Dexter: Faster Troubleshooting of Misconfiguration Problems Using System Logs

Rukma Talwadker
NetApp Inc.

22nd May 2017

Presented at SYSTOR 2017
Software Misconfigurations are a big deal!

- There is a much sprawl: nearly 2K options in Firefox and over 36K options in LibreOffice
- Complex interaction between configuration settings and execution environment renders the configuration effort difficult and error-prone
- Significant portion of the operational cost (over 70%) is made up by people costs
  - Change of hands via escalations
  - Communication logistics
Misconfiguration troubleshooting workflow

Disk shelf ID’s
- 12.234.56.7
- 12.345.6.7
- AS45698 WD
- 12.36.78.6

Suspicious entry

Misconfiguration Root Cause

System Logs
- “Error: Invalid Shelf ID”
- “Warning: Delay in I/O response”
- “Error: Exceeded the connection limit”
- “Warning: Node clock skew detected”

Problem indicator

Resolution
- “Shelf ID should be a 32-bit numeric address written as four numbers separated by periods.”

*Dexter is a misconfiguration troubleshooting tool!*
Dexter Spotlights
Dexter’s role is two-fold

Problem Spotlights

Ranked list of system log messages, one or more of which correlate with the ongoing misconfiguration problem.

Resolution Spotlights

Heuristically derived possible solutions to the given misconfiguration problem by mining the command history logs of the system.

Note: AutoSupport® is a digital exhaust of system operations from various customer installations, sent back to NetApp periodically.
Problem Spotlights
Extracting System Log features (metrics)

Dexter considers the following features:

1. Reported timestamp
2. software module and/or sub system which logged the message
3. The actual error message

A sample Apache HTTP server log

Finding Problem Spotlights

Metrics that matter

Log Metrics defined by Dexter:

- **Message Recency**: messages temporally correlated with the problem are potential clues
- **Message Severity**: messages with higher severity are more concerning than a warning message
- **Message Frequency**: more frequently appearing error message which is also recent is possibly correlated

Ranked list of log messages is presented to the support engineer
Resolution Spotlights
Resolution Spotlights

A heuristics approach for offline resolution prediction component

1) Dexter checks for disappearance of problem spotlights post case closure in the system logs.

2) Correlates disappearance of problem spotlights with:
   - Execution of a new command
   - Re-execution of the new argument or an option with a new value.
   - And time correlation of the command execution with the message disappearance.
Dexter Workflow

online and offline components

Misconfiguration Support case

Case opened

Online workflow

Process logs & extract features

Extract Problem spotlights

publish

Top 10 Problem spotlights

Offline workflow

Estimate the Problem indicating spotlights

Correlate Commands with Logs temporally

Signature database

{problem spotlight, command list}
Results

Validation on a sample set of 600 cases.
Q&A
An example of resolution prediction.
Evaluation Criteria

Effectiveness of Dexter

Problem Spotlights

Dexter posts the top 10 problem spotlights to the misconfiguration support case

- AutoSupports were enabled and available
- The problem indicating log message:
  - was contained (Recorded) in the logs
  - was ranked within the top 10 log messages for the system at the time Dexter was invoked.

Resolution Spotlights

All (- pre-filtered) correlated commands are presented as a possible solution

- AutoSupports were enabled and available
- The solution indicating command log:
  - was contained (Recorded) in the command history logs.