Event Analysis Toolset
Incident & Problem Management in large Zabbix monitored cloud

Konstantin Yakovlev
Dmitry Shchemelinin, Ph.D.
Sergey Smirnov, MBA
Eugene Prisyazhniy
Artem Akinchits
Zabbix@RingCentral

- 4 Zabbix Servers
- 36 Proxies
- 200+ Host types
- 3K+ Templates
- 10K+ Hosts
- 1.3M+ Items
- 400K+ Triggers (300K+ visible for NOC, 100K for Dev purposes)
- 2 maintenance windows daily

- 2K Events visible for NOC /24h average
- 2K Events /1h during outages
- 500+ Active alerts simultaneously during outages
'A few statistics

💡 Maintenances are noisy even if you put maintained hosts to Zabbix MW
A few statistics

Monitoring may be unreadable during infrastructure outages on large environments

EVENT SPAWNING DURING OUTAGE
Old solution demo

- Demo of what we saw during outages on a dashboard showing active visible alarms. [Saved HTML]
Problems (repetitive Events of any nature) provide a lot of noise

Causes:
• Development issues
• Operations issues
• Trigger expression misconfiguration
A few statistics

Exclamation mark Flapping events, a lot of flapping events

EVENT QUANTITY SATURATION BY AGE

- Age above 120s: 56%
- Age under 120s: 44%
Issue processing

💡 Precise issue logging is a must to have transparent Operations

The cost of manual issue processing is pretty high on large environments
So what do we have?
And what do we want?

• Availability, **five nines** of it

• Decrease overall events quantity, considering growing environment
• Increase monitoring transparency and quality
• Decrease issue processing costs and MTTR
Decision making background

People do things differently
Skills:
• A lot of experience as Monitoring operator
• Experience in issues investigation
• Passion to make life easier

Technologies:
• Zabbix 2.2 API
• DB: **MySQL 5.6 Galera Cluster**
• App: **Laravel 5.1 - PHP 5.6**
PSP Scheme
Problem management

So we decided to fight excessive alarming. Here are some requirements:

• Tool to collect Zabbix Events
• Tool to search through Events History
• Tool to track Problems and integrate with issue tracking and Events History
• Daily reporting on top alerting events with RCA and Problem Cases updates
Event Collector

event.get is good for Event analysis, but we wanted more

• Infrastructure relations
• Age, we want to know how fast our alerts are being resolved
Event Search

Features we need:
• Statistics visualization for selection
• Flexible search of events by properties
• Ability to navigate through history and narrow the selection by click
• Max delay to real-time data - 60s
• No entry barriers for investigation
Event Search demo
Event Report demo
How does Problem Management look like
Problem Management Dash demo

Inside Problem management case

PS Case #234: Several request from PAS to CPX were failed

- Case status: Investigation ONG
- JIRA / ZMR / INC: CPS-63002
- Affected components: PAS, CPX
- Affected layers/pods: 100, 200, 300, 400, 500, 600, 700, 800, 900
- Tagged events: 1829
- Created: 2016-02-01 11:48:00
- Cases merged in: 322

Related triggers (suppressed by 1 rules)

<table>
<thead>
<tr>
<th>Trigger description</th>
<th>Component</th>
<th>Severity</th>
<th>Events</th>
<th>Suppress</th>
<th>Remove</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPX response time too long</td>
<td>PAS</td>
<td>1</td>
<td>153</td>
<td>false</td>
<td>false</td>
</tr>
<tr>
<td>HTTP request rejected</td>
<td>CPX</td>
<td>2</td>
<td>95</td>
<td>false</td>
<td>false</td>
</tr>
<tr>
<td>PAS request failed</td>
<td>CPX</td>
<td>3</td>
<td>95</td>
<td>false</td>
<td>false</td>
</tr>
<tr>
<td>Several request from PAS to CPX were failed</td>
<td>PAS</td>
<td>4</td>
<td>1767</td>
<td>false</td>
<td>false</td>
</tr>
</tbody>
</table>

©2012 RingCentral, Inc. All rights reserved. RingCentral Confidential
Incident Management

Event processing upon receipt
Sometimes OK event may not be generated in Zabbix due to different causes
Monitoring Console demo

Demo with comparison of our old solution and new one on outage case and not only

### NOC Monitoring Console

<table>
<thead>
<tr>
<th>Count</th>
<th>Trigger</th>
<th>Ack</th>
<th>Jira</th>
<th>KE</th>
<th>Esc</th>
<th>PM</th>
<th>First seen</th>
<th>Last seen</th>
<th>Location</th>
<th>Layer</th>
<th>Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>iad01-c01-r001: Packet loss at VoIP WAN c001-r002-X0:133583</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1m 59s</td>
<td>1m 59s</td>
<td>IAD01</td>
<td>C01</td>
<td>RTR</td>
</tr>
<tr>
<td>1</td>
<td>UPTC_POD1: UPTC Inbound SMS Test FAILED [11592495341 -&gt; 12054150568]</td>
<td>✔</td>
<td></td>
<td>✔</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>zm01-c01-mon03: Free swap size &lt; 35%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>sp01-c01-x0301: AUS: Total amount of pending requests was increased</td>
<td>✔</td>
<td></td>
<td>✔</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>am01-c01-r001: am01-c01-r002: xe-0/0/0 (am01-c01-r002:xe-0/0/0) link is down: xe-0/0/0 (am01-c01-r002:xe-0/0/0) link is down</td>
<td>✔</td>
<td></td>
<td>✔</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>iad01-c01-lbr02: Certificate SSL on LBR is missing or expired</td>
<td>✔</td>
<td></td>
<td>✔</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>sp01-puppet.ringcentral.com: LMD545657 warn - sp01-puppet.ringcentral.com AAR-sp01-c001 Status</td>
<td>✔</td>
<td></td>
<td>✔</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>✔</td>
<td>INC-17639</td>
<td>✔</td>
<td></td>
<td>✔</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>✔</td>
<td>iad01-p01-vhs03: DELL server overall status is critical</td>
<td>✔</td>
<td></td>
<td>✔</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

©2012 RingCentral, Inc. All rights reserved. RingCentral Confidential
The last but not the least

Going back to problems

Problems

Real
- Operations issue

False positive
- Development issue
- Trigger misconfiguration
Retrospective trigger analyzer

Features:
• Ability to model trigger behavior on existing Item values across multiple hosts
• Ability to compare existing trigger to new one.
• Help in decision making for monitoring fixes.
Trigger Analyzer Scheme

Flow:
• Get payload
• Configure TEST host/items/triggers
• Get data from Production Zabbix
• Wait till items are ready to receive data
• Push values via zabbix_sender to TEST instance
• As values push is complete – collect resulting events and present to client
Trigger analyzer demo
Conclusion

Using approaches described above we managed to:

• Provide Ops with convenient event history navigation
• Provide NOC with event pre-analysis and routine automation
• Increase overall monitoring visibility
• Hold event count growth - having 7% event count versus 30% trigger count growth
• Decrease MTTR on most sensitive parts of the Incident management: Ack to Escalation and Investigation.
Thanks!