AUTOMATIC ROOT CAUSE ANALYSIS IN MOBILE NETWORKS

Julio Arpírez CTO @ Zhilabs jarpirez@zhilabs.com



OUTLINE

- The Problem
- The State
- The Analysis
- The Conclusion

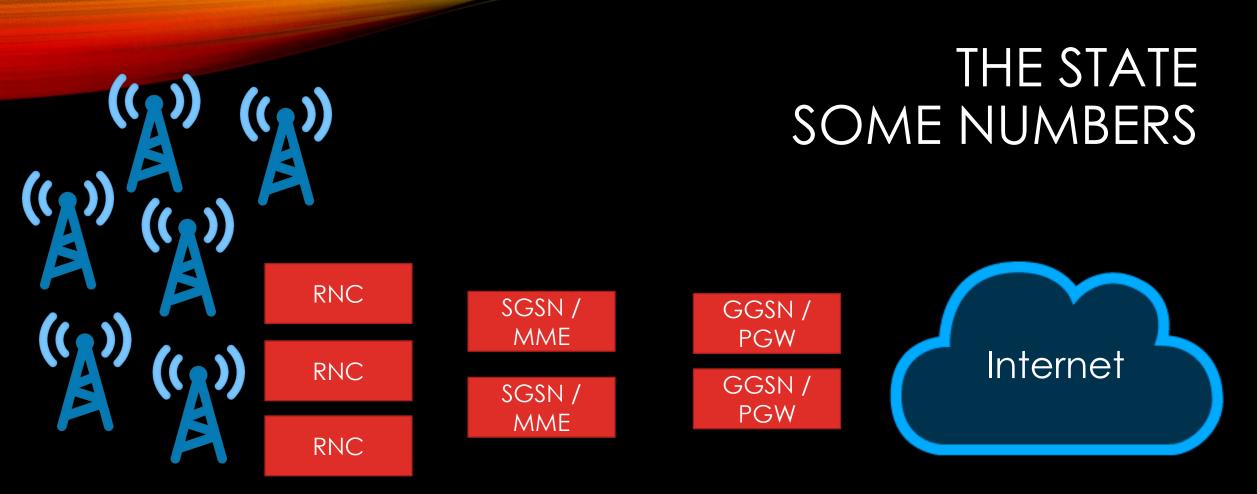
THE PROBLEM

- Today's mobile networks should ideally
 - Be fully reliable and accessible (ideally 100% of the time)
 - Be usable for a wide variety of services (e.g. CS Voice, SMS, PS, VoLTE)
 - Serve subscribers and machines through different technologies (2G/3G/4G) and devices
 - Provide perfect communications for person-to-person, person-to-server (e.g. OTT services), M2M and IoT
- A network is a very complex system and consequently problems arisen are often difficult to diagnose and correct
 - All of us know the WAR-ROOM to WAR-ROOM game
- Problems in the network have a <u>huge direct financial impact</u> (revenue loss and the resources associated to operate and troubleshoot the network) as well as <u>indirect</u> (loss of credibility and impact on the brand)

THE STATE

Network State (Huge Dimension ality)

Tolerable Operation Universe Network State (Huge Dimension ality) Problematic Operation Universe → Root Cause Analysis and Fixing



50000 e/nodeB x 100 RNC x 20 SGSN/MME x 10 GGSN/PGW = 1000 Million Service Paths 10 State Variables x 1000 Million = <u>10000 Million Options to Explore</u>

THE STATE ADD MORE ITEMS...

Network Related

- Technology Type (2G/3G/4G)
- Policy servers
- OCS Systems
- DPI Nodes
- Optimizers

 \bullet

Device & OTTs

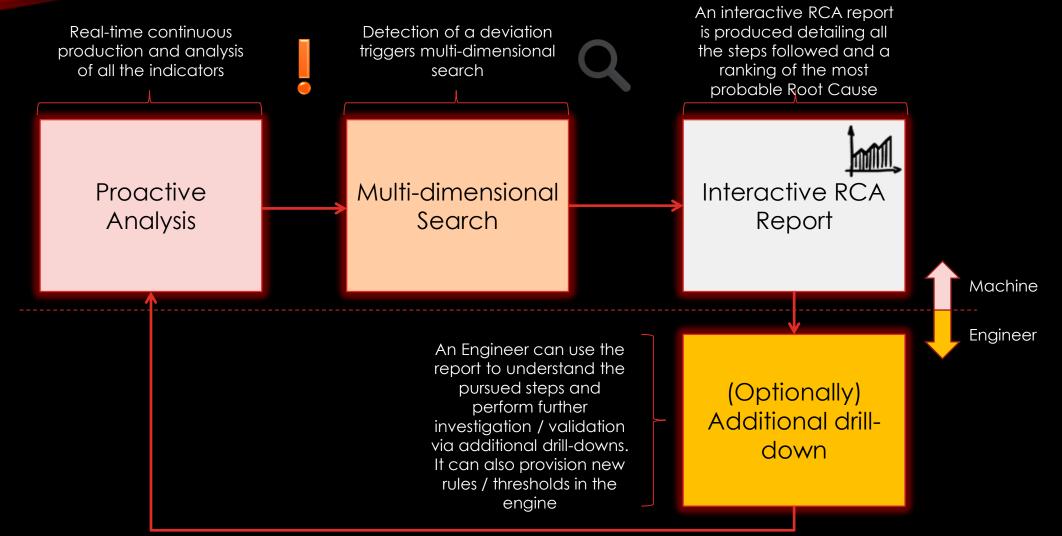
- Device types & models (typically a small network has around 10000 active models)
- Internet servers providing fancy services (e.g. Google, Facebook, YouTube, Periscope, Twitter)
- OTT Apps (e.g. WhatsApp, SnapChat)

Keep on multiplying... The number of combinations to explore in quest of failures is unmanageable © Zhilabs 2015

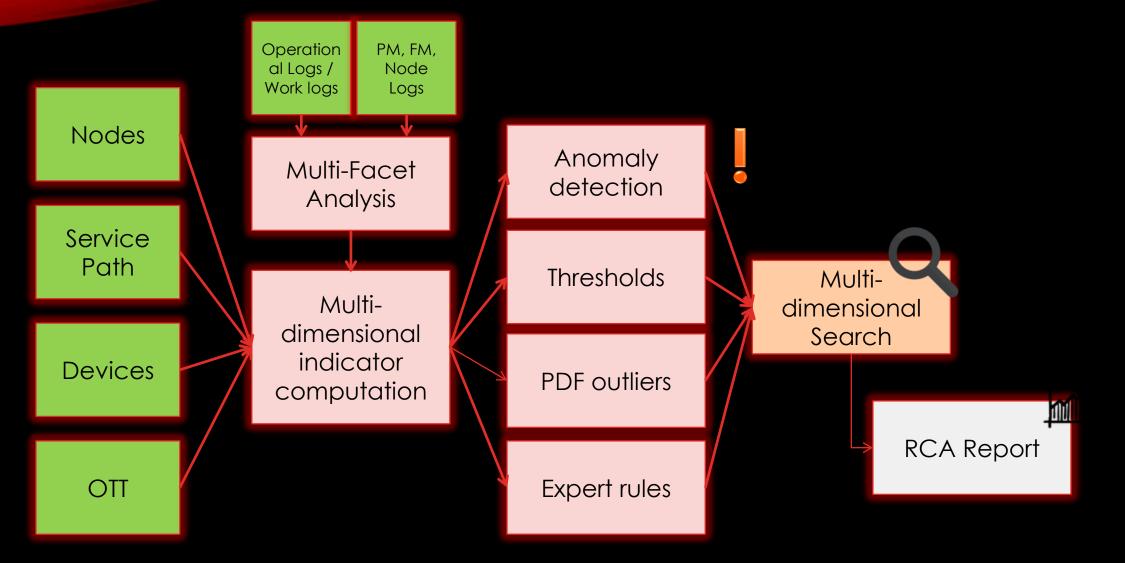
THE STATE WHAT'S NORMAL

- Given the state, we need to identify the boundaries of the Tolerable Operation Universe
- Typically, this **Tolerable Operation Universe** is delimited by a
 - **Baseline for certain KQIs/KPIs** (dimensionality reduction) Typical KQIs include accessibility, retainability, speed, etc.
 - + Confidence Interval
- The <u>Confidence Interval</u> may be
 - Hard SLAs (typically, human-defined)
 - **Soft** Automatic Baseline Computation
- Deviations from the baseline + confidence interval are usually due to problems (deviations from normal behavior)

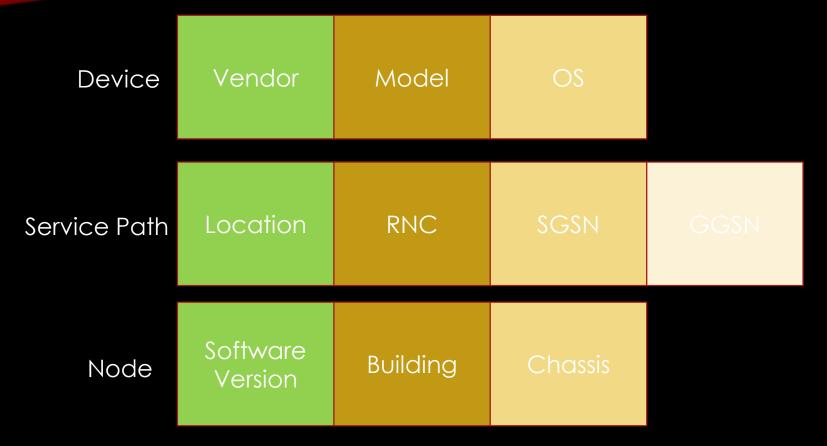
ROOT CAUSE ANALYSIS THE FLOW



AUTOMATIC RCA: INTERNALS

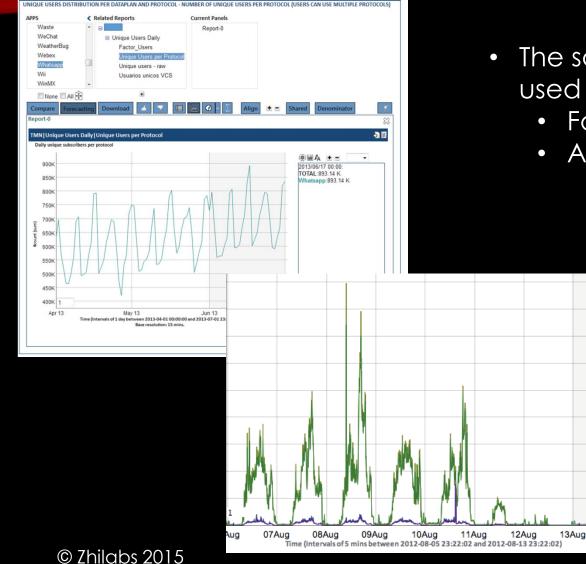


ROOT CAUSE ANALYSIS MULTI-DIMENSIONAL SEARCH: SOME EXAMPLES

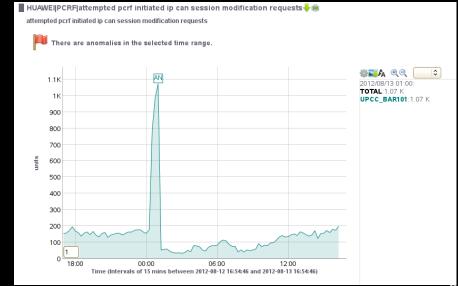


- Indicators are produced for all the dimensions
- A search of the values at the different dimensions allows to isolate origin of the deviations

AUTOMATIC RCA: ANOMALY DETECTION



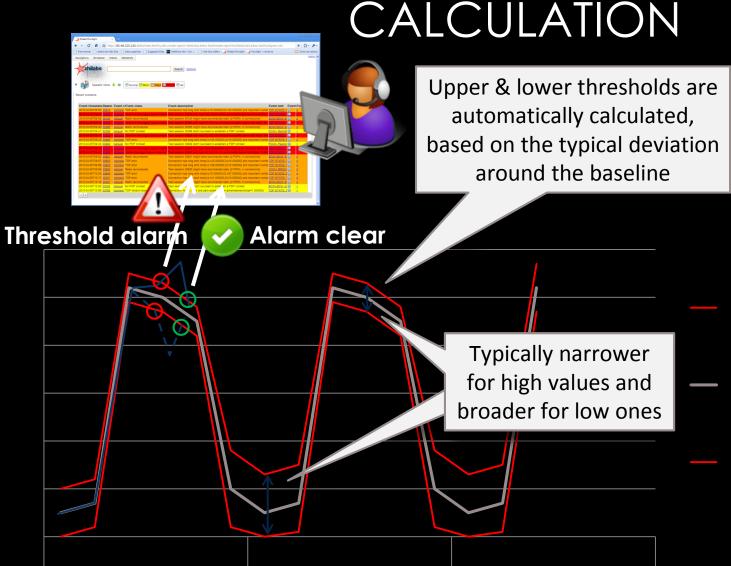
- The same algorithms available in FlowSight are used by the engine under the covers:
 - Forecasting: ARIMA
 - Anomalies: cyclostationary analysis, AR



FORECASTING AND ANOMALY DETECTION AUTOMATIC THRESHOLDS - BASELINE

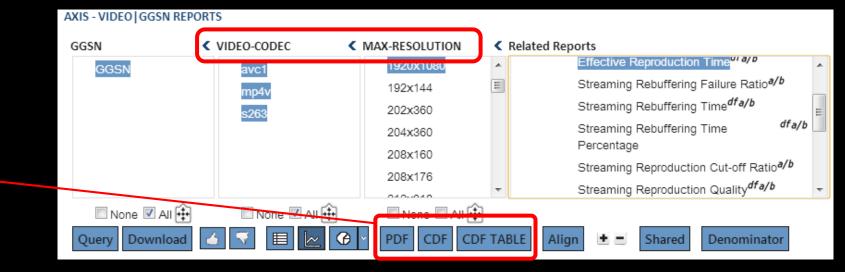
 AR models are good at detecting anomalies

 ARIMA Methods take into consideration seasons and trends

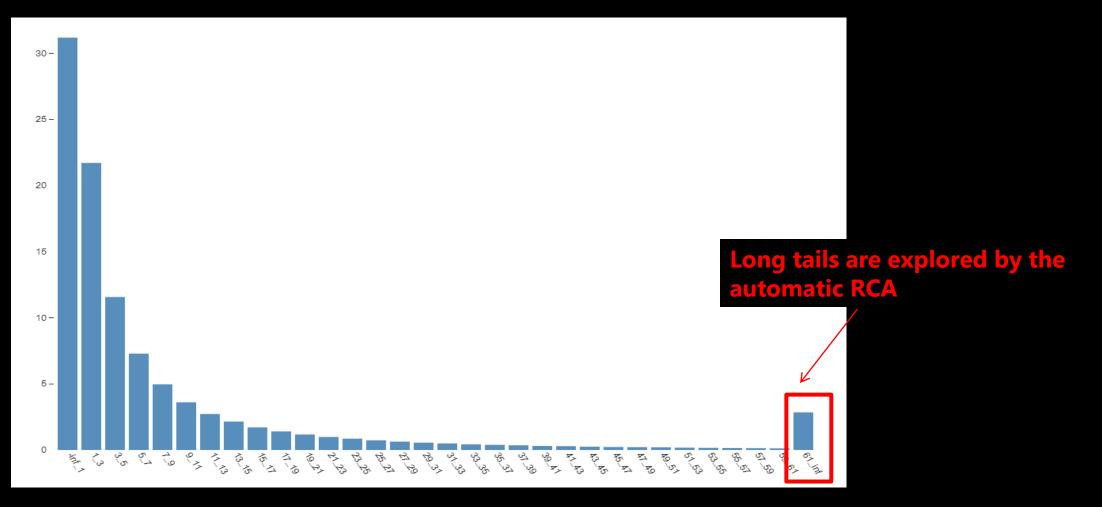


MULTI-DIMENSIONAL ANALYSIS AND PDFS

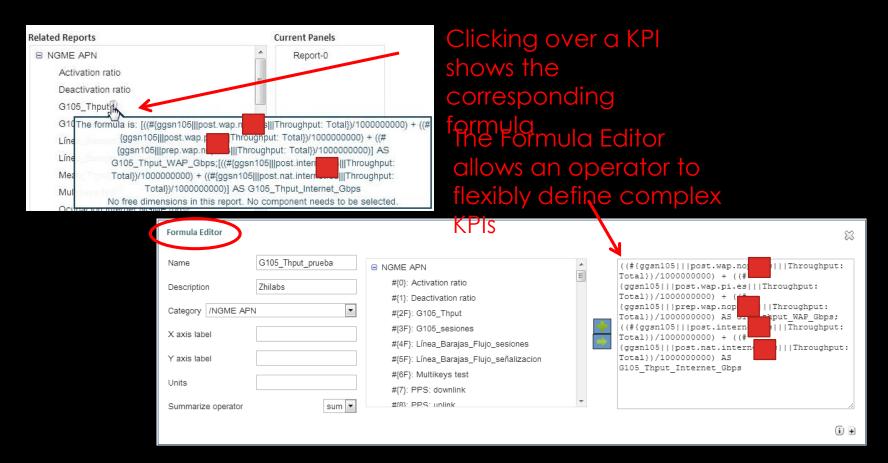
The system computes automatically **Distribution Functions (PDF and CDF)**.



PROBABILITY DISTRIBUTION FUNCTIONS:



EXPERT RULES: FORMULAS¹⁵



AUTOMATIC RCA: DETAILS¹⁶

1. Proactive Analysis:

- Input: Nodes, Service Path, Devices, OTT
- Additional inputs (multi-facet analysis):
 - Operational logs / Work logs (detailing changes performed in the network)
 - Performance Mgmt / Fault Mgmt / Logs from the different nodes
- Multi-dimensional Indicator production
- Four elements:
 - Anomaly detection (autoregressive and cyclostationary algorithms)
 - Thresholds
 - PDF outliers
 - Expert Rules

2. Multi-dimensional Search:

- Multi-dimensional analysis of deviation to locate commonalities
- Determination of most probable root-cause (ranking process)

3. Interactive RCA report:

- Details steps (1) and (2) via clickable reports
- Allows an engineer to validate the RCA and/or further investigate from the machine-produced baseline

ROOT CAUSE ANALYSIS MULTI-DIMENSIONAL SEARCH

d1

d3

d2

d

- 1. Pattern similarity:
 - Uses time-range/indicator that signaled the anomaly as an input
 - Computes similarity for that scope in other dimensions
 - Correlation used as a similarity function between the time series
- 2. Ranking:
 - Dimensions effectively become a loosely-coupled service graph graded by the similarity function
 - A variation of Personalized PageRank algorithm is applied:
 - Random walk
 - Similarity score used to modulate transitions and teleportation

Intuitively equivalent to several Engineers drilling-down through the different dimensions driven by similarities in the time series.

THE ANALYSIS

Automatic Root Cause Analysis Demos

AUTOMATIC ANOMALY DETECTION APPLE SERVICES DOWN

🗲 🔇 Time Reports Apps Setup			😽 Logged a	admin Logout
Query Input Pivot				
m Time from=2015-03-03T00:00:00 to=2015-03-13T23:59	9:59 sourcetype=report.transactions-site-900-timeline chart annotated-timeline http.clicks anomaly tra	ain 1 -order=48	BB	Q Search
Zoom m 3m 6m YTD 1y All	From Mar 3, 2015 To Mar 13, 2015	🕹 Export		
		TIMESTAMP	GLOBAL	SUM(COMMON
	200k	2015-03-03T00:00:00	GLOBAL	120.72K
Λ , Λ , Λ		2015-03-03T01:00:00	GLOBAL	109.11K
		2015-03-03T02:00:00	GLOBAL	103.84K
		2015-03-03T03:00:00	GLOBAL	102.02K
		2015-03-03T04:00:00	GLOBAL	103.89K
				106.21K
Saturday. M		2015-03-03T06:00:00	GLOBAL	124.46K
		2015-03-03T07:00:00	GLOBAL	157.96K
	3 75k	2015-03-03T08:00:00	GLOBAL	180.78K
	50k	2015-03-03T09:00:00	GLOBAL	191.09K
		2015-03-03T10:00:00	GLOBAL	190.86K
3. Mar 4. Mar 5. Mar 6. Mar 7. Ma	25k ar 8. Mar 9. Mar 10. Mar 11. Mar 12. Mar 13. Mar	2015-03-03T11:00:00	GLOBAL	190.69K
		2015-03-03T12:00:00	GLOBAL	191.14K
I 4. Mar 6. Mar	8. Mar 10. Mar 12. Mar	2015-03-03T13:00:00	GLOBAL	184.93K ¥
4. Mai 0. Mai			🔹 1-50 of 264 🕟	H

RCA: USE CASES INTRODUCTION









ANOMALIES

RCA WALL

🕅 🔇 Time 🛛 Re	aports Apps Setup	Logged as admin	Logout
Query Input	Pivot		
Time Current	query: Network Anomaly in HTTP Transactions Indicator at 20140222194500	Q Sea	rch

Anomalies are presented in the main WALL by Anomaly Detector, showing:

- Anomaly Level (e.g Network)
- Indicator (e.g HTTP Transactions)
- Anomaly Date (e.g at

20140222194500)

WALL

Anomalies
Network Anomaly in HTTP Transactions Indicator at 20140222194500 (admin) ×
Network Anomaly in PDP Context Activation Failure Ratio Indicator at 20140823034500 (admin) ×
Network Anomaly in RTT Client Indicator at 20140823210000 (admin) ×
Network Anomaly in Time-To-Stream-Start Indicator at 20141014220000 (admin) ×

Top Contributors

admin (Last update @ Wed Dec 10 23:33:28 GMT+100 2014)

RCA NETWORK ANOMALY IN PDP CONTEXT ACTIVATION FAILURE RATIO INDICATOR

🕻 🔇 Time Reports Apps Setup	Logged as admin Logout
Query Input Pivot	
🛗 Time Current query: Network Anomaly in PDP Context Activation Failure Ratio Indicator at 20140823034500	Q Search
國 Network Anomaly in PDP Context Activation Failure Ratio Indicator at 20140823034500	
Zoom 1m 3m 6m YTD 1y All	When accesing to the anomaly report, a timeline is shown for a longer period
	rather than only the anomaly itself in
	order to check the signal evolution of the
ANOMALY	presented indicator
$\overline{\mathbf{A}}$	
h	
23. Aug 03:00 06:00 09:00 12:00 15:00 18:00 21:00	24. Aug 03:00 06:00 09:00 12:00 15:00 18:00 21:00 25. A

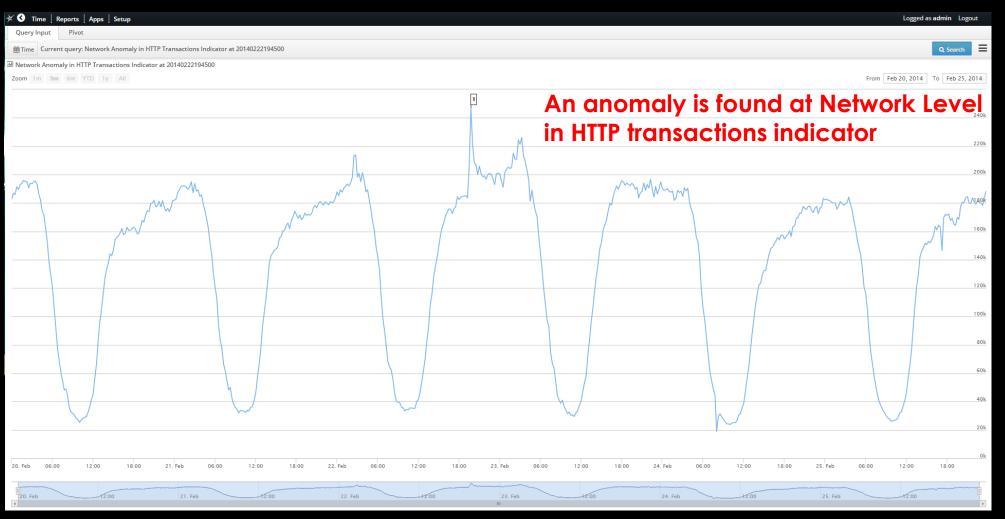
RCA NETWORK ANOMALY IN PDP CONTEXT ACTIVATION FAILURE RATIO INDICATOR

🗲 🔇 Time 🛛 Reports	Apps Setup	Logged as admin Logout
Query Input Pivo	ot	
🛗 Time Current query	: Network Anomaly in PDP Context Activation Failure Ratio Indicator at 20140823034500	Q Search
Network Anomaly in P Zoom 1m 3m 6m	DP Context Activation Failure Ratio Indicator at 20140823034500	By clicking the 'anomaly flag', a popup. 2014 To Aug 25, 201
Saturday, Aug 23, 03:45 GLOBAL PDP Context Creation Fail	ure Ration higher than 17%	panel appears listing all anomalies for the selected period and also allowing to
(Anomaly Information	•
	List of anomalies for GLOBAL on Sat Aug 23 03:45:32 GMT+200 201 1: PDP Context Creation Failure Ration higher than 17% Perform Root Cause Analysis	perform the Root Cause Analysis of this anomaly
$\overline{\mathcal{N}}$		
	hh	
23. Aug 03:00	06:00 09:00 12:00 15:00 18:00	21:00 24 Aug 03:00 06:00 09:00 12:00 15:00 18:00 21:00 25. A

RCA NETWORK ANOMALY IN PDP CONTEXT ACTIVATION FAILURE RATIO INDICATOR

🗲 🔇 Time 🛛 Reports	Apps Setup					Logged as admin Logou	t
PROACTIVE ANALYS	SIS	In this case, RAT-TYPE =	Root c	ause an	alysis is perf	ormed and	
ACT #1: PDP Context C	Creation Failure Ratio higher than 17% from 2014-08-23T03:45:00 t	° 2014-08-2GERAN shows the			· ·		
ULTI-DIMENSIONAL A	ANALYSIS		systen	n snows d	as a dashbo	ara me	
ACT #1: PDP Context C	Creation Failure Ratio higher than 17%	highest correlation	FACTS	and the	Possible Iss	ues found	
ltem	Value	PDP Context Cretion Failure Ratio (%)		anomal			
Dimension	RAT				~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~		~
Affected Values	GERAN	4 0 4 04:00 08:00 12:00 16:00	20:00 24. Aug	04:00 08:0	00 12:00 16:00	20:00 25. Aug	
Time Range	2014-08-23T03:45:00 to 2014-08-25T03:59:59	-Fail	GER	AN - UTRAN			
ltem	Creation Failure Ratio with RAT=GERAN higher than 20%	PDP Context Creation Failure Ratio by affected APN's (%)	1 1 1		G	tem drills-do JERAN for the mensions an	e Apn
Dimension	APN				73,21	nensions an	a muas
Affected Values	POSTPAID3 PREPAID MESSAGING	29.52				issues	
Time Range	2014-08-23T03:45:00 to 2014-08-25T03:59:59	0 5 10 15 20 25 30 35	40 45 50 !	5 60 65 70	75 80 85 90	95 100 MESSAGING	
					PDP-Context-Creation-Failu	re-Ratio-(%)	
ROOT-CAUSE ANALYSIS	5						
						And APNs fo	und by
Ranking #1	Possible issue at	Ranking #2 Possible issue at		Ranking #3	Possible issue at	are present	
Dimension	RAT & APN	Dimension RAT & APN		Dimension	RAT & APN	•	
Affected Values	GERAN & POSTPAID3	Affected Values GERAN & PREPAID		Affected Values	GERAN & MESSAGING	of ranke	a result
Time Range	2014-08-23T03:45:00 to 2014-08-25T03:59:59	Time Range 2014-08-23T03:45:00 to 2014-08-25T03:59:59		Time Range	2014-08-23T03:45:00 to 2014-08-25	T03:59:59	

RCA NETWORK ANOMALY IN HTTP TRANSACTIONS INDICATOR



RCA NETWORK ANOMALY IN HTTP TRANSACTIONS INDICATOR

Network Anomaly in HTTP Transactions Indicator at 20140222194500 Zoom Im 3m 6m YTD Iv All Saturday, Feb 22, 19:45 GLOBAL Abrupt HTTP Transactions Indicator variation found Anomaly Information List of anomalies for GLOBAL on Sat Feb 22 19:45:00 GMT+100 2014 1: Abrupt HTTP Transactions Indicator variation found	Search ====================================
	Eeb 25, 2014
Zoom 1m 3m 6m YTD 1y All Saturday, Feb 22, 19:45 From Feb 20, 2014 To Abrupt HTTP Transactions Indicator variation found Anomaly Information Image: Comparison of the compari	250k
GLOBAL Abrupt HTTP Transactions Indicator variation found Anomaly Information List of anomalies for GLOBAL on Sat Feb 22 19:45:00 GMT+100 2014 1: Abrupt HTTP Transactions Indicator variation found	250k
Abrupt HTTP Transactions Indicator variation found Anomaly Information List of anomalies for GLOBAL on Sat Feb 22 19:45:00 GMT+100 2014 1: Abrupt HTTP Transactions Indicator variation found	
List of anomalies for GLOBAL on Sat Feb 22 19:45:00 GMT+100 2014 1: Abrupt HTTP Transactions Indicator variation found	
List of anomalies for GLOBAL on Sat Feb 22 19:45:00 GMT+100 2014 1: Abrupt HTTP Transactions Indicator variation found	225k
1: Abrupt HTTP Transactions Indicator variation found	225k
	,
Perform Root Cause Analysis	200k
Man Mar	00.1
	175k
	·V
	150k
	125k
	1238
	100k
	75k
	ſ
	50k
	ſ
	25k
20. Feb 08:00 16:00 21. Feb 08:00 16:00 22. Feb 08:00 16:00 23. Feb 08:00 16:00 24. Feb 08:00 16:00 25. Feb 08:00 16:0	0k

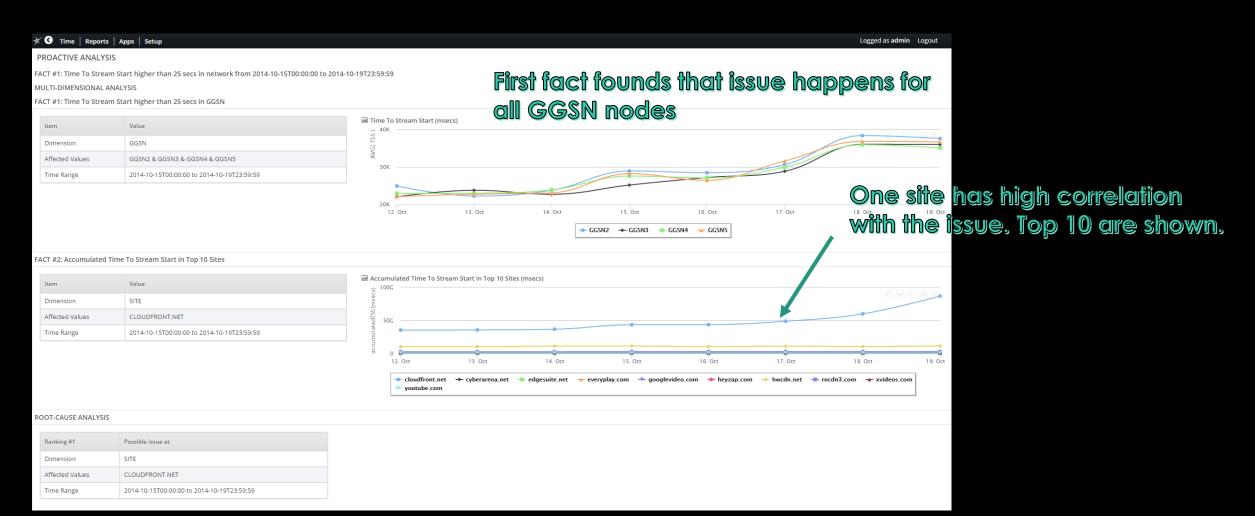
RCA NETWORK ANOMALY IN HTTP TRANSACTIONS INDICATOR

🗲 🔇 Time Reports	Apps Setup					Logged as admir	1 Logout
PROACTIVE ANALYSIS	<u>ا</u>						le le le contrate la serie de
FACT #1: Volume variatior	on found from 2014-02-22T19:45:00 to 2014-02-22T19:59:59					it is posi	<mark>ble</mark> to see how volume for
MULTI-DIMENSIONAL ANA	ALYSIS					- h h	
FACT #1: Volume variatior	on found for whatsapp.net & viber.com sites				W	narsap	<mark>o a</mark> nd viber increase during
ltem	Value	Total Volume for v	vhatsapp.net & viber.com (bytes)				he anomaly period
Dimension	SITE	(bytes		Maria			
Affected Values	whatsapp.net & viber.com	0 20. Feb	12:00 21. Feb 12:00 22. Feb 12:	00 23. Feb 12:00	24. Feb 12:00	25. Feb 12:00	Moun
Time Range	2014-02-22T19:45:00 to 2014-02-22T19:59:59	D 20. Feb			24. FED 12:00	25. Feb 12:00	
·		4	- viber.c	om 🔶 whatsapp.net			
FACT #2: Variation in HTTF	IP Transactions indicator found for whatsapp.net						
ltem	Value	HTTP transactions	for whatsapp.net				
Dimension	SITE	IONS					
Affected Values	whatsapp.net	лок — Sok					
Time Range	2014-02-22T19:45:00 to 2014-02-22T19:59:59	- TRAT		hunn			
		20. Feb	12:00 21. Feb 12:00 22. Feb 12:00	23. Feb 12:00	24. Feb 12:00	25. Feb 12:00	
FACT #3: Variation in Data	a Call Access Failure Ratio found for whatsapp.net site						
ltem	Value	Data Call Access F	ailure Ratio for whatsapp.net			seems fo	be whatsapp, suffering a
Dimension	SITE	SATIO		My			
Affected Values	whatsapp.net	- 25		- Mp	SVS	lem dov	<mark>vnti</mark> me during the anomaly
Time Range	2014-02-22T19:45:00 to 2014-02-22T19:59:59	o			0//0		
		TTTT 20. Feb	12:00 21. Feb 12:00 22. Feb 12:00	23. Feb 12:00	24. Feb 12:00	25. Feb 12:00	period
ROOT-CAUSE ANALYSIS							
Ranking #1	Possible issue at	Ranking #2	Possible issue at	Ranking #3	Possible issue at		
Dimension	SITE	Dimension	SITE	Dimension	SITE		
Affected Values	whatsapp.net & viber.com	Affected Values	whatsapp.net	Affected Values	whatsapp.net		
Affected Indicator	Total Volume	Affected Indicator	HTTP Transactions	Affected Indicator	Data Call Access Failure Ratio	,	
Time Range	2014-02-22T19:45:00 to 2014-02-22T19:59:59	Time Range	2014-02-22T19:45:00 to 2014-02-22T19:59:59	Time Range	2014-02-22T19:45:00 to 2014	-02-22T19:59:59	

RCA NETWORK ANOMALY IN TIME-TO-STREAM-START INDICATOR

C S Time Reports Apps Setup	Logged as admin Logout
Query Input Pivot	
篇 Time Current query: Network Anomaly in Time-To-Stream-Start Indicator at 20141014220000	Q Search
Network Anomaly in Time-To-Stream-Start Indicator at 20141014220000 Zoom 1m 3m 6m YTD 1y All	From Oct 12, 2014 To Oct 19, 2014
	36
An anomaly is found at Network Lev	el
in Time To Stream Start Indicator	32
Wednesday, Oct 15, 00:00 GLOBAL	igher than 25 secs in network
	Anomaly Information
	List of anomalies for GLOBAL on Wed Oct 15 00:00:00 GMT+200 2014 1: Time To Stream Start higher than 25 secs in network Perform Root Cause Analysis
12. Oct 08:00 16:00 13. Oct 08:00 16:00 14. Oct 08:00 16:00	22 15. Oct 08:00 16:00 16. Oct 08:00 16:00 17. Oct 08:00 16:00 18. Oct 08:00 16:00 19. O

RCA NETWORK ANOMALY IN TIME-TO-STREAM-START INDICATOR



RCA NETWORK ANOMALY IN RTT CLIENT INDICATOR

🗚 🔇 Time Reports Apps Setup		Logged as admin Logout
Query Input Pivot		
Time Current query: Network Anomaly in RTT Client Indicator at 20140823210000		Q Search
細 Network Anomaly in RTT Client Indicator at 20140823210000		
Zoom 1m 3m 6m YTD 1y All	From At	ug 23, 2014 To Aug 23, 2014
An anomaly is found at Network Level	Anomaly Information O	00 0.675
	List of anomalies for GLOBAL on Sat Aug 23 21:00:00 GMT+200 2014	than 650 msecs 0.65
in RTT-Client Indicator	1: RTT Client higher than 650 msecs	0.625
	Perform Root Cause Analysis	0.6
		0.0
		0.575
		0.55
		0.525
		à.5
		0.475
		0.45
		0.425
		0.4
		0.375
		0.35
04:00 05:00 06:00 07:00 08:00 09:00 10:00 11:00 12:00 13:00 14:00 15:00	16:00 17:00 18:00 19:00 20:00 21	0.325

RCA NETWORK ANOMALY IN RTT CLIENT INDICATOR



THE CONCLUSIONS

- Root Cause Analysis can be seen as a search in a huge multi-dimensional space
- Different approaches are possible
 - Human experience (e.g. rule-based systems)
 - Fully automatic root-cause analysis (ML-aided systems, like time-series correlation techniques)
 - Mixed
- Once fully automatic tools are in place, the network can become selfrepairing and adaptive, realizing the vision of self-organizing networks
- Root Cause Analysis applications do not only cover the network but also individual subscribers (e.g. Customer Care where call-dispatch time should be as short as possible)

THE CONCLUSIONS TOWARDS THE FUTURE

- Machine-Learning to the rescue!
 - To speed up troubleshooting
 - To detect offending elements before they become a real problem
- Close the Loop in the Network automatically
 - To fix problems
 - And also, to make your network a Chameleon! The socalled Self Organizing Network (SON)!
 - The network will adapt to the changing conditions and adapt itself by using the required resources (NFV)!



THANK YOU!

