# Chorus Logging

Network Computer Event Logging and Machine Intelligence

# Reporting Problem

- Student machines make up almost half of Engineerings 2,400 managed computers
- Students typically don’t report problems – they just try other PCs
- How do we support them in a meaningful way
  - Ignoring them would be a disservice

# Problems

- Some issues since Nexus moved to private domain, also some bad user configs and hardware issues
  - Had periods of poor login reliability last March, apparently dependant on load
  - Hard time tracking down the issues – working with IST on issure resolution
  - Difficult to chase rumours
  - Could not provide us or IST with useful data describing problems in their areas or ours

# Inspiration : Eric Praetzel of E&CE

- Mr. Praetzel had been syslogging Nexus Windows event logs since Windows 2000 days
- He provided solid field numbers of some failures that helped isolate issues
  - Using grep on the syslogs
- What if we were to collect and analyze data on a larger scale and in more detail
- What if we could see inside boot process, not just after network enabled

# Starting Small

- We began a project to collect event logs, starting with 6 PCs
- Including historical data, as DHCP/boot issues not visible while network down
- When network is up, server is on average 1 second behind event logs – close to real time collection

# Growing the System

- Collection value increases with number of computers – nonlinearly
- Within days we deployed to about 500 PCs in labs and a few in offices
- Added machine learning through Bayesian Analysis
- Within minutes, it was spitting out issues galore – some that no one had spotted before
- Web GUI interface for exploration

# A Word About Bayesian Analysis

- Association is not causality
  - Eg. top 6/7 Winter Olympic teams come from countries with socialized medicine and gun control.  True, but not necessarily the cause.
- Creates fingerprints that tell you what systems have in common
- Sometimes there is a direct unambiguous correlation that you can infer one flag guarantees another condition – though not cause

# What we found?

- Only 750 distinct error types – mostly just repetition
- Google finds cause of most of the events
- Registry errors – from prior aborted logins – complete with userids for remediation
- Software errors, such as needing different .Net libraries
- Machines which had physical problems
- Etc.

# What is a SIEM?

- the product capabilities of gathering, analyzing and presenting information from network and security devices
- identity and access-management applications
- vulnerability management and policy-compliance tools
- operating-system, database and application logs
- external threat data
- A key focus is to monitor and help manage user and service privileges, directory services and other system-configuration changes; as well as providing log auditing and review and incident response.

# What we desire?

- SIEM is focused on security events
  - We can detect every privileged use on a machine
  - We can detect login failures ON WORKSTATIONS
  - We USED to scan login activity this on Nexus servers, but not since IST took responsibility for the servers
  - Workstation access is not the usual place for account compromise
- We are interested in operational aspects of the event logs
  - Measuring and improving service delivery
  - Detecting issues in real time

# Where are we heading - short term?

- Now processing logs every 5 or 10 minutes
- Still classifying some undocumented events (unknown to Google)
- We now look to event logs, Bayes analysis early in debug cycle of problems
  - Even if only to spark ideas
- We are now looking for problems before they are reported, based on Bayes
- Will soon send self-service messages to people who have login problems
  - or will repair accounts if known bad

# Questions