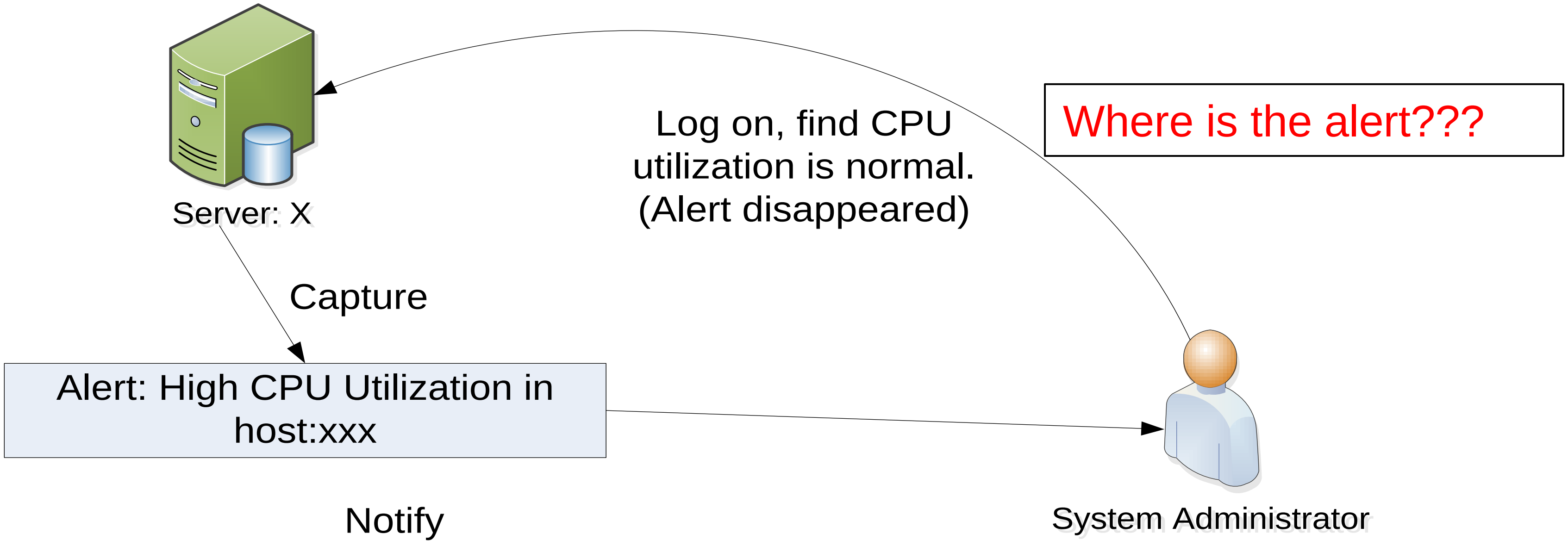


	Status	Situation Name	isplay Iter	ourc	Impact	Local Timestamp	
	Open	Linux_High_Zombies	sh	t...	Process	09/19/05 15:31:46	1 Day, 6 H
	Open	Linux_Process_stopped	sh	t...	Process	09/19/05 15:16:46	1 Day, 6 H
	Open	Linux_Process_stopped	netstat	t...	Process	09/19/05 15:16:46	1 Day, 6 H
⊞	Open	NT_Invalid_Logon_Attempt	SYSTEM	...	System	09/20/05 17:01:48	4 H
	Open	NT_Log_Space_Low	Security	...	System	09/19/05 15:16:45	1 Day, 6 H

Status	Name	splay It	gin Nc	Global Timestamp	Local Timestamp	
Open	KTM_Missing_Metadata	We...	xwi...	09/20/05 20:16:46	09/20/05 20:16:46	HU
Open	NT_Invalid_Logon_Attempt	SY...	Pri...	09/20/05 17:01:48	09/20/05 17:01:48	HU
Open	NT_Invalid_Logon_Attempt	SY...	Pri...	09/20/05 16:46:48	09/20/05 16:46:48	HU
Open	KTM_Missing_Metadata	Fra...	dep...	09/20/05 10:01:46	09/20/05 10:01:46	HU
Open	NT_Invalid_Logon_Attempt	SY...	Pri...	09/19/05 18:16:48	09/19/05 18:16:48	HU
Open	NT_Invalid_Logon_Attempt	SY...	Pri...	09/19/05 18:01:48	09/19/05 18:01:48	HU
Open	Linux_High_Zombies	sh	tita...	09/19/05 15:31:46	09/19/05 15:31:46	HU
Open	Linux_Process_stopped	net...	tita...	09/19/05 15:16:47	09/19/05 15:16:46	HU
Open	Linux_Process_stopped	sh	tita...	09/19/05 15:16:46	09/19/05 15:16:46	HU
Open	NT_Log_Space_Low	Sec...	Pri...	09/19/05 15:16:45	09/19/05 15:16:45	HU

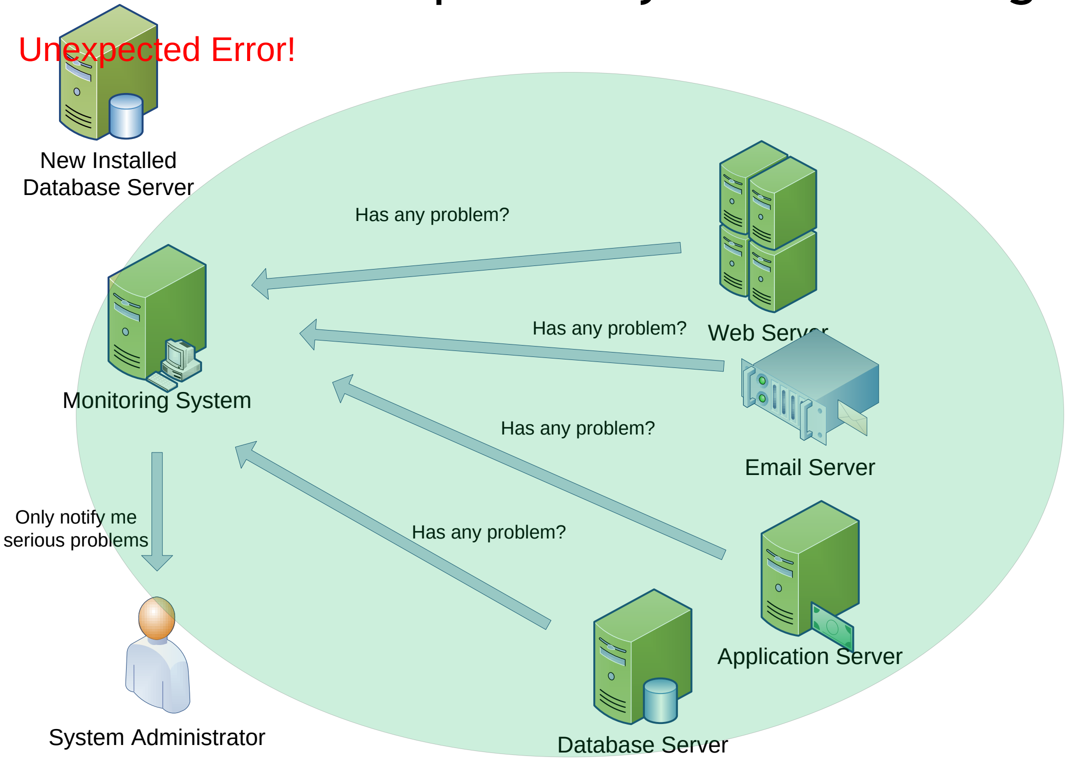
What is Misconfiguration? (1)

- **False Positive:**
  - Too Conservative threshold (CPU utilization < **50%**).
  - Transient Alert(Automatically disappear in a short time).


Server: X
Capture
Alert: High CPU Utilization in host:xxx
Notify
System Administrator
Log on, find CPU utilization is normal. (Alert disappeared)
Where is the alert???

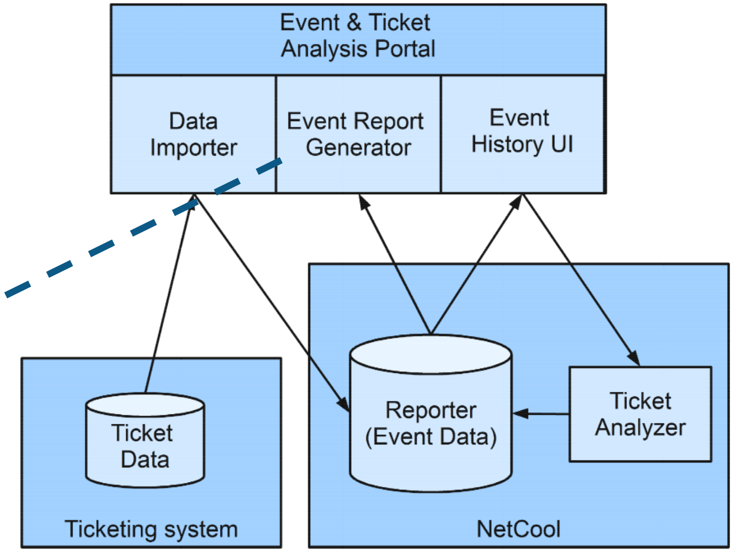
What is Misconfiguration? (2)

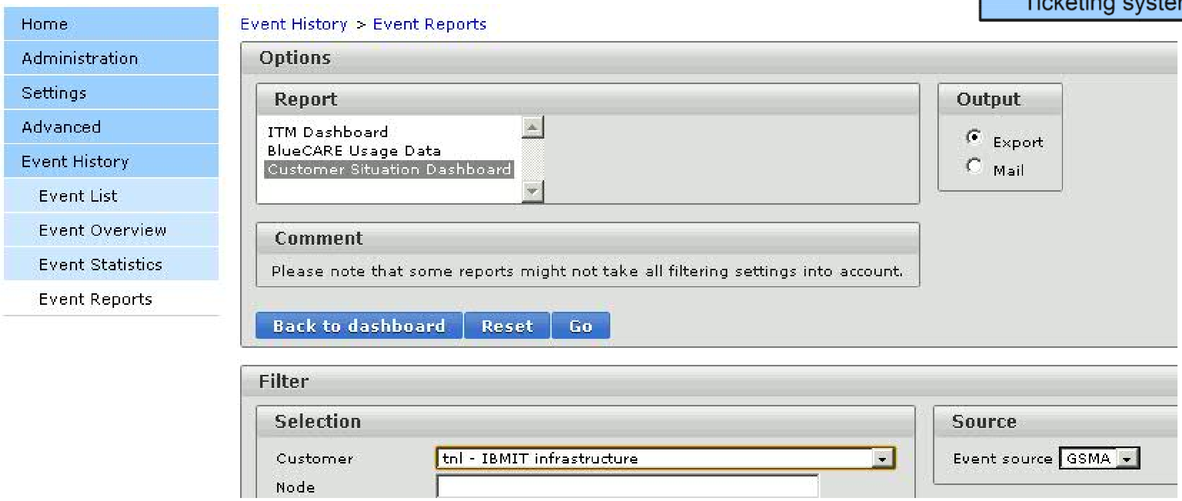
- **False Negative:**
  - Installed a new database server, but forget to add it into the monitoring situation. If this server has a problem, it would not be captured by the monitoring system.



Event and Ticket Analysis Portal

Our solution: Develop a system module in *Event & Ticket Analysis Portal* to analyze the monitoring events with tickets and to help system admin to correct misconfigurations.





Event History > Event Reports

Options

Report

ITM Dashboard
BlueCARE Usage Data
Customer Situation Dashboard

Output

● Export
○ Mail

Comment

Please note that some reports might not take all filtering settings into account.

Filter

Selection

Customer: tnl - IBMIT infrastructure

Node:

Source

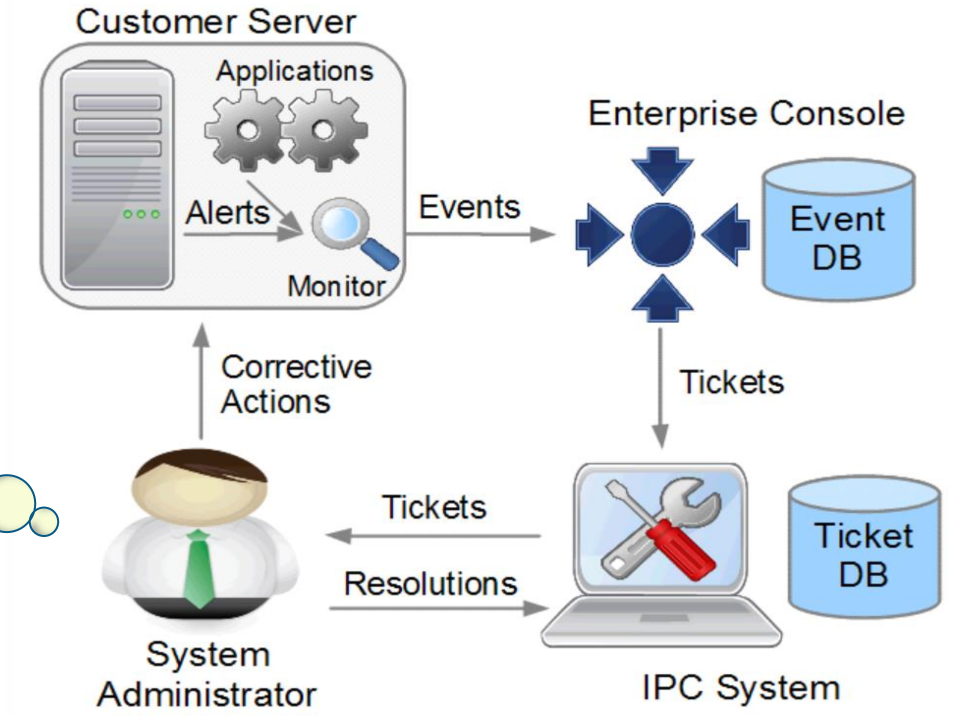
Event source: GSMA

How to Detect False negative and False positive?

- **Ticket data** is the ground truth (labeled data) and created by the human.

Human labor cost is very high!!!

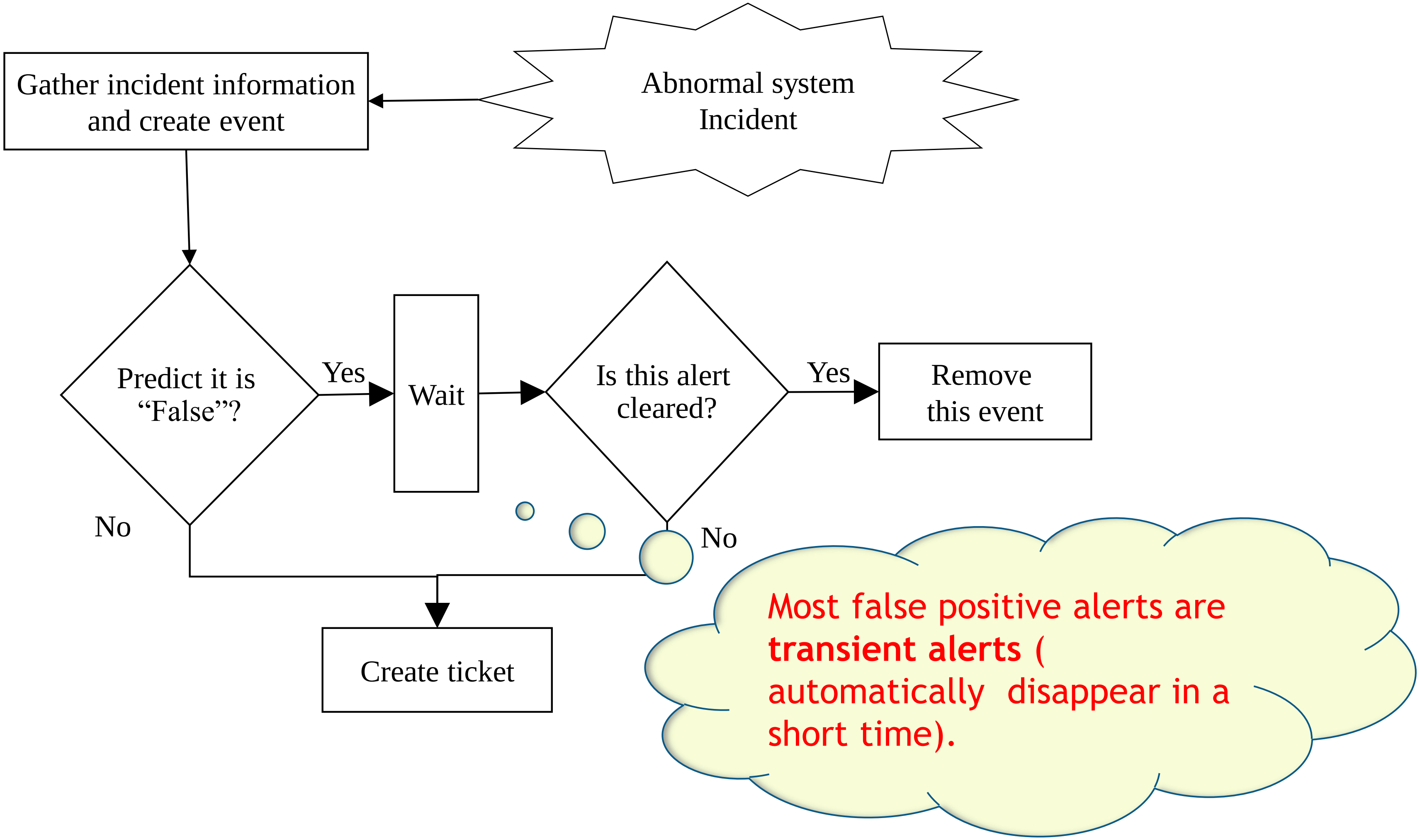
Can we use their knowledge to improve the monitoring?



Eliminating False Positive (1)

- A straightforward solution: **Binary classifier**
  - label “1” means a real alert, “0” means a false alert.
  - features are system event attributes
    - process name
    - CPU time
    - number of threads.
- Limitations:
  - We can NOT miss any real alert (would cause system crash or data loss).
  - No classification algorithm can guarantee 100% accuracy.

Eliminating False Positive (2)



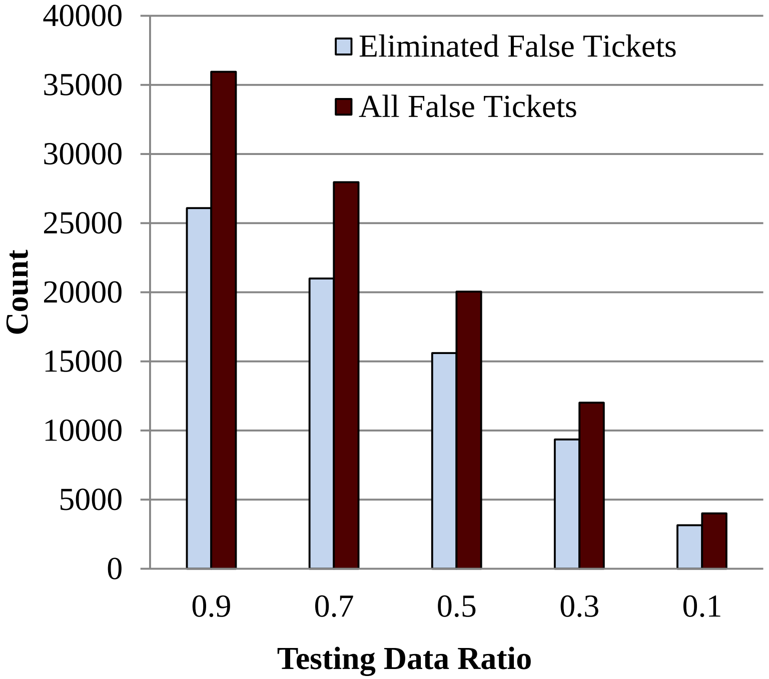
Most false positive alerts are **transient alerts** (automatically disappear in a short time).

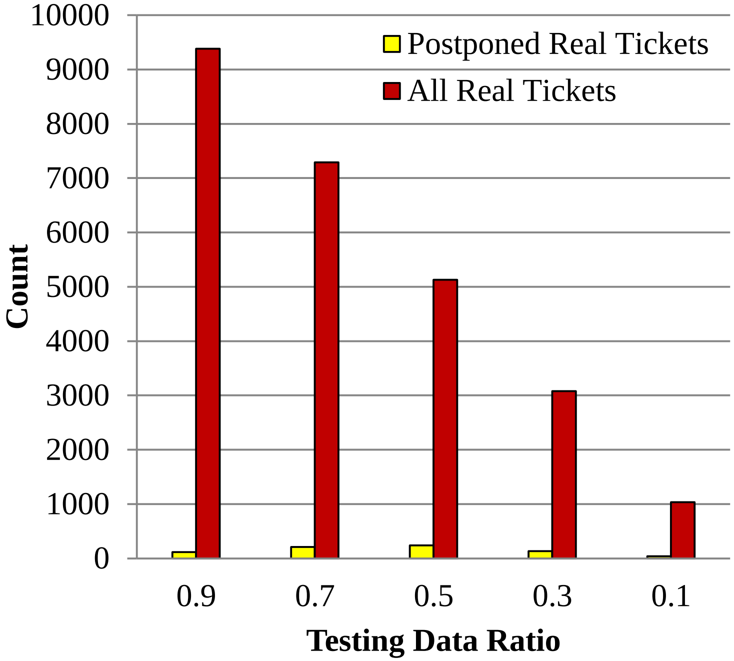
Eliminating False Positive (3)

- The rules generated by a classifier can be directly translated into monitoring situations:
  - If PROC_CPU_TIME > 50% and PROC_NAME = ‘Rtvscan’, then it is false.
- *Waiting time* is the polling interval of a monitoring situation in IBM Tivoli Monitoring.

We do NOT have to build another system to deploy our classifier

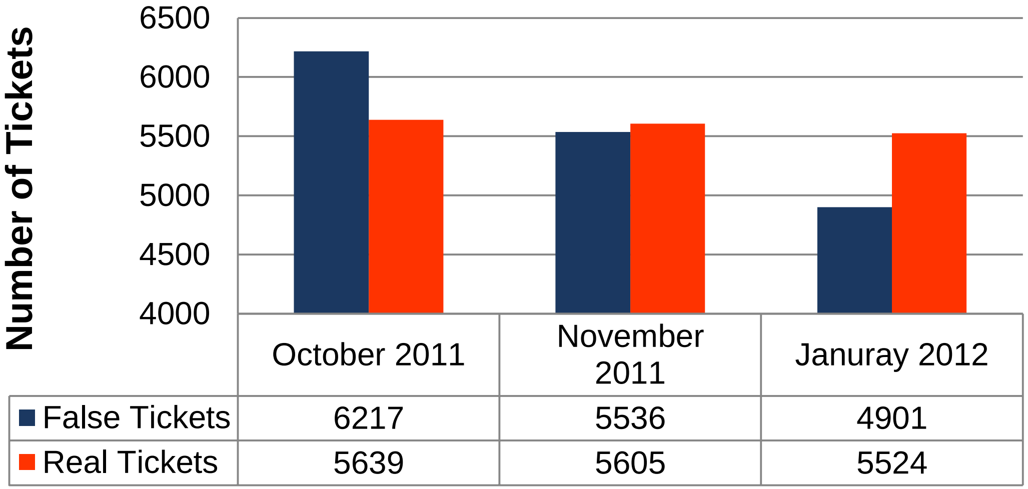
Offline Evaluation on Testing Data





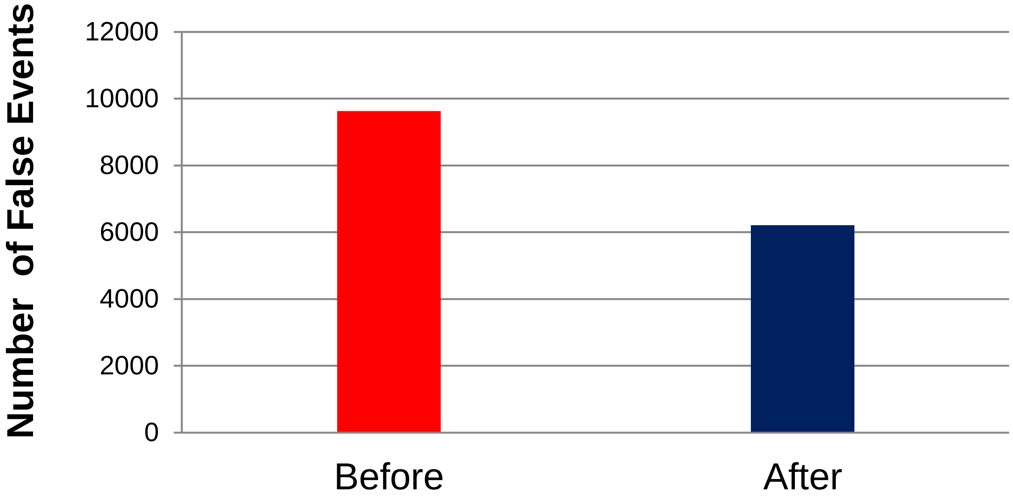
Ratio of the testing data size and training data size

Online Evaluation



	October 2011	November 2011	Januray 2012
🟦 False Tickets	6217	5536	4901
🟥 Real Tickets	5639	5605	5524

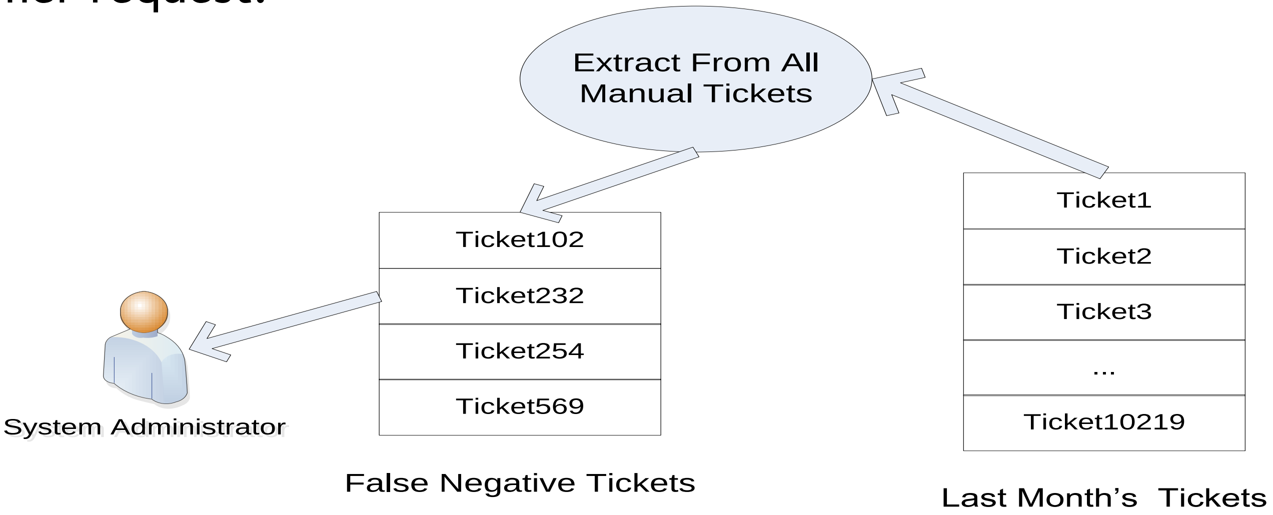
A large financial company.



An internal account in IBM.

Eliminating False Negative (1)

- How to eliminate false negatives (missed alerts)?
  - False negative are quite few (less than 20-40 tickets for a situation). No need an automatic approach to correct it.
- False negatives are missed alerts. Where can we track them?
  - **Manual Tickets (captured by human).**
  - However, manual tickets contain other kinds of tickets, such as customer request.



Eliminating False Negative (2)

- **Problem Definition:** Find missed alerts from manual tickets

- **Challenges:**
  - Not enough labeled data.

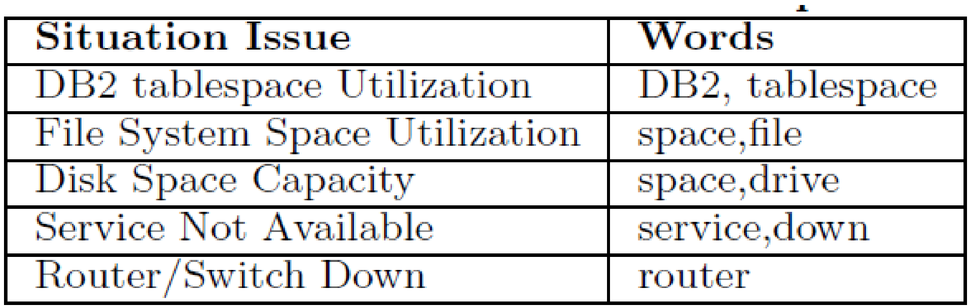
We cannot hire an expert to label the ticket every day…

  - Highly Imbalanced data: few false negative alerts, large amount of other manual tickets.

We randomly select a subset of tickets to label. Maybe none of them is false negative alert. It is bad for training.

Selective Labeling in Highly Imbalanced Data

- Use some *domain words* to narrow down the training ticket scope

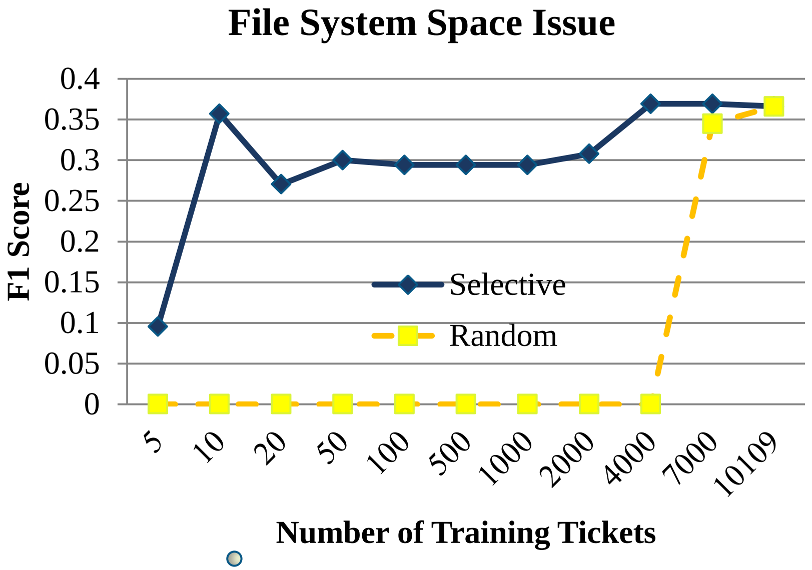


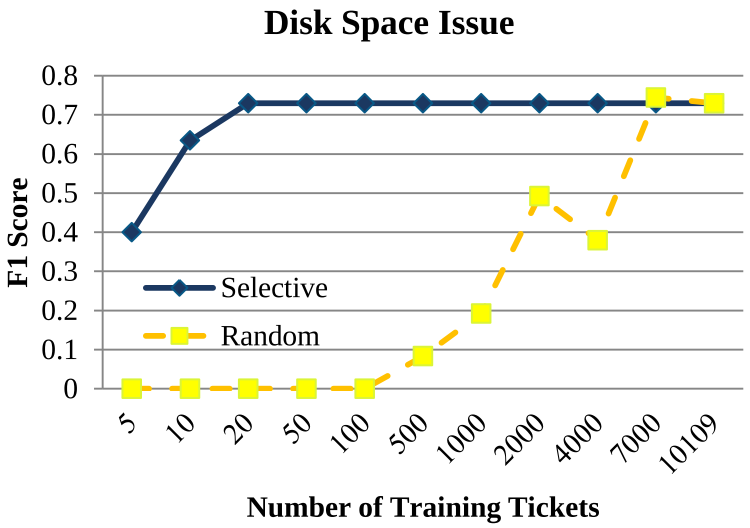
Situation Issue	Words
DB2 tablespace Utilization	DB2, tablespace
File System Space Utilization	space,file
Disk Space Capacity	space,drive
Service Not Available	service,down
Router/Switch Down	router

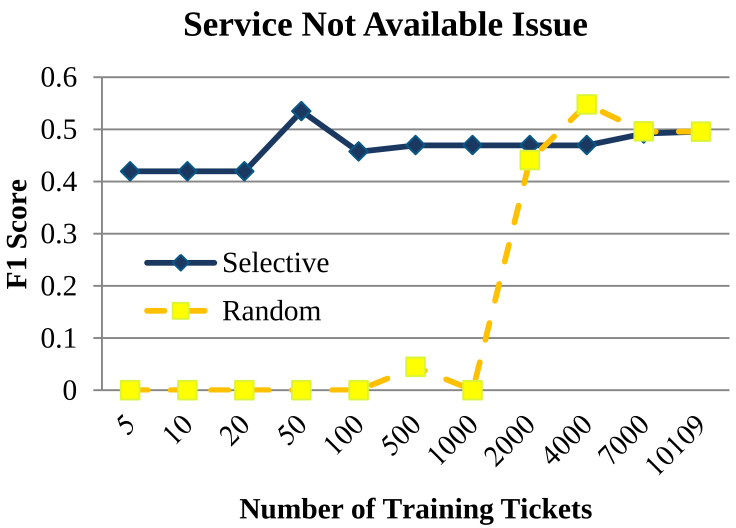
These words can be obtained from system admins or system documents

- Build a binary classifier (SVM) on selected tickets.
  - Given a ticket, label “1” means this ticket is a false negative. Label “0” means it is not.

Selective Labeling vs Random Labeling







Easy to learn!!! Not many variations of discriminative words

A Case Study

Discovered False Negatives (Missed alerts)

Situation	Ticket
dsp_3ntc_std	*Please clear space from E drive xxxx-fa-ntfwwfdb Please clear space from E drive xxxx-fa-ntfwwfdb.it is having 2 MB free...*
fss_rlzc_std	*/opt file system is is almost full on xxx Hi Team@/opt file system is almost full. Please clear some space /home/dbasso>df -h /optFilesystem...*
svc_3ntc_std	*RFS101681 E2 Frontier all RecAdmin services are down Frontier RecAdmin services are not running on the batch server Kindly logon to the server : xxx.xxx.155.183/xxx ...*
…	…

I will add these devices into Tivoli monitoring configuration.

System Administrator

Summary

- Analyzed the main types of misconfiguration of monitoring systems in large IT infrastructures.

- Proposed a framework to integrate system events and tickets for improving the configurations of monitoring systems (IBM Tivoli monitoring).

- Conduct offline and online experiments for the proposed framework.

- Develop and deployed the module in Event and Ticket Analysis Portal in IBM IT service platform.

End

- Thank you!

- Any question?