European Journal of Computer Science and Information Technology
Vol.6, No.5, pp.51-65, November 2018
Published by European Centre for Research Training and Development UK (www.eajournals.org)

A COMPREHENSIVE SURVEY OF CLOUD MONITORING

Shahzada Khurram and Osman Ghazali

Department of Computer Science, School of Computing, Universiti Utara Malaysia, Kedah, Malaysia

ABSTRACT: *Cloud monitoring is an undertaking that helps to maintain applications active and perform constantly well. Cloud monitoring services collect a range of information about applications' performance, storage, servers and networks within a provider's ecosystem. A well-designed monitoring infrastructure, which does the aforementioned prerequisites, must be scalable and support various important performance parameters. Cloud monitoring is essential for both cloud users and providers. Along with, it is a central tool for managing software and hardware infrastructure. Moreover, it furnishes information and key performance indicators for cloud layer services. In cloud computing, monitoring has two types namely high-level and low-level monitoring. The high-level monitoring is for virtual platform status while the low level is related to the physical infrastructure. Furthermore, two types of cloud network monitoring are underlay and overlay. In this paper, the authors present a comprehensive survey and discussion in respect of high level, low level, underlay and overlay cloud monitoring. A generalize metric grouping for high level, low level, underlay, and overlay cloud monitoring were established.*

KEYWORDS: Cloud, Monitoring, Metrics, High Level, Low Level, Underlay, Overlay

INTRODUCTION

In cloud computing, a large collection of remote servers is networked to facilitate on-demand computing resources that are available to everyone over the Internet. Cloud computing is made up of three layers namely Software-as-a-Service (SaaS), Platform-as-a-Service (PaaS) and Infrastructure-as-aService (IaaS). SaaS provides access to application software that runs on distant virtual machines (VM) via the Internet. PaaS provides a computing platform, which usually includes operating systems for execution of required applications. IaaS provides computing resources including servers, storage, networking and data center space on pay per use basis. To acquire cloud layers services, a Service Level Agreement (SLA) between the cloud provider and the customer is required. The acquired service based on the SLA will be monitored continuously.

Monitoring is the process of observing and tracking applications as well as resources at run time. Fatema et al. provides a brief definition of the term monitoring as *"a process that fully and precisely identifies the root cause of an event by capturing the correct information at the right time and at the lowest cost in order to determine the state of a system and to surface the status in a timely and meaningful manner"*(Fatema, Emeakaroha, Healy, Morrison, & Lynn, 2014). A multi-tenant nature of cloud can be challenging for smooth management in term of performance constraints and quality of service because the services of the cloud are scalable, flexible and on demand. Monitoring plays an important role in the utilization of cloud resources at all layers.

Print ISSN: 2054-0957 (Print), Online ISSN: 2054-0965 (Online)

In this paper, we provide a comprehensive survey and discussion about cloud monitoring in term of high level, low level, underlay networks and overlay network monitoring. The remainder of this paper proceeds as follows. Section overlay network vs. network virtualization explains the differences between overlay network and network virtualization in cloud computing environment. Next section describes the type of cloud monitoring. Discussion section related to analysis studies that deal with monitoring in cloud computing environment and analysis of cloud metrics. Finally, concludes the paper and briefs future research direction in cloud monitoring.

OVERLAY NETWORK VS. NETWORK VIRTUALIZATION

Overlay networks are similar to network virtualization but with different functionality. Overlay network and network virtualization terms often are used interchangeably when to discuss the demand for multi-tenancy or occupancy and address space solution, point to overlay network like VXLAN (Mahalingam et al., 2014), NVGRE (M. Sridharan , A. Greenberg , N. Venkataramiah , Y. Wang & I. Ganga , G. Lin , M. Pearson, 2012), and STT (Davie & Gross, 2012). Since the growth of cloud

computing, cloud data centers expanded their services with standalone physical computers and network devices. The growing demand and dependency on internet resources required virtualization of data centers. Network virtualization allows running of isolated logical networks on a shared physical network. It consists of a combination of multiple network resources, capabilities and functionalities into a single unit known as a virtual network. It is the solution for expanding data center devices that connect each other within virtualized environment. Virtual private network (VPN) and virtual local area network (VLAN) are examples of network virtualization. However, network virtualization has some limitations and issues. Cloud data centers growth may increase the number of virtual machines deployment. Though, VLAN can only identify a maximum of 4094, IP addressing and VLAN are often assigned to virtual machines. However, advance and large cloud networks have thousands of switches that interconnect thousands of VMs that making it insufficient for identifying users on large Layer-2 cloud networks. To overcome this limitation, overlay network technology which can encapsulate layer-2 frames into layer-3 was introduced into a cloud environment. Overlay network technology is used in cloud data centers to effectively isolate multiple tenants and automate network-wide virtual machine migration that fully satisfy the requirements of large cloud service providers and enterprises.

Types of Monitoring In Cloud

In this section, we discuss types of cloud monitoring. There are four categories of cloud monitoring: high-level monitoring, low-level monitoring, underlay network monitoring, and overlay network monitoring. Figure 1, presents the types of monitoring with relevant cloud layers.

Print ISSN: 2054-0957 (Print), Online ISSN: 2054-0965 (Online)

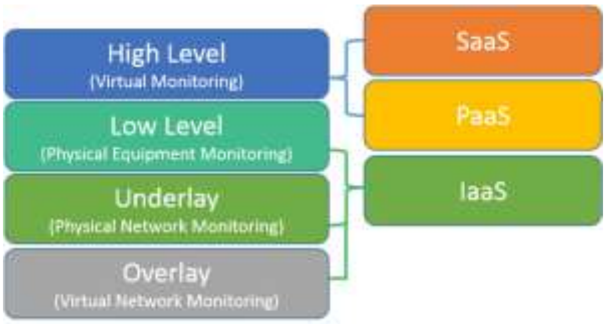

Figure. 1: Types of Cloud Monitoring

High Level

High-level monitoring collects data that is relevant to virtualized platform in cloud computing environment. In cloud layers, monitoring SaaS and PaaS are considered as high-level monitoring. In SaaS environment, users can access cloud applications through the Internet. Cloud host applications can perform an extensive range of tasks for consumers. Facebook, webmail and instant messaging are examples of SaaS. Consumers are able to access the services through any internet enabled device. On the other hand, developers can build applications according to their requirements using PaaS environment. Web servers, App servers and Database management are examples of PaaS.

Low Level

Low-level monitoring collects data that is relevant to the physical infrastructure of the cloud. In cloud monitoring, IaaS layer is considered as low-level monitoring. IaaS provides data center resources containing storage, servers, and datacenter space on a pay per use basis. Low-level monitoring is divided into two categories namely computation based and network based (Mei, Liu, Pu, & Sivathanu, 2010). A collection of computation based matrices is called low-level monitoring. Meanwhile, network-based monitoring in IaaS is divided into underlay network monitoring and overlay network monitoring.

Underlay

A collection of metrics relevant to the physical network in cloud computing environment is called underlay network monitoring. Underlay monitoring is also a part of IaaS services, which includes monitoring of layer 2 and 3 switches, routers, firewall, IDS, IPS and other physical network resources in cloud network infrastructure.

Overlay

An overlay network is a kind of network virtualization. The goal of an overlay network monitoring is to improve productivity and efficiency in large cloud computing environment. In cloud computing environment, VMs can move from one host to another for various reasons, for instance, distribution of workloads or host failure. These move the required basic configuration of VLAN trunking in cloud networked switches. Advance and large cloud networks have thousands of switches that interconnect thousands of VMs. They are being limit by the broadcasts domains in layer 2 due to the limitation of 4094 VLANs. To overcome this problem, overlay tunneling techniques, which can encapsulate layer 2 frames into layer 3 IP

Print ISSN: 2054-0957 (Print), Online ISSN: 2054-0965 (Online)

packets, was introduced in a cloud environment. A few standards have been proposed to enable overlay networks, which include Virtual extensible LANs (VXLAN), Network Virtualization with GRE (NVGRE) and Stateless Transport Tunneling Protocol (STT). These overlay protocols use different encapsulation techniques to overcome the current network limitations.

DISCUSSION

We selected and presented the most relevant works on monitoring systems and architectures. A few studies deal with monitoring in cloud computing environment. These studies do not address the detail of low level, high level, underlay and overlay network monitoring metrics.

Academic research works for cloud monitoring

In (Bardhan & Milojicic, 2012) proposed a prototype to monitor QoS measurement in a federated cloud environment. The author used a basic time-based mechanism to represent and measure QoS continuously for both individual service and composite services. However, the proposed prototype monitored only a single metric which is availability. The metric is based on servers' uptime and downtime that are hosted on different clouds. The limitation of the proposed mechanism is the parameter of the low-level monitoring, which considered only one metric.

Lee et al. proposed a service level QoS measurement for SOA environment. The author discussed high level performance monitoring of web services and provided mathematical formulas of different metrics of QoS measurement which included throughput, availability, accessibility and successability (Lee, 2010). However, it was lacking the implementation of any proposed metrics.

Liu et al. proposed a lightweight system to monitor and to detect SLA violation in a cloud environment (Duo Liu, Utkarsh Kanabar, 2013). The proposed system was based on SNMP (Stallings, 1998) protocol. It monitored and obtained dynamic resource information for memory usage, CPU and link utilization for each VM for analyzing and detecting SLA violation. The proposed system only monitored some high-level parameters but lacking low level and overlay monitoring.

In (Grati & Boukadi, 2014) proposed a framework for IaaS to SaaS monitoring of Business Process Execution Language that is processed in the cloud. The authors proposed manager and agent-based model for monitoring all cloud layers from both customer and cloud service provider point of view. However, the framework only monitored response time for low-level monitoring which was implemented in the given framework. While (Alhamazani, Ranjan, Rabhi, Wang, & Mitra, 2012) discussed the importance of dynamic monitoring of QoS in cloud layers. He described an ongoing Ph.D. work that developed methods to monitor the quality of service of cloud layers using Simple Network Management Protocol (SNMP) (Stallings, 1998). The focuses were high level, low level and underlay network monitoring. Likewise, in (Alhamazani et al., 2014) the author proposed cross-layer federated cloud application monitoring as a service framework. It is based on server and agent using standard protocol SNMP for data collection at all cloud layers or multi-cloud environment. The focuses are high level, low level and underlay network monitoring but not overlay network monitoring.

Print ISSN: 2054-0957 (Print), Online ISSN: 2054-0965 (Online)

In (Ye, Zhang, Shi, & Du, 2012), authors proposed SLA verification framework that leveraged a third-party auditor. They developed an effective testing algorithm that can monitor and detect SLA violations of the physical memory size of the virtual machine in cloud computing. They only focus on low-level monitoring and violation detection for VM memory size but no focus on other required parameters in cloud layers.

Shao et al. proposed a runtime model based monitoring approach in a cloud environment. The author focused on basic cloud monitoring parameters in cloud layers. The model can organize cloud monitoring data which was collected by different methods and presented in a well-organized way. In addition, the presented model was capable of attaining a balance between runtime overhead and monitoring of cloud services (J. S. J. Shao, Wei, Wang, & Mei, 2010). Generally, a high-level and low-level parameters were discussed but not implemented.

In (Adinolfi, Cristaldi, Coppolino, & Romano, 2012) authors proposed a portable architecture for QoS monitoring in the cloud. He outlined the design and implementation of QoS-MONaaS framework.

The downside of the framework was its dependable monitoring facility being the part of another project and only high-level monitoring for SaaS.

In (Dhingra, Lakshmi, & Nandy, 2012) proposed a distributed monitoring framework for cloud computing. They presented a generic framework that could support customized monitoring using open source software. They focused on low-level monitoring in IaaS cloud layer but not high level and overlay network monitoring. While (Kai, Weiqin, Liping, & Chao, 2013) proposed SCM: a design and implementation of a monitoring system for CloudStack. The authors presented a scalable and flexible monitoring system architecture to monitor the Apache CloudStack platform. The proposed prototype only monitored low-level parameters in IaaS cloud layer. However, (Rodrigues, Granville, & Tarouco, 2012) authors presented features of different frameworks for cloud computing operating at IaaS level. They investigated the cloud monitoring process in main cloud platforms such as Eucalyptus, OpenNebula and Nimbus. The authors only observed low-level monitoring process and discussed the management challenges.

In (Lakshmanan, Keyser, Slominski, Curbera, & Khalaf, 2010) authors proposed an outline that allows end-to-end monitoring of combined business applications. They developed and implemented a monitoring prototype to accomplish viewing of end-to-end monitoring data for various services. However, the focus was only on the high-level monitoring but no other monitoring type. While (Katsaros et al., 2012) proposed a cloud monitoring framework for measuring the quality of services in different cloud layers from application to infrastructure level. The proposed model was able to collect and combine monitoring data of runtime resources allocation and decision making in cloud computing environment. The focuses were basic high-level and low-level monitoring but not overlay monitoring. However, (He et al., 2013) authors proposed a framework to design a monitoring model for simulation cloud. The focus was on high-level and low-level monitoring in a federated cloud environment that collect performance information metrics for cloud monitoring containing physical and virtual resources of the federated cloud. The proposed framework could detect real time anomalous behaviors of cloud resource consumption.

A. Meera and S. Swamynathan proposed an agent-based resource monitoring system in IaaS cloud environment (Meera & Swamynathan, 2013). The authors introduced the idea of how monitoring agents collect virtual machine resource usages like CPU and memory utilization.

Print ISSN: 2054-0957 (Print), Online ISSN: 2054-0965 (Online)

However, it was limited to low-level monitoring but not high level and overlay monitoring. While (Katsaros, Kübert, & Gallizo, 2011) proposed cloud monitoring framework to monitor physical and virtual resources of the cloud. Following REST architecture principle, it collected and stored monitoring information of physical and virtual resources of cloud infrastructure using open source software such as Nagios and Libvirt. The focuses were basic low-level, high-level and underlay network monitoring but not overlay network monitoring. However authers proposed an architecture for real-time monitoring in distributed cloud environment (Smit, Simmons, & Litoiu, 2013). The proposed architecture has scalability and fault tolerant capability. Monitoring performance parameters were managed by streams that were available to the users via push notifications. The focuses were high-level, low-level and underlay network monitoring but not overlay network monitoring.

In (Montes, Sánchez, Memishi, Pérez, & Antoniu, 2013) proposed an architecture for generic cloud monitoring in cloud computing environment. Author implemented the proposed cloud monitoring architecture which covered all three layers of cloud computing and it showed improved results. The author presented low-level, high-level and underlay network monitoring metrics but not overlay monitoring. While (Povedano-Molina, Lopez-Vega, Lopez-Soler, Corradi, & Foschini, 2013) proposed a distributed architecture for monitoring resources utilization in multi-clouds. The proposed architecture could monitor different performance metrics for cloud monitoring including efficiency, reliability and multitenancy, without incurring a significant overhead. The monitoring involved low-level, high-level and underlay network monitoring but not overlay network monitoring.

Suciu et al. proposed a model for network management and monitoring for cloud systems (Suciu, Halunga, Ochian, & Suciu, 2014). He provided the way out to monitor cloud services using an open source monitoring tool named Nagios and many other add-ons and plugins. The focuses were on lowlevel and underlay network monitoring but not overlay network monitoring. While (Andreolini, Colajanni, & Pietri, 2012) authors proposed a scalable real-time monitoring architecture for large information systems with the goal for guaranteeing scalability and availability. The focus was the monitoring component of the logical cluster and physical racks for large-scale network infrastructures hosted in data centers. It only monitored underlay network but not overlay monitoring.

Kim et at. proposed a monitoring performance architecture for application services in multi-clouds (Kim et al., 2013). The key features of the architecture were flexible integration with external agents and separation of output channels such as real-time publish/subscribe and DBMS-based offline friendly repository for analysis of monitoring metrics. The focuses were low level and underlay network monitoring but not overlay network monitoring.

In (Nair & Gopalakrishna, 2009) proposed a framework for monitoring cloud computing services [24]. They implemented a prototype for a few high-level monitoring parameters in SaaS layer using standard SNMP protocol. The drawback of the framework was lacking configuration management. While (Liu & Wood, 2015) proposed a distributed software-based network monitoring framework for cloud data centers . The authors presented a deployment strategy for a software-based packet monitoring system in the network to measure video contents using Software Defined Networking. Once data was captured, it was aggregated and sent to a processing engine. The focuses were underlay network monitoring but not overlay network monitoring. However, (Mann, Vishnoi, & Bidkar, 2013) proposed a network service monitoring solution for cloud infrastructure. He identified problems within virtual switches for permitting flow-based network monitoring. They implemented and analyzed flow monitoring

Print ISSN: 2054-0957 (Print), Online ISSN: 2054-0965 (Online)

protocols such as NetFlow and sFlow (Phaal, Panchen, & McKee, 2001) on physical switches and virtual switches in a cloud environment for underlay network monitoring.

Deri and F. Fusco proposed a cloud monitoring architecture that described the real time traffic problems in heterogeneous cloud infrastructure (Deri & Fusco, 2013.). The proposed architecture was based on network probe, which was able to analyze classifying traffic into flows and to separate flows according to user requirements. Underlay monitoring information according to the user requirements was accessible via external software. While (Scharf et al., 2012) proposed a distributed monitoring system which collected the monitoring data using the Application Layer Traffic Optimization (ALTO) protocol (Alimi, Yang, & Penno, 2014.). They demonstrated that the ALTO protocol was very suitable to monitor the information for networked cloud infrastructure. They presented low level and underlay network monitoring but not overlay network monitoring in a cloud environment. However, (Mullins & Bagula, 2013) proposed a lightweight network management protocol for monitoring community clouds. It was an agent based monitoring tool using open source software SIGAR (“SIGAR,” n.d.) that provided lightweight monitoring of community networks for cloud infrastructures. The proposed agent-based architecture was more flexible than the SNMP protocol in the case of data storage collected from the network, deployment and reduction of the traffic load in the network. The proposed prototype had lowlevel and underlay network monitoring capability but not overlay networks in the cloud.

Madan and M. Mathur proposed cloud network management model for monitoring cloud traffic (Madan & Mathur, 2014). The proposed model modified the traditional SNMP protocols to manage cloud traffic with more accurate results. They presented only underlay network monitoring but not high level, low level and overlay network monitoring for the cloud. While (Clayman, Galis, & Mamatas, 2010) proposed a monitoring framework for resources utilities in virtualized cloud environments. The proposed framework could gather monitoring data such as memory, CPU and network utilization for each of the virtual execution environment within the cloud. The focuses were low level and underlay monitoring but not overlay monitoring. However, (Ward & Barker, 2013) proposed a monitoring system for a large cloud environment that is scalable and having ability to collect and analyze monitoring data for cloud infrastructure. The proposed architecture described a technique for gathering the monitored parameters and also provided a mechanism to analyze these parameters in the large cloud. The focuses were low level and underlay network monitoring for IaaS layer.

In (Peng & Chen, 2011) proposed an SNMP protocol based on cloud virtual infrastructure monitoring framework. The authors introduced scalable SNMP agents to monitor VM resources in heterogeneous virtualized cloud environment. The proposed framework could manage monitoring resources according to the user requirements. They presented low level and underlay network monitoring in the cloud. While (Moses, Iyer, Illikkal, Srinivasan, & Aisopos, 2011) authors proposed a prototype to monitor shared resource and throughput optimization for a cloud environment. They focus on low-level monitoring and discussed why shared cache contention is a serious problem in cloud virtualized environment and suggest a solution to them. However, (Ma, Sun, & Abraham, 2012) authors proposed a lightweight framework for monitoring public cloud. The proposed design is less resource intensive and based on client-server architecture. They used open source plugins to monitor public cloud. A few high level and low-level monitoring metrics were discussed on SaaS and IaaS cloud layers.

In (De Chaves, Uriarte, & Westphall, 2011) authors proposed an architecture for private cloud monitoring resources. The proposed architecture was flexible and extensible. It can be adapted

Print ISSN: 2054-0957 (Print), Online ISSN: 2054-0965 (Online)

for centralize monitoring parameters to help improve QoS in a private cloud environment. They focused on low-level and underlay network monitoring on IaaS cloud layer. While (J. Shao & Wang, 2011) proposed performance guarantee approach for cloud applications based monitoring. They presented a cloud monitoring framework that collects real time data for PaaS metrics monitoring using a series of machine learning algorithms for future performance prediction in a cloud environment. They focused on highlevel monitoring.

In (Emeakaroha, Ferreto, Netto, Brandic, & De Rose, 2012) authors proposed SaaS layer monitoring for SLA violation detection in cloud computing environments. He used standard protocol SNMP for efficient high-level monitoring and detection of SLA violation in SaaS cloud computing. While (Rak, Venticinque, Máhr, Echevarria, & Esnal, 2011) authors proposed a cloud monitoring application using the mOSAIC framework. The proposed solution composed of four main components namely mOSAIC API, mOSAIC platform, mOSAIC provisioning system, and semantic engine. These components enabled the structure of custom monitoring systems for cloud monitoring using the mOSAIC API. The author also discussed high level, low level and underlay monitoring techniques using standard SNMP protocol to collect the data at different layers of cloud but lacking discussion on the implementation.

Open source software for cloud monitoring

OpenNebula ("OpenNebula | Flexible Enterprise Cloud Made Simple," n.d.) is an open source management tool for heterogeneous cloud infrastructures. The functionality covered by OpenNebula as a cloud infrastructure manager was extensive. It manages the physical resources, virtual machines, virtual networks and storage capacity. It also collects monitoring data via probes which have been installed on the systems. The main features are the provision of scalability and adaptability. OpenNebula has the ability to monitor high level, low level, underlay and virtual networks.

Nagios ("Nagios Core. Nagios Open Source Project.," n.d.) is a well-known open source monitoring software that support monitoring of heterogeneous cloud computing environment. Nagios was based on the centralized client-server architecture to monitor cloud infrastructure. Nagios core architecture was designed for flexibility and scalability of monitoring. It provides several APIs to allow it's feature-set to be easily extended through additional plugins. It is also used for monitoring OpenStack. Nagios have the capability to monitor high level, low level, underlay and virtual networks.

Nimbus ("Nimbus," n.d.) is an open source monitoring software that provides an efficient IaaS cloud monitoring solution. It allowed cloud monitoring for both consumer point of view and as well as cloud provider point of view. It is highly scalable monitoring software which could help the developers to customize monitoring according to users' requirements. Nimbus can monitor high level, low-level underlay networks.

Many other open source monitoring systems like GMonE, DARGOS, Lattice, PCMONS, mOSAIC and CASViD have been described previously.

Cloud provider and licensed monitoring software

Cloud providers have proprietary monitoring software that is available for consumers. It is provided according to the SLA defined metrics which could be used to monitor different layers of cloud. For example, Amazon Cloud Watch ("Nimbus," n.d.) could monitor applications running on AWS and Amazon EC2. AzureWatch ("Monitoring and Autoscaling features for

Print ISSN: 2054-0957 (Print), Online ISSN: 2054-0965 (Online)

Windows Azure with AzureWatch indepth - AzureWatch," n.d.) could monitor Azure-based resources including web applications, windows instances, SQL databases and windows storage. Licensed cloud monitoring software is available for monitoring several cloud platforms at once. For instance, Nimsoft ("Nimsoft," n.d.) can be used to monitor Rackspace cloud, Google App Engine, Google Apps, S3 Web Services, Amazon EC2, Salesforce CRM and Microsoft Azure. Meanwhile, CloudKick (Patel & Meniya, 2013) can be used to monitor Rackspace cloud, GoGrid and Amazon EC2. These tools have the abilities to monitor low-level and underlay networks metrics.

ANALYSIS OF CLOUD MONITORING

The discussion of monitoring in cloud computing environment indicates many solutions have been proposed in academic research, open source and proprietary software. Each of them has specific abilities which might be a good option for certain users depending on the users' monitoring requirements. Table 1, provides the analysis of monitoring focus in the reviewed works, which can be classified in respect of high level, low level, underlay and overlay cloud monitoring.

Table 1: Monitoring focus in cloud monitoring systems.

Work	Cloud Layers	High Level	Low Level	Underlay	Overlay
(Adinolfi et al., 2012),(Lakshmanan et al., 2010),(Nair & Gopalakrishna, 2009),(Emeakaroha et al., 2012)	SaaS	□			
(J. Shao & Wang, 2011)	PaaS	□			
(Lee, 2010)	SaaS, PaaS	□			
(J. S. J. Shao et al., 2010),(Katsaros et al., 2012),(Katsaros et al., 2012),(Ma et al., 2012)	SaaS, PaaS, IaaS	□	□		
(Alhamazani et al., 2012), (Alhamazani et al., 2014), (Katsaros et al., 2011), (Povedano-Molina et al., 2013), (Rak et al., 2011), ("OpenNebula.," n.d.), ("Nimbus," n.d.), ("Nagios Core. Nagios Open Source Project.," n.d.), ("Amazon CloudWatch - Cloud & Network Monitoring Services," n.d.)	SaaS, PaaS, IaaS	□	□	□	
(Dhingra et al., 2012), (Kai et al., 2013), (Suciu et al., 2014), (Kim et al., 2013), (Scharf et al., 2012), (Mullins & Bagula, 2013), (Clayman et al.,	IaaS		□	□	

Print ISSN: 2054-0957 (Print), Online ISSN: 2054-0965 (Online)

2010), (Peng & Chen, 2011)					
(Bardhan & Milojicic, 2012), (Duo Liu, Utkarsh Kanabar, 2013), (Grati & Boukadi, 2014), (Ye et al., 2012), (Rodrigues et al., 2012), (Meera & Swamynathan, 2013), (Moses et al., 2011), (De Chaves et al., 2011), (Grigoras et al., 2011)	IaaS		□		
(Andreolini et al., 2012), (Liu & Wood, 2015), (Mann et al., 2013), (Deri & Fusco, n.d.), (Mamta Madan, 2014)	IaaS			□	

In the table, it is clear that different monitoring efforts may focus on different aspects of cloud monitoring. Generally high-level, low-level and underlay monitoring have been used in monitoring SaaS, PaaS and IaaS. However, none of them monitor overlay cloud computing services or performance. The reason for the lack of focus on overlay cloud monitoring could be the application of overlay concept in cloud computing services is just at the beginning.

Based on the reviewed work, monitoring metrics can be furthered classified based monitoring focus as depicted in Table 2. High-level monitoring has been done at SaaS and PaaS layers while low level and underlay monitoring have been done at IaaS layer only. Several monitoring metrics were used to monitor cloud computing performance and services. The detail list of the monitoring metric is provided in Table

Table 2: Generic cloud monitoring metrics

Monitoring Types	Layer	Monitoring Metrics
High level	SaaS	delay, loss, availability, utilization, bytes read and write
	PaaS	system uptime, system description, system processes, system services, memory utilization, bytes read and write
Low Level	IaaS	CPU speed, temperature, workload, utilization, disk/memory throughput, uptime, response time, number of memory page exchange, number of VM and VM startup time
Underlay	IaaS	Jitter, throughput, response time, one-way delay, round-trip time, packet loss, capacity and traffic volume.
Overlay	IaaS	No information found

Print ISSN: 2054-0957 (Print), Online ISSN: 2054-0965 (Online)

CONCLUSION

This paper surveyed and discussed the state of the art of different types of cloud monitoring and generic cloud monitoring metrics. Cloud monitoring services collect a range of performance data on servers, storage and other services within a provider's ecosystem. Unfortunately, cloud network monitoring features are often limited, which means they can miss out on major performance issues in the cloud environment. To keep a close watch on cloud networks and catch potential problems, a cloud network monitoring tool to track and report more in-depth on performance is needed. This paper provides a survey of cloud monitoring metrics which can guide us for developing a cloud monitoring solution for consumers and service providers. A comprehensive framework of cloud network monitoring model will be the next stage of our future study.

REFERENCES

Adinolfi, O., Cristaldi, R., Coppolino, L., & Romano, L. (2012). QoS-MONaaS: A portable architecture for QoS monitoring in the cloud. *8th International Conference on Signal Image Technology and Internet Based Systems, SITIS 2012r*, 527–532. https://doi.org/10.1109/SITIS.2012.82

Alhamazani, K., Ranjan, R., Mitra, K., Jayaraman, P. P., Huang, Z., Wang, L., & Rabhi, F. (2014). CLAMS: Cross-layer Multi-cloud Application Monitoring-as-a-Service Framework. *2014 IEEE International Conference on Services Computing*, (iii), 283–290. https://doi.org/10.1109/SCC.2014.45

Alhamazani, K., Ranjan, R., Rabhi, F., Wang, L., & Mitra, K. (2012). Cloud monitoring for optimizing the QoS of hosted applications. *CloudCom 2012 - Proceedings: 2012 4th IEEE International Conference on Cloud Computing Technology and Science*, *1*, 765–770. https://doi.org/10.1109/CloudCom.2012.6427532

Alimi, R., Yang, Y., & Penno, R. (n.d.). Application-Layer Traffic Optimization (ALTO) Protocol. Retrieved from https://tools.ietf.org/html/rfc7285

Amazon CloudWatch - Cloud & Network Monitoring Services. (n.d.). Retrieved from https://aws.amazon.com/cloudwatch/

Andreolini, M., Colajanni, M., & Pietri, M. (2012). A Scalable Architecture for Real-Time Monitoring of Large Information Systems. *2012 Second Symposium on Network Cloud Computing and Applications*, 143–150. https://doi.org/10.1109/NCCA.2012.24

Bardhan, S., & Milojicic, D. (2012). A mechanism to measure quality-of-service in a federated cloud environment. *Proceedings of the 2012 Workshop on Cloud Services, Federation, and the 8th Open Cirrus Summit - FederatedClouds '12*, 19. https://doi.org/10.1145/2378975.2378981

Clayman, S., Galis, A., & Mamatas, L. (2010). Monitoring virtual networks with Lattice. *Network Operations and Management Symposium Workshops (NOMS Wksps), 2010 IEEE/IFIP*, (ii), 239– 246. https://doi.org/10.1109/NOMSW.2010.5486569

Davie, B., & Gross, J. (2012). A stateless transport tunneling protocol for network virtualization (STT). *Draft-Davie-Stt-06*, 1–21. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:A+Stateless+Transport+Tunne ling+Protocol+for+Network+Virtualization+(STT)#0

Print ISSN: 2054-0957 (Print), Online ISSN: 2054-0965 (Online)

De Chaves, S. A., Uriarte, R. B., & Westphall, C. B. (2011). Toward an architecture for monitoring private clouds. *IEEE Communications Magazine*, *49*(12), 130–137. https://doi.org/10.1109/MCOM.2011.6094017

Deri, L., & Fusco, F. (n.d.). MicroCloud-based network traffic monitoring (p. 1418). Proceedings of the 2013 IFIP/IEEE International Symposium on Integrated Network Management (IM 2013) : 27-31 May 2013, Ghent, Belgium.

Dhingra, M., Lakshmi, J., & Nandy, S. K. (2012). Resource usage monitoring in clouds. *Proceedings - IEEE/ACM International Workshop on Grid Computing*, *012*, 184–191. https://doi.org/10.1109/Grid.2012.10

Duo Liu, Utkarsh Kanabar, C.-H. L. (2013). A light weight SLA management infrastructure for cloud computing. *Canadian, Ieee Of, Conference Engineering, Electricalcomputer*, 1–4.

Emeakaroha, V. C., Ferreto, T. C., Netto, M. a S., Brandic, I., & De Rose, C. a F. (2012). CASViD: Application level monitoring for SLA violation detection in clouds. *Proceedings - International Computer Software and Applications Conference*, 499–508. https://doi.org/10.1109/COMPSAC.2012.68

Fatema, K., Emeakaroha, V. C., Healy, P. D., Morrison, J. P., & Lynn, T. (2014). A survey of Cloud monitoring tools: Taxonomy, capabilities and objectives. *Journal of Parallel and Distributed Computing*, *74*(10), 2918–2933. https://doi.org/10.1016/j.jpdc.2014.06.007

Grati, R., & Boukadi, K. (2014). A framework for IaaS-to-SaaS monitoring of BPEL processes in the Cloud : design and evaluation, 557–564. https://doi.org/10.1109/AICCSA.2014.7073248

Grigoras, C., Voicu, R., Tapus, N., Legrand, I., Carminati, F., & Betev, L. (2011). MonALISA-based grid monitoring and control. *European Physical Journal Plus*, *126*(1), 1–7. https://doi.org/10.1140/epjp/i2011-11009-9

He, Y., Wang, X., Chen, Y., Du, Z., Huang, W., & Chai, X. (2013). A simulation cloud monitoring framework and its evaluation model. *Simulation Modelling Practice and Theory*, *38*, 20–37. https://doi.org/10.1016/j.simpat.2013.06.007

Kai, L., Weiqin, T., Liping, Z., & Chao, H. (2013). SCM: A design and implementation of monitoring system for CloudStack. *Proceedings - 2013 International Conference on Cloud and Service Computing, CSC 2013*, 146–151. https://doi.org/10.1109/CSC.2013.30

Katsaros, G., Kousiouris, G., Gogouvitis, S. V., Kyriazis, D., Menychtas, A., & Varvarigou, T. (2012). A Self-adaptive hierarchical monitoring mechanism for Clouds. *Journal of Systems and Software*, *85*(5), 1029–1041. https://doi.org/10.1016/j.jss.2011.11.1043

Katsaros, G., Kübert, R., & Gallizo, G. (2011). Building a service-oriented monitoring framework with REST and nagios. *Proceedings - 2011 IEEE International Conference on Services Computing, SCC 2011*, 426–431. https://doi.org/10.1109/SCC.2011.53

Kim, Y. M., Lee, K. S., Uhm, J. C., Kim, S. C., Lee, C. G., Song, M., & Woo, H. (2013). Architecture of a network performance monitor for application services on multi-clouds. *International Conference on Ubiquitous and Future Networks, ICUFN*, 594–599. https://doi.org/10.1109/ICUFN.2013.6614888

Lakshmanan, G. T., Keyser, P., Slominski, A., Curbera, F., & Khalaf, R. (2010). A business centric endto-end monitoring approach for service composites. *Proceedings - 2010 IEEE 7th International Conference on Services Computing, SCC 2010*, 409–416. https://doi.org/10.1109/SCC.2010.45

Lee, Y. (2010). QoS metrics for service level measurement for SOA environment. *2010 6th International Conference on Advanced Information Management and Service (IMS)*, 509–514. Retrieved from http://ieeexplore.ieee.org/articleDetails.jsp?arnumber=5713503

Liu, G., & Wood, T. (2015). Cloud-Scale Application Performance Monitoring with SDN and NFV. *2015 IEEE International Conference on Cloud Engineering*, 440–445. https://doi.org/10.1109/IC2E.2015.45

M. Sridharan , A. Greenberg , N. Venkataramiah , Y. Wang, K. D., & I. Ganga , G. Lin , M. Pearson, P. T. and C. T. (2012). NVGRE: Network virtualization using generic routing encapsulation. Retrieved from https://buildbot.tools.ietf.org/html/rfc7637

Ma, K., Sun, R., & Abraham, A. (2012). Toward a lightweight framework for monitoring public clouds. *Proceedings of the 2012 4th International Conference on Computational Aspects of Social Networks, CASoN 2012*, 361–365. https://doi.org/10.1109/CASoN.2012.6412429

Madan, M., & Mathur, M. (2014). Cloud Network Management Model A Novel Approach to Manage Cloud Traffic. https://doi.org/10.5121/ijccsa.2014.4502

Mahalingam, M., Hutt, D. G., Duda, K., Agarwal, P., Kreeger, L., Sridhar, T., … Wright, C. (2014). VXLAN: A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks, (October), 1–23. Retrieved from http://tools.ietf.org/pdf/draft-mahalingam-dutt-dcops-vxlan-09.pdf

Mamta Madan, M. M. (2014). Cloud network management model -A novel approach to manage cloud traffic. *International Journal on Cloud Computing: Services and Architecture (IJCCSA)*, *4*(October), 9–20.

Mann, V., Vishnoi, A., & Bidkar, S. (2013). Living on the edge: Monitoring network flows at the edge in cloud data centers. *2013 Fifth International Conference on Communication Systems and Networks (COMSNETS)*, 1–9. https://doi.org/10.1109/COMSNETS.2013.6465540

Meera, a., & Swamynathan, S. (2013). Agent based Resource Monitoring System in IaaS Cloud Environment. *Procedia Technology*, *10*, 200–207. https://doi.org/10.1016/j.protcy.2013.12.353

Mei, Y., Liu, L., Pu, X., & Sivathanu, S. (2010). Performance measurements and analysis of network I/O applications in virtualized cloud. In *Proceedings - 2010 IEEE 3rd International Conference on Cloud Computing, CLOUD 2010* (pp. 59–66). IEEE. https://doi.org/10.1109/CLOUD.2010.74 Monitoring and Autoscaling features for Windows Azure with AzureWatch indepth - AzureWatch. (n.d.). Retrieved January 1, 2016, from https://www.paraleap.com/AzureWatch

Montes, J., Sánchez, A., Memishi, B., Pérez, M. S., & Antoniu, G. (2013). GMonE: A complete approach to cloud monitoring. *Future Generation Computer Systems*, *29*(8), 2026–2040. https://doi.org/10.1016/j.future.2013.02.011

Moses, J., Iyer, R., Illikkal, R., Srinivasan, S., & Aisopos, K. (2011). Shared resource monitoring and throughput optimization in cloud-computing datacenters. *Proceedings - 25th IEEE International Parallel and Distributed Processing Symposium, IPDPS 2011*, 1024–1033. https://doi.org/10.1109/IPDPS.2011.98

Mullins, T., & Bagula, A. (2013). Monitoring community clouds: The lightweight network management protocol. *Proceedings - IEEE 10th International Conference on Ubiquitous Intelligence and Computing, UIC 2013 and IEEE 10th International Conference on Autonomic and Trusted Computing, ATC 2013*, 678–684. https://doi.org/10.1109/UIC-ATC.2013.31

Nagios Core. Nagios Open Source Project. (n.d.). Retrieved from https://www.nagios.org/

Nair, M. K., & Gopalakrishna, V. (2009). "CloudCop": Putting network-admin on cloud nine: Towards cloud computing for network monitoring. *2009 IEEE International*

Conference on Internet Multimedia Services Architecture and Applications, IMSAA 2009. https://doi.org/10.1109/IMSAA.2009.5439474
Nimbus. (n.d.). Retrieved from http://www.nimbusproject.org/
Nimsoft. (n.d.). Retrieved from http://www.nimsoft.com
OpenNebula | Flexible Enterprise Cloud Made Simple. (n.d.). Retrieved from http://opennebula.org/
Patel, H. A., & Meniya, A. D. (2013). A Survey on Commercial and Open Source Cloud Monitoring. *International Journal of Science and Modern Engineering (IJISME)*, *1*(2).
Peng, Y.-S., & Chen, Y.-C. (2011). SNMP-based monitoring of heterogeneous virtual infrastructure in clouds. *13th Asia-Pacific Network Operations and Management Symposium*, 1–6. https://doi.org/10.1109/APNOMS.2011.6077010
Phaal, P., Panchen, S., & McKee, N. (2001). InMon Corporation’s sFlow: A Method for Monitoring Traffic in Switched and Routed Networks, (3176), 31. Retrieved from http://www.rfceditor.org/rfc/rfc3176.txt
Povedano-Molina, J., Lopez-Vega, J. M., Lopez-Soler, J. M., Corradi, A., & Foschini, L. (2013).
DARGOS: A highly adaptable and scalable monitoring architecture for multi-tenant Clouds.
Future Generation Computer Systems, *29*(8), 2041–2056. https://doi.org/10.1016/j.future.2013.04.022
Rak, M., Venticinque, S., Máhr, T., Echevarria, G., & Esnal, G. (2011). Cloud application monitoring: The mOSAIC approach. *Proceedings - 2011 3rd IEEE International Conference on Cloud*
Computing Technology and Science, CloudCom 2011, 758–763. https://doi.org/10.1109/CloudCom.2011.117
Rodrigues, C., Granville, L. Z., & Tarouco, L. M. R. (2012). Network and Services Monitoring : A Survey in Cloud Computing Environments, (c), 7–13.
Scharf, M., Voith, T., Roome, W., Gaglianello, B., Steiner, M., Hilt, V., & Gurbani, V. K. (2012). Monitoring and abstraction for networked clouds. *2012Proceedings of the Intelligence in Next Generation Networks, ICIN 2012*, 80–85. https://doi.org/10.1109/ICIN.2012.6376038
Shao, J. S. J., Wei, H. W. H., Wang, Q. W. Q., & Mei, H. M. H. (2010). A Runtime Model Based Monitoring Approach for Cloud. *Cloud Computing (CLOUD), 2010 IEEE 3rd International Conference On*. https://doi.org/10.1109/CLOUD.2010.31
Shao, J., & Wang, Q. (2011). A performance guarantee approach for cloud applications based on monitoring. *Proceedings - International Computer Software and Applications Conference*, 25–30.
https://doi.org/10.1109/COMPSACW.2011.15
SIGAR. (n.d.). Retrieved from https://support.hyperic.com/display/SIGAR/Home
Smit, M., Simmons, B., & Litoiu, M. (2013). Distributed, application-level monitoring for heterogeneous clouds using stream processing. *Future Generation Computer Systems*, *29*(8), 2103–2114. https://doi.org/10.1016/j.future.2013.01.009
Stallings, W. (1998). SNMP,SNMPV2,Snmpv3,and RMON 1 and 2. Retrieved from http://dl.acm.org/citation.cfm?id=521036
Suciu, G., Halunga, S., Ochian, A., & Suciu, V. (2014). Network Management and Monitoring for Cloud Systems, 1–4.
Ward, J. S., & Barker, A. (2013). Varanus: In Situ Monitoring for Large Scale Cloud Systems. *2013 IEEE 5th International Conference on Cloud Computing Technology and Science*, *2*, 341–344. https://doi.org/10.1109/CloudCom.2013.164

Ye, L., Zhang, H., Shi, J., & Du, X. (2012). Verifying Cloud Service Level Agreement. *2012 IEEE Global Communications Conference (GLOBECOM)*, 777–782. https://doi.org/10.1109/GLOCOM.2012.6503207