Network Health Assessment – Using Big Data to Perform Network Diagnosis and Predict

Qin Wu (bill.wu@huawei.com)

Liang Zhang (zhangliang1@huawei.com)

Agenda

- What is “Network Health Assessment”
- Architecture Overview
- Network Health Indicator Use Cases
- Network Health Indicator: Network Diagnostics and Analytics Components
- Conclusion

What is “Network Health Assessment”




Blood Test Report

		Patient Value		Normal Value	
Indicator 1:	Sulphates	70.00	mg/ahmv	15 - 20	mg/ahmv
Indicator 2:	Phosphates [Total]	60.00	Gramm/ahmv	60 - 80	Gramm/ahmv
Indicator 3:	Chloride	46.00	mg/ahmv	55 - 75	mg/ahmv
Indicator 4:	Pottassium	60.00	mg/ahmv	25 - 45	mg/ahmv

Customer Complaints

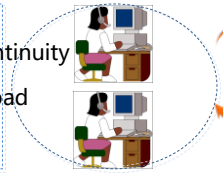
Customer Center

Slow network

Video service discontinuity

Web Page fails to load

QQ fail to log in



No pre-processing tool


Where is the problem ? How to quickly handle customer complaints

Headache

Network Barometer

Overall Objectives

High Scale Network

Proliferation of applications and services

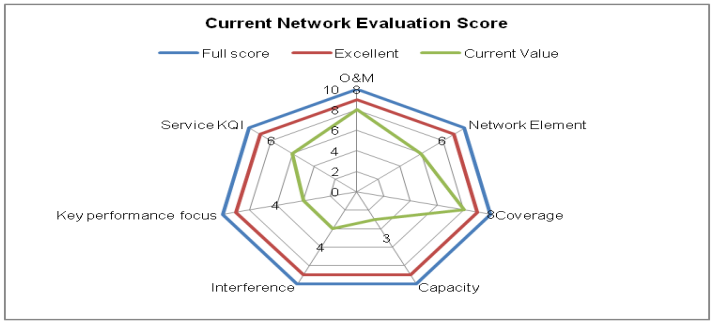
Customer high expectation on the service quality for various applications

increased network incidents and unregulated network changes,unsatisfied network quality

Network Awareness:	Perceive Network Performance and detect unregulated event
Service Assurance:	Verify Network SLA acceptance, better understand customer feel on application(e.g.,web browsing, AV application)
Metric Calculation:	use network health indicator to reflect unsatisfactory level of network impairment
Network Re-Optimization:	Operate and optimize the network to meet on demand service requirement.
Performance Monitoring:	Troubleshoot is hard, tracing the traffic in the network consume tremendous network and server resource.
Trend Analysis:	Event correlation , anticipant network event, forecast short term change and risk in the network

Network Health Indicator vs MoS indicator

The network health indicator provides a Numerical indication of the network anomaly Degree from underlying network impairment Parameters based On big Data Analytics and Diagnosis.



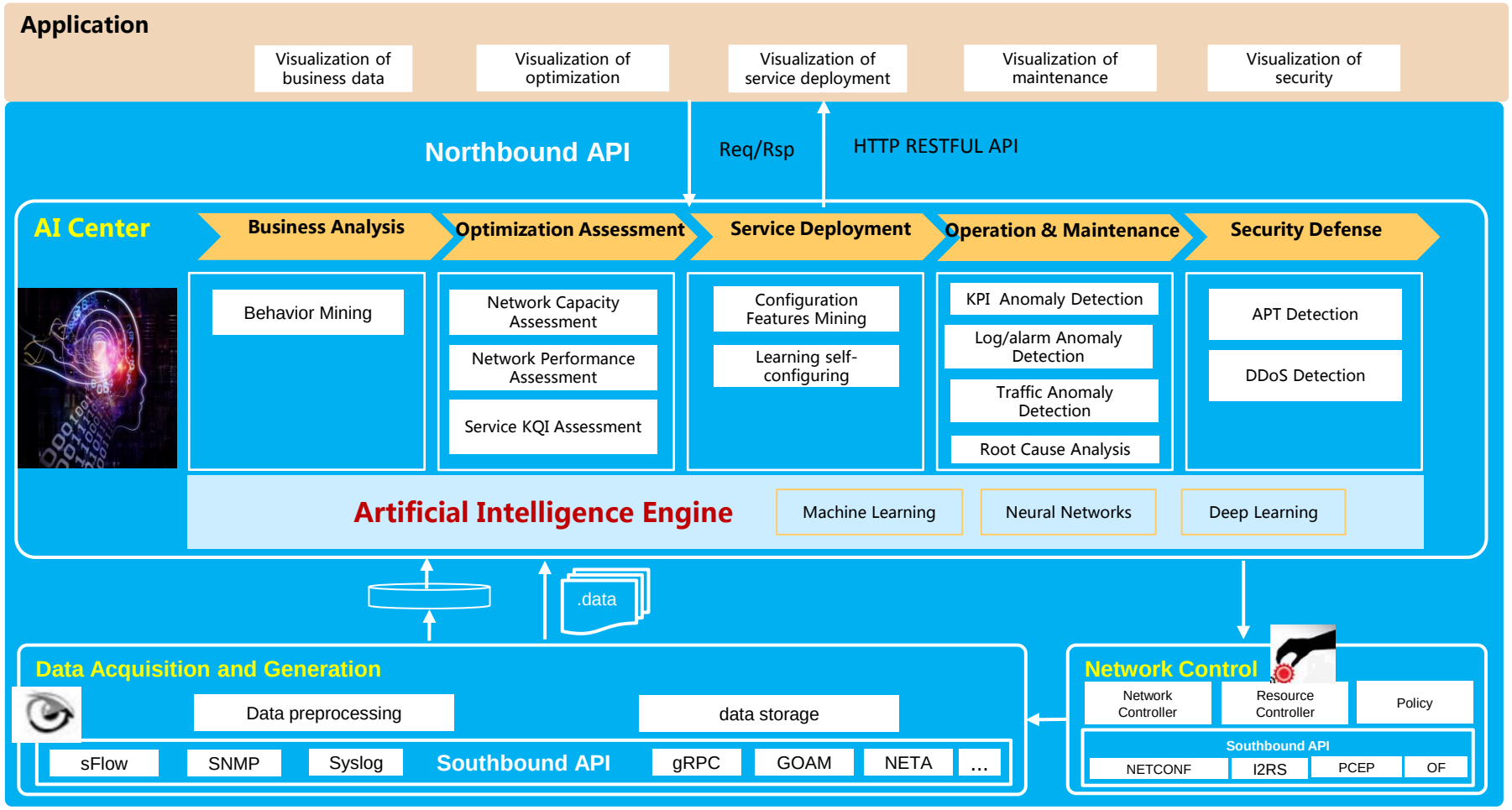
The MOS provides a numerical indication of the perceived quality from the users' perspective of received media after compression and/or transmission

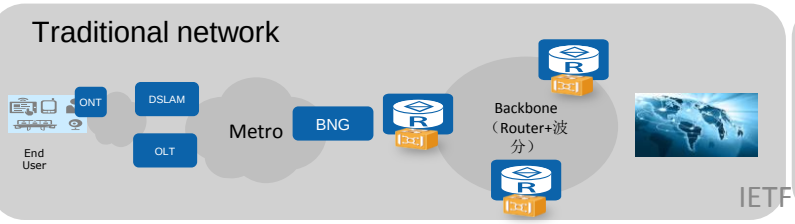
Mean opinion score (MOS)

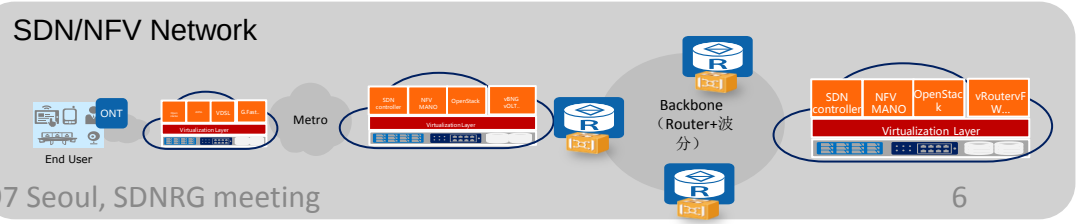
MOS	Quality	Impairment
5	Excellent	Imperceptible
4	Good	Perceptible but not annoying
3	Fair	Slightly annoying
2	Poor	Annoying
1	Bad	Very annoying

	Network Health Indicator	Application Specific MoS
Network-or application layer monitoring	Network layer monitoring	Application layer monitoring
Purpose	Network Anomaly Evaluation And accurate network diagnosis (delimitate to specific network portion, network element, network module) And root cause analysis, fault prediction	Service assurance assessment And coarse granularity application layer diagnosis (delimitate to specific network portion)
Calculation algorithm	Anomaly detection, correlation degree analysis, conformity degree analysis, data consistency check, root cause algorithm	QoE algorithm, e.g., MoS calculation algorithm and other media quality assessment algorithm
Usage	Network planning, dimension, network monitoring and diagnosis	Service monitoring
Assessment Model	Network specific model(could be application independent model or application specific model), defined in the context a specific network and for specific service	Application specific model, defined in the context of specific application and specific usage session
Contributing factors	Network layer parameter, e.g., Network Log data, Network Warning data, Network configuration data, etc.	Application specific parameter(e.g.,GOP, bitrate, I frame loss, PCR, PTS error) Terminal specific parameter (e.g.,Jitter buffer)

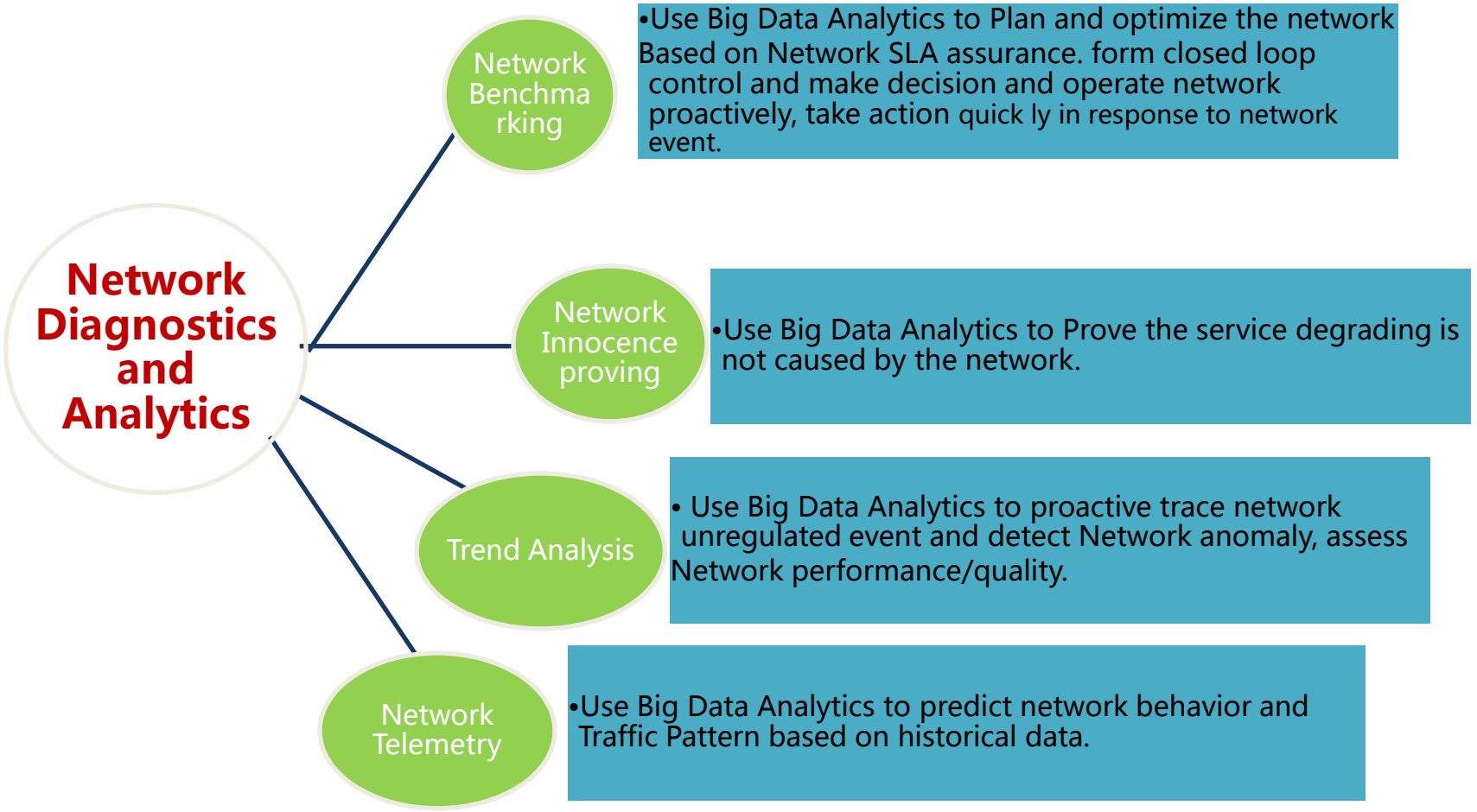
Architecture Overview





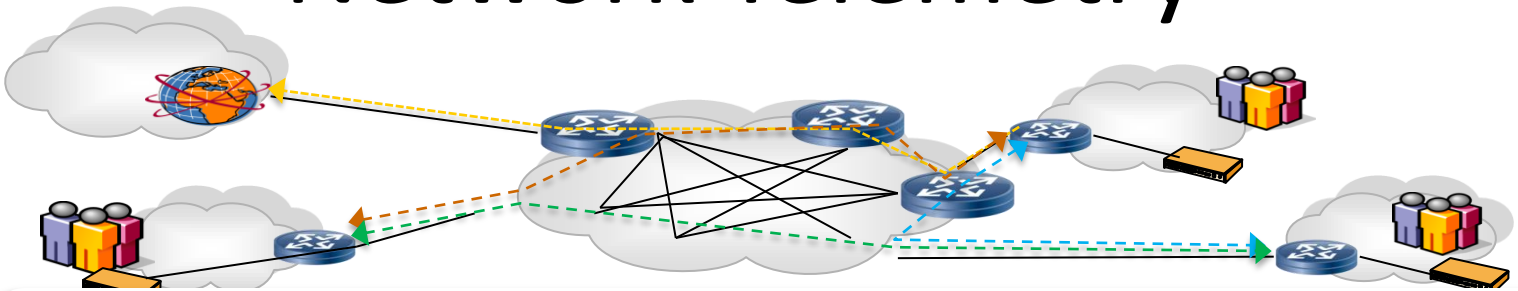


Network Health Assessment Use Cases

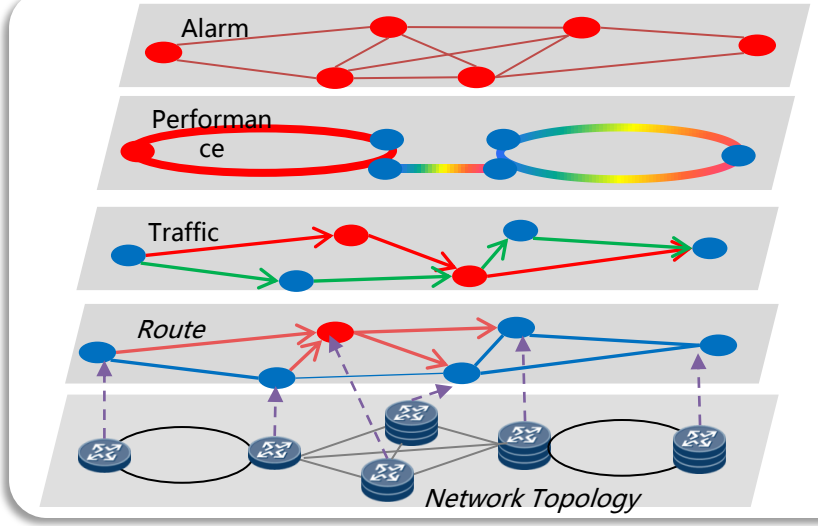


For the first three use Cases, see ITU-T SG12 proposal https://www.itu.int/md/T13-SG12-C-0368/en

Network Telemetry



Network Telemetry Data



Alarm	Quickly locate network fault with alarm and event information
Performance	Locate anomaly device in the topology based on Path KPI value change using KPI statistics learning algorithm
Traffic	Locate anomaly traffic and device based on statistics learning traffic model
Route	Locate Anomaly device based on route behavior and interaction assessment model
Network Topology	Topology model with layering and service classification and provide basic infrastructure for network fault diagnosis

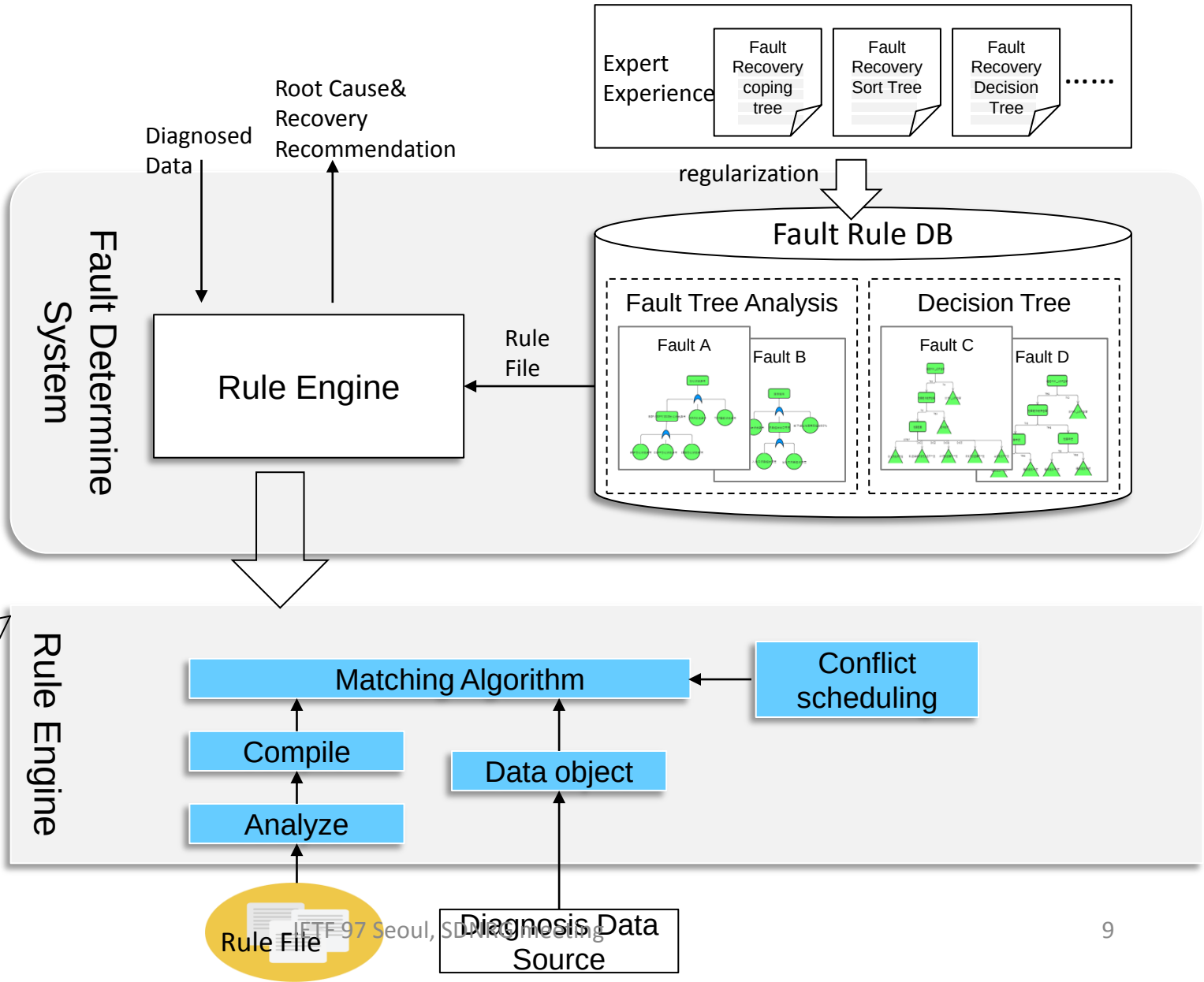
Using Syslog, SNMP,NetFlow as data source to collect traffic statistics, route behavior, performance information, Warning information.

- various network telemetry method are proposed
  - Syslog
  - SNMP
  - Data probing proposed by Facebook
  - In Network Telemetry(INT)
  - In band OAM proposed by Cisco
  - gRPC proposed by Google
- Limitations of these methods
  - Scalability of trace collection
  - Limitation of passive tracing for some methods
  - Data format lack efficiency in the wire
  - Lack pub-sub capability

Diagnostics and Analytics System

Rule and Engine separation

The Rule Engine parses and compiles rule file, generates data object based on diagnosed data, and then apply rule matching, recursive reason, conflict scheduling



Diagnostics and Analytics: Anomaly Analysis

- **Spatial Dimensions Anomaly Analysis：NE configuration parameters comparison**
  - When normal NE and malfunctioned NE are running at the same time, the event type is different, the occurring frequency of the same event is also different.

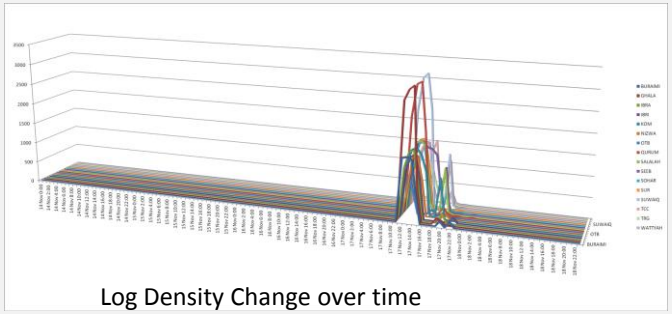
Each Row represents network events per NE, the value of network Event represent Network Event Type

The Network Event Type of Malfunctional is different from one from Normal NE

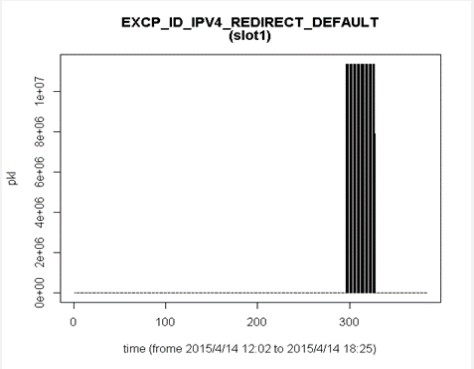
Buraimi_NE80	IERA_NE80	Ibri_NE80	KOM_NE80	Nizwa_NE80	SUR-	RPR-NE80	TRG-RPR-NE80
36	36	36	0	36		36	36
37	37	37	0	37		37	37
0	0	0	0	0		122	122
0	0	0	0	0		123	123
0	0	0	0	0		124	124
0	0	0	0	0		125	125
24	24	24	24	24		24	24
47	47	47	47	47	47	47	47
0	0	0	0	0	0	126	126
0	0	0	0	0	0	127	127
0	0	0	0	0	0	128	128
25	25	25	25	25	25	25	25
	0	0	0	0	0	129	0
51	51	51	51	51	51	51	51
63	63	0	0	0	63	0	63
30	30	30	30	30	30	30	30
0	0	0	76	0	0	76	76
29	29	29	0	29	29	29	29
0	0	0	0	104	0	0	0
34	34	34	34	34	34	34	34
35	35	35	35	35	35	35	35
0	0	0	0	0	0	0	132
0	0	0	0	0	0	0	133
0	0	0	0	0	0	0	134
0	0	0	0	0	0	0	135
60	60	60	60	60	60	60	60
0	0	0	71	0	0	0	0
0	0	0	72	0	0	0	0
0	0	0	73	0	0	0	0
0	0	0	74	0	0	0	0
64	64	64	64	64	64	0	64
65	0	0	0	0	65	0	65
66	0	0	0	0	66	0	66
67	0	0	0	0	67	0	67
0	0	0	0	0	0	0	136

NE Event Type Statistics

- **Time Dimension Anomaly Analysis：NE historical data comparison**
  - The change of the performance measurement results and network event occurring frequency in malfunctioned NE is different from ones in Normal NEs.



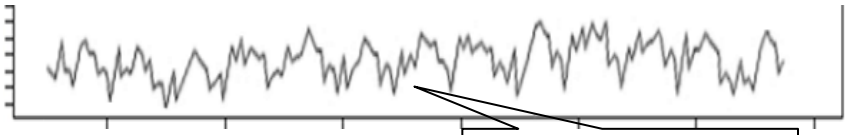
Log Density Change over time



Packet Loss KPI change over Time

Diagnostics and Analytics : KPI Anomaly Prediction

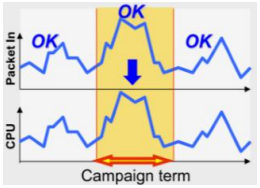
- Single KPI Anomaly Prediction： long resource leaking hide time, strong concealment, anomaly is easy to be concealed by Normal data
- Multiple KPI Anomaly Prediction: Need to consider correlation between KPIs, without its impact on single KPI threshold detection, mis-report, wrong report can be generated



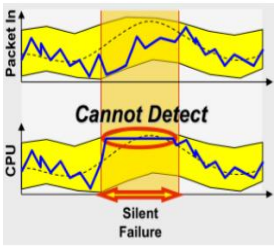
Leaked fault is hard to precieve

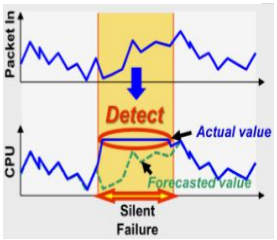
$Y = T + S + R$

- The time sequence Y can be decomposed as: Y= tendency variable+ season variable+random variable
- Given Y, adopt moving average(MA) to get Tendency variable T, Use season len as MA interval to offset season impact
- Linear fitting on tendency var T, i.e., T=a+Kt
- The gradient K, indicate tendency var change, with K to anticipant memory leak issue

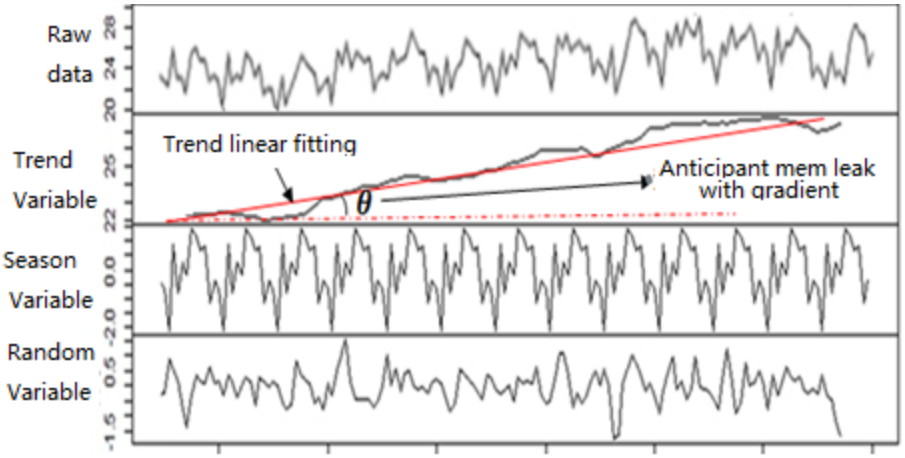


Num of Packets sent to CPU is increased, CPU utilization is also increased, this is normal, but it is easy to be wrong-reported by traditional method





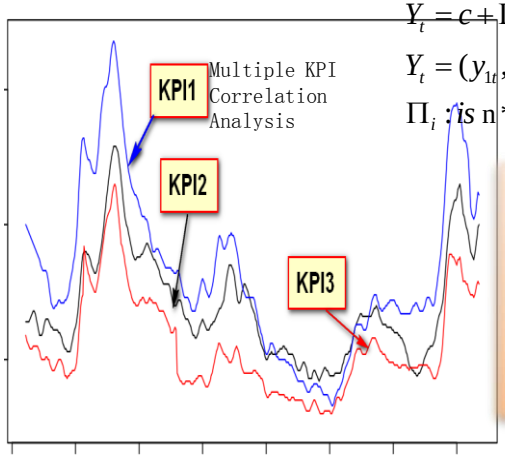
Num of Packets sent to CPU is decreased, CPU utilization is also increased, this is abnormal, which is easy to be mis-reported



$$Y_t = c + \Pi_1 Y_{t-1} + \Pi_2 Y_{t-2} + \ldots + \Pi_p Y_{t-p} + \varepsilon_t$$

$Y_t = (y_{1t}, y_{2t}, \ldots y_{nt})'$ is n dimension timeserial vector

Π_i : is n * n coefficience matrix



- Assume KPI1 in t1 is related to KP2 in t-1, t-2,..t-p
- With KPIs correlation in consideration, more accurate
- Compare predicted value with actual value to detect anomaly, reduce mis-report and wrong report.

Metric Definition: Network Profile

Network Profile: Describe network Constraints, Characteristics

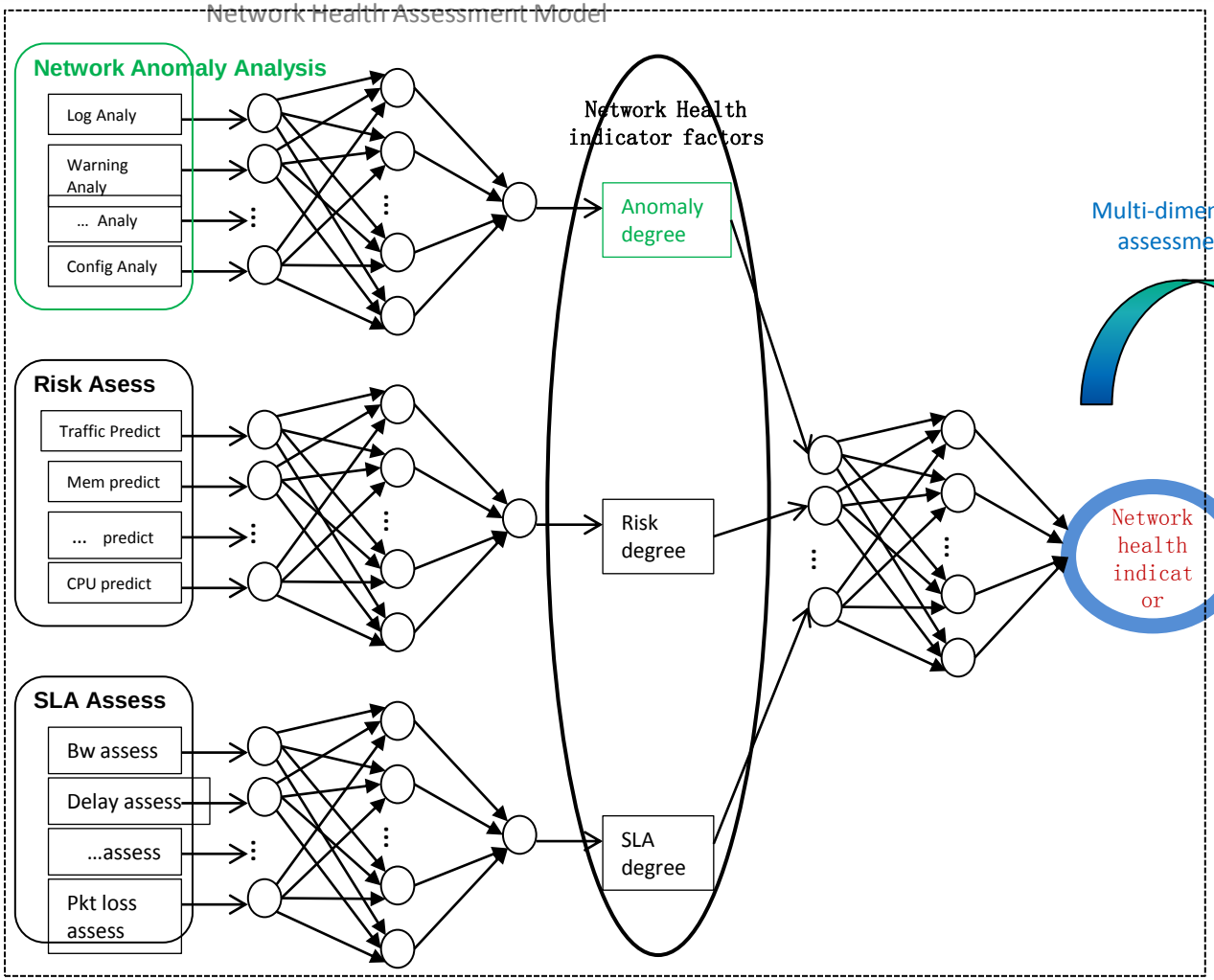
Network Category	Attributes
Network Type	LAN, WAN, WLAN, MAN, SAN, PAN, EPN & VPN
Network technology	MPLS Tech, IP Tech,Segment Routing Tech, etc.
Network Coverage	FBB, MBB, Home Broadband, Corporate Lease line
Network Segment	Access, Aggregation, Edge, Core
Application support	Data Service, Storage ,Video, Audio, Real time service, Data Center
Transport Protocol	UDP, TCP, HTTP
Bearer	Ethernet, Optical,
Network Access Mode	Wireline, Wireless
Routing Tech	BGP/ISIS/OSPF/RIP/Static routing
Network Topology	Hub spoke, Full Mesh
Network QoS	Total Bandwidth, Resvered bandwidth, Bandwidth Utilization, Packet Loss, Jitter, Delay, Max-Route, Throughput, CoS Value.
Network Multicast	Unicast, Multicast, Broadast
Network Security	Authentication, Encryption, Integration Protection, etc.
Network OAM	BFD, LSP Ping, IP OAM, Ethernet OAM, PW OAM
Network Tunnel	Tunnel Type, Tunnel Technology
Network Protection	Link Protection, Node Protection,Link and Node Protection, Repair time
Network Protocol complexity	Disruption frequency, Disrupt time, Protocol parameters consistency
Network Topology complexity	Primary path, backup path, Domain Diversity, Link Diversity, Node Diversity

Metric Calculation

- ITU-T SG12 Q16 has been tasked to work on Network Health assessment standards in the new study period 2016-2020.

  "Network Health Assessment Using Big Data Fault Analytics"

Network Health Assessment Model



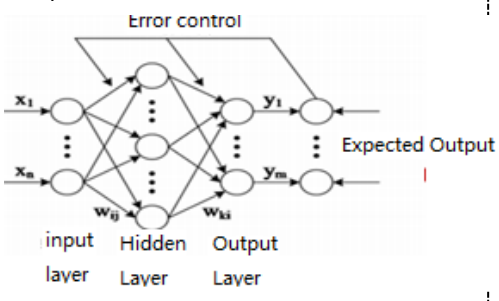
Multi-Dimension correlation assessment model:

Adopt Neural networking method to model network health from several dimension, calculate network health score through training with historical assessment data sampling

Input: Log Anomaly, Warning Anomaly, Config Anomaly, traffic risk,CPU risk, mem Risk,BW ,delay, pkt loss,etc;

Middle level parameters: anomaly degree, risk degree, SLA degree;

Output : Network Health Indicator



3 layer BP Neural Network Architecture (n-3-1)

Conclusion

- Troubleshooting is hard
  - Protocol misbehave
  - Mis-configuration
  - Packt loss
  - End to end latency
  - Load balancing, etc.

- Network health indicator is the key for Network Diagnostic and Analytics
  - Build Closed loop echo-system
  - Schedule network resource based on service requirements from customer

- Network telemetry is the key to troubleshooting
  - Network telemetry provide data for diagnostic and analytics
  - various network telemetry method are proposed
    - Syslog
    - SNMP
    - Data probing proposed by Facebook
    - In Network Telemetry(INT)
    - In band OAM proposed by Cisco
    - gRPC proposed by Google