

Event Analysis Toolset

Incident & Problem Management in large Zabbix monitored cloud

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

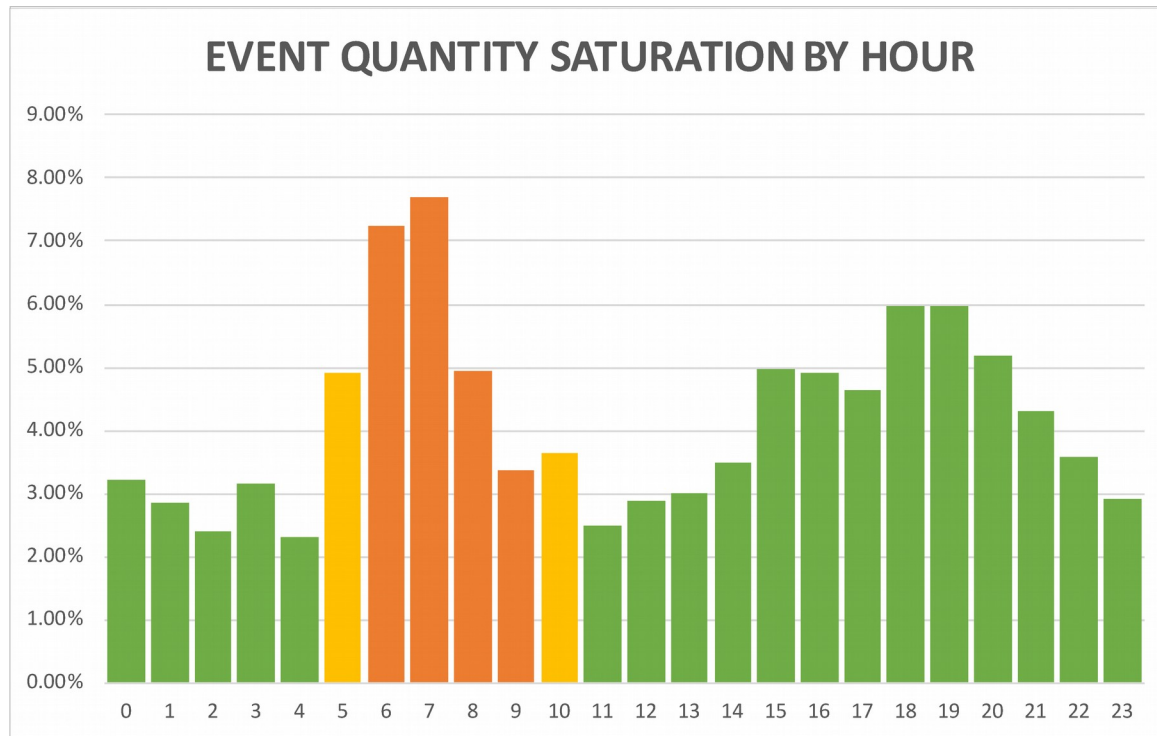
Zabbix@RingCentral

- 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

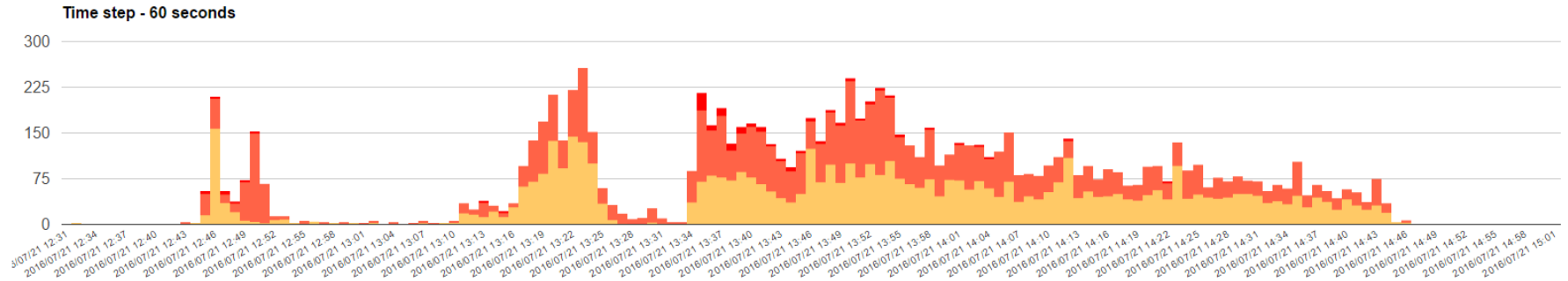
☞ Maintenances are noisy even if you put maintained hosts to Zabbix MW



A few statistics

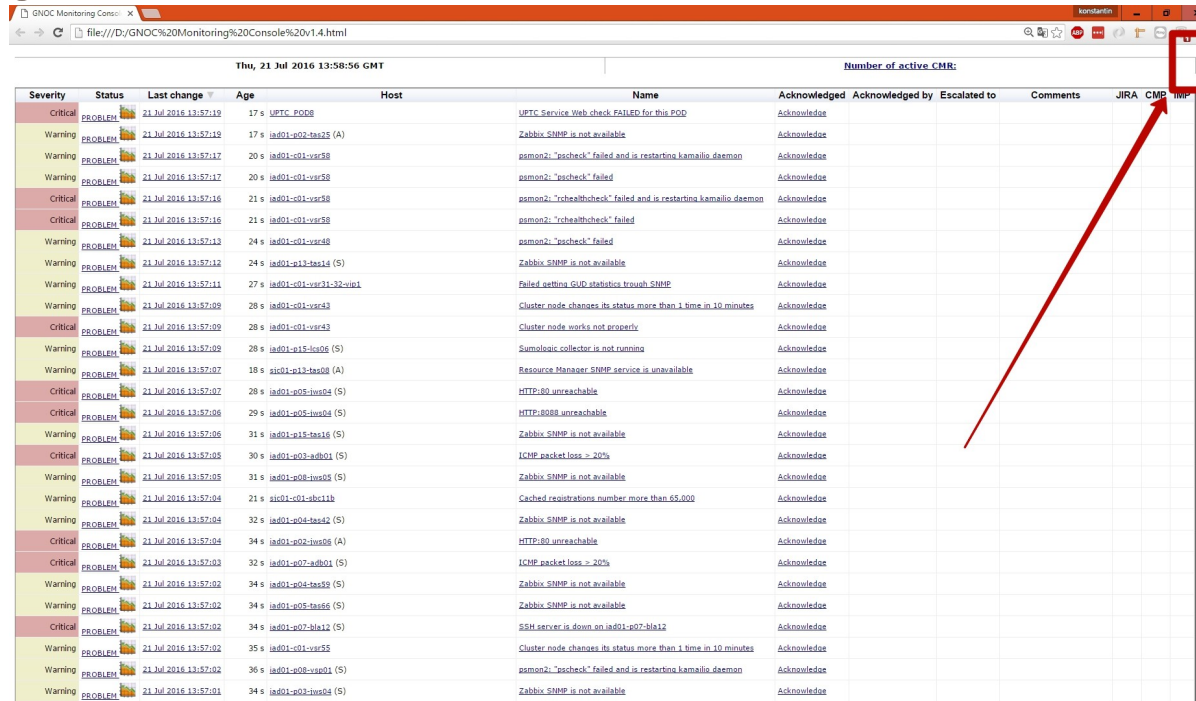
☞ Monitoring may be unreadable during infrastructure outages on large environments

EVENT SPAWNING DURING OUTAGE



Old solution demo

- Demo of what we saw during outages on a dashboard showing active visible alarms. [Saved HTML]



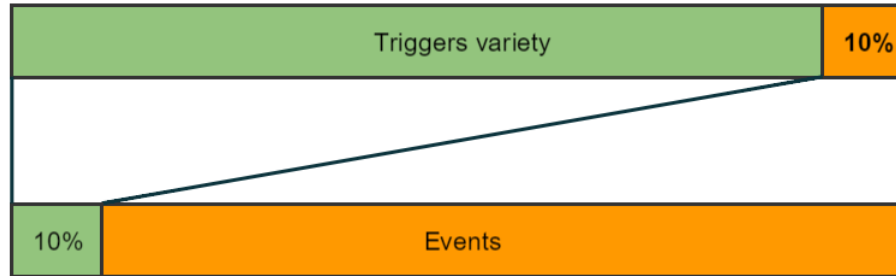
Thu, 21 Jul 2016 13:58:56 GMT

Number of active CMR:

Severity	Status	Last change	Age	Host	Name	Acknowledged	Acknowledged by	Escalated to	Comments	JIRA	CMP
Critical	PROBLEM	21 Jul 2016 13:57:19	17 s	UPTC-POD8	UPTC Service Web check FAILED for this POD	Acknowledged					
Warning	PROBLEM	21 Jul 2016 13:57:19	17 s	ia01-e02-tas28 (A)	Zabbix SNMP is not available	Acknowledged					
Warning	PROBLEM	21 Jul 2016 13:57:17	20 s	ia01-e01-vsr38	psmon2: "pscheck" failed and is restarting kamailio daemon	Acknowledged					
Warning	PROBLEM	21 Jul 2016 13:57:17	20 s	ia01-e01-vsr58	psmon2: "pscheck" failed	Acknowledged					
Critical	PROBLEM	21 Jul 2016 13:57:16	21 s	ia01-e01-vsr38	psmon2: "rhealthcheck" failed and is restarting kamailio daemon	Acknowledged					
Critical	PROBLEM	21 Jul 2016 13:57:16	21 s	ia01-e01-vsr58	psmon2: "rhealthcheck" failed	Acknowledged					
Warning	PROBLEM	21 Jul 2016 13:57:13	24 s	ia01-e01-vsr48	psmon2: "pscheck" failed	Acknowledged					
Warning	PROBLEM	21 Jul 2016 13:57:12	24 s	ia01-e01-tas14 (S)	Zabbix SNMP is not available	Acknowledged					
Warning	PROBLEM	21 Jul 2016 13:57:11	27 s	ia01-e01-vsr31-32-vip1	Failed setting GUD statistics through SNMP	Acknowledged					
Warning	PROBLEM	21 Jul 2016 13:57:09	28 s	ia01-e01-vsr43	Cluster node changes its status more than 1 time in 10 minutes	Acknowledged					
Critical	PROBLEM	21 Jul 2016 13:57:09	28 s	ia01-e01-vsr43	Cluster node works not properly	Acknowledged					
Warning	PROBLEM	21 Jul 2016 13:57:09	28 s	ia01-e01-tas06 (S)	Sumologic collector is not running	Acknowledged					
Warning	PROBLEM	21 Jul 2016 13:57:07	18 s	ia01-e01-tas08 (A)	Resource Manager SNMP service is unavailable	Acknowledged					
Critical	PROBLEM	21 Jul 2016 13:57:07	28 s	ia01-e05-ivr04 (S)	HTTP:80 unreachable	Acknowledged					
Critical	PROBLEM	21 Jul 2016 13:57:06	29 s	ia01-e05-ivr04 (S)	HTTP:8088 unreachable	Acknowledged					
Warning	PROBLEM	21 Jul 2016 13:57:06	31 s	ia01-e05-tas16 (S)	Zabbix SNMP is not available	Acknowledged					
Critical	PROBLEM	21 Jul 2016 13:57:05	30 s	ia01-e03-ad001 (S)	ICMP packet loss > 20%	Acknowledged					
Warning	PROBLEM	21 Jul 2016 13:57:05	31 s	ia01-e08-ivr03 (S)	Zabbix SNMP is not available	Acknowledged					
Warning	PROBLEM	21 Jul 2016 13:57:04	21 s	ia01-e01-ibc11b	Cached registrations number more than 65,000	Acknowledged					
Warning	PROBLEM	21 Jul 2016 13:57:04	32 s	ia01-e04-tas42 (S)	Zabbix SNMP is not available	Acknowledged					
Critical	PROBLEM	21 Jul 2016 13:57:04	34 s	ia01-e02-ivr06 (A)	HTTP:80 unreachable	Acknowledged					
Critical	PROBLEM	21 Jul 2016 13:57:03	32 s	ia01-e07-ad001 (S)	ICMP packet loss > 20%	Acknowledged					
Warning	PROBLEM	21 Jul 2016 13:57:02	34 s	ia01-e04-tas59 (S)	Zabbix SNMP is not available	Acknowledged					
Warning	PROBLEM	21 Jul 2016 13:57:02	34 s	ia01-e05-tas66 (S)	Zabbix SNMP is not available	Acknowledged					
Warning	PROBLEM	21 Jul 2016 13:57:02	34 s	ia01-e07-bla12 (S)	SSH server is down on ia01-e07-bla12	Acknowledged					
Warning	PROBLEM	21 Jul 2016 13:57:02	35 s	ia01-e01-vsr55	Cluster node changes its status more than 1 time in 10 minutes	Acknowledged					
Warning	PROBLEM	21 Jul 2016 13:57:02	36 s	ia01-e08-ivr01 (S)	psmon2: "pscheck" failed and is restarting kamailio daemon	Acknowledged					
Warning	PROBLEM	21 Jul 2016 13:57:01	34 s	ia01-e03-ivr04 (S)	Zabbix SNMP is not available	Acknowledged					

A few statistics

👉 Problems (repetitive Events of any nature) provide a lot of noise



Causes:

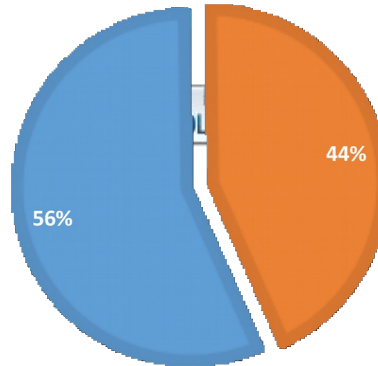
- Development issues
- Operations issues
- Trigger expression misconfiguration

A few statistics

👉 Flapping events, a lot of flapping events

EVENT QUANTITY SATURATION BY AGE

■ Age above 120s ■ Age under 120s



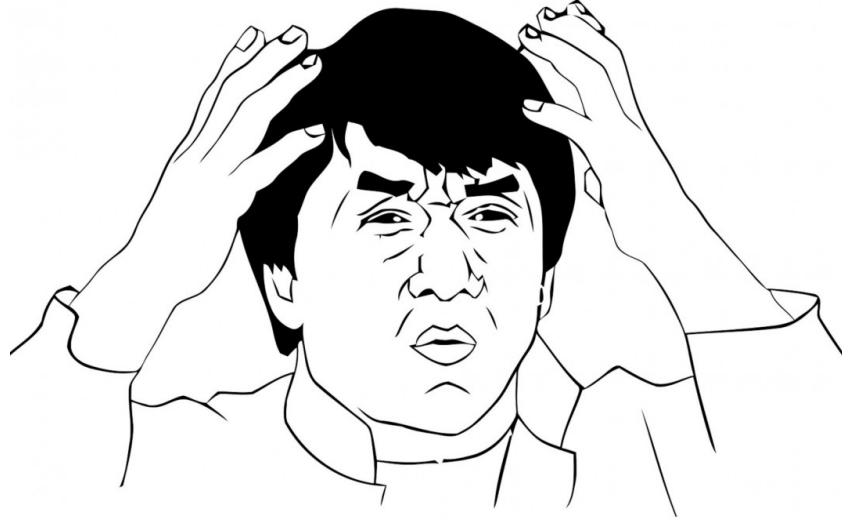
Issue processing

☞ 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?



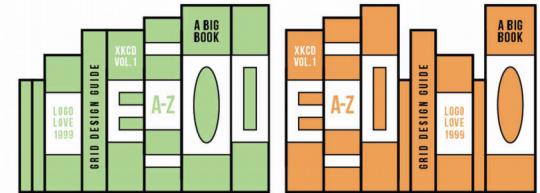
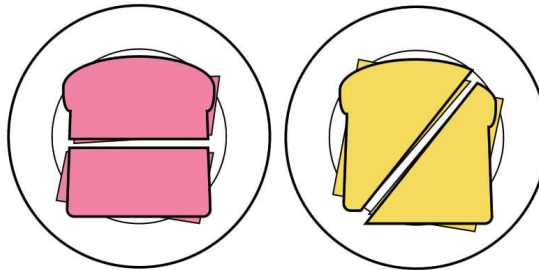
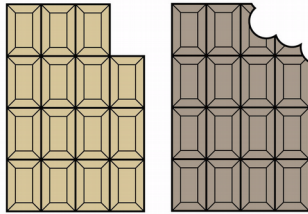
And what do we want?

- 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



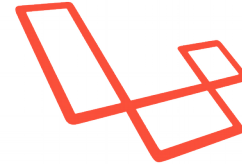
Armory

Skills:

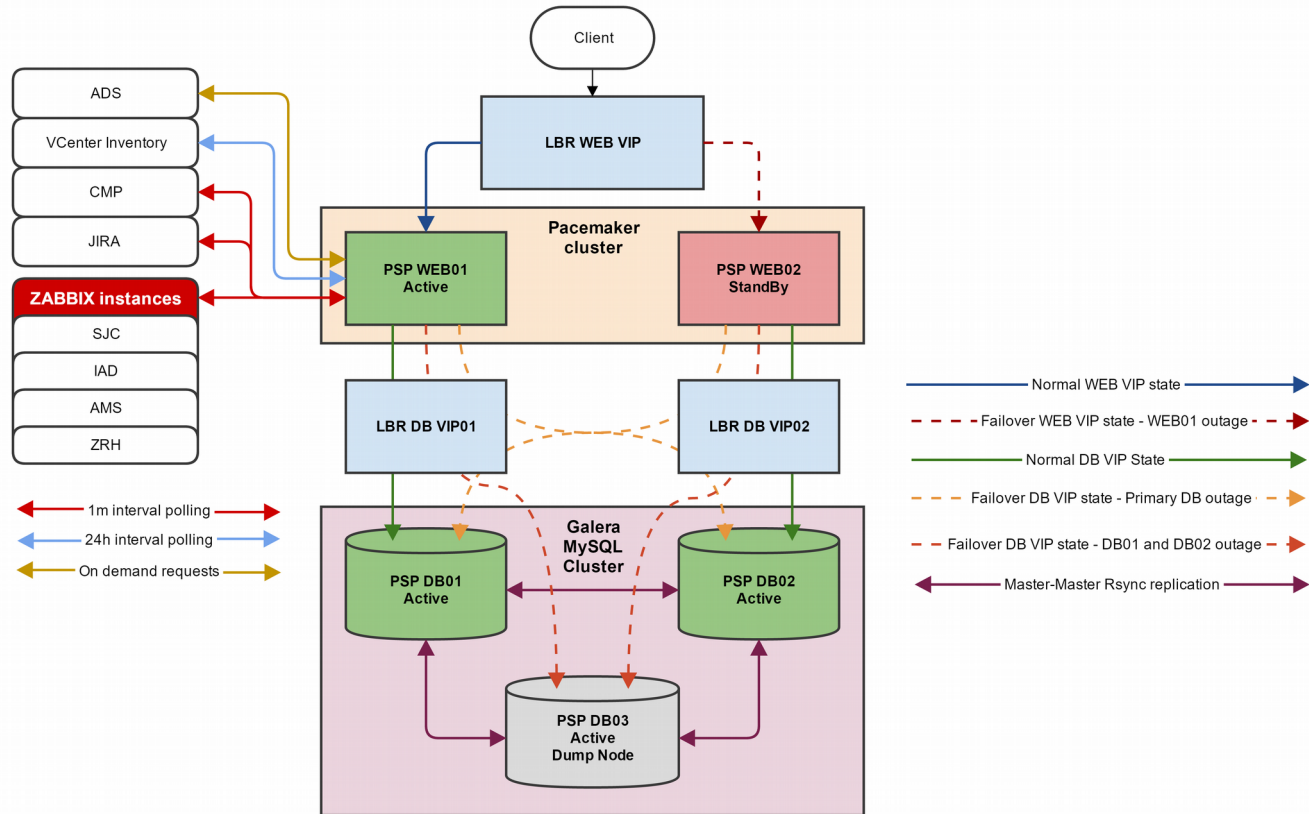
- A lot of experience as Monitoring operator
- Experience in issues investigation
- Passion to make life easier

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

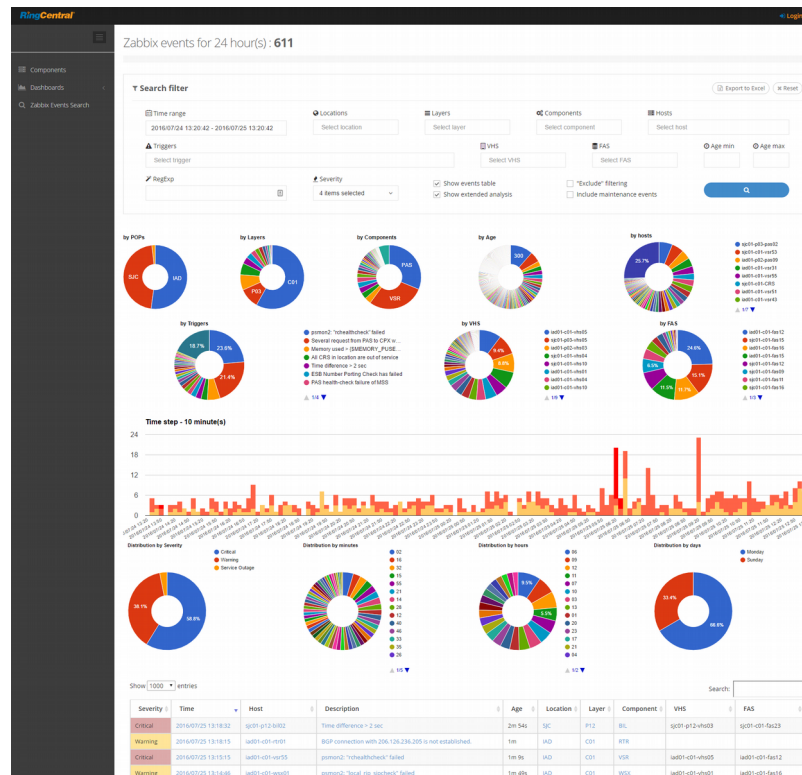
- Infrastructure relations
- Age, we want to know how fast our alerts are being resolved

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

RingCentral

Components

Dashboards

Zabbix Events Search

Internal Tools

PS Tools

konstantinyakovlev

Actions

26 July 2016

Visible events

Informational events

Component	Severity	Trigger	IAD	SJC	AMS	ZRH	N/A	Total	Comment
PAS	Warning	Several request from PAS to CPX were failed	158	145				303	OPS-63022
VSR	Critical	psmon2: "rchealthcheck" failed	75	23				98	Add comment
ALL (15)	Critical	ICMP packet loss > 20%	14	7	53		4	78	Add comment
PAS	Critical	PAS health-check failure of FCC	42	10				52	Add comment
WSX	Warning	psmon2: "local_rip_sipcheck" failed	51					51	Add comment
N/A	Warning	UPTC SBC FQDN - sip219.ringcentral.com is DOWN					49	49	Add comment
N/A	Warning	UPTC SBC FQDN - sip311.ringcentral.co.uk is DOWN					49	49	Add comment
PAS	Critical	PAS health-check failure of Bitly			30	19		49	Add comment
N/A	Warning	UPTC SBC FQDN - sip411.ringcentral.co.uk is DOWN					47	47	Add comment
N/A	Warning	UPTC SBC FQDN - sip115.ringcentral.com is DOWN					45	45	Add comment
N/A	Warning	UPTC SBC FQDN - sip113.ringcentral.com is DOWN					45	45	Add comment
N/A	Warning	UPTC SBC FQDN - sip114.ringcentral.com is DOWN					45	45	Add comment
N/A	Warning	UPTC SBC FQDN - sip111.ringcentral.com is DOWN					45	45	Add comment
N/A	Warning	UPTC SBC FQDN - sip118.ringcentral.com is DOWN					44	44	Add comment
N/A	Warning	UPTC SBC FQDN - sip117.ringcentral.com is DOWN					44	44	Add comment
								3055	

2015-2016 © PSP @ 10.13.32.52 / psp.int.ringcentral.com

psupport@nordiggy.ru

Found bug? Please submit a JIRA

version: 1.5.5.4 | revision: fb11ac2aa558

Problem Management Dash demo

How does Problem Management look like

RingCentral

Components

Dashboards

Zabbix Events Search

Internal Tools

PS Tools

Problem Management

Create case

Export to Excel

Status	Name	Severity	Component	JIRA ticket	JIRA status	Age	Due date	Assignee	Events for week	Suppress	No update	
Development	JWS config.getParameters request isn't processed	Warning	JWS SC5	UIA-44466	Open	52 w		Tatyana Chaykovskaya	0	515	2/2	25 d
Development	Publishing to Pubnub is slow	Warning	CNS CSG	PLA-8607	Open	91 w			1713	672	1/5	6 d
Investigation OPS	Several request from PAS to CPX were failed	Warning	PAS CPX	OPS-63022	In Progress	2 w		Pavel Timofeev	2075	821	1/4	1 d
Development	UPTC TAS Testing Failed	Warning	TAS	UPTC-920	Open	21 w		Christopher Palce	48	0	0/1	48 d
Investigation OPS	Tail drops in US DataCenters	Warning	DSW	OPS-59160	Open	12 w		Igor Enyushin	9	212	31/31	11 d
Investigation OPS	Busy regular workers threshold exceeded	Warning	CNS	OPS-59801	In Progress	10 w		Iurii Miroshnichenko	623	0	2/2	12 d
Development	TAS crashes and restarts related events	Critical	TAS RNG CNV OEP AWS FAX JWS HMP PAS BLA	CNV-17035	Closed	20 w		Irina Tomilova	1730	0	0/13	17 d
Waiting for deploy	Multiple PAS machines experiencing rsyslog crash	Warning	PAS	OPS-61532	Resolved	5 w		Alexander Yakovlev	12	0	0/2	8h 32m
Investigation PS	Zabbix Agent is not available	Warning	RNG TAS CNV OEP AWS FAX JWS	CNV-1703	Closed	248 w		Andrey Emelin	908	0	0/7	1709 d
Open	CSG: Several requests to PAS failed	Warning	CSG	PLA-19450					0	4120	1/1	

2015-2016 © PSP @ 10.13.32.52 / psp.int.ringcentral.com

Found bug? Please submit a JIRA

version: 1.5.3.4 | revision: fb11ac2aa558

Problem Management Dash demo

Inside Problem management case

RingCentral

Components

Dashboards

Zabbix Events Search

Internal Tools

PS Tools

konstantin.yakovlev

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

MergeEdit

SUMMARY

Case status: Investigation OPS

JIRA / CMR / INC: OPS-63022

Affected components: PAS CPX

Affected layers/pods: P01 P02 P03 P04 P05 P07 P08 P09 P10

Tagged events: 1829

Created: 2016-02-01 16:46:06

Cases merged in: #223

Related triggers (suppressed by 1 rules)

Trigger description	Component	Severity	Events	Suppress	Remove
CPX response time too long	PAS	393	193 +162		
HTTP request rejected	CPX	62	0 0		
PAS request is failed	CPX	82	95 +31		
Several request from PAS to CPX were failed	PAS	368 32	1787 +126		

COMMENTS

Type a message here...

artem.akinchits at 2016-07-13 09:04:08

Suppression requested in TLS-8538

konstantin.yakovlev at 2016-07-11 15:15:04

Created new OPS-63022 as we have the issue still, requested suppression

sergey.smirnov at 2016-06-07 12:11:50

No updates, re-requested

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

20

RingCentral®

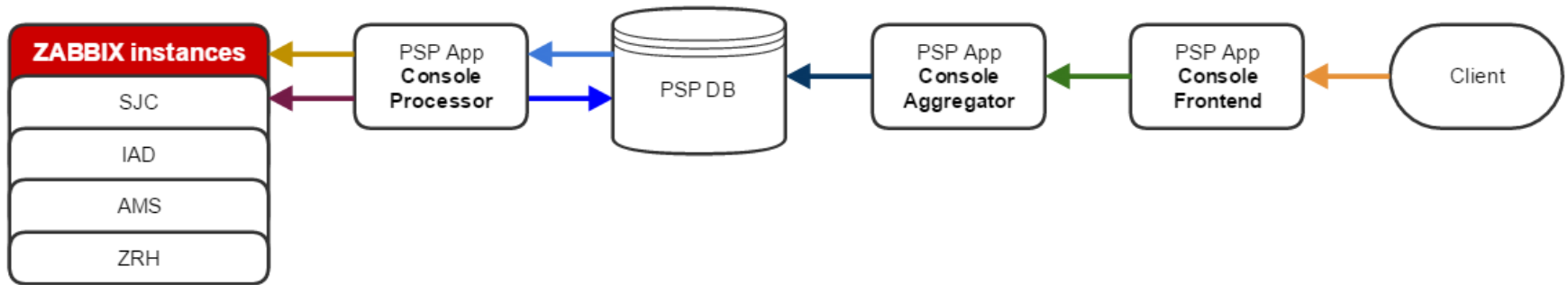
Incident Management

Event processing upon receipt



Console Scheme

☞ 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

Wed, 27 Jul 2016 12:20:16 GMT  less than a minute ago

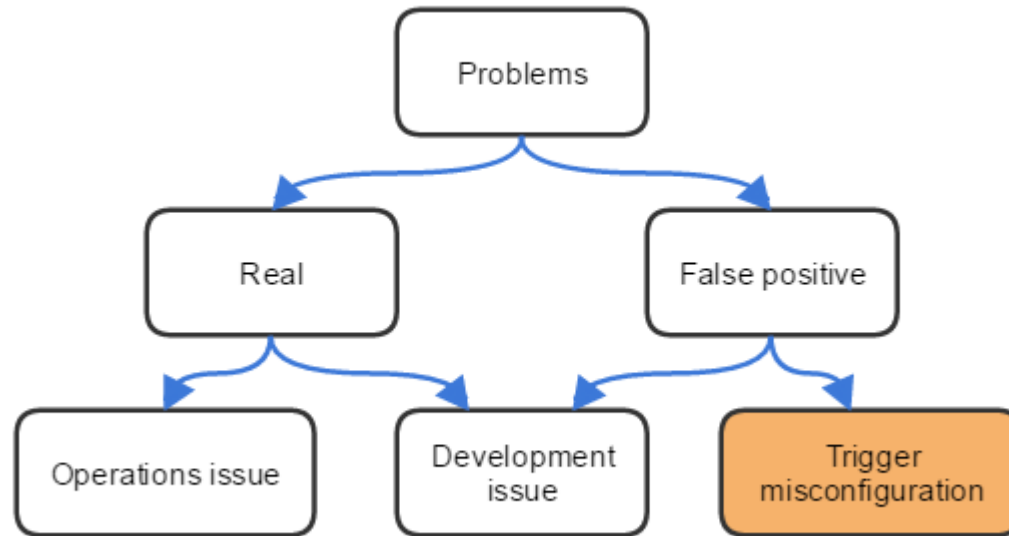
 **Warning** 4 **Critical** 8 **Service outage** 0 **Disaster** 0 **Acknowledged** 11/12

Bulk **Expand** 

Count	Trigger	Ack	Jira	KE	Esc	PM	First seen	Last seen	Location	Layer	Components
1	iad01-c01-rtr01: Packet loss at VoIP WAN_ord01-rtr02_XO:133683 NET						1m 59s	1m 59s	IAD01	C01	RTR
1	UPTC_POD1: UPTC Inbound SMS Test FAILED [16502496311 -> 12054195068]						54m 43s	54m 43s	N/A	N/A	N/A
1	zrh01-c01-mon03: Free swap size < 35%						2h 19m	2h 19m	ZRH01	C01	MON
1	sjc01-c01-ibs01: AUS: Total amount of pending requests was increased						3h 18m	3h 18m	SJC01	C01	IBS
1	ams01-c01-rtr02: ams01-c01-rtr02: xe-0/0/0 (ams01-c01-rtr02:xe-0/0/0) link is down: xe-0/0/0 (ams01-c01-rtr02:xe-0/0/0) link is down NETWORK RELATED						1d 4h	1d 4h	AMS01	C01	RTR
1	iad01-c01-lbr02: Certificate SSL on LBR is missing or expired NET						1d 4h	1d 4h	IAD01	C01	LBR
1	sjc01-puppet.ringcentral.com: LMD384067 warn - sjc01-puppet.ringcentral.com AAR-sjc01-ci05 Status LOGIC MONITOR						3d 13h	3d 13h	SJC01	N/A	N/A
4	INC-17639						5d 17h	5d 2h	IAD01, AMS01	P12, C01	BLA MSW
1	iad01-p13-vhs03: DELL server overall status is critical S VHS						5d 22h	5d 22h	IAD01	P13	VHS

The last but not the least

Going back to problems



Retrospective trigger analyzer

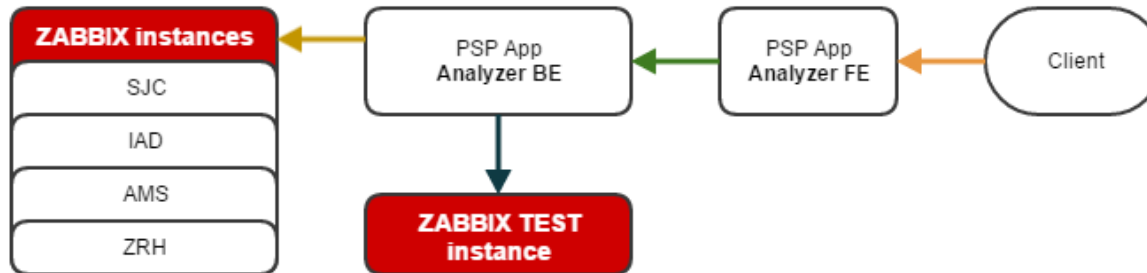
Features:

- 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

RingCentralkonstantin.yakovlev

Components

Dashboards

Zabbix Events Search

Internal Tools

PS Tools

Retrospective Zabbix Trigger Analyzer

Create new trigger expression > Select hosts and items > **Select trigger from existing one** > Get analyzed data

Input zabbix trigger expression:
{A.last()}>0
A block of help text that breaks onto a new line and may extend beyond one line.

Common items: A: Unique items: :

Next step

Select component:
JWS

Select hosts:
Filter hosts
sjc01-p13-jws03
sjc01-p13-jws06
sjc01-p13-JWS
sjc01-p09-jws05
sjc01-p09-jws06
sjc01-p10-jws06
sjc01-p10-jws05
sjc01-p06-jws05
sjc01-p06-jws06

Selected hosts:
Filter hosts
sjc01-p11-jws03
sjc01-p11-jws04
sjc01-p11-jws05
sjc01-p11-jws06

Next step

Select common items:
A Failed step of scenario "S1".

Next step

Select trigger:
JWS config.getParameters request isn't processed

Next step

Select timeframe:
2016/7/21 12:37:46 - 2016/7/27 12:37:46

Analyze

2015-2016 © PSP @ 10.13.32.52 / psp.intrincentral.com psupport@nordigy.ru Found bug? Please submit a JIRA version: 1.5.5.4| revision: fb11ac2aa558

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

27

RingCentral

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!

