Journal of Advanced Zoology

ISSN: 0253-7214
Volume 44 Issue S-6 Year 2023 Page 524:530

Life cycle of KPIs in Cloud Networks

Manjunatha S[#1], Dr Prabhakar M[*2]

[1,2] School of Computer Science and Engineering, REVA University
[1]r18tcs07@cit.reva.edu.in
[2]prabhakar.m@reva.edu.in

*Corresponding author's E-mail: r18tcs07@cit.reva.edu.in

Article History

Received: 06 June 2023
Revised: 05 Sept 2023
Accepted: 29 Nov 2023

CC License
CC-BY-NC-SA 4.0

Abstract

Key Performance Indicators are the benchmarks for measuring the health of a given network like Performance, Availability and Resiliency, etc. The Virtualization and Cloud computing transformed how the network operates and delivers services. This led to a new paradigm shift in how we measure the KPIs of Cloud Networks. The KPI management is a big challenge for cloud service providers. This paper describes KPI management for Cloud Networks through the Life cycle of KPI. The KPI Life cycle contains the systematic procedure for KPI Identification, Monitoring, Storage, Visualization and Analysis.

Keywords: *Performance Indicator (KPI), Gateway (GW), Input & Output (I/O), Quality of Service (QoS), Service Level Agreement (SLA)*

1. Introduction

The network which adheres to SLA is the most successful among other several deployed networks. End-user or consumer perspective of a good network is robust, reliable, and the promised performance. These network characteristics are realized as Metrics, which are nothing but quantifiable measurements. The metrics measure the operational state and properties of fundamental resources of any computing unit i.e., CPU, Memory, I/O and Network. As the CPU advanced from Single- Core to Multi-core, Clock speed from MHz to GHz, Memory from KB to GBs, I/O from Kbps to Gbps, the metrics are also evolved.

Table 1 CPU, Memory, Input/Output Metrics

CPU Usage		
# Cores	Busy	Idle
Core-1	80 %	19 %
Core-2	50 %	45 %

Memory Usage		
# Process/Task	Usage (KB)	Memory leak (KB)
Process-1	40	2
Process-2	100	0
Process-n	25	1

I/O Usage		
#Parameters	Network-1	Network-2
Packet Drops	50	100

Peak Data Rate	10Mbps	8Mbps
Average Data Rate	5Mbps	4Mbps

The metrics shown in Table 1 are primitive in nature and provides very basic resource usage measurements, not adequate to identify the system, performance issues and its rectification. Modern networks carry a wide variety of traffic, originated from numerous applications. These applications are real-time and non-real time, delay sensitive with various QoS requirements. The primitive metrics are not enough to ensure the required QoS requirements are met, system and network operating at optimal efficiency. The EMS, NMS and FCAPS subsystem works together to ensure the required QoS requirements are met, and the system/network is in good health.

The primitive metrics aggregated and evolved into many composite metrics like KPI and SLA. Eventually, SLA and KPI are monitored and measured to evaluate the system/network health. The current paper discusses the framework which focuses on deriving composite metrics from primitive metrics. The most desirable framework, which is flexible, scalable, and agnostic of the system/network. The flexibility is achieved when SLA parameters are successfully mapped into metrics and vice-versa. Scalability helps in enhancing the metrics from rudimentary to composite and hierarchical. Agnostics ensure the framework is smooth across different network topology, configuration, and deployment.

2. Literature Review

Hierarchical definition of KPI maps SLA attributes into KPI, [1] suggests consolidating multiple metrics into a single performance value. SLA management & monitoring [2] is the most significant mechanism to monitor and verify whether SLA contractual agreements are met or not. Vendor-specific and generic cloud resource monitoring tools are described in [3]. Efficient storage and retrieving of metrics improve performance [4].

KPI Framework Model

The KPI framework model is explained through the Life cycle of KPI as shown in Fig. 1.

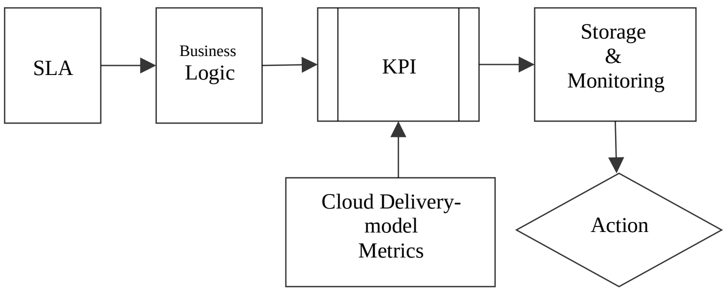

Fig. 1 KPI Life cycl

KPI framework & life cycle comprises of sub-systems:

1) SLA & Business Logic

2) Derivation of KPI, Cloud Layer metrics

3) Monitoring

4) Storage

5) Visualization

6) Analysis & Action

SLA and Business Logic

SLA is vague and obscure as they lie in the problem domain. High rewarded skills are required to derive and map SLA into measurable numerical identities like KPIs. System Architects and Requirements team gets SLA as their input. SLA is the contractual agreement. Typical SLA template show in Table 2.

Table 2 SLA Template

SLA Template	
Attributes	**Values**
Availability	95%
Data Rate per user	UL: 2 Mbps, DL: 20 Mbps
Date Rate per GW	UL: 1Gbps DL: 40 Gbps
Price	Rs 100 per day
Simultaneous connection per GW	1 Lakh user connection
Power consumption	Watts

Business Logic is the first step from the problem domain towards the solution domain. Business Logic represents the System Requirements in a more formal way. Domain experts write down the system requirements.

Deriving KPIs

SLA attributes are verified through measurable numeric KPIs. System architects, Domain Experts and Traffic Engineers study SLA and Business Logic then derives KPIs. Deriving KPIs classified into two types: (a) Direct Metrics (b) Hierarchical Metrics.

Direct Metrics: The SLA attributes are directly mapped. E.g., Availability of service and network uptime.

Hierarchical Metrics: These are derived from other SLA attributes or Cloud layer metrics. E.g., Data Rate per GW. These are collected at various interfaces in a GW and current user connection and their data transfer rate. Hierarchical metrics are defined by Dependency graphs. In Fig. 2 Dependency graphs identifies, how given KPI is derived from multiple other metrics. The KPI definition template is shown in Table 3.

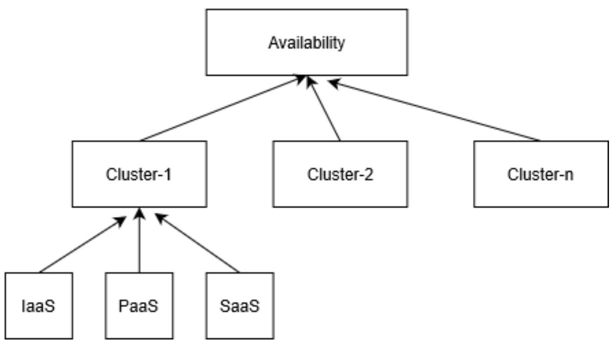

Fig 2. Hierarchical Metrics

Table 3 KPI Definition Template

KPI Definition					
SLA Attributes	**KPI Name**	**Type**	**Value Range**	**Storage** **IMDB/Disk**	**Derived** **(yes/no)**
Availability	Uptime	Numeric	Hours	Both	No

Cloud Layer Metrics

The traditional Network metrics collected at functional units like per Network Link, User sessions etc. In Cloud Networks, the metrics are defined based on the Cloud layers such as IaaS, PaaS, and SaaS. The primitive metrics of CPU, Memory, Network, Storage are in the IaaS layer. The Middleware, Pre-configured platform metrics are in PaaS. The Application and Service metrics are in SaaS. The cloud layered architecture separates the metrics into different scopes. The KPIs are derived from these Cloud layer metrics.

Table 4 Example of KPI Definition

KPI Definition					
Operation Metrics	**KPI Name**	**Type**	**Value Range**	**Storage IMDB/Disk**	**Derived (yes/no)**
Busy Hour call	BHC	Hours/ Numeric	Integer	Both	Yes
Unauthorized User	Call Drops	Numeric	Integer	Both	Yes
No resource	Call Drops	Numeric	Integer	Both	Yes

Monitoring

The Monitoring brings life to the metrics. The Monitoring comprises a collection of metrics (statistics & counters). The KPI framework discusses real-time and non-real time monitoring of SLA, KPI and Cloud Layer metrics.

(a) SLA Monitor: The Cloud Service Provider, Cloud Service Consumer and Traffic Engineering team monitors the SLA. This is the firsthand information that mirrors the health of the service and backhaul network. The SLA monitor periodically retrieves the underlying KPIs from the framework to verify any violations in the contract agreement. The revenue and profit are directly depending on meeting the SLAs. The SLA monitor minimizes the penalty need to pay by the network operator on not meeting the contract agreement. The Autonomous Monitoring agents periodically do SLA audits and log the results.

`

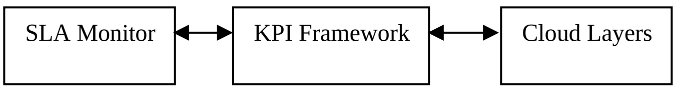

Fig 3. SLA Monitor

(b) Cloud Layer Metrics Monitor:

The Cloud Layer metrics are defined and gathered on IaaS, PaaS and SaaS layer as shown in Table 5.

Table 5 Cloud Layer Model

Cloud Layer	**Resource Type**	**Metrics**
SaaS	Applications, Service Realization	Availability, Bytes Read/Write, Delay
PaaS	Pre-configured Platform, Middleware & Other IT Resources	Availability, Restarts, Delay, Data Loss, Memory
IaaS	Physical Resources	CPU: Utilization

	CPU, Memory, Storage, Networking	Memory: Utilization, Free Storage: Utilization, Free Network: Throughput

The Cloud Layer Metrics Monitor is provided by Cloud Service providers as listed in [3]. 'CloudWatch' is Amazon's monitoring tool for EC2 Could.

Metric Storage and Visualization

The Monitor gathers the metrics in real-time as well as non-real time. The metrics are Current, Cumulative and Historical and gathered metrics are stored in-memory or in disk storage.

Features of Storage:

1. Data Store: It is important to choose the right data store for storing the collected metrics. The efficiency of read and write operations depends on the chosen data storage, as highlighted in [4]. For real-time analysis preferred store is 'In-memory database (IMDB)', such as Redis is used widely in industries. The number of transactions per second is the deciding factor in choosing the right data store. For non-real-time and historical visualization traditional DBMS is used.

2. Purge: Compared to Read and Write operations, the Purge is very less in number. Data store shall stagger the Purge operation.

3.Transport Mechanism: Push and Pull method is used to transport stored metrics into other subsystems if needed for further analysis. SFTP is the preferred protocol.

In IoT (Internet of Things), the metrics are stored locally as well as in Cloud-based IoT deployment model. The Edge Computing principle is to store data as close to edge nodes where data is required to make decisions locally. This avoids the round-trip delay.

The GUI (Graphical User Interface) is the most popular way of visualizing the metrics with various types of Tables, Rows, Columns and Charts, Bars. With visualization, the correlation between metrics can be drawn graphically and reveals the hidden information which would have been not noticed. Traditional networks used proprietary GUI developed on their choice of Programming Languages, Widgets and Controls. Modern IoT systems uses open source-based dash boards. The monitoring tools display metrics in the integrated dashboards.

Many open-source platforms are available such as "Thingsboard" (thingsboard.io) and Grafana. The Cloud-Native Initiative (CNI) uses Grafana and Prometheus for metrics management. Big Data pushed for many open-source visualization tools. Big Data can be structured, semi-structured or un-structured which undergo various data preparation process prior to visualization [5]. The 'Availability' metrics dependency graph shows all the primitive and composite metrics which are aggregated to get 'Availability'.

Analysis & Action

The Analysis and Action go hand in hand. The SLA, KPI and Cloud Layer Metrics are analysed by various parties. The Analysis defines the next Actions to be performed. If Analysis founds that SLA has deviated, then alarms are raised to get network administration attention. Table 6 shows the various stake holders analysing respective metrics. The real-time analysis quickly raises alerts and alarms to notify undesired conditions in Cloud Network. Non-real time analysis in Historical & Cumulative metrics identifies the pain points in the network such as congestion.

Table 6 Metrics Analysis

SLA	Cloud Service Provider Cloud Service Consumer Third Party Audit
KPI	Cloud Server Provider Third Party Audit
Cloud Layer Metrics	Cloud Service Provider Cloud Vendor

Challenges in Analysis:

False-Positives are very difficult to find, the domain expertise and network engineering experience are needed to find them.

Action: The actions are the controllers which steer the network towards the desired state such as meeting the SLA contractual agreements.

The Corrective Actions:

(1) No Action: False-Positive cases no action required.

(2) Manual Intervention: Maintenance window to upgrade the software when the network traffic is low and stable.

(3) Policy Based: Policy based actions are autonomous. On stated conditions are met, actions are performed automatically, and reports are generated.

E.g.:

1. Admission Control: Dropping traffic if not enough resources.

2. Triggering of traffic shaping algorithms.

The Preventative Actions:

The monitored KPIs are shared with System Architect and traffic engineering team to address issues in the future. The KPI monitoring and analysis identifies the pain points and bottlenecks that impact network performance.

E.g.: Online shopping during ‘The Great Indian Summer Sale’, and ‘Thanksgiving Sale’. These are recurring issues in the future. The historical KPIs becomes an input for the next capacity planning and commissioning.

Observability

The modern Cloud-Native applications are developed & deployed as Microservices. The Service Mesh is an infrastructure layer that provides Routing/Inter-service communication, Security and Observability for the Cloud- Native application. The Observability infers the external output and measures the internal states of a system. Observability is PaaS component of cloud and it is not merely a Monitoring.

The modern Cloud-Native Applications are more dynamic and complex. Observability should be employed to better understand modern applications and its performance. The health status check & metrics via observability framework is discussed in [6]. The industry adopted service meshes are Istio, Linkerd and Consul Connect. The Istio provided observabilities are (1) Metrics (2) Distributed Traces (3) Access Logs. These observabilities are exported to dashboards and accessed via the web. Istio Metrics exported to Prometheus, Distributed Traces to Jaeger Tracing. Observability wins over traditional monitoring systems in web scalability and use of open-source solutions.

4. Conclusion

The primitive metrics not sufficient to get end-to-end network health. This paper discussed the generic KPI Framework and life cycle of KPIs which maps the SLA attributes in Cloud Networks. Big Data is a new paradigm shift which scale up Metrics storage, analysis, and visualization. The SLA and KPIs are handwritten down by System Architects, Domain Experts and Traffic Engineers. Adapting AI/ML model can identify KPIs, which the traffic engineering team could not write down. Cloud-Native is a breakthrough in how applications are developed and deployed, provides a wide range of PaaS components for KPI collection and analysis.

Future Scope of Work

The ITU-T defined Network Data Analytics Function (NWDAF) in 3GPP 5G [7-13]. This is a breakthrough in standardization by drafting specifications and getting away with vendor specific methods. The 5G mobile network is Cloud-Native Microservice accessed over REST/HTTP. The standardization would be expanding to any web service which is developed and deployed through REST/HTTP.

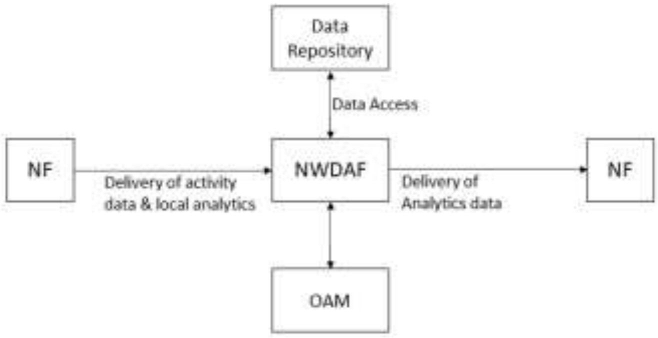

Fig 4. NWDAF in 5G

The Cloud-Native community contributing continuously with various software and tools across IaaS, PaaS, and SaaS. Adoption of these software and tools helps reducing human error in defining, deriving and analysis of KPIs.

References:

1. Illias Leontiadis, Joan Serra, Alessandro Finamore, Giorgos Dimopoulos, Konstantina Papagiannaki. 2017 IEEE 33rd International Conference on Data Engineering. "The good, the bad, and the KPIs: how to combine performance metrics to better catpure underperforming sectors in mobile networks". 2375-026X/17, DOI 10.1109/ICDE.2017.89.
2. Marios Touloupou, Evgenia Kapassa, Chrysostomos Symvoulidis, Panagiotis Stavrianos, Dimosthenis Kyriazis. Departmentof Digital Systems, University of Piraeus, Greece. An Integrated SLA Management Framework in a 5G Environment. 22nd Conference on Innovation in Clouds, Internet and Networks (ICIN2019).
3. Khalid Alhamazani, Rajiv Ranjan, Karan Mitra, Fethi Rabhi, Samee Ullah Khan, Adnene Guabtni, Vasudha Bhatnagar: An overview of commerical cloud monitoring Tools: Research Dimensions, Design Issue, and State-of-the-Art.
4. Baris Kurt, Engin Zeydan, Utku Yabas, Ilays Alper Karatepe, Gunes Karabulut Kurt, Ali Taylan Cemgil: "A Network Monitroing System for High Speed Network Traffic". DOI: 10.1109/SAHCN.2016.7732965.
5. Nicolas Marie-Magdelaine, Toufik Ahmed, Gautheir Astruc-Amato, Univ Bordeaux, Bordeaux INP, CNRS, LaBRI, UMR5800, F-33400 Talence, France: Demonstration of an Observability Framework for Cloud Native Microservices.
6. V Sankhivel, R Mythregai, M Priyadarshni. 2018 International Conference on Soft-computing and Network Security (ICSNS): A SURVEY ON WATCHDOG TOOLS FOR CLOUD AND BIGDATA.
7. 3gpp TS 29.520, 3rd Generation Partnership Project; Technical Speicifcationi Group Core Network and Terminals; 5G System; Network Data Analytics Services; Stage 3.
8. Sudhan Murugan Bhagavathi, Anitha Thavasimuthu, Aruna Murugesan, Charlyn Pushpa Latha George Rajendran, A Vijay, Raja Laxmi, Rajendran Thavasimuthu, Weather forecasting and prediction using hybrid C5.0 machine learning algorithm International Journal of Communication Systems, Vol. 34, Issue. 10, Pp. e4805, 2021.
9. PM Surendra, S Manimurugan, A New Modified Recurrent Extreme Learning with PSO Machine Based on Feature Fusion with CNN Deep Features for Breast Cancer Detection, Journal of Computational Science and Intelligent Technologies, Vol. 1, Issue. 3, Pp. 15-21, 2020.
10. PK Sadineni, Comparative Study on Query Processing and Indexing Techniques in Big Data, 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), pp. 933-939, 2020.
11. AH Omar Baabood, Prajoona Valsalan, Tariq Ahmed Barham Baomar, IoT Based Health Monitoring System, Journal of Critical Reviews , Vol. 7, Issue. 4, pp. 739-743, 2020.
12. Sajay KR, Suvanam Sasidhar Babu, Vijayalakshmi Yellepeddi, Enhancing The Security Of Cloud Data Using Hybrid Encryption Algorithm, Journal of Ambient Intelligence and Humanized Computing, 2019. https://doi.org/10.1007/s12652-019-01403-1
13. Bindhia K Francis, Suvanam Sasidhar Babu, Predicting academic performance of students using a hybrid data mining approach, Journal of Medical Systems, 43:162, 2019. https://doi.org/10.1007/s10916-019-1295-4