Czech University of Life Sciences Prague

Faculty of Economics and Management

Department of Information Technologies

Diploma Thesis

Analysis of cloud based monitoring tools

Bc. Murodjon Ganiyev

CZECH UNIVERSITY OF LIFE SCIENCES PRAGUE

Faculty of Economics and Management

DIPLOMA THESIS ASSIGNMENT

Murodjon Ganiyev

Informatics

Thesis title

Analysis of cloud based monitoring tools

Objectives of thesis

The diploma thesis is focused on the use of cloud solution for monitoring business applications. Main goal of the thesis is to analyse and select the optimal solution for monitoring of application in the cloud for a company.
Partial goals are:
- to make a literature review of the current state of the art of cloud computing, monitoring tools and administration;
- to analyse a particular case of selected company that will move its applications to the cloud;
- to identify possible solutions for monitoring of applications in cloud and make a decision.

Methodology

The literature review will be based on well-known international books, white papers, journals, articles and international actual online statistical sources. Methodology of the thesis is based on study and analysis of well know monitoring tools. In the practical part, a case study about application of selected monitoring tool will developed. Based on synthesis of theoretical knowledge and the results of practical solutions the conclusion of the thesis will be formulated. Author will use Multiple Attribute Decision Making methods in the thesis.

The proposed extent of the thesis

60 – 80 pages

Keywords

Cloud computing, SaaS, PaaS, IAAS, service, hybrid cloud, private cloud, public cloud, virtualisation, operation, monitoring data centre

Recommended information sources

1. FRED VAN DER MOLEN (LEAD AUTHOR) a [TEXT ED.: COLIN BRACE .. ET AL.]. Get ready for cloud computing: a comprehensive guide to virtualization and cloud computing. 1st ed. Zaltbommel: Van Haren Publishing, 2010. ISBN 9789087536404.
2. AHSON, Syed a Mohammad ILYAS. 2011. Cloud computing and software services: theory and techniques. Boca Raton, FL: CRC Press, xiv, 442 p. ISBN 9781439803158.
3. SMOOT, Stephen R a Nam Kee TAN. 2012. Private cloud computing: consolidation, virtualization, and service-oriented infrastructure. Waltham, MA: Morgan Kaufmann, xvii, 399 p. ISBN Private cloud computing.
4. MOYER, Christopher M. 2011. Building applications in the cloud: concepts, patterns, and projects. Upper Saddle River, NJ: Addison-Wesley, xiii, 326 s. ISBN 978-0-321-72020-7.
5. Talkincloud.com [online]. [cit. 2015-10-14]. Dostupné z: http://talkincloud.com/
6. Docs.splunk.com [online]. [cit. 2015-10-14]. Dostupné z: http://docs.splunk.com/

Expected date of thesis defence

2015/16 SS – FEM

The Diploma Thesis Supervisor

Ing. Miloš Ulman, Ph.D.

Supervising department

Department of Information Technologies

Electronic approval: 2. 11. 2015

Ing. Jiří Vaněk, Ph.D.

Head of department

Electronic approval: 11. 11. 2015

Ing. Martin Pelikán, Ph.D.

Dean

Prague on 24. 03. 2016

Declaration

I declare that I have worked on my diploma thesis titled "Analysis of cloud based monitoring tools" by myself and I have used only the sources mentioned at the end of the thesis. As the author of the bachelor thesis, I declare that the thesis does not break copyrights of any third person.

In Prague on 20.03.2016 Bc.Ganiyev Murodjon _____

Acknowledgement

I would like to thank and acknowledge my thesis supervisor Ing. Miloš Ulman, Ph.D. for his incredible contribution, guidance, suggestions and friendly co-operation during my thesis compilation. From the choice of thesis topic to the final completion of the thesis, it always gave me a sensitive leadership and tireless support. My supervisor gave me great help in the research, not only helped me to correct errors, but also gave me valuable advice. I would like to take this opportunity to express my heartfelt gratitude and deepest respect.

Second thanks goes to my direct technical team leader Menno Kloos for his brilliant cooperation, motivation and support in this diploma thesis. And all other my colleagues, NN engineers especially for Erwin Horjus for his warm cooperation in it.

Lastly, I would like to extend my gratefulness and appreciation to my farther Ganiyev Sodiqjon, my mother Shoira Yuldasheva, especially for my lovely wife Zebo (I would not make it without her moral support), my brothers , my sister and all friends and others that have contributed to my development for their financial, moral and general support in bringing me thus far in my life. I feel blessed to have this group of people around me.

Analysis of cloud based monitoring tools

Summary

The goal of this research is to find out optimum cloud monitoring solution and proposing to technical implementation for a company. Thesis focuses on a specific topic in a cloud monitoring area and goes into more details.

Firstly, author makes critical literature review of the topic in general cloud computing, service and deployment models, cloud computing monitoring, taxonomy of monitoring areas. Secondly, analysis company's ambitions, visions, goals, short/long term strategic plans in general by making brief case study.

Thirdly, author starts analysis to find out best cloud monitoring solutions to monitor a large enterprise applications by applying empirical methods.

Fourthly, author designs and technically implements new solution as a proposal.

Author, include in his analyses new log-file based monitoring tool called Splunk and proposes as a best monitoring tool in this company's on-premise and hybrid cloud environments. At the end author proposes to implement Splunk by making component design of this tool.

This research is targeted also to:

- Public, Private or Hybrid cloud solution integrated enterprises.
- Enterprises that suppose to migrate data into cloud.
- Enterprises that use very large scalable embedded systems.

Keywords: Cloud computing, SaaS, PaaS, IaaS, service, hybrid cloud, private cloud, public cloud, virtualisation, operation, monitoring data centre

Analýza cloudových monitorovacích nástrojů

Souhrn

Cílem tohoto výzkumu je nalezení optimálního cloudového monitorovacího řešení a jeho technické implementace do společnosti. Diplomová práce se zaměřuje na specifickou problematiku v oblasti cloud monitoringu a rozebírá ji do hloubky.

Za prvé, autor rozebírá odbornou literaturu a popisuje problematiku cloud computingu, služeb a modelů nasazení, monitoring cloud computingu a taxonomii monitorovacích oblastí.

Za druhé, akademická práce obsahuje analýzu ambicí společnosti, vizí, cílů, krátkodobých a dlouhodobých strategických plánů v obecné rovině na příkladu stručné případové studie.

Za třetí, autor provádí analýzu k nalezení nejlepšího cloudového monitorovacího řešení k monitorování velkých podnikových aplikací použitím empirických metod.

Za čtvrté, autor navrhuje řešení a možnou technickou implementaci tohoto návrhu. Autor zahrnuje do analýzy nový monitorovací nástroj založený na zaznamenávání informací do logu zvaný Splunk a navrhuje ho jako nejlepší monitorovací nástroj ve zkoumané společnosti za předpokladu hybridního cloudového prostředí. Na konci autor navrhuje implementovat Splunk prostřednictvím návrhových komponent tohoto nástroje.

Tento výzkum se zaměřuje rovněž na:

- Veřejné, privátní nebo hybridní cloudová řešení integrovaná do společností.
- Společnosti, které uvažují o migraci svých dat do cloudu.
- Společnosti, které používají velké škálovatelné vestavné systémy.

Klíčová slova: Cloud computing, SaaS, PaaS, IaaS, servis, hybridní cloud, privátní cloud, veřejný cloud, virtualizace, provoz, monitorování datových center

Table of content

List of tables

List of Figures

List of Annexes

Acronyms

1. IOPS- Input/output Operations Per Second
2. ETF- Enterprise Technology Framework
3. VPC- Virtual Private Cloud
4. ALM- Application life cycle management
5. AWS Amazon Web Services
6. CSP- Cloud Service Providers
7. CSA- Cloud Security Alliance
8. LSA- Local Security Account
9. MFA- Multi Factor Authentication
10. SOA - Service Oriented-Architecture
11. CDN - Content Delivery Network
12. COTS- Commercial off-the-shelf (COTS)- is a term used to describe the purchase of products that are standard manufactured products rather than custom products
13. CIO- Cheaf Exequitive Office
14. A-OSG- Operational Security Guidelines (of an application)
15. BCP- Business Continuity Plan
16. BIA-Business Impact Assessment
17. CIA-Confidentiality Integrity Availability-basic security requirements
18. CMDB-Configuration Management Database
19. DTAP-Development, Test, Acceptance and Production environments
20. IT RAM - IT Risk Assessment Methodology
21. LPAD-Logical and Physical Architectural Design or Solution architecture (of an infrastructure service)
22. OSG-Operational Security Guidelines (of an infrastructural service)
23. RCEC-Registration & Certification of External Connections
24. RPO-Recovery Point Objective: the point in time to which work should be restored following a business continuity emergency event, incident or crisis.
25. RTO Failure - Recovery Time Objective Failure: the time by which the application and its dependencies must be recovered after a failure.
26. RTO Disaster Recovery Time Objective Disaster: the time by which the application and its dependencies must be recovered after a disaster.
27. SAP-Security Action Plan (part of the Sprint IT Risk Management methodology)
28. SLA-Service Level Agreement
29. SRA-Security Risk Assessment (part of the Sprint IT Risk Management methodology)
30. SRL-Security Requirements List
31. SCCM- Software Configuration and Change Management - controlling the evolution of a software product
32. JMS- Java Messaging Services
33. SIEM- Security Information and Event Management

1 Introduction

As water and an electricity, IT services, systems and technology infrastructure, websites, applications, servers, networks, sensors, mobile devices and all generate massive amounts of machine data's are becoming our major day to day consumptions. The emergence of Cloud Computing has ushered in a new era of Internet-based service provisioning opportunities. The cloud is no longer an emerging technology. it's an essential one for businesses. Today, is very rare to find company which is not using cloud technologies in some way.

70% of organisations in the world are already using the cloud[49] 72% of businesses polled expect to put more than half of their workloads in the cloud by 2017[49]. More than 50% enterprises using private cloud, 33% enterprises using public cloud 22% are adopted into hybrid cloud [49].

NN Group (we will define as NN in the next chapters) is also one the biggest finsurance company which going to transform its datacentre into hybrid and at the end into public cloud. NN is one of the largest insurance company in the Netherlands. It's famous as insurance and asset management company active in more than 18 countries, with a strong presence in a number of European countries and Japan. Recently company decided to migrate there entire infrastructure into cloud. Author is monitoring specialist on-premise infrastructures of this company and he challenged to find out best monitoring solution to monitor NN applications on cloud.

The size of the program embraces 1100 applications, divided over 14 countries and more than 20 business and functional units. When, company reaches in's goals allow than 1100 applications will be managed via cloud.

Main goal of the thesis is to find out the best possible monitoring tool which can help to NN during the migration as an hybrid scenario and as a long term to monitor their applications on cloud.

Why is important for NN?

- Digital is disruptive for traditional industries
- NN needs agility to be successful
- DC is really costly

- Development and Test not need anymore
- At the end NN group should not have its own hosting

Monitoring solution must be able monitor at least:

- High-level cloud monitoring issues
- Large scale envviremnts
- Must have capapbilies to monitor all cloud service models
- Ability to monitor in all cloud deployment models

Cloud Computing is characterised by the provision of resources as general utilities that can be leased and released in an on-demand manner. Consequently, IT resources represent an operational rather than a capital expenditure. A broad variety of pricing models can be applied to Cloud resources, from simple fixed rental schemes to pay-as-you-go models. Monitoring techniques are indispensable in order to manage large-scale Cloud resources and enforce quality of service for consumers. Given the multi-tenant nature of Cloud environments, efficient management in the face of quality of service and performance constraints can be a challenge. Monitoring tools have an important role to play in these areas by allowing informed decisions to be made regarding resource utilisation. Automated monitoring of physical and virtual IT resources allows for the identification and resolution of issues with availability, capacity, and other quality requirements. The benefits of automated monitoring have long been recognised, even in non-Cloud environments. The importance of monitoring has been widely addressed in the literature in various contexts, such as: system/network, distributed systems/Grid application and Cloud. For Cloud environments, appropriate monitoring is crucial as usage based billing and elastic scaling are impossible to implement in the absence of relevant metrics. Currently, a variety of Cloud monitoring tools is applied in an ad-hoc and non-systematic way, everywhere from low-level, general-purpose infrastructure monitoring to high-level application and service monitoring. The purpose of this paper is to comprehensively review these tools to assess whether they are adequate in satisfying the essential objectives for measuring intrinsic Cloud behaviours. The focus of our work is to capture the evolutionary adaptation of monitoring tools' capabilities from general purpose to Cloud monitoring and to present a full capability analysis with respect to practical Cloud

operational areas that would help Cloud providers and customers in making an informed choice of an appropriate monitoring tool. The monitoring platforms considered in this paper have been chosen based on literature reviews and perceived industrial acceptance.

Monitoring tools all about capabilities delivered as a service with a clear boundary between the provider of the service and the consumer.

Capabilities based on Figure 4

Cloud computing promises us to increase economies of scale, agility, flexibility, speed, and infinitely elastic and innovative advantages over traditional IT in that certain time frame.

2 Objectives and Methodology

2.1 Objectives

The diploma thesis is focuses on the use of cloud solution for monitoring business applications. Main goal of the thesis is to analyse and select an optimal solution for monitoring of application in the cloud for a company.

Partial goals are:

- to make a literature review of the current state of the art of cloud computing, monitoring tools and administration;
- to analyse a particular case of selected company that will move its applications to the cloud;
- to identify possible solutions for monitoring of applications in cloud and make a decision.

2.2 Methodology

The literature review will be based on well-known international books, white papers, journals, articles and international actual online statistical sources. Methodology of the thesis is based on study and analysis of well know monitoring tools. In the practical part, a case study about application of selected monitoring tool will be developed. Based on synthesis of theoretical knowledge and the results of practical solutions the conclusion of the thesis will be formulated. Author will use one of the most known Multiple Attribute Decision Making scoring method in the thesis, and makes SWOT analysis chosen tool.

The author then identifies the cloud monitoring capabilities. These capabilities are generalized groups of features commonly required by this class of cloud based monitoring tools. Each capability is assigned a level of importance in fulfilling that particular need; some sets of features are more important than others, depending on the use case being evaluated. Use cases are collected requirements from NN cloud engineers.

Each monitoring tools are evaluated in terms of how well it delivers each capability, on ten points ratings. These ratings are displayed side-by-side for all monitoring tools, allowing easy comparisons between the different sets of features.

To determine an overall score for each capability in the use cases, the tools ratings are multiplied by the weightings to come up with the product score in use cases. Author implemented them by MADM using scoring method.

The prioritized capabilities which author selected from NN egnineers are not represent all capabilities for any monitroing tools; therefore, may not represent those most important for a specific use situation or business objective. Consumers should use a prioritized capabilities analysis as one of several sources of input about a product before making a monitoring tools decision.

3 Literature Review

In this part firstly, author makes a critical literature review of the topic in general cloud computing, service and deployment models, cloud computing monitoring, taxonomy of monitoring areas. At the end of this chapter, cloud service providers, that will be subjects of decision making in the analytical part, are briefly introduced and described at the end of the chapter.

3.1 Cloud Computing

Cloud computing [28]promises us to increase economies of scale, agility, flexibility, speed, and infinitely elastic and innovative advantages over traditional way of working.

According to National Institute of Standards and Technology (NIST) cloud computing is a " Model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (network, servers, storage, applications, services) that can be rapidly provisioned and released with minimal management effort or service provider interaction" [1].

Cloud service models, deployment models, hosting, and roles are some of the important concepts and essential characteristics related to cloud computing technologies defined by NIST[1] and elaborated in, open source and commercial cloud service providers(in the below CSP) including Amazon Web Services (AWS), Microsoft Azure, Salesforce.com, Google App Engine and others offer the cloud consumers options to deploy their applications over a network of infinite resource pool with practically no capital investment and with modest operating cost proportional to the actual use. For example, Amazon EC2 cloud runs around half million physical hosts, each of them hosting multiple virtual machines that can be dynamically invoked or removed [23].

3.1.1 Strength of cloud monitoring

- Fully virtualised and self-service base
- Fully automated and integrated
- Fully international standardized and necessity
- Monitoring and visibility even more important in the cloud

3.1.2 Main characteristics of cloud computing

Essential characteristics of cloud computing, is actually “cloud” one of the general utility for computing services, sharing storage, network and other IT operations. One of the well-known publisher Peter Mell says "When agencies or companies use this definition, they have a tool to determine the extent to which the information technology implementations they are considering meet the cloud characteristics and models," says Peter Mell in NIST[1] As other general unities, cloud is a best in economies of scale cost and energy saving IT utility.

NIST, also summarizes[8] definitions with five common essential characteristics of cloud computing.

There are:

1. **On-demand self-service:** It’s automation of cloud commuting which is consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service provider. IT industries now a days requires energy of human resources than any other inductors and automation of IT services becoming challenge task todays IT companies.
2. **Broad network access:** “*Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, tablets, laptops and workstations)”*.[8] As other communal unities like power supply, oil, gas - IT services also becoming common general utility which consumers willing to pay per use. Businesses finding out why small reality broker has to worry about their server, file utilisation or small e-shop owner thinks about server malware, SQL injections live attacks to their e-shop web site? Such questions are actually essential characterises of cloud environment.
3. **Resource pooling:** “*The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand.”* [8] There is of course location independence policies that consumer generally has no control over the exact location of the provided resources but they can able to

specify location at a higher abstraction levels (e.g., country, state or datacentre). For instances, resources include processing, storage, memory and network levels.

4. **Rapid elasticity:** Is one of the magic characteristics of cloud computing. In one word we can say "capability". Cloud computing gives us chance to be elastic as possible and to be provisioned and released, in most cases it's automatically, there is also definition "*to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be appropriated in any quantity at any time.*"[8]
5. **Measured service:** *"Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth and active user accounts). Resource usage can be monitored, controlled and reported, providing transparency for the provider and consumer"*.[8]

NIST also informs:[1] The definitions are intended to serve as a means for broad comparisons of cloud services and deployment plans, and to provide a baseline for discussion from what is cloud computing to how to best use cloud computing.

3.2 Service models

To be able to get full advantages of cloud computing, we must closely look into well-known cloud computing service models. Recently, we can read almost in all cloud researches one very common word called "SPI[1]" acronym which intended as a most common cloud computing service models Software as a Service(SaaS), Platform as a Service(PaaS) and Infrastructure as a Service(IaaS).

These service models[9] are explained as layered based as first SaaS then PasS, IaaS. In these below chapters we will take a look into most common and brief literature view of all tree most common cloud computing models.

[1] SPI - Acronym for cloud commuting models as SaaS, PaaS and IaaS

3.2.1 Software as a Service (SaaS)

Research Director in Gartner, Dennis Smith clearly pointing out that Gartner survived of indicated Fifty-five percent of CIOs will structure more than half of their applications as SaaS or manage them in a public cloud infrastructure by end of 2020[14]

SaaS service model is software as service approach becoming one of the most common feature model of cloud computing. due to its usability. PCMAG encyclopaedia explains SaaS as following *"Software that is rented rather than purchased. Instead of buying software and paying for periodic upgrades, SaaS is subscription based, and all upgrades are provided during the term of the subscription. When the subscription period expires, the software is no longer valid."*[11]

There is also some disadvantages of this model that all SaaS users should be worried. SaaS environments are often comes with potential risks like security leaks, data breaches from hackers and some other internal errors. Even SaaS providers, they belong political groups or in worth cases even governments may predict consumers behaviour, and they can forcast consumed government future plans.

3.2.2 Platform as a Service (PaaS)

Gartner explaining PaaS as *"A **platform as a service (PaaS)** offering, usually depicted in all-cloud diagrams between the SaaS layer above it and the IaaS layer below, is a broad collection of application infrastructure (middleware) services (including application platform, integration, business process management and database services). However, the hype surrounding the PaaS concept is focused mainly on application PaaS (aPaaS) as the representative of the whole category."*[12]

3.2.3 Infrastructure as a Service (IaaS)

Is one of the biggest and older cloud computing service that provides complete infrast ructure level of computing area. Gartner points out IaaS as *"**Infrastructure as a service (IaaS)** is a standardized, highly automated offering, where compute resources, complemented by storage and networking capabilities are owned and hosted by a service provider and offered to customers on-demand. Customers are able to self-provision this infrastructure, using a Web-based graphical user interface that serves as an IT operations*

management console for the overall environment. API access to the infrastructure may also be offered as an option." [13]

3.3 Deployment models

This chapter is conducted to give brief explanation and literature view of main and common cloud computing deployment models.

3.3.1 Private cloud

Private cloud is one of the cloud computing platform that implemented within corporate firewall, under the control of the IT department.[15] A private cloud is focused to offer almost the same features and benefits of public cloud models, but it removes a number of objections to the cloud computing models for example controlling over enterprise and customer data, security issues, and issues connected to regulatory compliance agreements[15]

Interoute explains private cloud model as: *"A private cloud is a particular model of cloud computing that involves a distinct and secure cloud based environment in which only the specified client can operate. As with other cloud models, private clouds will provide computing power as a service within a virtualised environment using an underlying pool of physical computing resource. However, under the private cloud model, the cloud (the pool of resource) is only accessible by a single organisation providing that organisation with greater control and privacy."*[16]

3.3.2 Public cloud

Ventilation, grid technology and fast internet capabilities gave us chance to share our IT infrastructure and basically we are calling this offer or service as a public cloud. In most recognisable service model of cloud computing to consumers is the public cloud service model, which is all cloud services will be provided in a virtualised environments, constructed using pooled shared physical resources, and accessible over a public network such as the internet. To some extent they can be defined in contrast to private clouds which ring-fence the pool of underlying computing resources, creating a distinct cloud platform to which only a single organisation has access. Public clouds, however, provide services to multiple clients using the same shared infrastructure [17].

As other cloud computing service models, public cloud offers following benefits and features.[17]

- **Ultimate scalability:** Any limitation on resource pooling, running applications can respond seamlessly to fluctuations in activity.
- **Cost effective:** Due to shared infrastructure, pooled resources approach public cloud is relatively cheaper than any models and at the same time it can benefit from the largest economies of scale. Infrastructure management, server provisioning, space, swap utilisation, application shut down, or network hiccup all can be managed automatically. But we should also point out one fact that cost effectiveness of the public cloud is not the main reason that consumers are willing to choose it, senior Gartner editor Rob van der Meulen submitted article in Stamford conference called "Gartner Highlights the Top 10 Cloud Myths"[18] he clearly mentioning this point that only 14% of survived consumers pointed out to migrate into public cloud due to cost effectiveness.
- **Pay as you go charging style: "***public cloud services often employ a pay-as-you-go charging model whereby the consumer will be able to access the resource they need, when they need it, and then only pay for what they use; therefore avoiding wasted capacity."*[17]
- **Reliability**; the sheer number of servers and networks involved in creating a public cloud and the redundancy configurations mean that should one physical component fail, the cloud service would still run unaffected on the remaining components. In some cases, where clouds draw resource from multiple data centers, an entire data center could go offline and individual cloud services would suffer no ill effect. There is, in other words, no single point of failure which would make a public cloud service vulnerable[17]
- **Flexibility**; there are a myriad of IaaS, PaaS and SaaS services available on the market which follow the public cloud model and that are ready to be accessed as a service from any internet enabled device. These services can fulfill most computing requirements and can deliver their benefits to private and enterprise clients alike. Businesses can even integrate their public cloud

services with private clouds, where they need to perform sensitive business functions, to create hybrid clouds

- **Location independence**; the availability of public cloud services through an internet connection ensures that the services are available wherever the client is located. This provides invaluable opportunities to enterprise such as remote access to IT infrastructure (in case of emergencies etc) or online document collaboration from multiple locations.

3.3.3 Hybrid cloud

A hybrid cloud is an integrated cloud service utilising both private and public clouds to perform distinct functions within the same organisation. All cloud computing services should offer certain efficiencies to differing degrees but public cloud services are likely to be more cost efficient and scalable than private clouds. Therefore, an organisation can maximise their efficiencies by employing public cloud services for all non-sensitive operations, only relying on a private cloud where they require it and ensuring that all of their platforms are seamlessly integrated.[20]

Financial services, government organisations, banking area and all other sensitive data holders are going to move into hybrid cloud environment due to the security and complains policies.

Hybrid cloud models can be implemented in a number of ways:[20]

- Separate cloud providers team up to provide both private and public services as an integrated service
- Individual cloud providers offer a complete hybrid package
- Organizations who manage their private clouds themselves sign up to a public cloud service which they then integrate into their infrastructure

In practice, an enterprise could implement hybrid cloud hosting to host their e-commerce website within a private cloud, where it is secure and scalable, but their brochure site in a public cloud, where it is more cost effective (and security is less of a concern). Alternatively, an Infrastructure as a Service (IaaS) offering, for example, could follow the hybrid cloud model and provide a financial business with storage for client data within a private cloud, but then allow collaboration on project planning documents in the public cloud - where they can be accessed by multiple users from any convenient location.

As other cloud computing models hybrid environment also gives scalability(i), cost efficiencies(ii) security(iii) and flexibility(iv) benefits and features[20]

3.3.4 Community cloud

Gartner explains community cloud as *"computing refers to a shared cloud computing service environment that is targeted to a limited set of organizations or employees (such as banks or heads of trading firms). The organizing principle for the community will vary, but the members of the community generally share similar security, privacy, performance and compliance requirements. Community members may wish to invoke a mechanism that is often run by themselves (not just the provider) to review those seeking entry into the community.*"[27]

At the same time techopedia explains in a little different manner. "A community cloud is a cloud service model that provides a cloud computing solution to a limited number of individuals or organizations that is governed, managed and secured commonly by all the participating organizations or a third party managed service provider."[19]

Community clouds are often designed for businesses and organizations working on joint projects, applications, or research, which requires a central cloud computing facility for building, managing and executing such projects, regardless of the solution rented.

3.4 Cloud monitoring in general

In this chapter, author critically reviews literatures, which are explained more or less about general cloud monitoring approaches. Due to our best knowledge, there is only few related academic white paper[35,42,40,41] which analysed top commercial and open source cloud monitoring tools. However, there is no any academic, significant research or white paper which is analysed cloud monitoring tools in a specific conditions as in practice of large enterprises, and largely, highly distributed systems[2]. Author starts reviewing architecture of the general cloud monitoring tools[35] and then briefly explains taxonomy of cloud monitoring. As we discussed our main goal of this thesis is to analyse cloud monitoring tools. Author, analyses then base in there features and capabilities in the NN cloud environments. In this chapter we will review desirable capabilities of cloud monitoring tools[35] and cloud monitoring operational areas[3,5,6]

Under this section, we present the basic components, phases and layers of application architecture on clouds. Also, this section will present the state of the art in cloud monitoring as well as how it is conceptually correlated to QoS and SLA.

Weaknesses of traditional way of monitoring

- Work only designed within corporate data centres
- Substantial build and deploy time > many months
- Poor elasticity and dynamic provisioning
- Local data collection based on agent monitoring

In the below Figures we can compare traditional and Cloud monitoring operations steps.

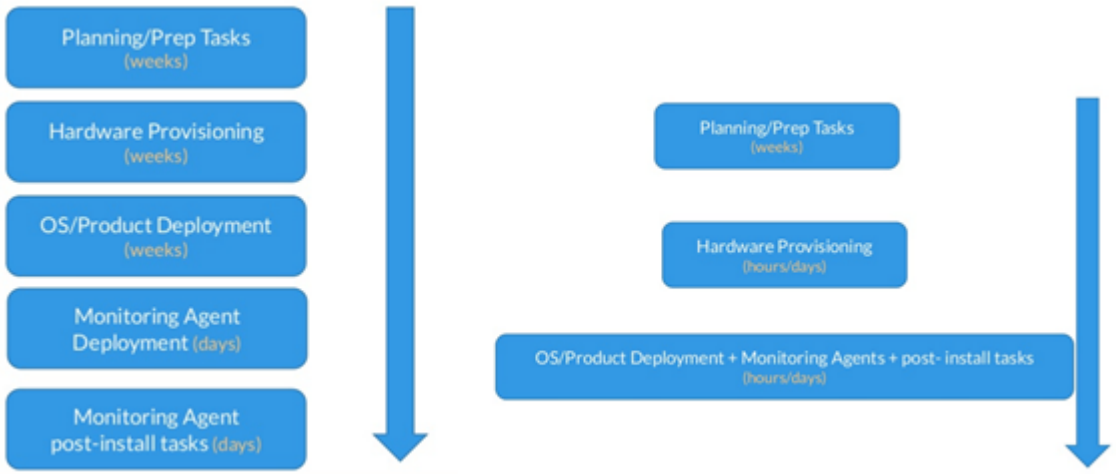


Figure 1 Traditional IT & Cloud monitoring life cycle
Source:[10]

In clouds, monitoring is essential to maintain high system availability and performance of the system and is important for both providers and consumers [2,5]. Primarily, monitoring is a key tool for (i) managing software and hardware resources, and (ii) providing continuous information for those resources as well as for consumer hosted applications on the cloud. Cloud activities like resource planning, resource management, data center management, SLA management, billing, troubleshooting, performance management, and security management essentially need monitoring for effective and smooth operations of the system. Consequently, there is a strong need for monitoring looking at the elastic nature of cloud computing.

Monitoring can be of two types: high-level and low-level. High-level monitoring is related to the virtual platform status[38]. The low-level monitoring is related to information collected about the status of the physical infrastructure [38]. Cloud monitoring system is a self-adjusting and typically multi-threaded system that is able to support monitoring functionalities. It comprehensively monitors pre-identified instances/resources on the cloud for abnormalities. On detecting an abnormal behaviour, the monitoring system attempts to auto-repair this instance/resource if the corresponding monitor has a tagged auto-heal action. In case of auto-repair failure or an absence of an auto-heal action, a support team is notified. Technically, notifications can be sent by different means such as email, or SMS[38].

3.4.1 Cloud monitoring architectures

In cloud monitoring, the network and system related information is collected by the systems. For example, CPU utilization, network delay and packet losses. This information is then used by the applications to determine actions such as data migration to the server closest to the user to ensure that SLA requirements are met. Typically, network monitoring can be performed on centralized and de-centralized network architectures.

Centralized

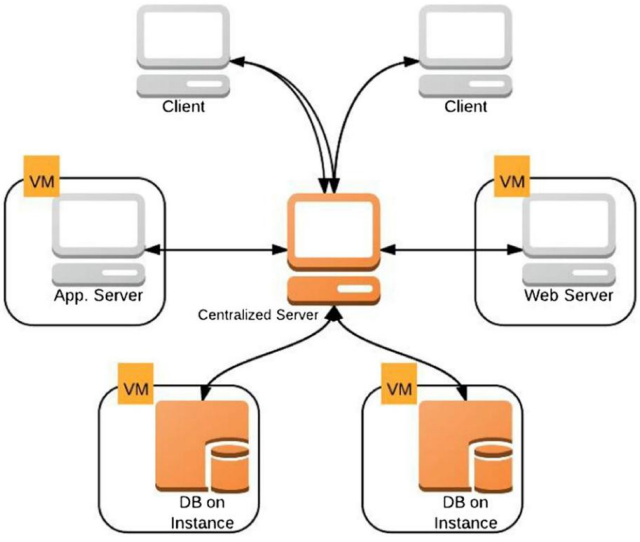


Figure 2: Centralized cloud monitoring
Source: [39]

In centralized architecture shown in [Figure 2] the PaaS and IaaS resources send QoS status update queries to the centralized monitoring server. In this scheme, the monitoring techniques continuously pull the information from the components via periodic probing messages. In , the authors show that a centralized cloud monitoring architecture allows better management for cloud applications. Nevertheless, centralized approach has several design issues, including:

• Prone to a single point of failure;

• Lack of scalability;

• High network communication cost at links leading to the information server (i.e., network bottleneck, congestion); and

• Possible lack of the required computational power to serve a large number of monitoring requests.

Decentralized

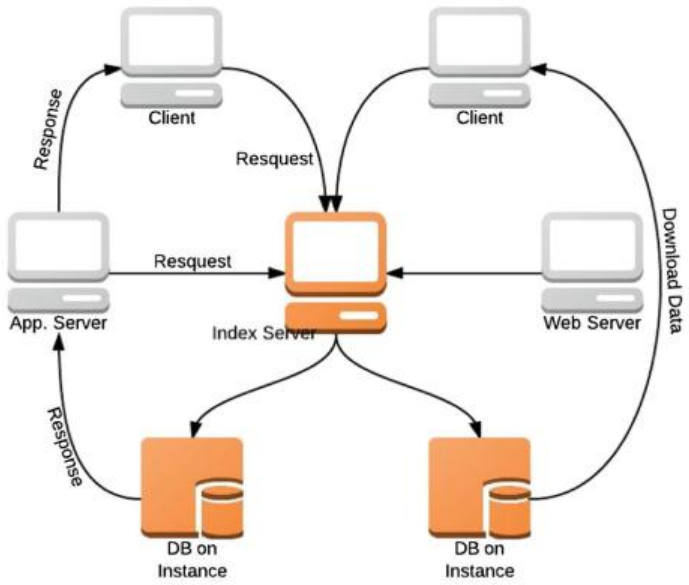


Figure 3: Decentralized cloud monitoring
Source: [39]

Recently, proposals for decentralized cloud monitoring tools have gained momentum. Figure 3 shows the broad schematic design of decentralized cloud monitoring system. The decentralization of monitoring tools can overcome the issues related to current

centralized systems. A monitoring tool configuration is considered as decentralized if none of the components in the system is more important than others. In case one of the components fails, it does not influence the operations of any other component in the system.

Structured peer-to-peer Looking forward to have a network layout where a central authority is defused has lead to the development of the structured peer-to-peer net-works. In such a network overlay, central point of failure is eliminated. Napster is a popular structured peer-to-peer system [39].

Unstructured peer-to-peer Unstructured peer-to-peer networks overlay is meant to be a distributed overlay but the difference is that the search directory is not centralized unlike structured peer-to-peer networks overlay which, leads to absolute single point failure in such network overlay. Gnutella is one of the well-known unstructured peer-to-peer systems [39].

Hybrid peer-to-peer Is a combination of structured and unstructured peer-to-peer net-works systems. Super peers can act as local search hubs in small portions of the network whereas the general scope of the network behaves as unstructured peer-to-peer system.

Hierarchical

Our monitoring technique is strictly a software monitoring system. No special hardware support is required to use our monitoring system as it is the case in hardware and hybrid monitoring. One of the basic requirements of our design is to assure portability and flexibility of our monitoring scheme across different platforms and to minimize the economic cost.

Event-driven

Our system is an event-driven monitoring system since we obtain information about the monitored objects based on the interesting events occurred during the application execution. This approach has the below advantages over the alternative time-driven moni-toring approach: (1) it provides a dynamic view of the appli-cation activity since only the information about the changes in the system status are collected, (2) unlike time-

driven, it encounters less overhead since monitoring information is supplied without an explicit periodic acquisition

3.4.2 A taxonomy of cloud monitoring

Monitoring is a term currently used in several fields of study to represent various processes. In the context of computing, there are some definitions for this term relating it to specialised areas such as Grid, Service Oriented-Architecture (SOA) and Distributed Systems [3,4,5]. However, these definitions are not complete and do not reflect the full characteristics of a modern monitoring system in our opinion. Therefore, we propose a precise and concise definition of the term monitoring so that we can contextualise what we believe to be the important capabilities associated with the monitoring process.

These capabilities can then be used as the basis to assess the monitoring tools. Recent researches define[35] 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*’’.

This kind of definition views monitoring from a capabilities perspective. especially, terms like lowest cost, systems statuses and timely, indicate how the monitoring process should be exploited operationally in the service of managing complex systems like Clouds. The key characteristics of Cloud, such as, agility, low cost, device and location independence, multi-tenancy, high reliability, high scalability, security and sustainability [35], also identify some capabilities a monitoring tool should possess to adapt to Cloud environments. There are many Cloud operational areas, such as SLA and configuration management and security, within which monitoring plays an important role in servicing Cloud provider and Cloud consumer objectives.

3.4.3 Desirable capabilities of cloud monitoring

This section presents some important capabilities of an efficient Cloud monitoring tool. Interpreting our definition of monitoring in the context of Cloud, and from the key Cloud characteristics, we identify the following list of prevalent capabilities:

Scalability: Cloud deployment can be of very large scale, consisting of thousands of nodes. In order to manage these re-sources, a monitoring tool needs to be scalable to

deliver the monitored information in a timely and flexible manner. The importance of this capability has been discussed in the literature [40,41]. Developers are currently striving to achieve high scalability in terms of resource and application management in Clouds.

Portability: Cloud environment incorporates heterogeneous platforms and services. Therefore, the portability of monitoring tools, i.e., the ability to move the tool from one platform to another, is indispensable to facilitate efficient management of such environments and to achieve wide penetration across multiple Clouds [35,41,40].

Non-intrusiveness: In Clouds, there are large numbers of resources to be monitored, hence the computational power consumed by the monitoring tools to monitor all of these resources might have a big impact on the performance of the overall sys-tem. To cater for such environments, a monitoring tool should consume as little resource capacity as possible on the monitored systems so as not to hamper the overall performance of the monitored systems [35,42].

Robustness: Clouds represent a frequent and dynamically changing environment. It is important that the monitoring tool detects changes in circumstances, such as the addition or removal of tenants and resources [35, 42]. A monitoring tool needs the ability to adapt to a new situation by continuing its operation in the changed environment, which helps to mitigate faults and to provide accurate monitored information.

Multi-tenancy: Clouds may offer a multi-tenant environment where multiple tenants share the same physical resources and application instances. A number of works in the literature have discussed the necessity of this functional requirement, especially in guaranteeing service level agreements and virtual machine monitoring [35, 41]. To support multi-tenancy provisioning, the Cloud monitoring tool should maintain con-currency, i.e., multiple customers being able to get common monitored information and isolation, i.e., tenants only being able to access the information that is addressed to them. Efficient monitoring tools should embody this capability.

Interoperability: Currently, Cloud environments may include dozens of independent, heterogeneous data centres operating mostly as stand-alone resources. Many business analysts have predicted the need for interoperable federated Clouds. Interoperability is a prerequisite for Cloud bursting and for the creation of federated offerings from multiple

providers. A modern monitoring tool should be capable of sharing monitoring information between heterogeneous Cloud components for managing collaborative operations.

Customizability: There are presently numerous Cloud ser-vice offerings and many providers are seeking ways to de-liver unique services to their customers by allowing them high customisation flexibility. Considering the large number of customers, providers must be able to manage the service customisation of each customer, for example, by granting customers the ability to choose the metrics to be monitored for their service. Thus, to realise this goal, efficient monitoring tools should possess this capacity.[35]

Extensibility: With the rapid growth of Cloud computing, there are continuous changes and extensions to technlogies especially in the area of management. Since monitoring techniques are fundamental to Cloud management, the monitoring tools need to be extensible and be able to adapt to new environments, such as being able to incorporate new monitoring metrics [40,41, 35, 42].

Shared resource monitoring: Cloud technology uses virtualization of physical machine resources to achieve usage isolation in the form of virtual machines. The virtual machines share the underlying resources while multiple applications share the re-sources of virtual machine. Thus, to avoid resource contention among the virtual machines or manage resources shared by applications in a virtual machine, efficient monitoring is needed. A Cloud monitoring tool needs the capability of supervising shared resources to manage such an environment [41].

Usability: Usability is one of the critical issues facing the adoption of Cloud computing. Fitness for purpose is an important factor when evaluating usability since the intended goal of a monitoring tool determines the usability judgement. As a consequence, any monitoring tool that is designed to support Cloud management needs to be easily useable [42]. To be highly use-able a monitoring tool should facilitate deployment, maintenance and human interaction.

Affordability: One of the reasons behind the popularity of Cloud adaptation is the reduction of cost. Cost effectiveness of a monitoring tool (e.g., being open source) impacts on its wide spread acceptance [35]. We rate affordability by considering both the cost of monitoring agent and the back end server component.

Achievability: The availability of historical data can be useful for analysing and identifying the root cause of a problem in the long term [35]. In order to serve this purpose, a monitoring tool should possess a means of storing historical data.

3.4.4 Cloud monitoring operational areas

Cloud stakeholders, such as providers and consumers, have varying motivations for gaining insight into Cloud operations. In this section, we present some Cloud operational areas that can be supported by monitoring. Later on, we present a taxonomy of the corresponding capabilities that a monitoring tool needs to possess in order to support these Cloud operational areas. Cloud computing offers a new style of computing that allows consumers to pay only for the services used and frees them from the management overhead of the underlying infrastructure. This enables low initial set-up cost for business owners. In spite of the pricing advantage, consumers are still sceptical about Cloud offerings [43]. For assurance, they may require insight into areas such as (i) Cloud usage information to confirm the correctness of their bills, (ii) SLA enforcement mechanisms ensuring their QoS objectives and (iii) security and privacy policies guiding the storage and transfer of their data. Surveys show that the possible lack of security and loss of control over data in Clouds are among the major concerns of Cloud consumers. Monitoring plays a role in detecting security breaches and hence can provide assurance of security maintenance.[43] Whilst Cloud consumers are free from the worry of maintenance overhead, providers on the other hand have the responsibility of maintaining and managing the underlying infrastructure. Monitoring is an essential part of Cloud management and serves various objectives of Cloud providers, such as (i) provisioning re-sources/services, (ii) optimal capacity planning, (iii) assuring SLAs,(iv) configuration management, (v) billing and (vi) security/privacy assurance. The above discussion highlights the Cloud operational areas that are facilitated by monitoring.

In the following section, we describe the Cloud operational areas that can benefit from monitoring. In the process, we reflect on the desirable capabilities of a monitoring tool that would make it fit for the purpose in question.

Accounting and billing: The notion of providing computing as a utility service relies heavily on the ability to record and account for the Cloud usage information on which billing schemes are based. Accurate accounting and billing relies on the ability to

capture the consumption and allocation information of virtual resources as well as that of applications (e.g. compute hour used, bandwidth used) [44]. This is a capability that monitoring can provide. Furthermore, the provision of a transparent billing system that is able to record data in a verifiable and trustworthy manner, to ensure protection against forgery and false modifications, requires robust and secure cloud monitoring capabilities [35].

SLA management: A Service Level Agreement (SLA) represents a contract signed between a service provider and a customer specifying the terms of a service offering including quality of service (QoS), pricing and penalties in case of violating the agreed terms [35]. SLA management is an area of great importance for Cloud providers since the assurance of SLA enforcement is inevitable for customer satisfaction and hence is a driving force for the continuity and growth of a Cloud business. The providers are expected to meet the QoS requirements as well as the Key Performance Indicators (KPI) for services in order to enforce their agreed SLA terms. Monitoring is essential to achieve these goals. The cloud monitoring capabilities required to support operations in this area include the ability to measure QoS parameters, storing and analysing data, resource consumption measuring and SLA parameter assessment. These capabilities are expected from a monitoring tool for the purpose of SLA management.

Service/resource provisioning: Service/resource provisioning involves the allocation of resources optimally in order to match the workload [44]. It is an essential requirement for providing Cloud elasticity. Provisioning can be implemented in two ways: (1) static provisioning where VMs are created with a specified size and then consolidated onto a set of physical servers. The VM capacity does not change; and (2) dynamic provisioning: VM capacity is dynamically adjusted to match workload fluctuations. The ability to measure the overall resource consumption of a system, along with the ability to measure per service resource consumption (which identifies the amount of resources each individual service needs), is essential for efficient provisioning. Furthermore the ability to assess risk and QoS is needed for effective provisioning decisions, such as whether to allocate/release resources to ensure that the quality is not compromised or resources are not wasted [35, 44].

Capacity planning: Capacity planning is an important domain in Cloud computing, especially for the provider. It ensures adequate resource availability to meet the capacity demand necessary for securing a level of service quality and for serving various Cloud operational management activities, e.g., disaster recovery and maintaining backups [44]. The ability to measure capacity usage enables operations such as predicting the need for more resources or determining resource wastage. Furthermore, the ability to detect Cloud node availability is necessary to maintain a required level of resource limits [44]. Cloud monitoring capabilities such as component status identification play an important role in facilitating these goals.

Configuration management: Configuration is a set of parameters and values that determine the behaviour of devices and software [35]. While a Cloud provider may operate a multi-tenant environment it needs to manage customer-specific configuration. The initial configurations may contain the minimal set of resources required for a certain service. Resources may be added or released depending on varying load resulting into reconfiguration at run.

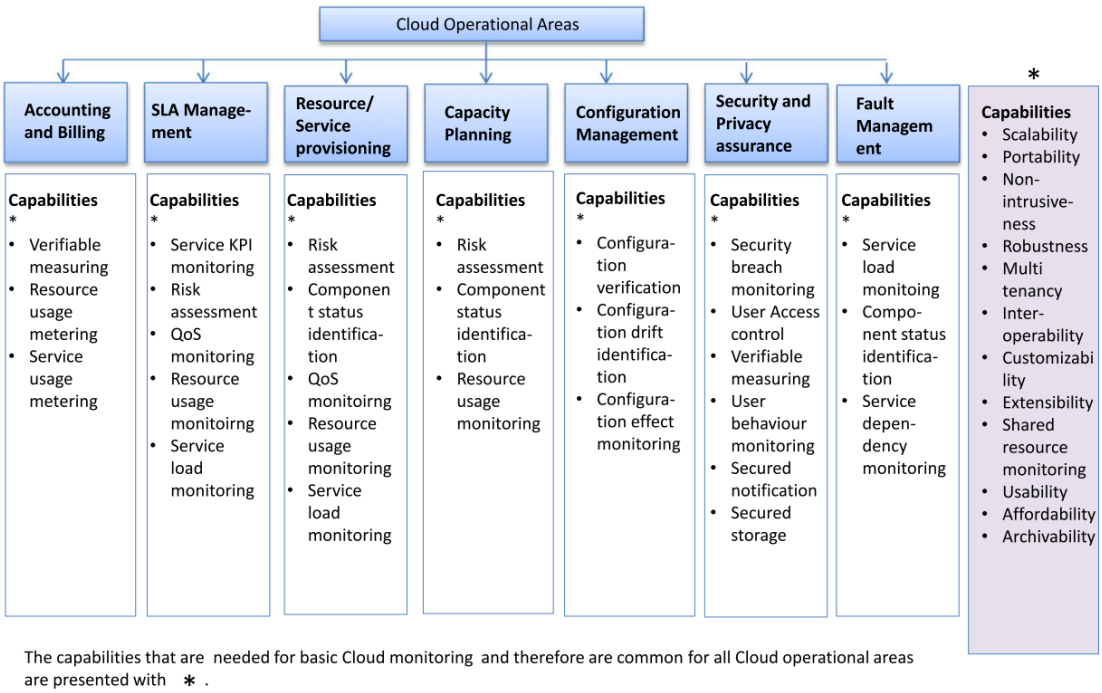


Figure 4. Taxonomy of cloud monitoring capabilities based on Cloud objectives.
Source: [35]

3.4.5 Monitoring tools

This section presents analysis of the monitoring tools described in[35]. Analysis is based on the monitoring capabilities taxonomy. In line with their approach, authors of this article partitioned the set of tools into two groups:
those that are general purpose and those that are Cloud specific. Table 1 presents the analysis of the general purpose monitoring tools and Table 2 presents the analysis for Cloud specific monitoring tools. The 2nd columns of the tables show the weighted average percentage of the implemented capabilities by the tools. In the calculations, they assigned ‘‘1’’ if a tool has a particular capability and ‘‘0’’ if it does not. There is the assignment of ‘‘0.5’’ if the tool partly implements such a capability. The sum of these values is used to calculate the assessment percentage. According to Table 1 and Table 2, a ‘‘1’’ is equivalent to ‘‘Yes’’, a ‘‘0’’ implies ‘‘no’’ and ‘‘0.5’’ represents ‘‘limited’’. The capabilities that have scored ‘‘0’’ for all the tools are excluded from the calculation of weighted average percentage of capabilities covered by each tool which are presented in the last rows of the tables. For tools with multiple versions—for example, Nagios, Opsview, Hyperic, CloudKick, Nimsoft and Monitis, the capabilities are identified based on the superset of features of all versions. The determinations of the capability implementation by the tools are based on literature reviews and evaluations of the tools. Some of the capabilities such as non-intrusiveness and usability are subjective, hence they are evaluated based on the weight of opinion found in the reviewed literature. In the below table we can see in more details.

Table 1: General purpose monitoring tool analysis.

Capability/features	Percentage implemented	Nagios	Collectd	Opsview	Cacti	Zabbix	Open NMS	Ganglia	Hyperic	IBM Tivoli	Kiwi Monitor	DAMS	RDT
Scalability	46%	no	no	limited	no	yes	no	yes	yes	yes	no	yes	no
Portability	79%	limited	limited	yes	limited	yes	yes	yes	yes	yes	no	yes	yes
Non-intrusiveness[a]	50%	limited	limited	yes	no	yes	yes	limited	limited	yes	no	no	no
Robustness	33%	no	no	no	no	no	no	yes	yes	yes	no	no	yes
Multi-tenancy	33%	yes	no	yes	no	no	no	no	yes	yes	no	no	no
Interoperability	25%	no	yes	no	no	yes	yes	no	no	no	no	no	no
Customizability	100%	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Extensibility	100%	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Shared resource monitoring	42%	yes	yes	yes	no	yes	no	no	yes	no	no	no	no
Usability[a]	92%	no	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Affordability	79%	limited	yes	limited	yes	yes	yes	yes	limited	no	yes	yes	yes
Archivability	67%	yes	no	yes	yes	yes	yes	yes	yes	yes	no	no	no
Verifiable measuring	0%	no	no	no	no	no	no	no	no	no	no	no	no
Resource usage metering	75%	yes	yes	yes	yes	yes	yes	yes	yes	yes	no	no	no
Service usage metering	50%	yes	yes	yes	no	yes	no	no	yes	no	yes	no	no
Service KPI monitoring	0%	no	no	no	no	no	no	no	no	no	no	no	no
QoS monitoring	50%	yes	no	yes	no	yes	no	no	yes	yes	yes	no	no
Risk assessment	58%	yes	no	yes	yes	yes	yes	no	yes	yes	no	no	no
Component status identification	100%	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Service load monitoring	50%	yes	yes	yes	no	yes	no	no	yes	no	yes	no	no
Configuration verification	75%	yes	yes	yes	yes	yes	yes	yes	yes	yes	no	no	no
Configuration drift identification	75%	yes	yes	yes	yes	yes	yes	yes	yes	yes	no	no	no
Configuration effect monitoring	75%	yes	yes	yes	yes	yes	yes	yes	yes	yes	no	no	no
Security breach monitoring	33%	yes	no	no	no	yes	no	no	yes	yes	no	no	no
User access control	50%	no	no	yes	yes	yes	yes	no	yes	yes	no	no	no
User activity	17%	no	no	no	no	no	no	no	no	yes	yes	no	no
Secured notification	17%	no	no	no	no	no	yes	no	no	yes	no	no	no
Secured storage	17%	no	no	no	no	no	no	no	yes	yes	no	no	no
Service dependency	21%	no	no	no	no	no	no	no	yes	no	no	yes	yes
Percentage covered by tools		61%	52%	70%	46%	78%	59%	50%	**85%**	78%	33%	30%	30%

Source: [35]

There are hundreds of monitoring tools which can help us to monitor our infrastructure, middleware and all other levels of application chains. General purpose monitoring tools are adding more features to be able to monitor even on cloud environment. But, when we think about cloud based monitoring tools by it is capabilities only few of them are capable to monitor on all level of cloud service models and all level of cloud deployment models.

In the below table we will review another table from with ten cloud based monitoring tools which focuses on cloud environments.

Table 2: Cloud based monitoring tool analysis

Capability/features	Percentage implemented	CloudKick	Nimsoft	Monitis	Amazon Cloud Watch	Azure Watch	PCMONS	Boundary app. monitor	mOSAIC	CASViD
Scalability	78%	yes	yes	yes	yes	yes	no	yes	no	yes
Portability	56%	limited	yes	yes	no	no	no	yes	yes	yes
Non-intrusiveness[a]	94%	yes	yes	yes	yes	yes	limited	yes	yes	yes
Robustness	67%	yes	yes	yes	yes	yes	no	yes	no	no
Multi-tenancy	44%	yes	yes	no	yes	yes	no	no	no	no
Interoperability	33%	no	no	no	no	no	yes	no	yes	yes
Customizability	100%	yes	yes	yes	yes	yes	yes	yes	yes	yes
Extensibility	100%	yes	yes	yes	yes	yes	yes	yes	yes	yes
Shared Resource monitoring	78%	yes	yes	yes	yes	yes	no	yes	no	yes
Usability[a]	94%	yes	yes	yes	yes	yes	limited	yes	yes	yes
Affordability	50%	limited	limited	limited	no	no	yes	no	yes	yes
Archivability	78%	yes	yes	yes	yes	yes	yes	yes	no	no
Verifiable measuring	0%	no	no	no	no	no	no	no	no	no
Resource usage metering	100%	yes	yes	yes	yes	yes	yes	yes	yes	yes
Service usage metering	100%	yes	yes	yes	yes	yes	yes	yes	yes	yes
Service KPI monitoring	0%	no	no	no	no	no	no	no	no	no
QoS	89%	yes	yes	yes	yes	yes	no	yes	yes	yes
Risk assessment	92%	yes	yes	limited	yes	yes	yes	yes	yes	yes
Component status identification	100%	yes	yes	yes	yes	yes	yes	yes	yes	yes
Service load monitoring	100%	yes	yes	yes	yes	yes	yes	yes	yes	yes
Configuration verification	78%	yes	yes	yes	yes	yes	yes	no	no	yes
Configuration drift identification	78%	yes	yes	yes	yes	yes	yes	no	no	yes
Configuration effect monitoring	100%	yes	yes	yes	yes	yes	yes	yes	yes	yes
Security breaches monitoring	56%	yes	yes	no	yes	yes	no	yes	no	no
User access control	67%	yes	yes	yes	yes	yes	no	yes	no	no
User activity monitoring	0%	no	no	no	no	no	no	no	no	no
Secured notification	33%	no	yes	no	yes	yes	no	no	no	no
Secured storage	33%	no	yes	no	yes	yes	no	no	no	no
Service dependency	11%	no	no	no	no	no	no	yes	no	no
Percentage covered by tools		78%	**87%**	70%	85%	85%	52%	73%	54%	69%

Source:[35]

Note that some of these capabilities are somewhat subjective. Author of this table maintained in mind the evaluation presented here is based on the weight of opinion as reflected in the reviewed literature. But, authority of this article analysed all capabilities as the same priorities. They did not categorised by necessity of individual capabilities. In the practical part we will implement this table in our analysis in a specific environment especially by collecting priorities from company's engineers.

3.5 Chosen cloud service providers

In this chapter we will focus on basic overview, architecture and pricing strategy of cloud computing providers. NN already decided to migrate their data into cloud on AWS and Microsoft Azure averments. Companies first goals is to use IaaS and later PaaS solution on both cloud environment.

In the below Figure 5 Garner confirms top leading cloud service providers via magic quadrant figure. As you can see in the public cloud storage services in worldwide AWS, and Microsoft Azure are two of the top leaders. Due to this reason we have chosen only these two top cloud service providers.

Figure 5: Magic Quadrant for Public Cloud Storage Services, Worldwide
Source: Gartner [29]

3.5.1 Amazon Web Services (AWS)

One of the biggest top cloud provider in the world is AWS. It is very popular by using PaaS and IaaS cloud service models. AWS started offering its technology

infrastructure platforms in 2006. Since then, hundreds of thousands of customers across 190 countries using AWS in every imaginable way.

"We have developed considerable experience operating at scale. We've also innovated and delivered at a very rapid pace (delivering 159 significant features and services in 2012 and 280 in 2013). Expect this focus on rapidly delivering what customers want to continue"[22].

Amazon Web Services (AWS) provides trusted, cloud-based solutions to help us meet our business needs. Running our solutions in the AWS Cloud can help us get our applications up and running faster while providing the same level of security that international organizations and standards like ISO or Pfizer, Intuit, and the US Navy rely on.

AWS also provides resources around the world, so we can deploy our solutions where our customers are[22]. The AWS Cloud makes a broad set of services, partners, and support options easily available to help make sure that we can focus on what will make our solution a success.

AWS Architecture

Below figure is general overview of AWS on 3 tier level.

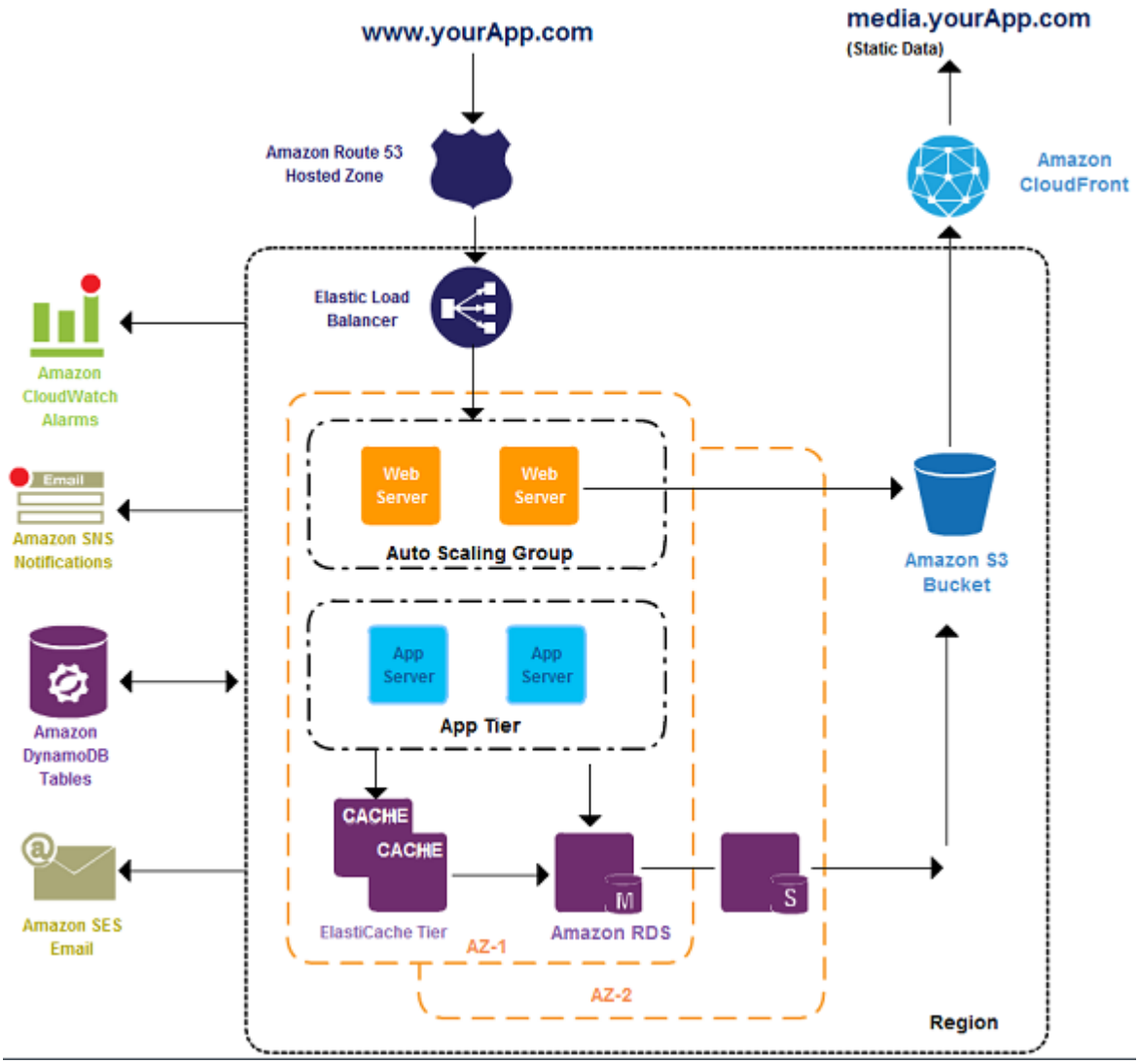


Figure 6 AWS Auto-scalable Web Application Architecture
Source: [23]

Pricing EC2

In the below table we can see the general use, compute optimized, and memory optimized virtual machine types offered by Amazon Web Services and Microsoft. Prices shown are for standard Linux machines running in datacentres on the Central United States. The hourly price represents on demand instance pricing (no contract) while the one year price includes the hourly discount for a 1 year 100% use contract. There are significant savings for contract pricing over on demand instances if we can accurately estimate your long term needs.

Machine Name	vCPU Cores	Ram (GB)	Hourly Price	One Year
m3.medium	1	3.75	$0.07	$379.68
m3.large	2	7.5	$0.14	$546.12
m3.xlarge	4	15	$0.28	$870.24
m3.2xlarge	8	30	$0.56	$1,500.96

Table 3: Amazon EC2 prices for standard compute machines
Source: [22]

3.5.2 Microsoft Azure

With Microsoft Azure, we can spin up new Windows Server and Linux virtual machines in a very quick time and adjust our usage as your needs change. With their pay-as-you-go approach, we only pay for what we use and there are never any penalties for changing our virtual machine configurations.

Microsoft Azure enterprise architecture

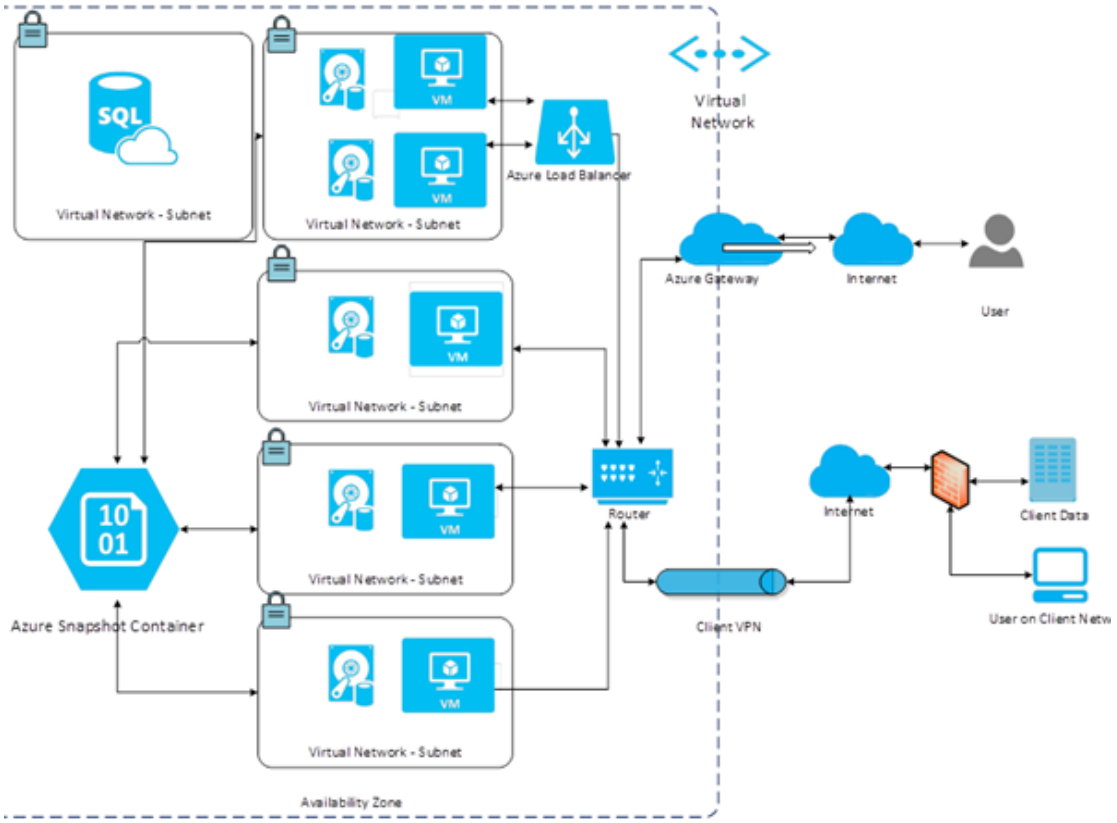


Figure 7: Microsoft Azure enterprise architecture
Source:[45]

Pricing (standard)
In the table below standard compute machines prices are given.

Machine Name	vCPU Cores	Ram (GB)	Hourly Price	One Year
Extra Small (A0)	0.5	0.75	$0.02	$135.78
Small (A1)	1	1.75	$0.06	$611.01
Medium (A2)	2	3.5	$0.12	$1,222.02
Large (A3)	4	7	$0.24	$2,444.04
Extra Large (A4)	8	14	$0.48	$4,888.08

Table 4: Azure prices for standard compute machines
Source:[author]

4 Practical part

This part presents the analysis a particular case in selected company that is supposed to move its applications, infrastructure and all datacentres to the cloud;
As a one of the partial goal, author identifies possible solutions for monitoring of applications in cloud and makes a decision.

Firstly, we will make brief case study of the company by defining it's cloud transformation road maps, plans, it's main goals during the migration period, scope of migration, corporate visions and benefits of cloud migration.

Secondly, we compare major cloud service providers which are company is going to move on (AWS, Microsoft Azure) by making collection of their features, core services, database and other additional capabilities.

Thirdly, we compare general comparison of chosen tools and we identify the cloud monitoring capabilities that are essential for facilitating complex management activities in Cloud environments. From this we construct and present a taxonomy of cloud monitoring capabilities in the context of specific Cloud operational areas. Meanwhile we will make quick look into AWS and Micrasoft Azure

Fourthly, author makes a component design of chosen(winner – best tool from the list of analysis) tool as a proposal for a company and makes his conclusion.

4.1 Brief case study (NN Group)

4.1.1 Executive overview of the company

NN Group is an insurance and asset management company operates in more than 18 countries, with a very strong presence in a number of European countries and Japan. Companies roots lie in the Netherlands, with a rich history that stretches back 170 years[25].

NN Group includes Nationale-Nederlanden, NN (previously known as ING Insurance) and NN Investment Partners (previously known as ING Investment Management). NN is formerly part of ING Group, NN Group listed as an independent

stand-alone company on Euronext Amsterdam on 2 July 2014. NN Investment Partners offers its products and services globally through offices in several countries across Europe, the United States, the Middle East and Asia, with the Netherlands as its main investment hub.

Overview	Details
Type	**Public**
Industry	**Insurance**
Founded	**1963**
Headquarters	**The Hague, Netherlands**
Products	**Life Insurance** **Property insurance** **Casualty insurance** **Mortgages**
Revenue	**€5.6 billion**
Net income	**€1.4 billion**
Employees	**11500**
Operates in	**18+ countries**
Website	**www.nn.nl**

Table 5: NN Group overview [author]

4.1.2 The Roadmap

NN targets to achieve 'Fast IT' by transforming along three pillars: people, technology and processes. Pillars can be found in the left side of the pillar.

In the below [Figure 8. Brief migration roadmap], we can see company's roadmap which was started transformation from 2014. Author, focus mostly targeted to migration and preparation pillars. Because, pillar People only about human resource, trainings and some other administration activates.

Until January 2016, company succeeded to make trial period on AWS and Microsoft Azure and successfully implemented network connectivity between NN datacentre and cloud service providers. As we can see from figure Office365 cloud migration already started in January 2016. From February 2016 company started to migrate their infrastructure as PaaS on both Microsoft and Azure environments.

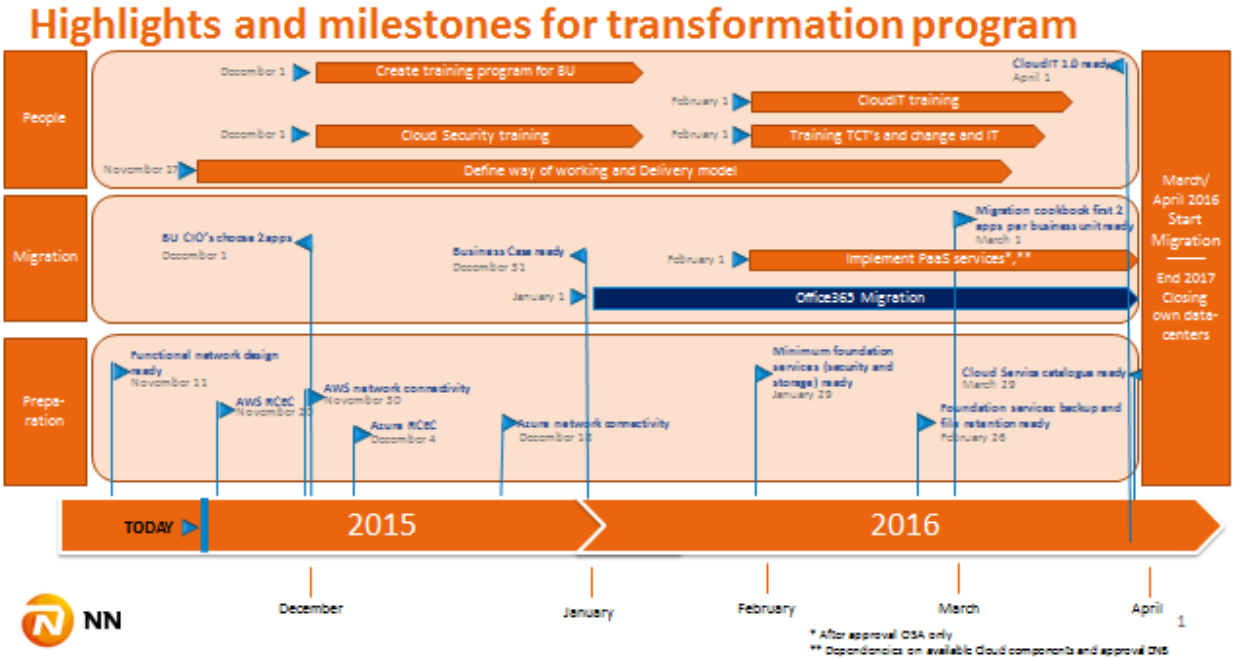


Figure 8. Brief migration roadmap
Source: Company's intranet page

4.1.3 Main goals during cloud migration period

This migration program includes **three main goals**.

First goal: is to have all company's applications (1100 applications) in scope in a migration program

Second goal: enhanced with a minimum level of automation by the 1st of January 2018

Third goal: NN desired to introduce a new way of working. The end-state within the other business units and countries is still to be determined due to the primary focus on CIO.

4.1.4 Scope of migration

Company's scope is both NN Group and NN Insurance International, where they are focusing on migrating and automating their current X86-landscape to Amazon Web Services and Microsoft Azure. Mainframe, AS400 and P-series (ZIP) and Japan are out of scope of the cloud transformation program, as well as their current SaaS landscape (as this is already running in the cloud). The size of the program embraces **1100 applications**, divided over 14 countries and more than 20 business and functional units.

Below, table briefly explains in scope corporate units, belonging countries and business units. In scope of NN Insurance International and NN Group following countries will be migrated.

In Scope	Countries	Business Unit
NN Insurance International NN Group NN HQ	Czech Republic & Slovakia	SD&C
	Greece	Bank
	Turkey	Non-life
	Luxembourg	Pensions
	Spain	Services
	Bulgaria	N/A
	Hungary	CFO
	Belgium	CRO
	Romania	HR
	Poland	N/A

Table 6 In scope (Countries & BUs)
Source: [author]

Although, NN Investment Partners is currently out of scope, however they are free to consume all material provided by the program.

4.1.5 Corporate vision in cloud migration

IT co-driving business growth and innovation through 'Fast IT' is one of the main corporate vision in cloud transformation. Company aims to provide their customers with an excellent experience that inspires them to recommend to their family and friends over any other financial companies. They want to offer value for money as an integral part of this.

NN customers' expectations have changed significantly with the rise of digital disruptors, companies which are able to iterate quickly to match customer needs. As David Knibbe stated about the changing market: "*Our customers demand transparency, clear information and an excellent digital service.*"

Head of Cloud team in NN, Vincent Snijder says *"NN using IaaS especially in the first migration base to deliver as the same service as we Now have with in our data centre. what we have in Azure and AWS. NN Operation System are stored in cloud. In the cloud*

our Oracle build, our Windows, our Linux build." [50]. To meet their customer's requirements and to enable NN to adapt to the changing market NN need to act with fast and encourage innovations. Company's products need to be digital, relevant and personal, and their organization needs to become more agile. We believe IT is key in this transformation.

*"We envision IT co-driving business growth and innovation through 'Fast IT', enabling an excellent digital, personal and relevant experience for our customers." [Source: F*rom company's internal source]

NN believes the Cloud transformation program supports IT to move beyond a supporting role in the business strategy to co-driving business growth, where in IT is vital to adding value to the quality of our customers experience. Company will implement an innovative foundation on which they can transform IT into 'Fast IT': delivering agility, innovation and cost efficiency, enabling us to give the business more speed, flexibility and dynamic costs.

Transformation of the current IT value chain is inevitable. With the three enablers – an agile way of working, automation of the IT value chain, Cloud usage – they will create a stable and future proof organization that is able to deliver service quicker, improving business agility and enabling innovation against optimal costs.

4.1.6 Benefits

- With Cloud Technology NN will transform the IT infrastructure, increasing their speed and flexibility and allowing them to pay for what they use only.
- With Automation of the IT Value chain company will reap the full benefits of Cloud by optimizing cloud usage, eliminating repetitive manual processes and deployments, ensuring operational activities can run more efficient and effectively.
- Cloud is dynamically changing new way of working, they will create small, self-managing and autonomous teams to increase performance of our IT organization.
- By supporting the IT transformation implementing this three-stage rocket-model they can deliver 'Fast IT' for NN and for their customers: delivering innovative, digital, relevant and personal products, helping customers to secure their financial future with NN.

- NN aim to provide its customers an excellent experience that inspires them to recommend them to their family and friends over any other financial company but also the market is changing rapidly and NN has to be able to adapt to the changing market. Therefore they need to optimize it's IT product.

In addition, NN need to optimize IT products that need to create a fast IT organization that accelerates innovation in the business. This results in a faster time to market and will decreases costs of failure in innovation, as maximum investment loss is limited, offset and paid off by big successes.

IT should enable acceleration of business innovation by optimizing the speed, enabling a competitive cost position and the flexibility of its own IT organization, through continuously reviewing its services and its operating model.

This creates a fast IT organization and ensure business continuity, one in which they are able to deliver speed, more flexibility against optimal costs. These business requirements are the foundation for our IT vision and result in the following enablers:

- Agile way of working
- Automation of the IT value chain
- Usage of Cloud technology

These three pillars of the transformation focus on people, processes and technology. The cloud migration is the accelerant for this transformation, however only by changing the three enablers we create the necessary impact and change the speed and flexibility of IT that enable the business to innovate.

The enabler Agile way of working focuses on providing the people with the ability to act which significantly improves the delivery speed of changes and innovation for the NN organization. The usage of cloud infrastructure technology enables the IT structure within the value chain to break free of their current infrastructure technology based limitations and facilitate cost dynamics, flexibility and increase our speed. Combining the cloud with automating the IT value chain allows to optimize cloud usage which will give them more business agility and cost optimization, and supports the way of working by allowing the teams to focus on adding value whilst ensuring operational activities run efficient and effectively.

4.1.7 Microsoft Azure initiation in NN

During analysis of NN, author found out company reached some major milestones to migrate there infrastructure into Azure as IaaS/PaaS solutions.

As of March 2016, the Azure team realized a couple of huge milestones in there road to the foundation of IaaS/PaaS services. They created a network from the NN data centres for the cloud environment in Azure.

Basically this means that company configured the most important parts of the network.

- The Express Route (the route from NN to the Azure Cloud has been configured)
- Ports have been opened for the first supporting services (security monitoring with Arcsight and technical state compliance monitoring with Nessus)
- The virtual networks for these first services are created.
- The Checkpoint virtual firewall has been installed and configured for connectivity between internal and external Azure services (private/public endpoint filtering). The virtual firewall was required because IaaS VMs need to access Microsoft PaaS services for basic functionality such as disk encryption and IAM[2]

Several components are implemented:

- Azure Foundation services
- Azure network
- Virtual firewall
- Express Route

In the following Figure, you can find as an overview.

[2] IAM- Identity Access Management

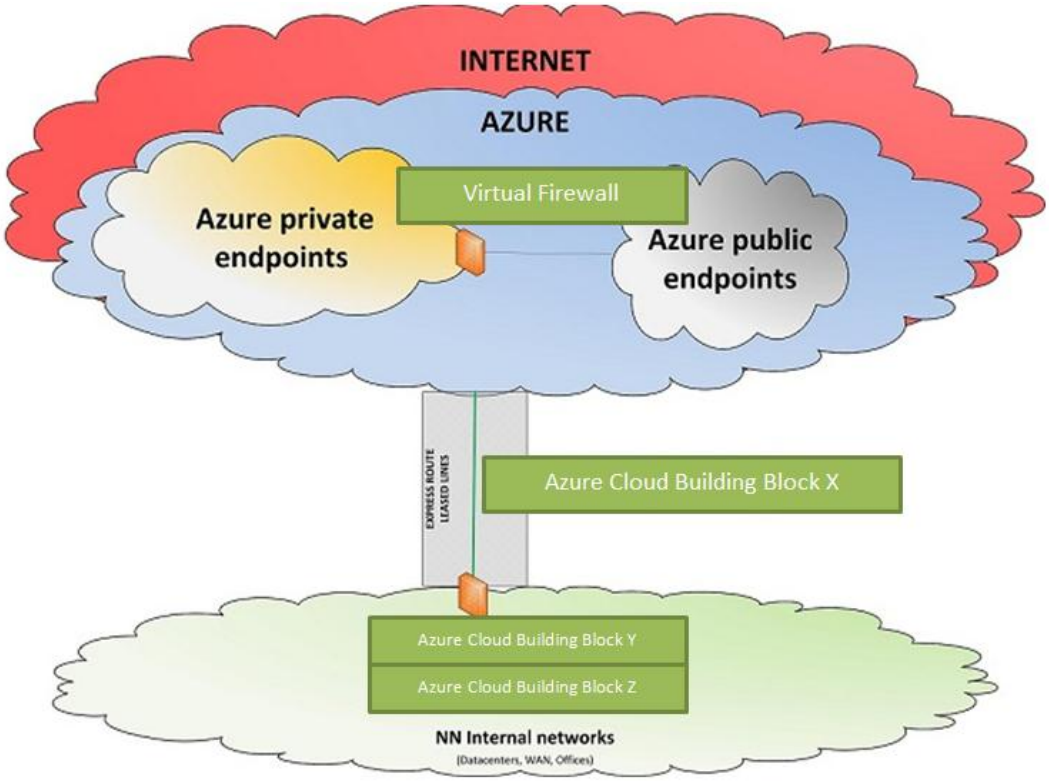


Figure 9: Draft overview of NN Azure as IaaS and PaaS solutions
Source: [author]

Next steps which company is going to implement is configuring Nessus and Arcsight on Red Hat Linux systems to be effectively implement security monitoring operations.

NN Azure Cloud engineers are mentioning as *"We put a lot of effort put in scripting to make implementation efficient and reliable, Not only for the network but also for other deployments like authorizations, resource groups, storage accounts, etc."*

As a stepping stones for the Cloud with CSA[3] and LSA[4] accounts have been realized and access with MFA[5] is in place. CSA accounts are used to manage Azure cloud native resources, LSA accounts are used to manage IaaS workloads within the environment With these steps company can continue to realize there foundation services in the DTAP environment. By April 2016 is will be a date for the first application migration to the Cloud.

[3] CSA-Cloud Security Alliance- one of the strong cloud security certification
[4] LSA- Local Security Account
[5] MFA- Multi Factor Authentication

At the end of this thesis author has few proposal for the company. They will be formulated in the conclusion part of the thesis.

4.1.8 AWS initiation in NN

As clearly explained in the [Figure 8. Brief roadmap" Amazon Web Services (AWS) is chosen as second cloud vendor in NN. Last year company implemented the Fiber connection (AWS Direct Connect) between the their datacentres and the AWS Cloud to a Virtual Private Cloud (VPC) in the formal AWS production account (Shared Services account). With this achievement, NN-Group is ready to utilize the fiber connections towards the Amazon Cloud. Use of this leased lines, brings in the advantages of having better security, latency stability and larger bandwidth.

In the earlier transformation period the line, called Direct Connect Link, was already operational for a couple of months, but it was not configured into an AWS account where it has to be configured into conform design and for later production. Now a days company succeeded with it. The implementation involves the creation of a Cloud Formation template (Infrastructure as code) to create the VPC, subnets and Security Groups in AWS, configuration of VLANs, BGP peering and NN firewalls in the datacentres. To test the connectivity, an EC2 instance was provisioned and used to successfully ping the network infrastructure in IT hosting.

Technically there are two physical lines of 5 Gb operational based, both connected to the AWS Region EU (Ireland). Lines can be upgraded in the future to 10 Gb.

4.2 Comparisons of CSPs (AWS & Azure)

In this chapter, author analyses top CSP's features and their technical capabilities. Author, chose only CSP's which are already chosen by selected company (NN Group). Due to the fact that NN already decided to migrate/transform there IT hosting and all datacentre infrastructure into major Cloud Service Providers AWS and Microsoft Azure we are not interested to analyse other cloud platforms.

4.2.1 Basic features

Features	AWS	Microsoft Azure
VM Sizes	Max CPUs: 40 Max Memory GB: 244	Max CPUs: 32 Max Memory GB: 448
SLA Terms	Credit for 1+ minutes downtime Max monthly credit: 30% Uptime SLA: 99.95%	Credit for 1+ minutes downtime Max monthly credit: 25% Uptime SLA: 99.95%
Certifications	CSA[6] FedRAMP[7] FISMA[8] HIPAA ISO 27001 ISO 27017 ISO 27018 PCI DSS1 SSAE16 SOC1 (Type II) SSAE16 SOC2 (Type II) SSAE16 SOC3 (Type II)	CSA FedRAMP FISMA HIPAA ISO 27001 ISO 27018 PCI DSS1 SSAE16 SOC1 (Type II) SSAE16 SOC2 (Type II) SSAE16 SOC3 (Type II)
Operating Systems	CentOS CloudLinux CoreOS Debian FreeBSD Gentoo Linux openSUSE Oracle Linux	CentOS CoreOS Debian FreeBSD openSUSE Oracle Linux RHEL SUSE

[6] CSA-Cloud Security Alliance- one of the strong international cloud security certification alliance

[7] FedRAMP- The Federal Risk and Authorization Management Program.

[8] FISMA- "*The Federal Information Security Management Act of 2002 ("FISMA", 44 U.S.C. § 3541, et seq.) is a United States federal law enacted in 2002 as Title III of the E-Government Act of 2002 (Pub.L. 107–347, 116 Stat. 2899). The act recognized the importance of information security to the economic and national security interests of the United States"*[31].

	RHEL SUSE Ubuntu Windows	Ubuntu Windows
Regions	ANZ (Australia/New Zealand) Asia Europe North America South America	ANZ (Australia/New Zealand) Asia Europe North America South America

Table 7: Basic features of chosen cloud service provides [author]

4.2.2 Core Services

Features	AWS	Microsoft Azure
Compute Services	Autoscaling Dedicated Hosts (virtual) Temporary VMs	Autoscaling
Network Services	CDN[9] Direct connect DNS Load Balancing VPN	CDN Direct connect DNS Load Balancing VPN
Storage Services	Archive Storage Block Storage File Storage Object Storage	Block Storage File Storage Object Storage

Table 8: Core Services features[author]

4.2.3 Database services features

Features	AWS	Microsoft Azure
Relational Databases	Amazon Aurora MariaDB Microsoft SQL Server MySQL Oracle	Microsoft SQL Server

[9] CDN- Content Delivery Network

	PostgreSQL	
Non RD	Hadoop NoSQL	Hadoop NoSQL
Other DBaaS	Caching Data Warehouse	Caching Data Warehouse

Table 9: Database services features[author]

4.2.4 Additional services

Features	AWS	Microsoft Azure
Application Services	Batch Processing Container as a Service Email Sending IoT[10] Machine Learning Microservices Push Notifications Queuing Search Stream Processing Transcoding/Encoding Workflow	Batch Processing Container as a Service IoT Machine Learning Microservices Push Notifications Queuing Search Stream Processing Transcoding/Encoding Workflow
Security & Identity	IAM Key Storage & Management Security Assessment	IAM Key Storage & Management Security Assessment

Table 10: Additional services features[author]

4.2.5 Computing model pricing (computing)

Table 11: Computing model pricing for computing

Options	AWS	Microsoft Azure
On-Demand Pricing	• Free Tier • Per Hour • No charge for "Stopped" • Pay for EBS volume	• Free Trial • Per-Minute • "Stopped" bills for VM, not SW • No charge for "Stopped (De-Allocated)
Discount Options	Reserved Instances • All upfront (largest discount) • Partial upfront • No upfront	• Through Resellers Enterprise agreement •Upfront monetary commitment to Azure •Consumed throughout the year by using any

[10] IoT- Internet of Things

	RI Volume Discounts • $500K-$4M = 5% • $4M-$10M = 10% • >$10M = contact AWS • Spot Instances • RI Marketplace	Azure services • Billed for overages at EA rate •MSDN (per month credit) •BizSpark

Source: [author]

4.3 Analysis of chosen tools

In the literature view[Table 2] researchers analysed top 10 cloud monitoring tools and author of this thesis selected only tools which has more than 85% full filled specific requirements of cloud computing. Our main goal is to find best of the best monitoring tool for this (NN Group) specific enterprise.

In this section, we analyse the Cloud specific monitoring tools using the previously described taxonomy. The goal of the analysis is to determine the strengths, drawbacks and challenges facing these tools. Table 14 presents the analysis. As shown in the table, all of the tools implement the customizability, extensibility, resource usage metering, service usage metering, component status identification, service load monitoring and configuration effect monitoring capabilities. This group of tools lacks on the implementation of portability, multi-tenancy, interoperability, secured notification, secured storage and service dependency capabilities. These tools are generally designed for monitoring in Clouds and many of them are commercial and proprietary, i.e., provider and platform dependent. This accounts for the low levels of interoperability and portability observed.

4.3.1 Amazon CloudWatch

Architecture overview

Amazon CloudWatch is one of the most popular commercial tools for monitoring the cloud. It is provided by Amazon to enable its consumers monitoring their resources residing on EC2. Hence, it does not support multi-cloud infrastructure monitoring. The technical approaches used in CloudWatch to collect data are implicit and not exposed to users. CloudWatch is limited in monitoring resources

across cloud layers. However, an API is provided for users to collect metrics at any cloud layer but requires the users to write additional code.

Amazon CloudWatch is basically a metrics repository. An AWS product—such as Amazon EC2—puts metrics into the repository, and we retrieve statistics based on those metrics. If you put our own custom metrics into the repository, we can retrieve statistics on these metrics as well. In the below we will see general architecture of CloudWatch.

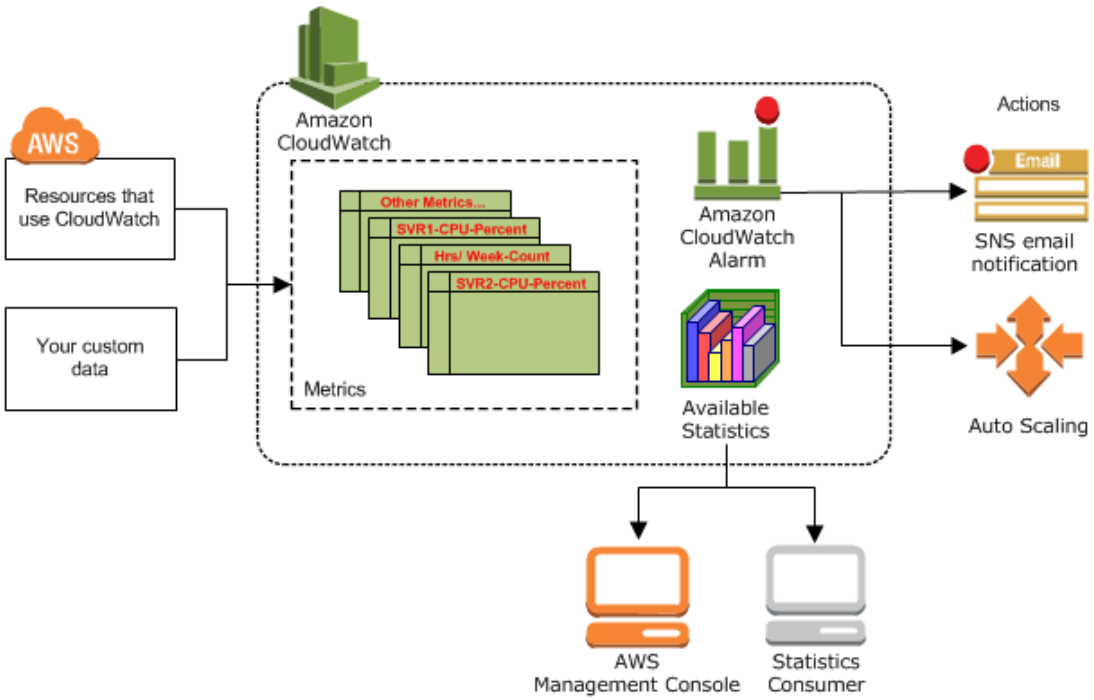


Figure 10: Amazon CloudWatch Architecture
Source: [32]

Conclusion

Amazon CloudWatch is a monitoring service for AWS cloud resources and the applications you run on AWS. One of the disadvantage of CloudWach it's AWS cloud resources only. A lot of companies now a days using hybrid cloud solutions and or at least two companied cloud vendors to be able to ready got any disaster recovery plans. Due the business continuity rules, it is good to have back app plan while using cloud services.

We can use Amazon CloudWatch to collect and track metrics too, collect and monitor log files, set alarms, and automatically react to changes in our AWS resources. Amazon CloudWatch can monitor AWS resources such as Amazon EC2 instances, Amazon DynamoDB tables, and Amazon RDS DB instances, as well as custom metrics

generated by our applications and services, and any log files our applications generate. We can use Amazon CloudWatch to gain system-wide visibility into resource utilization, application performance, and operational health. We can use these insights to react and keep your application running smoothly.

Last research shows [Table 2] that Amazon CloudWatch full fills all cloud monitoring capabilities by 85%. Author of this thesis has selected only equal or more than 85% full filled tools in the short list, and compared by company's prioritised ratings via Scoring and sequence methods.

4.3.2 Microsoft AzureWatch

Overview

AzureWatch is a cloud-based service dedicated to advanced monitoring and auto-scaling of Azure-based solutions. Every minute of every hour AzureWatch inspects performance of our Azure-based applications, notifies us of any problems, and dynamically adjust the number of compute resources dedicated to your applications according to real time demand. User-defined rules specify when to alert, scale up or scale down. Charts and reports are delivered on our browser, smartphone or RSS feed.

AzureWatch support verity of computing services, but they are only Microsoft Azure resources based, in the following point we briefly explain them.

AzureWatch can auto-scale:

- based on a schedule
- based on leading indicators (queue depths, rate of change in demand, etc.)
- based on trailing indicators (CPU utilization, requests/sec, bandwidth, disk space)
- based on historical performance
- according to user-defined upper and lower limits
- at the end or the beginning of a clock-hour to maximize costs
- or with any combination of the above

Cloud Services (Web/Worker Roles) options:

- Auto-Healing

- Alerts Visualization
- Performance counters
- Windows Event logs
- Azure Storage queues
- Azure Service Bus
- Server States

Virtual Machines

- performance counters
- Windows Event logs
- Azure Storage queues
- Azure Service Bus
- Azure Server States

Conclusion

Azure Authentication control: To gain high privileged access connection to the Azure environment is based on enforced Multi Factor Authentication (MFA). MFA is based on a soft-token, meaning UserID, Password and an authentication code which is send as SMS to your mobile phone Access control: Is one of the main controls to protect our services Security groups are the security perimeter to manage the access level towards application and its data Within Microsoft a security perimeter is created around each of Microsoft customers

Identity Management: is the way to ensure the right employee or identities are listed with the authorization they are entitled to. This must be verified on a regular bases (automated) Monitoring: Is based on cloud service provider monitoring solutions, which helps us to monitor and have alerts when incidents or changes happen and follow-up on events Logging: Is based on cloud service provider logging solutions, which helps us to analyze and monitor the events, which can be traced down to the individual or activity that happened and follow-up on events Audit assessment: Must happen on a regular bases to validate if the Cloud Service Provider is still in line with the contract, compliance with laws & Regulations, quality of services

Scalability: The services of Microsoft are scalable either horizontal or vertical. It is important that our applications can support this

Life Cycle: The cloud service provider has a high frequent life cycle (each 3 months), meaning our applications must support the life cycle frequency of the cloud service provider

4.3.3 Splunk

Architecture

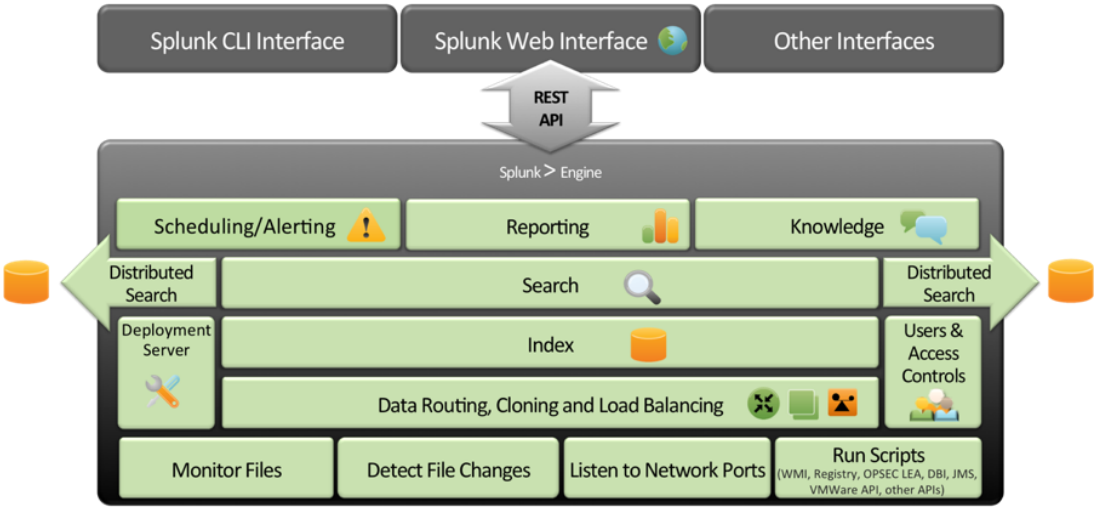


Figure 11: Architecture diagram
Source: [47]

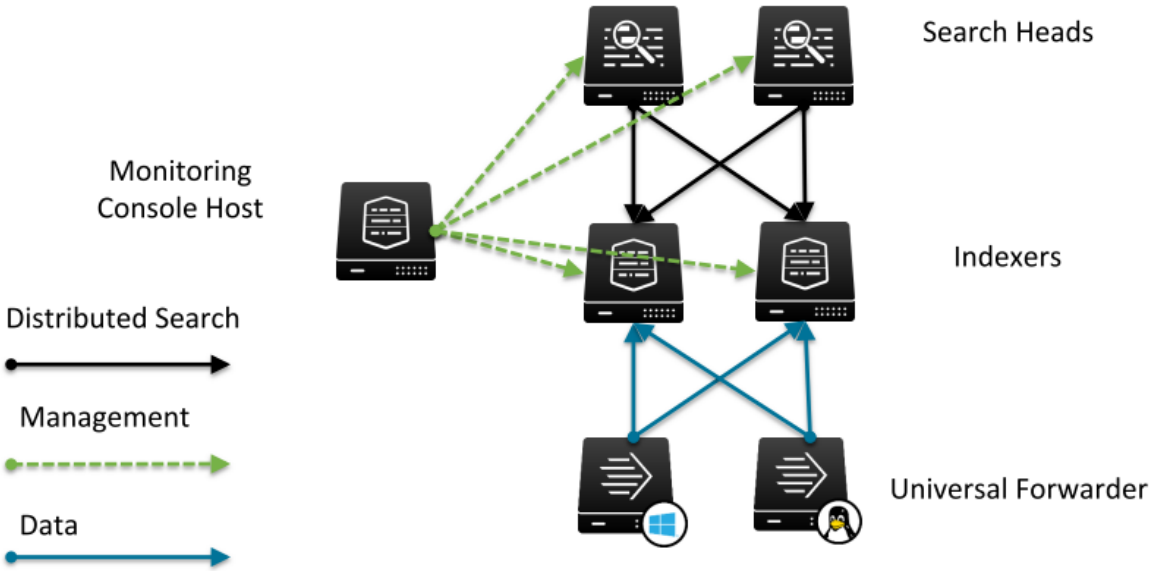


Figure 12: Splunk Monitoring Architecture
Source:[47]

Conclusion

As our consolidation effort continues we will be using Splunk to look into how an application is used in order to determine how it could be consolidated with other

applications of similar function. There is an amazing amount of information about usage patterns, what gets accessed and how often and who does the accessing.
These are all dependent on Splunk being able to get access to the log files. If Splunk has access to the log files, it can monitor these solutions. If Splunk can't get access, then

Splunk allows us access to that information. Here is where I need to bring up a big warning flag. Having access to the logs does not mean you can understand the logs. There are some errors where the team running a system is required to correctly interpret the logs, but in general having more eyes is a good thing. Some of the expertise can be developed over time, some more through developing dashboards and applications within Splunk.

Advantages of Splunk:

- 100% SLA agreement
- Can read and generate any data
- It's centralized data based
- ½ hour to troubleshoot
- Proactive alerts for issues
- Easy access to infrastructure data
- Real-time reporting
- Generic alerts. Ability to create alerts that work for systems that are not in the system yet. The ability to look at the entire environment as a single event stream is incredibly powerful.

Splunk is only exiting monitoring tool which guarantees 100% service level agreements. On SIEM[11] level it's one so far. Security information and event management (SIEM) tools are used to collect, aggregate and correlate log data for unified analysis and reporting. Typically, these tools can take logs from a large number of sources, normalize them and build a database that allows detailed reporting and analysis. While forensic analysis of network events may be a feature of a SIEM, it is not the only feature, nor is it the primary focus of the tool. On 9th of March 2016

[11] SIEM- Security Information and Event Management

4.3.4 Nimsoft

Architecture

Message Bus

Applications within the CA Nimsoft Monitor domain communicate by exchanging messages. The CA Nimsoft Monitor message bus provides the capabilities required to communicate across an entire enterprise infrastructure.

When a system within a CA Nimsoft Monitor domain has new data, it automatically publishes it via the message bus. All applications that subscribe to receiving updates for that system will automatically receive that update. Traditional point-to-point client-server systems require the sending of multiple copies of messages, one to each application, which is far less efficient. The CA Nimsoft Monitor architecture enables administrators to configure all publications and subscriptions via a single, native management console. All monitoring configurations can also be controlled programmatically via robust APIs—allowing cloud providers to completely automate the monitoring of their critical applications.

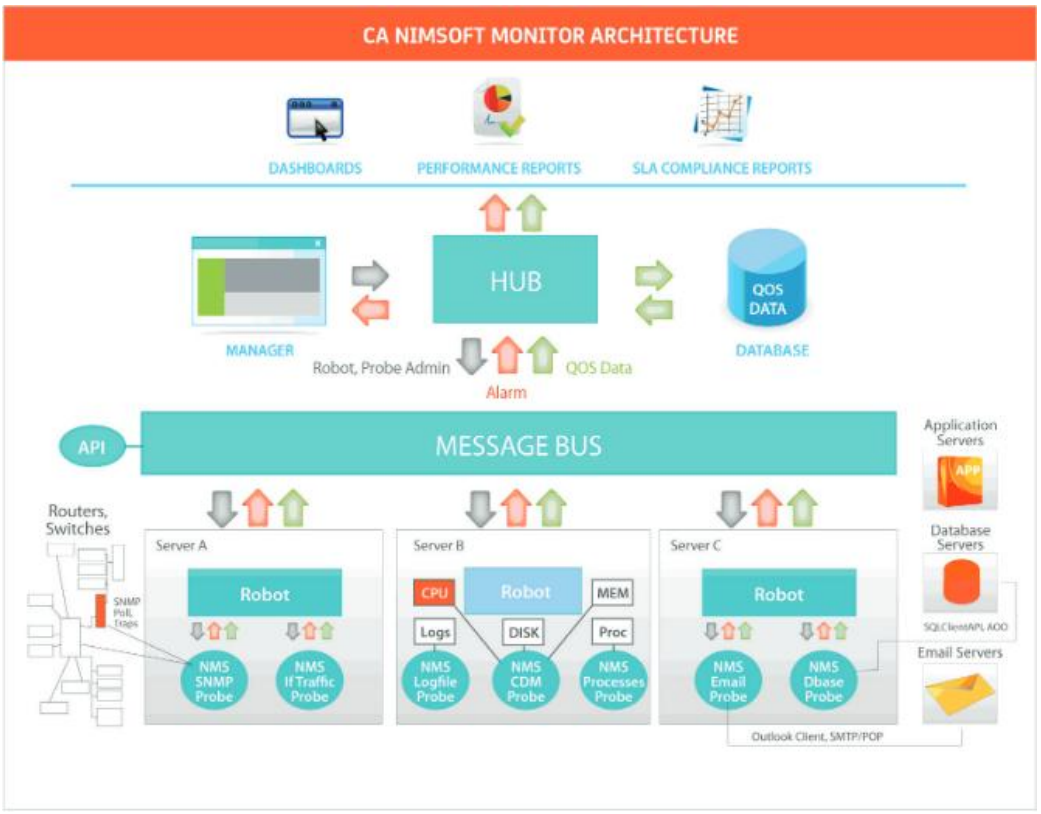


Figure 13: The CA Nimsoft Monitor Architecture
Source: [48]

Hub

A hub is a vital component within any CA Nimsoft Monitor deployment. A hub is basically a software component within a CA Nimsoft Monitor domain that enables components to connect to the message bus. A hub receives all messages posted by any client and distributes these messages to a set of subscribers of the publishing subject. A hub also keeps track of the addresses in the hub's domain, as well as information about each of the systems being monitored via CA Nimsoft Monitor robots (agents). A CA Nimsoft Monitor domain can have multiple hubs, which enables fail-over in the case of a communication disruption. Multiple hubs are also used to connect managed networks together across the Internet via SSL tunnels between the hubs—allowing service providers to connect remote customers to their operations centre.

Conclusion

Nimsoft Monitor Advantages

• Lightweight data collection

• Scalability to tens of thousands of devices with one instance

• Out-of-the-box-support for more than 140 device types—including servers, databases, applications, network devices and cloud services

• Single, multi-tenant portal view

• Customizable dashboards and reporting

Figure 14: Magic Quadrant for Security Information and Event Management

Source:[33]

The results of the 2015 Gartner SIEM Magic Quadrant are in line with feedback from Splunk reference customers who give us high satisfaction scores in the following areas:

- Big-data scalability
- Analytics-enabled security
- Real-time monitoring
- Advanced incident response
- Fast and flexible creation of correlation searches, alerts, reports and dashboards for security teams, business line owners, executives and auditors

4.3.5 Comparison of chosen tools

In this chapter we will compare Amazon Cloudwatch, Microsoft Azure watch, Nimsoft, Splunk by their technical features. In the Table 12, author made system , monitoring and alerting capabilities of chosen tools.

In the Table 11, author collects around 30 cloud monitoring well and known most prioritised features of cloud applications monitoring tools. We note that some of these capabilities are somewhat subjective. With that in mind the evaluation presented here is based on the weight of opinion as reflected in the reviewed literature. Author also participated in verity of official webinars, online and face to face meetings with Splunk sales engineers, NN cloud engineers and NN cloud operation leaders.

To the best of our knowledge, there is no any valuable academic research, or proven analysis which compares mentioned tool (Splunk) by applying empirical methods. After, case study of a company and face-to face meeting with NN infrastructure and cloud engineers, author summarised all features of Splunk as weight based. Weights of capabilities are based on individual meetings with engineers and they are targeted bases on NN needs and NN requirements.

4.3.6 General comparison

Table 12: General comparison of chosen tools (monitoring perspective)

Name of tools	Monitored resource	Open source	Operating system	Alerts	Messaging system	Implementation language	Reported limitations
Amazon Cloud Watch	AWS resources and applications and services running in AWS	No	Linux, Windows	E-mail	Amazon Simple Queue Service	Scripts in Windows PowerShell for Windows, Perl for Linux	Works for Amazon resources only, works in centralised models, does Not ensure availability, potential security threat due to Non-efficient use
Azure Watch	Azure based resources	No	Windows	E-mail, report to browser, smartphone, RSS feed	No	.Net	Works for Azure based resources only
Splunk	Anything that provides textual based logs. No limit as to what can be monitored	No	Linux, Windows, Mac	Email, links to some ticketing systems, can run custom scripts	Utilizes existing email infrastructure	C/C++, Python	Binary data (Not textual data), is not easy to get into Splunk, but there are ways to convert binary data to textual data, in order to digest if needed
Nimsoft	Various Cloud resources, OS, Network and applications	No	Linux, Netware, Unix, Windows, Mac	E-mail, RSS, text messages, instant messenger, Twitter	Nimsoft Message Bus	C/C++, Java, Perl, VB, and .Net	Does not cap resource consumption by the monitor, not fault-tolerant , does not support SLA compliance checking for data durability or location

Source: [author]

4.3.7 Capabilities fulfilment (in %)

Based on comparison [Table 2: Cloud based monitoring tool analysis] we will make another comparison by adding our new tool called Splunk. Values in the table are not weighted, they are based on Yes=1 or No=0 or limited=0.5 approaches. In the later analysis we will make weighted calculation based on MADM methods.

Note that some of these capabilities are somewhat subjective. Author of this table maintained in mind the evaluation presented here is based on the weight of opinion as reflected in the reviewed literature.

Table 13: Capabilities fulfilment in percentage

NO	Capability/features	Splunk	Nimsoft	Amazon Cloud Watch	Azure Watch
1	Scalability	Yes	Yes	Yes	Yes
2	Portability	No	Yes	No	No
3	Non-intrusiveness	Yes	Yes	Yes	Yes
4	Robustness	Yes	Yes	Yes	Yes
5	Multi-tenancy	Yes	Yes	Yes	Yes
6	Interoperability	Yes	No	No	No
7	Customizability	Yes	Yes	Yes	Yes
8	Extensibility	Yes	Yes	Yes	Yes
9	Shared Resource monitoring	Yes	Yes	Yes	Yes
10	Usability	Yes	Yes	Yes	Yes
11	Affordability	Yes	limited[12]	No	No
12	Achievability	Yes	Yes	Yes	Yes
13	Resource usage metering	Yes	Yes	Yes	Yes
14	Service usage metering	Yes	Yes	Yes	Yes
15	Service KPI monitoring	Yes	No	No	No
16	QoS	Yes	Yes	Yes	Yes
17	Risk assessment	Yes	Yes	Yes	Yes
18	Component status identification	Yes	Yes	Yes	Yes
19	Service load monitoring	Yes	Yes	Yes	Yes
20	Configuration verification	Yes	Yes	Yes	Yes
21	Configuration drift identification	Yes	Yes	Yes	Yes
22	Configuration effect monitoring	Yes	Yes	Yes	Yes
23	Security breaches monitoring	Yes	Yes	Yes	Yes
24	User access control	Yes	Yes	Yes	Yes
25	User activity monitoring	Yes	No	No	No

[12] Limited is half value – 0.5

26	Secured notification	Yes	Yes	Yes	Yes
27	Secured storage	Yes	Yes	Yes	Yes
28	Service dependency	Yes	No	No	No
29	Not own resource based	Yes	Yes	No	No
30	Application service – SLA	Yes	Yes	Yes	Yes
	Percentage covered by tools	**97%**	**85%**	**76%**	**76%**

Source:[author own processing, 35]

Weighted approach- Cloud capabilities fulfilment

This priority capabilities methodology requires author to identify the priority capabilities for a class of cloud monitoring capapbilies or services. Each capability is then weighted in terms of studied companies prioritized selections. Author focusing more into important capabilities which has collected from accountible engineers in the compnany, which were interviewed during the analysis. Next, selected montoring tools in the short list are rated in terms of how well they achieve each of the priority capabilities. A score that summarizes how well they meet the companies speisifc priority capabilities for each use case is then calculated for each cloud based monitoring tools.

Company's speisifc prioritized cloud capabilities are attributes that differentiate selected cloud monitoring tools as a class in terms of their quality and performance. Author callected them from NN cloud engineers which they considered as set of higher priority capabilities at some of the most important criteries for acquisition decisions.

In defining the cloud based application monitoring tools for evaluation, the author first identifies the leading uses for the monitoring tools in this market. Author, also considers What needs are end-users (NN in our case) looking to fulfill, when considering to choose cloud based applicaton monitoring tools in this market? The author matched his analyses with case study of the company and common client deployment scenarios. These distinct client scenarios defined in the case studies.

The author then identifies the cloud monitoring capabilities. These capabilities are generalized groups of features commonly required by this class of cloud based monitoring tools. Each capability is assigned a level of importance in fulfilling that particular need; some sets of features are more important than others, depending on the use case being evaluated. Use cases is are collected requiremnts from NN cloud engineers.

Each monitoring tools are evaluated in terms of how well it delivers each capability, on ten points ratings. These ratings are displayed side-by-side for all monitoring tools, allowing easy comparisons between the different sets of features.

To determine an overall score for each capapbliti in the use cases, the tools ratings are multiplied by the weightings to come up with the product score in use cases. Author implemented them by MADM usig scoring and sequence methods.

The prioritized capabilities which author selected from NN egnineers are not represent all capabilities for any monitroing tools; therefore, may not represent those most important for a specific use situation or business objective.

Author used a prioritized capabilities analysis as one of several sources of input about a product before making a monitoring tools decision. Our aanalysis by Multiple Attribute Decision Making using Scoring method. Scale between (1-10) 10 is best, 1 is not almost not fulfils.

point(weights) for each tool and capabilities gathered from interviewed engineers, company infrastructure IT operation leaders and cloud monitoring service vendors. We make averages of that given points. Note that, in the table below, point(weights) for Nimsoft, Amazon CloudWatch, and Azure Watch are based on this [35] research but we have added Splunk into table. Point(weights) for those 3 column author gathered via practical and theoretical knowledge's.

Table 14: Cloud capabilities fulfilment(1-10 scale)

NO	Capability/features	Splunk	Nimsoft	Amazon Cloud Watch	Azure Watch
1	Scalability	10	9	10	10
2	Portability	6	1	1	1
3	Non-intrusiveness	8	8	8	8
4	Robustness	9	9	9	8
5	Multi-tenancy	10	8	9	9
6	Interoperability	8	2	2	2
7	Customizability	10	9	10	10
8	Extensibility	9	8	9	8
9	Shared Resource monitoring	9	8	9	9
10	Usability	10	9	9	9
11	Affordability	6	3	5	5
12	Achievability	10	10	10	10
13	Resource usage metering	9	9	9	9
14	Service usage metering	8	7	9	9

15	Risk assessment	10	9	9	9
16	Component status identification	10	9	9	9
17	Service load monitoring	10	10	10	10
18	Configuration verification	6	7	8	8
19	Configuration drift identification	6	6	8	8
20	Configuration effect monitoring	10	8	8	8
21	Security breaches monitoring	10	7	9	8
22	User access control	10	8	10	10
23	Secured storage	9	8	9	9
24	Service dependency	10	8	3	3
25	Not own resource based	10	9	4	5
26	Application service - SLA	10	9	9	9
	SUM	**237**	**198**	**205**	**203**

Our comparison clearly shows that Splunk is becoming most suitable monitoring solution in this studied environment. It has maximum point(**237**) comparing to alternative solutions.

We can say second best option could be Amazon Cloudwatch with score **205,** and on the third list AzureWatch with 203. Last best option would be Nimsoft from CA company.

Table 15: Summary

Tools	Results	Most critical points
Splunk	**237**	Not portable and indexes ASCI based
Amazon CloudWatch	**205**	Monitors only AWS based services
Micrasoft AzureWatch	**203**	Monitors only Azure based services
Nimsoft	**198**	Not affordable, too expensive

Source:[author]

4.4 Technical implementation

In this chapter we will make technical implementation of this chosen tool by making a proposal as component design of Splunk in the real business environment. Due to companies security and complains priorities name of the servers, port numbers, gate numbers, and any other sensitive details are not mentioned or explained as demo names.

This technical implementation can be applied as architectural and component design of solution in the company, and the same time it can be reviewed as applications monitoring solution proposal in cloud environment.

4.4.1 Introduction

In this chapter, we will briefly explain purpose of this technical implementation and some specific component design decisions. As we have analysed and found out that Splunk as monitoring tool can be used in this business environment. In the next chapter we will gather all requirements from company and our implementation will be focus on those requirements.

As an outline of this implementation following standard engineering processes are used:

- All requirements are gathered
- We have analysed infrastructure requirements
- Requirements gathering
- Security requirements analysed
- Architectural design of the component is implemented

Purpose of this implementation

Purpose of these component design implementation is to make proposal of implementing Splunk in the enterprise as it's component based especially regarding the cloud infrastructure monitoring solution to a level that is sufficient to plan to implement this tool to the solution on the standard NN infrastructure building blocks. It provides the possibility to track the component design decisions to the requirements and principles on which they are based. Our proposal also regarding an Splunk infrastructure component that implements (part of) a solution building block from the Enterprise Technology Framework (ETF). It must provide sufficient detail to produce the software build used deploy the component. It provides the possibility to track the design decisions to the requirements and principles on which they are based.

Basic characteristics of chosen tool

Summary of Splunk can be found in this below table:

Application ID	Application-001
Application Name	Splunk
Application Architecture Type	Web Application/three tier
Application Development Type	COTS[0]
Application Vendor or Development Platform	**Vendor (reseller)** ABCD Technologies a.s. Prague, Czech Republic
Business Unit or Functional	CIO[13]

Unit	
Business Domain (BDM)	IT Tooling
Business Functionality	Analysing and Monitoring
Location in ETF	Splunk delivers services in multiple blocks in ETF[13]: Analytics services, Application services, all System Management components and Security Monitoring
Component Function	Analysing and monitoring
Vendor Support	Yearly term licence including maintenance and support by ABCD and Splunk.

Table 16: Basic characteristics of implementing tool[author]

In the above table name of the application vendor or development platform is anonymized as ABCD Technologies a.s. Prague, Czech Republic. It is not a real company.

Preconditions

Before implanting our tool in the company we should have agreement and defined strategies from security risk assessment management. A security risk assessment on Splunk and a resulting security action plan triggered the actions to create the basic implementations, like BIA[16], SRL[30], A-OSG[14] from scratch. Together with the increasing demands from business units the design of this implementation can be extended and improved to comply with the required ratings of confidentiality, integrity and availability. Our technical implementation proposal does not cover those above preconditions, they are out of scope of this thesis.

[13] ETF- Enterprise Technology Framework

Component design decisions

As a technical implementation of Splunk, it's is positioned as a monitoring tool, author prepared premade component design decisions and all are from tooling perspectives. Following specific component design decisions are made.

- In phase one the Universal Forwarders in the Cloud will connect to the Heavy Forwarders on-premise within the internal LAN. Because, company is going to move into Hybrid cloud environment, and that's why some applications are still on premise environment.
- The Universal Forwarders will get their configuration and apps from the deployment server on-premise within the internal LAN.
- Splunk ports of Universal Forwarders will be generally implemented within a security group, so No firewall changes have to be made when adding a Universal Forwarder.
- Deployment in the Cloud must be automated a much as possible. This also concerns Heavy and Universal Forwarders. They must be packaged completely including their configuration, and be automatically deployed without the need for a change (including their configuration = minimal: at minimum know how to reach the deployment server).

4.4.2 Requirements

Component Context

Splunk provides monitoring of applications, middleware and infrastructure in the NN IT Landscape. Splunk gathers data from these environments and indexes the data. This data is transformed into valuable information to proactively monitor the chain of applications by using searches, alerts, reports and dashboards. Splunk is positioned as a monitoring and analysing tool. Collecting the data can be done by a Splunk Universal Forwarder, query a database via the Splunk app DB Connect, executing SOAP requests, listening on a tcp/udp stream, listening on queues via JMS[14], etc.. Splunk can index all kind of data as long it is ASCII data. Before the data gets indexed it will be routed via the Heavy Forwarder. This component acts as a proxy, but is also being used to automatically delete, transform or

[14] **JMS**- Java Messaging Services

anonymize data before it gets indexed. This is used to save license costs and to prevent confidential data to get indexed.

When the data is indexed on the Indexers it can be search for with a web interface on a Search Head. Searching is done for troubleshooting and analyzing the data to get valuable information. These searches are the basis for alerts, reports and dashboards.

The following picture describes the Splunk environment from high level perspective and is independent of the type of datacenter or Cloud solution.

Graphical view

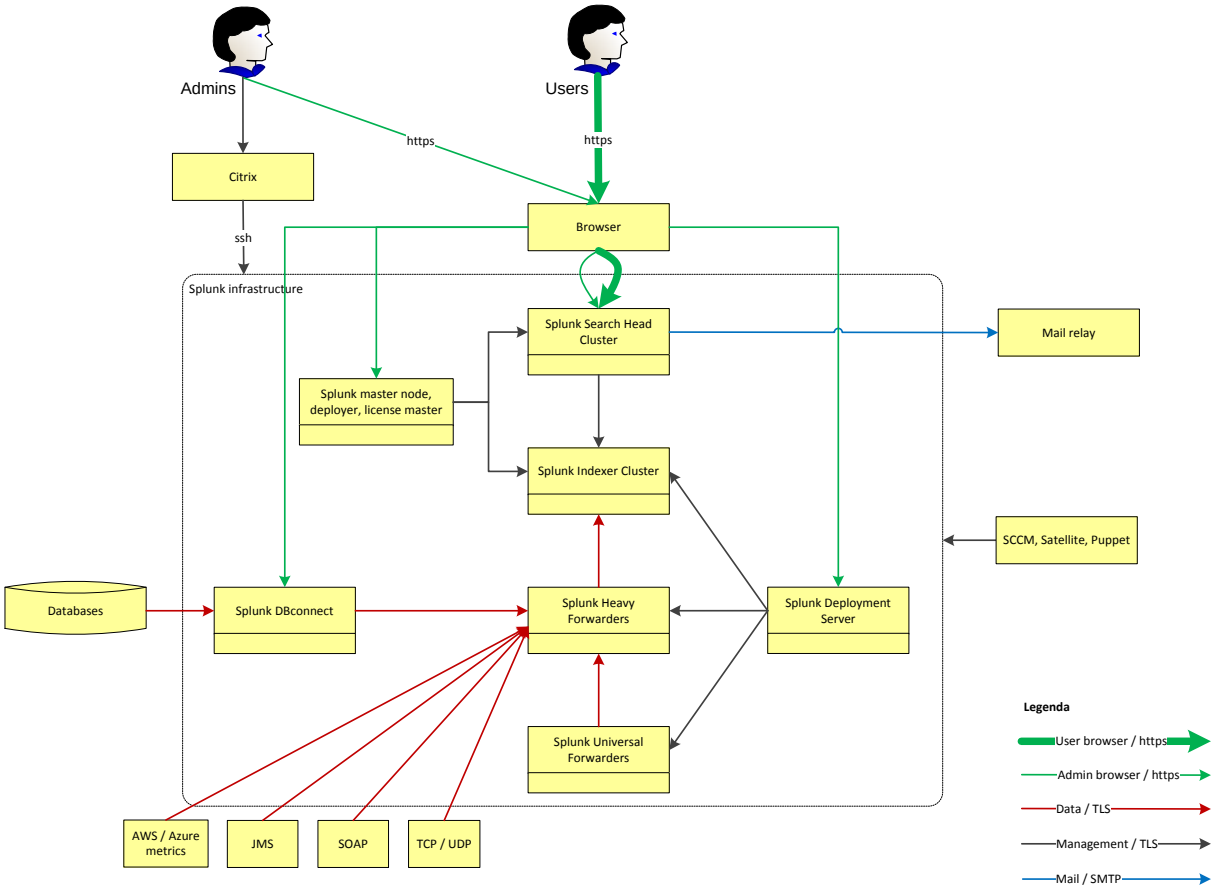


Figure 15: Splunk environment from high level perspective
Source:[author in cooperation with NN, 2016]

Table 17: Diagram description - Actors
Source:[author in cooperation interviewed with Engineers]

Actor	Description
NN employee – users	Functional Management (Creation and maintenance of searches, alerts, reports and dashboards)
NN employee - admins	Technical Management Infrastructure. Employees who setup and maintain the Splunk components and keep Splunk up and running.
Splunk Search Heads	A cluster of Search Heads which handle web requests and authentication of users and distribute the searches to the indexers. Reports, alerts and dashboards are stored on and distributed between the Search Heads.
Splunk Indexers	Data is collected, indexed and stored on the Indexers. The indexers actually execute the searches on the data and deliver the requested data to the Search Heads.
Splunk master node, deployer, license master	This component delivers three functions: Master node for the Index Cluster, deployer for the Search Head Cluster and will act as license master.
Splunk Heavy Forwarders	Acting as a proxy for Splunk Universal Forwarders and other data-sources like JMS, SOAP and TCP/UDP streams. Also used to automatically delete, transform or anonymize data before it gets indexed.
Splunk Universal Forwarders	Stream data from application hosts and infrastructure components to Splunk Heavy Forwarders.
Splunk DB connect	App on dedicated host to collect both technical and functional data from databases.
Splunk Deployment Server	Deployment of Splunk technical add-ons (apps), both apps from splunkbase.com and self-created apps.
SCCM, Satellite, Puppet	Default NN services to deploy Splunk software packages and perform initial necessary OS configuration.
Mail relay	Mail is send from a Search Head by an alert.

Non-functional requirements

Functional requirements have not been described because they will be different by each departments service catalogue. But author will specify only non-functional requirements as specific as possible.

Non-functional requirements overview:

- The Splunk server components must run on Linux. This is a prerequisite of the Splunk administrators.
- For ease of use and setup the Universal Forwarder and Heavy Forwarder must:
    - be part of a common security group to avoid having an extensive and time consuming process, change of documents, change of RCEC's and security approvals when adding such a Forwarder.
    - must be fully automated and must be a default option during provisioning of a host. If automation is not ready yet it will be done via a standard service request.

- use the Deployment Server to push configurations and Splunk apps.

Performance

- Searching and indexing is done on-premise by using an Indexer cluster and multiple Heavy Forwarders, there are always server components available to ingest data from the Universal Forwarders in the Cloud.
- Basic network throughput has to be available.
- The Universal Forwarders take a negligible amount of resources from the hosts installed on. No extra resources are needed on hosts where a Universal Forwarder is installed.
- Heavy Forwarders parse the data stream into events. This is not performance intensive. The requirement is 2 CPU cores, 4GB memory and 10GB storage.
- Heavy and Universal Forwarders are stable components and hardly fail. Most Forwarders are running for more than a year without failure.
- By using multiple Heavy Forwarders a more redundant environment can be achieved. Extra instances can be added any moment to improve performance and/or redundancy. Manual configuration is necessary to push adjusted configuration to the Universal Forwarders.

Operational

- Basic infrastructure monitoring of Splunk components must done by the CIO/SD&C standard monitoring tools for that (Cloud) environment.
- Monitoring of the application Splunk is done by Splunk itself.
- Service Window: See Splunk Service Catalogue.

4.4.3 Acceptance criteria

Operations

- Splunk administrators must have Linux access to the Splunk server components.
- Splunk administrators must have access to the web interface of Splunk components were the web interface is enabled.
- Splunk administrators must be able to centrally deploy Splunk apps and configurations to Splunk Universal Forwarders in the Cloud.
- Administrators of applications or hosts are not allowed to change configuration of the Splunk Universal Forwarder due to security and license usage.

Security and Risk

Risk and security requirements will be formulated based on companies local and international risks and compliance agreements. As every company has internal or external auditing obligations our implementing also follows into company's regulations. We will describe them as general as possible. Our formulated security ratings are very abstract.
The implementation of Splunk in the cloud must comply with Cloud Security Principles.

4.4.4 Functionality

The Figure 16 describes the Splunk environment from high level perspective and is independent of the type of datacenter or Cloud solution. Splunk will be implemented in the Cloud in steps. The end-goal is to move the entire Splunk infrastructure to the Cloud. Moving the Splunk server components to the Cloud can be done in several ways by using intermediate hybrid solutions. This phase is currently not in scope of this thesis.

The first goal is to get data from the Cloud (by Universal Forwarders) to the Heavy Forwarders on-premise. Only Universal Forwarders will be installed on the hosts of applications. No extra costs for virtual hardware have to be made. Data will only cross one border of a VPC[15] (AWS) or Vnet (Azure) to the on-premise Splunk infrastructure. By default, data is secured by TLS and compressed.

Depending on functional needs, load, amount of traffic or availability of other Splunk components a second step can be taken to make a Heavy Forwarder available in the Central VPC or Vnet. Extra virtual hardware is needed and data will cross multiple borders before it reaches on-premise Splunk infrastructure. This step will be a logical choice when moving Indexers to the Cloud, since than data will cross only one border. A third step can be taken to make a Heavy Forwarder available in individual application VPC's or Vnet. More virtual hardware is needed. The end user does not make use of the components situated in the Cloud.

4.4.5 Maintainability and Support

- All Splunk server components are long-lived instances, including Heavy Forwarders.
- Product Support is delivered by agreed teams. They need proper access on the Splunk web interfaces and Splunk server components. This will be described in the Splunk Authorization Matrix.

[15] VPC- Virtual Private Cloud

- Splunk administrators do not need access on hosts were Universal Forwarders are installed, regardless of short- or long lived instances. Configurations are retrieved from de Splunk Deployment Server by the Universal forwarder and are restarted automatically. Help of the administrator of the host is requested when necessary.

4.4.6 Capacity of environment

The capacity and performance of the Splunk environment is dependent on the amount of concurrent searches and the execution time of each individual search. A search can be created ad-hoc by a user, but most searches run in dashboard panels at regular interval or are scheduled for alerts and reports.

Capacity and performance is also impacted during collection of data, but this has less impact compared to the concurrent searches.

Indexer capacity

The indexers are hit by both collecting/indexing data and executing searches. These are both very disk intensive activities. The storage of the indexers should be capable of at least 700 IOPS. Because index clustering will be used, the SAN's of DC1 and DC2 do not have to replicate their Splunk data. Therefore the option for fast storage has been chosen.

Search Head capacity

The Search Head delivers the web front-end and orchestrates the searches. These are will be more CPU bound activities. The other components need less hardware. The specific requirements are will be described in the following table.

Requirements	Search Head	Indexer	Forwarder
Service recovery	Standard	Standard	Standard
Service window	24*7	24*7	24*7
Operating System	Red Hat Linux v6	Red Hat Linux v6	Red Hat Linux v6
Computing resource	Dedicated	Dedicated	Dedicated
vCPU	16 vCPU	12 vCPU	4 vCPU
Memory	12 GB	12 GB	8 GB
Storage	40 GB	4096 GB (> 800 IOPS[16])	10 GB
Network	ISBB	ISBB	2 within ISBB 2 within RSABB
Availability	>=98%	>=98%	>=98%
Node	Master Node	DBconnect	Deployment server
Service recovery	Fast Recovery	Fast Recovery	Fast Recovery
Service window	24*7	24*7	24*7
Operating System	Red Hat Linux v6	Red Hat Linux v6	Red Hat Linux v6
Computing resource	Dedicated	Dedicated	Dedicated
vCPU	4 vCPU	2 vCPU	2 vCPU
Memory	12 GB	8 GB	8 GB
Storage	40 GB	10 GB	40 GB

[16] IOPS- Input/Output Operations Per Second

Network	ISBB	ISBB	Management LAN
Availability	>=98%	>=98%	>=98%
Additional Service options	Not applicable	Not applicable	Not applicable
Middleware	Not applicable	Not applicable	Not applicable
Application software	Splunk enterprise	Splunk enterprise	Splunk enterprise

Table 18: Splunk environment required specifications [author]

Onetime/online workload

The online load profile of application Splunk is office hours. During office hours the load expected to be roughly twice and the load during the night and weekend it less then average. During office hours several teams start their dashboards and run ad-hoc searches. The expected growth of the online workload in the next 3 years is 20% per year.

Responsiveness of the tool

The expected response time for application Splunk depends on the amount of concurrent searches and the execution time of each individual search. If the timeframe of a search is set to 30 days, it will take longer to show the results compared to a timeframe of 4 hours. If a search delivers a lot of events, it will take longer to show the results compared to a search with a few events as a result.
The goals is to prevent queued and skipped searches. This is achieved by monitoring default available activity dashboards and teach employees to create efficient searches and dashboards.

4.4.7 Database size

The main advantages of Splunk monitoring is it's search functionality based on indexes and Splunk does not use a relational database like Oracle, DB2 or MSSQL or any other relational databases. But it stores its indexed data and other Splunk objects (like searches, alerts, reports and dashboards) right on the file system. The required file system sizes for the different Splunk components are described in the table below.

Component	File system	Size	Retention period	Growth/year
Search Head	/home/splunk	60 GB	Not applicable	**10%**
Indexer	/home/splunk /home/splunk/var	30 GB 5 TB	Not applicable 1 year	**0%** **20%**
Heavy Forwarder	/home/splunk	20 GB	Not applicable	**0%**
Splunk master Node, deployer, license master	/home/splunk	60 GB	Not applicable	**10%**
Universal Forwarders	/home/splunk	30GB	Not applicable	**0%**
Dbconnect	/home/splunk	30GB	Not applicable	**0%**
Deployment Server	/home/splunk	60GB	Not applicable	10%

Table 19: File archiving retention [author]

Retention of indexes on the Indexers will be configured within Splunk enterprise. At the same time we can specify above table as the storage requirements for application Splunk.

4.4.8 Conclusion

As a conclusion of this technical implementation of Splunk Enterprise in the specific company, author focused on company's technical and environmental requirements, short term cloud migration plans, company's cloud strategies. To author's best knowledge, there are No any academic papers or valuable scientific research that explains component design of Splunk Enterprise during cloud migration period. Author's proposal gives challenge to make a proper research about Splunk Enterprise by focusing transition period from on-premise to cloud environments.

In this particular case of the company Splunk Enterprise can monitor remaining applications on company's datacentre and the same time its capable to analyse and monitor all infrastructure and applications on AWS and Azure cloud environments. As clearly explained in the case study, currently NN using Hybrid Cloud environment but in the future their Cloud strategy may turn to use fully SaaS solution in this scenario company should be able to analyse and monitor around 1100 application on cloud environment.

As a future work, author will analyse Splunk Enterprise by implementing local Splunc Cloud environment. Author, also found very interest and relatively new advantages of

Security monitoring prospective of Splunk which can be alternative product for large scale applications as HP ArcSight.

5 Results and Discussion

Our comparison clearly shows that Splunk is becoming most suitable monitoring solution in this studied environment. It has maximum points (**237**) comparing to alternative solutions. We can say that the second best option could be Amazon Cloudwatch with score **205,** and on the **the** third place is AzureWatch with 203. Last option would be Nimsoft from CA company.`

Table 20: Summary

Tools	Results	Most critical points
Splunk	**237**	**Not portable, Little costly**
Amazon CloudWatch	**205**	**Monitors only AWS based services**
Micrasoft AzureWatch	**203**	**Monitors only Azure based services**
Nimsoft	**198**	**Not affordable, too expensive**

Source:[author]

5.1 Selection process of tools

As we have analysed and founded best tool in our seppsific enterprise enviremnt. We had below chalanges.

Selected montoring tools in the short list are rated in terms of how well they achieve each of the priority capabilities. A score that summarizes how well they meet the companies speisifc priority capabilities for each use case is then calculated for each cloud based monitoring tools.

Company's speisifc prioritized cloud capabilities are attributes that differentiate selected cloud monitoring tools as a class in terms of their quality and performance. Author callected them from NN cloud engineers which they considered as set of higher priority capabilities at some of the most important criteries for acquisition decisions.

The author then identifies the cloud monitoring capabilities. These capabilities are generalized groups of features commonly required by this class of cloud based monitoring tools. Each capability is assigned a level of importance in fulfilling that particular need; some sets of features are more important than others, depending on the use case being evaluated. Use cases is are collected requiremnts from NN cloud engineers.

Each monitoring tools are evaluated in terms of how well it delivers each capability, on 10 points ratings. These ratings are displayed side-by-side for all monitoring tools, allowing easy comparisons between the different sets of features.

To determine an overall score for each capapbliti in the use cases, the tools ratings are multiplied by the weightings to come up with the product score in use cases. Author implemented them by MADM usig scoring and sequence methods.

The prioritized capabilities which author selected from NN egnineers are not represent all capabilities for any monitroing tools; therefore, may not represent those most important for a specific use situation or business objective. Consumers should use a prioritized capabilities analysis as one of several sources of input about a product before making a monitoring tools decision.

All existing papers, researches are considered Splunk as analytical tool. But, based on it's incredible capabilities we can say, Splunk is not only data analytical tool but it is fully technically proved cloud or on-premises monitoring tool even better than any other monitoring tools.

5.2 Discussion

Splunk is only exiting monitoring tool which guarantees 100% service level agreements. On SIEM level it's one so far. Security information and event management (SIEM) tools are used to collect, aggregate and correlate log data for unified analysis and reporting. Typically, these tools can take logs from a large number of sources, normalize them and build a database that allows detailed reporting and analysis. While forensic analysis of network events may be a feature of a SIEM, it is not the only feature, nor is it the primary focus of the tool. On 9th of March 2016

Scmagazine magazine ammonised, about **Best SIEM Solution** Splunk Enterprise Security wins Best SIEM Solution and Splunk Enterprise Named Best Fraud Prevention Solution.

Splunk's strong presence in IT operations groups can provide security organizations with early hands-on exposure to its general log management, monitoring and analytics capabilities, monitoring deployment by operations for critical resources, and in-house operations support for expanded monitoring-focused deployments.

Splunk customers cite visualization and behavioural, predictive and statistical analytics as effective elements of advanced monitoring use cases, such as detecting anomalous user access to sensitive data.

Splunk has enhanced built-in support for a large number of external threat intelligence feeds from commercial and open sources.

The average of Splunk reference customer satisfaction scores for scalability and performance, effective and useful predefined rules and reports, rule and report customization features, report creation, ease and effectiveness of ad hoc queries, product quality and stability, and support experience is higher than the average scores for all reference customers in those areas.

5.3 Splunk SWOT analysis

Strengths

- Splunk the only available monitoring tool with 100% service level agreements in the market
- Not only on cloud based but also on-premises monitoring tool.
- Monitoring in all levels of cloud service models and deployments.

- Splunk can monitor any data from any machine. It supports any format and any amount of data, enables centralized log management
- Data is stored in its own index. No separate database requirements like in Oracle or SQL.
- Administrator time savings. Substantially better than manual processing of logs.

Weaknesses

- Not portable
- For the large enterprises it is quite costly. But, pay as you go pricing is offered.
- While indexing log files into only ASCI format is used.

Opportunities

- Splunk is "google" of enterprise data.
- It can generate any reports about even minor events such as a system hiccup.
- It gives opportunity as cloud monitoring tool in aggregating the log files. Splunk can convert complex log files to visual graphs and reports resulting simplified analysis, reporting and troubleshooting. It can easily monitor real-time incoming logs.
- Strong capabilities to monitor big data in enterprise.
- Real time operational intelligence.
- Best security, event management features.

Threats

- The Splunk App for Enterprise provides basic support for predefined correlations for user monitoring. Potential buyers should anticipate modifying those and building their own to implement more advanced user monitoring use cases.
- Workflow and case management functions lag behind those of competitors. Organizations with mature SOC[17] processes may require customization or integrations with third-party technologies for these functions.
- Splunks license model is based on data volume indexed per day. Customers report that the solution is more costly than others products where high data volumes are expected. But day by they are changing their pricing strategy.

[17] SOC- Security Operation Center

5.4 Future work

In the future research, we would continue our research to analyse following criticised gaps in cloud monitoring environment such as disaster recovery, horizontal scaling, load balancing, messaging services, snapshots, vertical scaling.

Next research will be focus to analyse by technical, financial, risk assessment perspectives.

As conclusion of this thesis, we would conclude our opinion about Amazon CloudWatch and Micrasoft AzureWatch. They are very capable and own resource based monitoring systems but, they does not support monitoring across different cloud infrastructures. Only this reason one the big advantage of these tools. The feature of multi-tenancy, interoperability of these tools is compared to for general purpose monitoring tools. Since this capability is indispensable for multi-tenant Cloud environment, thus, we argue that this area is a challenge for future research. The need for securing the monitoring tool itself is important for ensuring that the tool does not create any monitoring holes in the Cloud. As can be observed, capabilities associated with secure monitoring are lacking implementations. Therefore, future research efforts are required in this area.

Compared to the general purpose monitoring tools, the scalability, non-intrusiveness, robustness, multi-tenancy, shared resource monitoring, resource usage metering, per service resource consumption metering and service load monitoring capabilities are better addressed in the Cloud specific monitoring tools. This is reasonable as the demand for these capabilities is high in Clouds. In line with our descriptions, the last row of Table 5 shows the weighted average percentage of the capabilities implemented by each of the tools. The Splunk monitor emerges as the best with **97%** coverage. Second best option is Nimsoft with **85%** fulfilment of company's prioritised capabilities. The Amazon CloudWatch, Microsoft Azure Watch tool are the least with 76%. One interesting observation is that most of the capabilities are improved with Cloud based monitoring tool except for portability and affordability. This reinforces the fact that many of the Cloud monitoring providers should focus to develop portable, more flexible and affordable products.

For large enterprises Splunk products can be more beneficial while using as monitoring tool. Gartner confirms [Annex 8Annex 9Annex 10] Splunk is a leading technology even on security monitoring, event management and threat management. These are all dependent on Splunk being able to get access to the log files. If Splunk has access to

the log files, it can monitor these solutions. If Splunk can't get access, then monitoring is not available.

Interest on Splunk is incredibly growing[Annex 1,Annex 2],

6 Conclusion

The diploma thesis is focused to find optimum cloud monitoring solution for a company which were transforming their data into public cloud environments. We have achieved all our main and partial goals successfully, and we have proposed technical implantation of optimum solution as component design of the tool. Including to our SWOT analyses Splunk also helps:

- Big-data scalability
- Analytics-enabled security
- Real-time monitoring
- Advanced incident response

We can conclude our summarised analyses by Table 20 which shows that Splunk is optimum cloud monitoring solution for NN cloud environment by scoring 237 points which is highest in the selected short list. Amzon CloudWatch(205), AzureWach(203), and Nimsoft(198).

Fast and flexible creation of correlation searches, alerts, reports and dashboards for security teams, business line owners, executives and auditors

Our analysis shows that in our specific case, Amazon CloudWatch and Azure Watch are not full fills company requirements. They are own resource based – this capability is one of the biggest disadvantage of those tools. As we discussed in the earlier chapters, NN Group is migrating into top IaaS and PaaS cloud providers in the market, which are AWS and Microsoft Azure. In one hand Amazon CloudWatch becomes as first and main monitoring tool to monitor AWS, but it cannot monitor NN applications on Microsoft Azure, this is a one of the biggest reason that Amazon CloudWatch cannot be used in the NN cloud environment.

Azure Watch, due to localised resource limitations it cannot monitor AWS based applications.

All these 3 tools has less SLA agreements then Splunk. Splunk is only one existing tool which grantees 100% SLA options.

As a conclusion of this thesis, we strongly recommend Splunk as monitoring tool in any cloud environment based enterprises especially our analysed company in NN. Splunk defiantly visualises there goals, short and long term business strategies.

7 Bibliography

1. Mell P, Grance T (2011) The NIST definition of cloud computing (draft). NIST Spec Publ 800:145
2. MOLEN, F. V. (2010). *Get ready for cloud computing: a comprehensive guide to virtualization and cloud computing.* (C. Brace, Editor) Amsterdam: Van Haren Publishing.
3. AHSON, Syed a Mohammad ILYAS. 2011. Cloud computing and software services: theory and techniques. Boca Raton, FL: CRC Press, xiv, 442 p. ISBN 9781439803158.
4. SMOOT, Stephen R a Nam Kee TAN. 2012. Private cloud computing: consolidation, virtualization, and service-oriented infrastructure. Waltham, MA: Morgan Kaufmann, xvii, 399 p. ISBN Private cloud computing.
5. MOYER, Christopher M. 2011. Building applications in the cloud: concepts, patterns, and projects. Upper Saddle River, NJ: Addison-Wesley, xiii, 326 s. ISBN 978-0-321-72020-7.
6. Talkincloud.com [online]. [cit. 2015-10-14]. Dostupné z: http://talkincloud.com/
7. Docs.splunk.com [online]. [cit. 2015-10-14]. Dostupné z: http://docs.splunk.com/
8. *Inforisktoday.com* [online]. [cit. 2016-03-17]. Dostupné z: http://www.inforisktoday.com/5-essential-characteristics-cloud-computing-a-4189
9. *Slideshare* [online]. [cit. 2016-03-17]. Dostupné z: http://www.slideshare.net/Prolifics/best-practices-for-monitoring-your-cloud-environment-and-applications
10. *Prolifics.com: Monitoring* [online]. [cit. 2016-03-17]. Dostupné z: http://www.prolifics.com/search/Node/monitoring
11. *PCMAG: Encyclopedia* [online]. [cit. 2016-03-17]. Dostupné z: http://www.pcmag.com/encyclopedia/term/56112/saas
12. *Gartner: Gartner IT Glossary* [online]. [cit. 2016-03-17]. Dostupné z: http://www.gartner.com/it-glossary/platform-as-a-service-paas/
13. *Gartner: Gartner IT Glossary* [online]. [cit. 2016-03-17]. Dostupné z: http://www.gartner.com/it-glossary/infrastructure-as-a-service-IaaS/
14. *Gartner: Gartner IT Glossary* [online]. [cit. 2016-03-17]. Dostupné z: http://www.gartner.com/webinar/3250628?srcId=1-2924665417&cm_sp=wbnr-_-rr-_-top
15. *Webopedia: Private cloud* [online]. [cit. 2016-03-17]. Dostupné z: http://www.webopedia.com/TERM/P/private_cloud.html
16. *Interoute: What is private cloud?* [online]. [cit. 2016-03-17]. Dostupné z: http://www.interoute.com/cloud-article/what-private-cloud
17. *Interoute: Public Cloud* [online]. [cit. 2016-03-17]. Dostupné z: http://www.interoute.com/cloud-article/what-public-cloud
18. *Gartner: Gartner Highlights the Top 10 Cloud Myths* [online]. [cit. 2016-03-17]. Dostupné z: http://www.gartner.com/newsroom/id/2889217
19. *Techopedia: Techopedia explains Community Cloud* [online]. [cit. 2016-03-17]. Dostupné z: https://www.techopedia.com/definition/26559/community-cloud
20. *Gartner: Hybrid cloud* [online]. [cit. 2016-03-17]. Dostupné z: http://www.interoute.com/cloud-article/what-hybrid-cloud
21. *SoftpaNorama: HP Operations Manager* [online]. [cit. 2016-03-19]. Dostupné z: http://www.softpaNorama.info/Admin/HP_operations_manager/index.shtml

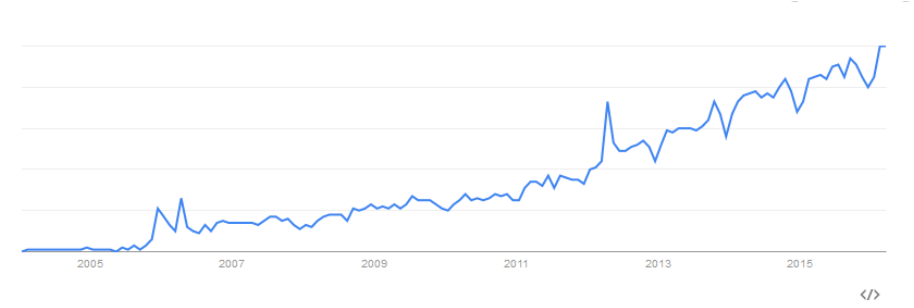
22. *Softwareinsider: Amazon Inc.* [online]. [cit. 2016-03-15]. Dostupné z: http://cloud-computing.softwareinsider.com/l/5/Amazon-Inc
23. *AWS: Products* [online]. [cit. 2016-03-21]. Dostupné z: https://aws.amazon.com/
24. Davis C, Neville S, Fernandez J, Robert J-M, Mchugh J (2008) Structured peer-to-peer overlay networks: ideal botnets command and control infrastructures? In: Computer security—ESORICS 2008, pp 461–480
25. *NN-group: Who we are* [online]. [cit. 2016-03-22]. Dostupné z: https://www.nn-group.com/Who-we-are/Who-we-are.htm
26. *NIST: Cloud Computing* [online]. [cit. 2016-03-23]. Dostupné z: http://www.nist.gov/itl/cloud/
27. *Gartner: community-cloud* [online]. [cit. 2016-03-17]. Dostupné z: http://www.gartner.com/it-glossary/community-cloud
28. *Cloud computing architecture: Cloud client platforms* [online]. [cit. 2016-03-23]. Dostupné z: http://america.pink/cloud-computing-architecture_1024785.html
29. *Gartner: Magic Quadrant for Public Cloud Storage Services, Worldwide* [online]. [cit. 2016-03-23]. Dostupné z: https://www.gartner.com/doc/3082618?ref=unauthreader
30. *Splunk: Get Two Gartner Reports: The 2015 Magic Quadrant and the 2015 Critical Capabilities for SIEM to See Why Splunk Was Named a Leader for the Third Straight Year* [online]. [cit. 2016-03-15]. Dostupné z: http://www.splunk.com/goto/SIEM_MQ
31. *NIST: Computer Security Devision* [online]. [cit. 2016-03-15]. Dostupné z: http://csrc.nist.gov/groups/SMA/fisma/overview.html
32. *AWS: Amazon CloudWatch* [online]. [cit. 2016-03-23]. Dostupné z: http://docs.aws.amazon.com/AmazonCloudWatch/latest/DeveloperGuide/cloudwatch_architecture.html
33. Gartner: Security and Event Management. *Gartner: Magic Quadrant* [online]. [cit. 2016-03-23]. Dostupné z: http://www.gartner.com/doc/reprints?id=1-2JNR3RU&ct=150720&st=sb
34. BAGIROV, Timur. *Splunk company overview april. 2015* [online]. [cit. 2016-03-24]. Dostupné z: http://www.slideshare.net/TimurBagirov/splunk-company-overview-april-2015-50798068
35. FATEMA, Kaniz, Vincent C. EMEAKAROHA, Philip D. HEALY a John P. MORRISON. *A survey of Cloud monitoring tools: TaxoNomy, capabilities and objectives.* , 2918–2933. DOI: 10.1016/j.jpdc.2014.06.007.
36. *Explore Splunk: Interest over time* [online]. [cit. 2016-03-19]. Dostupné z: https://www.google.cz/trends/explore#q=splunk
37. *Gartner: Critical Capabilities for Security Information and Event Management* [online]. [cit. 2016-03-28]. Dostupné z: https://www.gartner.com/doc/reprints?id=1-2O8Q585&ct=150929&st=sb
38. Caron E, Rodero-Merino L, Desprez F, Muresan A (2012) Auto-scaling, load balancing and monitoring in commercial and open-source clouds
39. ALHAMAZANI, Khalid, Rajiv RANJAN, Karan MITRA, Fethi RABHI a Prem PRAKASH JAYARAMAN. *An overview of the commercial cloud monitoring tools: research dimensions, design issues, and state-of-the-art: 1-21.* , 1-21. DOI: 10.1007/s00607-014-0398-5.
40. S. V. Gogouvitis, V. Alexandrou, N. Mavrogeorgi, S. Koutsoutos, D. Kyriazis and T. Varvarigou, "A Monitoring Mechanism for Storage Clouds," *Cloud and Green*

Computing (CGC), 2012 Second International Conference on, Xiangtan, 2012, pp. 153-159.
41. Giuseppe Aceto, Alessio Botta, Walter de Donato, Antonio Pescapè, Cloud monitoring: A survey, Computer Networks, Volume 57, Issue 9, 19 June 2013, Pages 2093-2115, ISSN 1389-1286, Dostupné z:http://dx.doi.org/10.1016/j.comnet.2013.04.001.
42. P. Hasselmeyer and N. d'Heureuse, "Towards holistic multi-tenant monitoring for virtual data centers," *Network Operations and Management Symposium Workshops (NOMS Wksps), 2010 IEEE/IFIP*, Osaka, 2010, pp. 350-356.
43. T. Lynn, P. Healy, R. McClatchey, J. Morrison, C. Pahl, B. Lee, The case for cloud service trustmarks and assurance-as-a-service, in: Proceedings of the 3rd International Conference on Cloud Computing and Services Science (CLOSER), 2013.
44. E. Elmroth, F.G. Marquez, D. Henriksson, D.P. Ferrera, Accounting and billing for federated cloud infrastructures, in: Grid and Cooperative Computing, 2009. GCC'09. Eighth International Conference on, IEEE, 2009, pp. 268–275.
45. *Deploying Sitecore on Microsoft Azure: Azure Architecture* [online]. [cit. 2016-03-20]. Dostupné z: http://www.xcentium.com/blog/2014/07/23/deploying-sitecore-on-microsoft-azure
46. *Best SIEM Solution* [online]. [cit. 2016-03-20]. Dostupné z: http://www.scmagazine.com/sc-awards-2016/section/5433/?publishDate=False×tamp=635852806671307920
47. *Splunk* [online]. [cit. 2016-03-29]. Dostupné z: http://www.splunk.com/
48. *CA NImsoft* [online]. [cit. 2016-03-29]. Dostupné z: http://www.ca.com/cz/~/media/Files/whitepapers/ca-nimsoft-monitor-delivering-a-unified-monitoring-architecture.PDF
49. *From Fad to Foundation: The Evolution of Cloud* [online]. [cit. 2016-03-29]. Dostupné z: http://research.toolkitcafe.com/content40504
50. *NN Group: Company's internal resources[online].*[cit. 2016-03-29*]*.

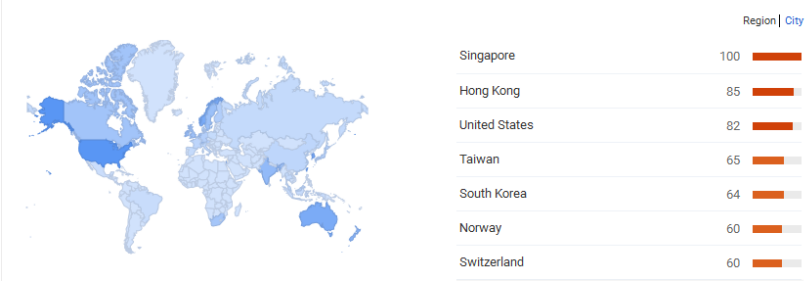
8 Annexes

8.1 Supplementary figures, charts and diagrams

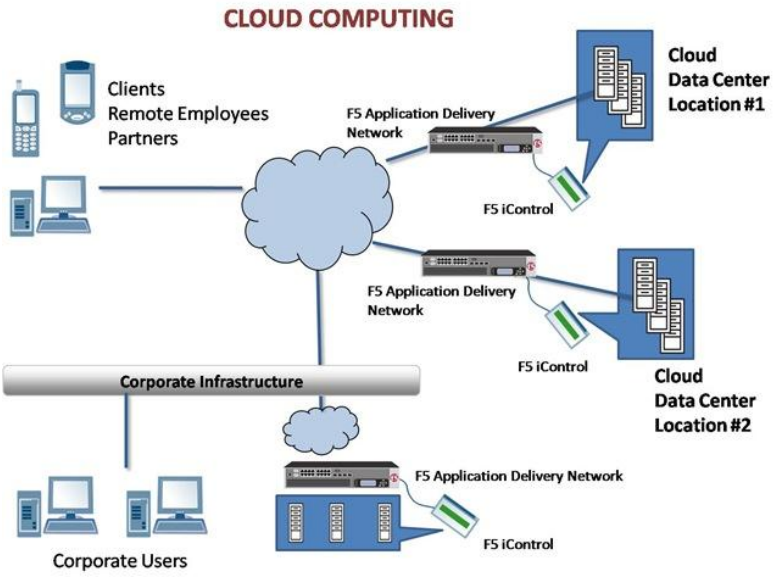
Annex 1: Google Trends. Splunk interest over time [36]



Annex 2: Google Trends. Splunk regional interest[36]

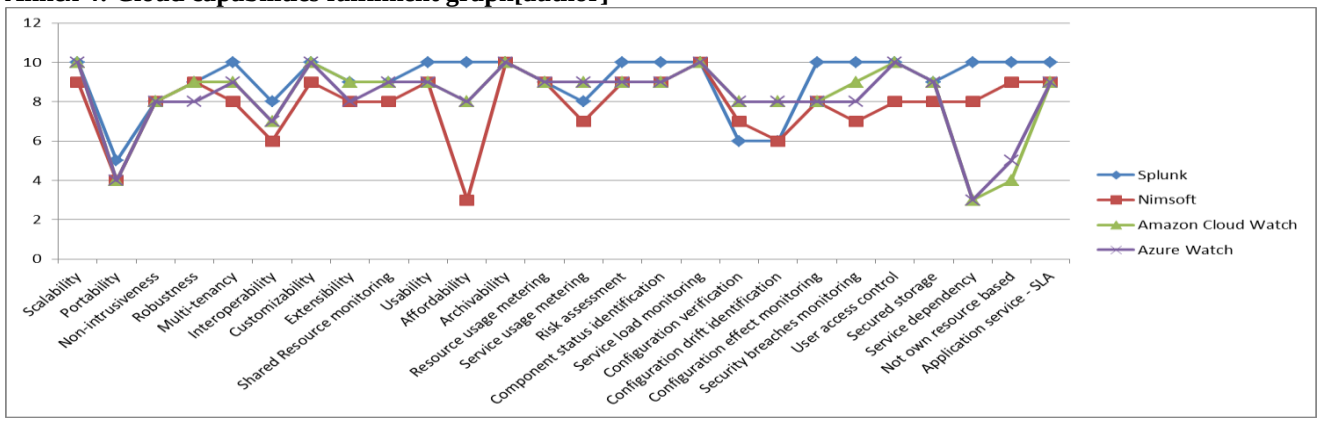


Annex 3: Cloud computing corporate and client platforms[28]

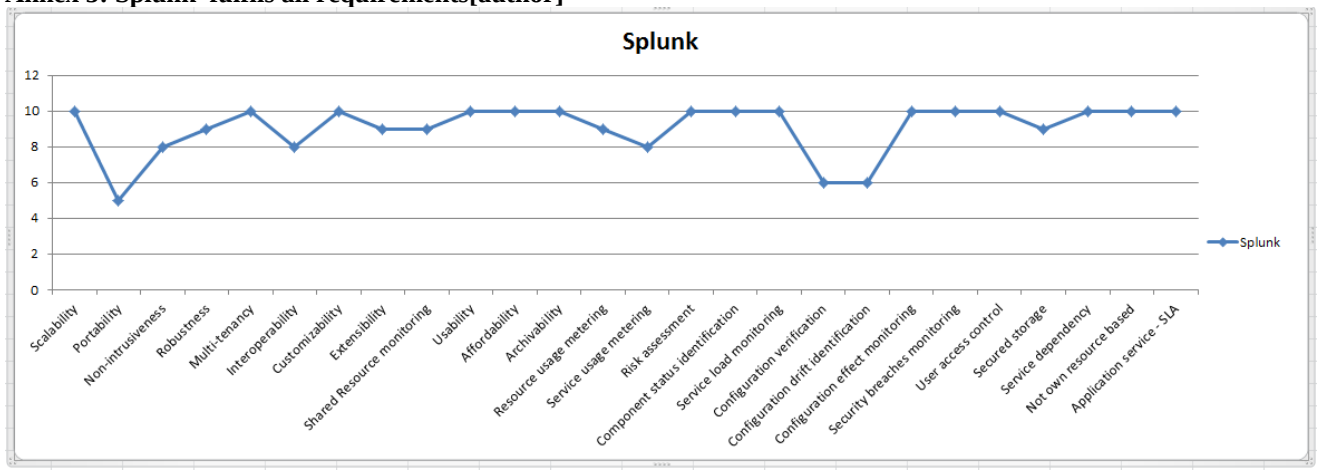


8.2 Supplementary figures- Splunk

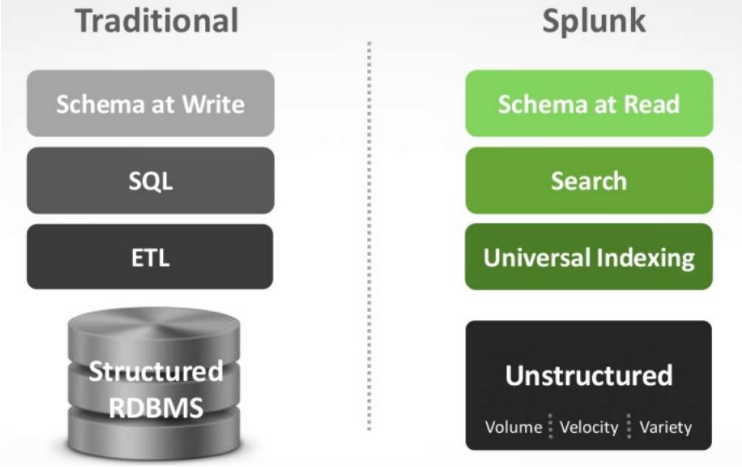
Annex 4: Cloud capabilities fulfilment graph[author]



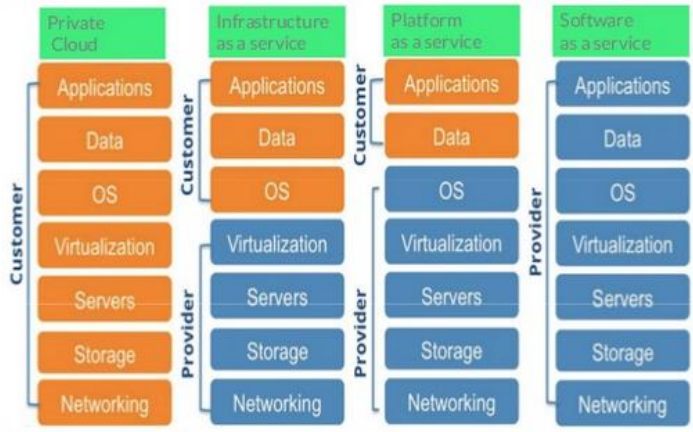
Annex 5: Splunk fulfils all requirements[author]



Annex 6: Traditional monitoring VS Splunk monitoring[34]

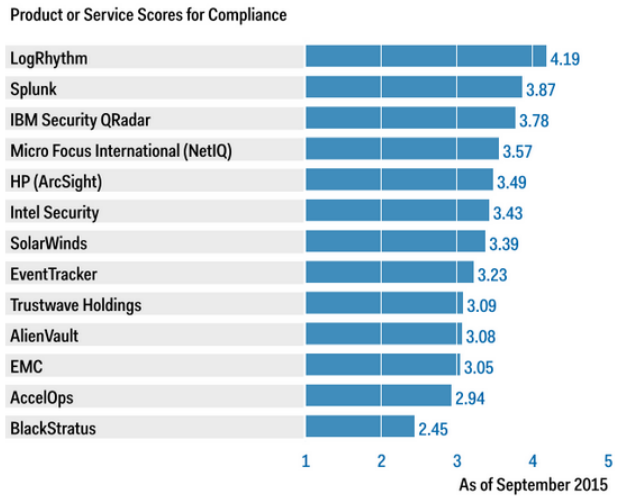


Annex 7: Cloud service models



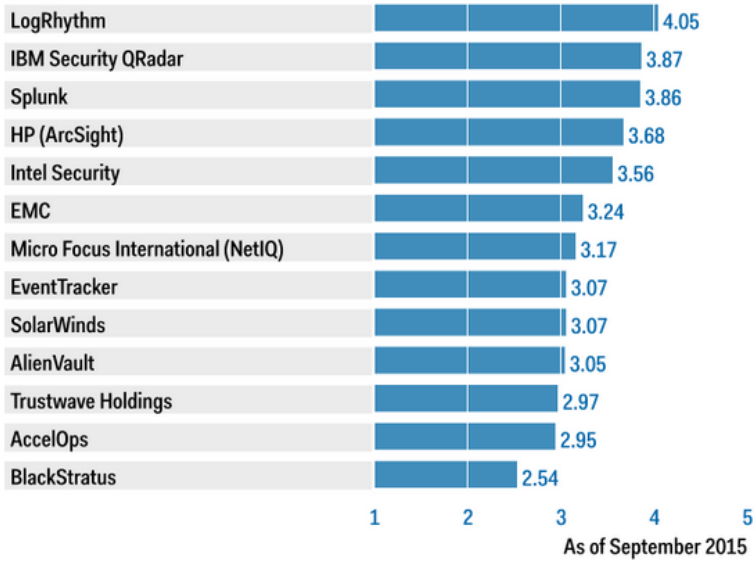
Source: [9]

Annex 8: Gartner vendor's Products Scores for the compliance use cases[37]



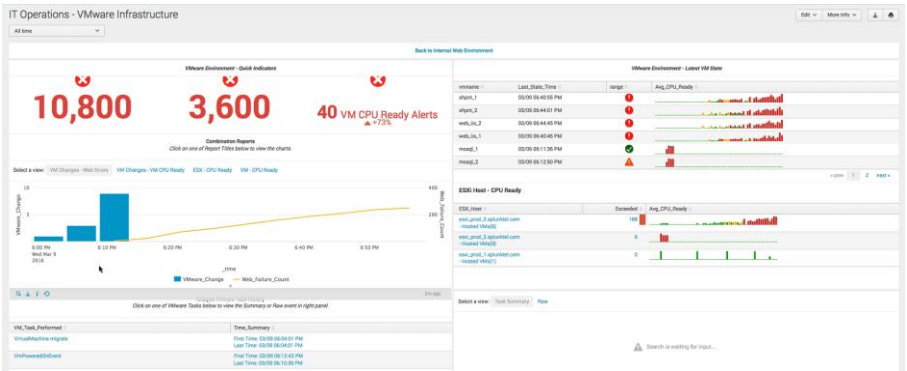
Annex 9: Vendors' Product Scores for the Threat Management Use Case[37]

Annex 10: Vendors' Product Scores for the SIEM Use Case[37]

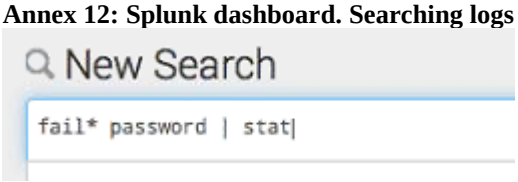


Annex 11: Splunk dashboard: IT Operations -VMware infrastructure

Source: [Author]

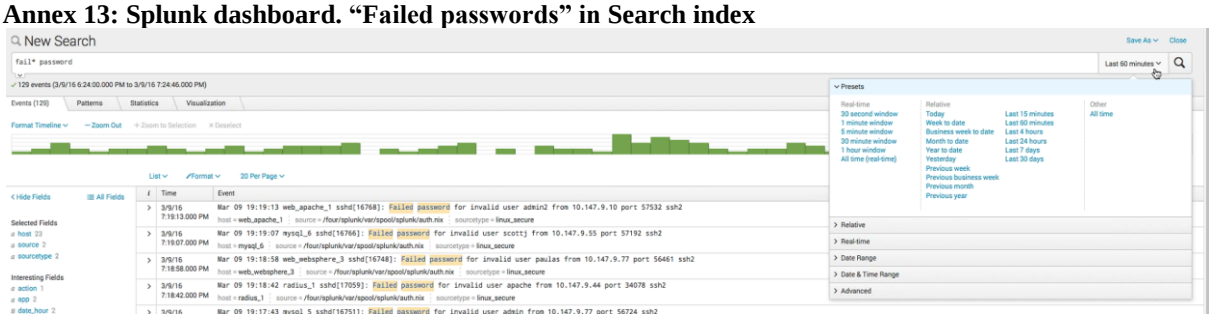


Annex 12: Splunk dashboard. Searching logs from Search field



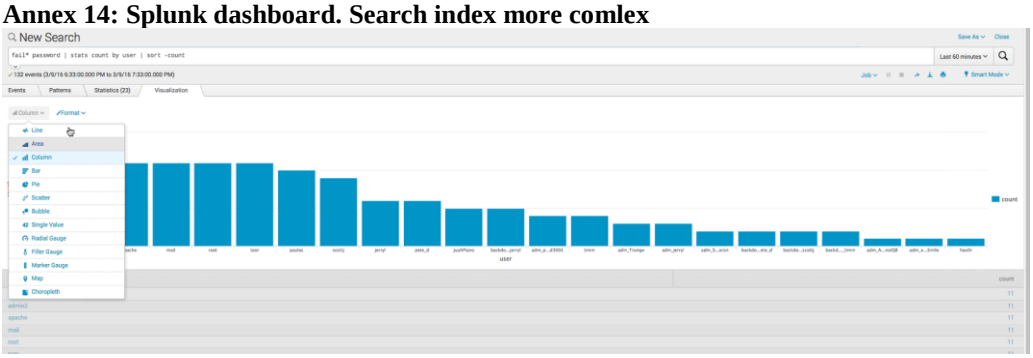
Source: [Author]

Annex 13: Splunk dashboard. "Failed passwords" in Search index

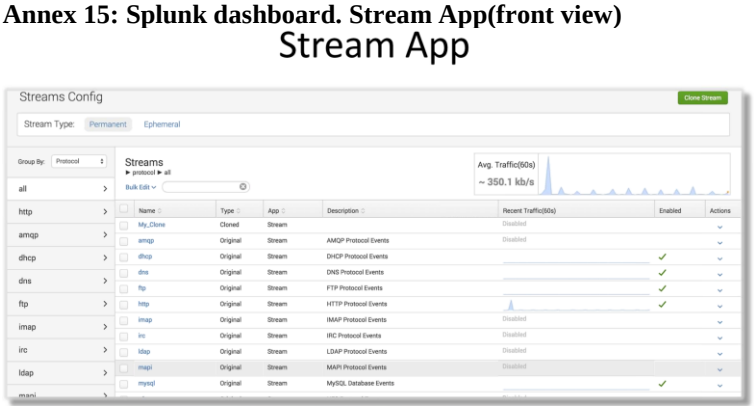


Source: [Author]

Annex 14: Splunk dashboard. Search index more comlex



Annex 15: Splunk dashboard. Stream App(front view)



Source: [Author]

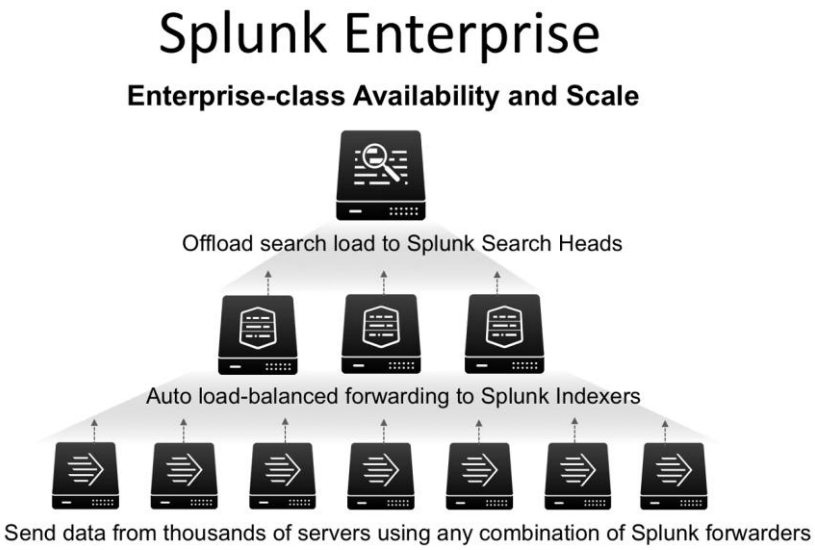
Annex 16: Splunk dashboard. Products

Source: [Author]

Annex 17: Splunk dashboard. ESA

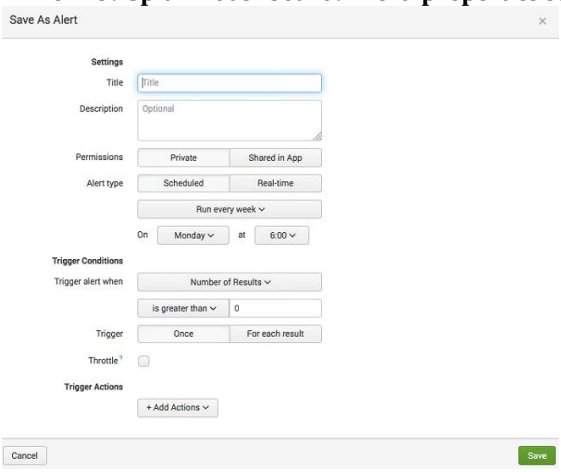
Source: [Author]

Annex 18: Splunk dashboard. SE



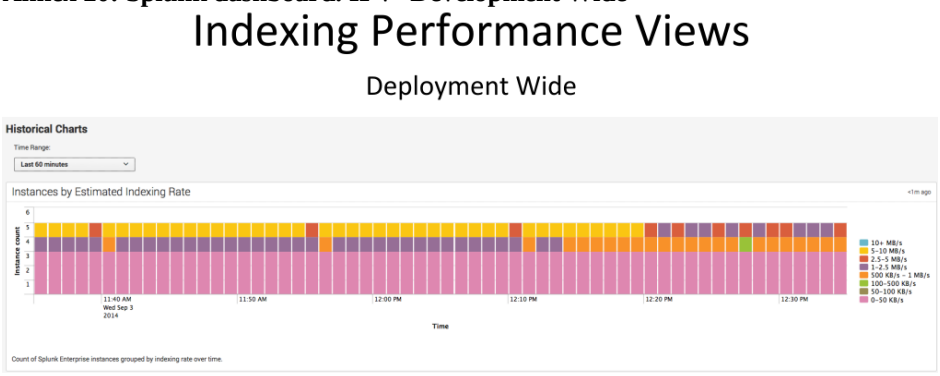
Source: [Author]

Annex 19: Splunk dashboard. Alert properties settings



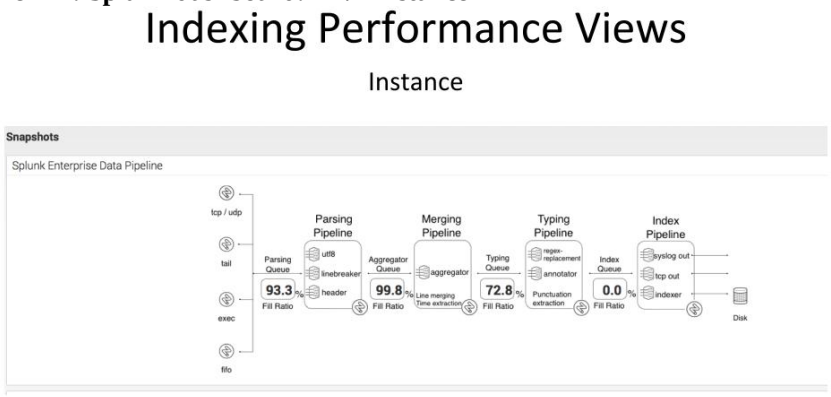
Source: [Author]

Annex 20: Splunk dashboard. IPV- Development Wide



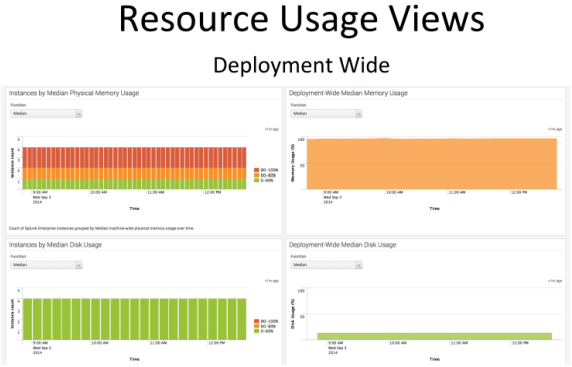
Source: [Author]

Annex 21: Splunk dashboard. IPV- Instance

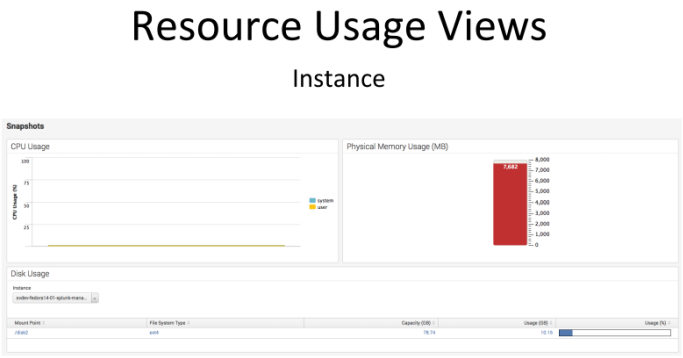


Source: [Author]

Annex 22: Splunk dashboard. RUV-Developemnt Wide

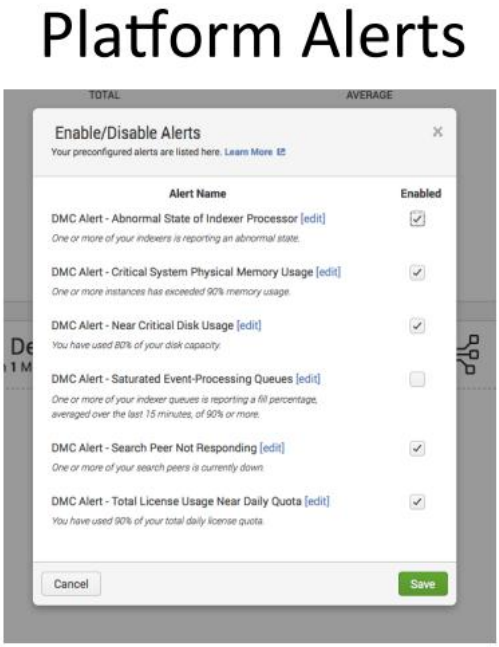


Source: [Author]



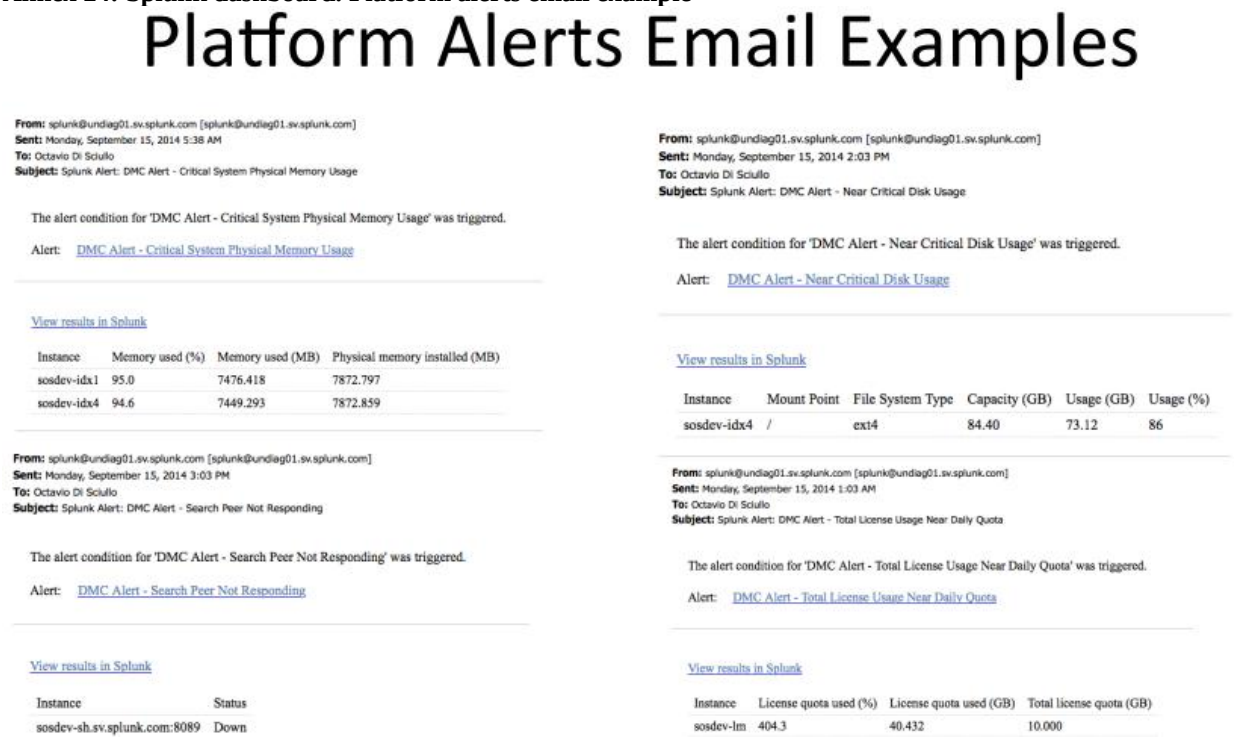
Source: [Author]

Annex 23: Splunk dashboard. Platform alerts view



Source: [Author]

Annex 24: Splunk dashboard. Platform alerts email example



Source: [Author]

8.3AzureWatch Dashboard examples

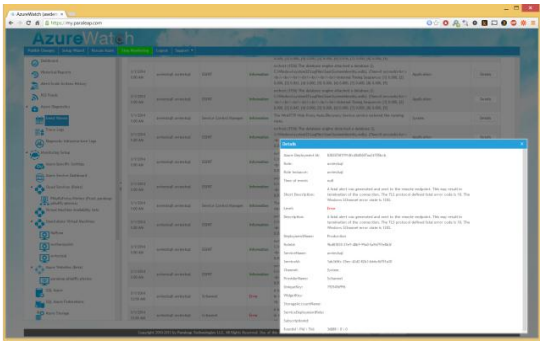
Annex 25: AzureWatch Dashboard

Source: [Author] - via my. paraleap.com

Annex 26: Azure Watch dashboard 2

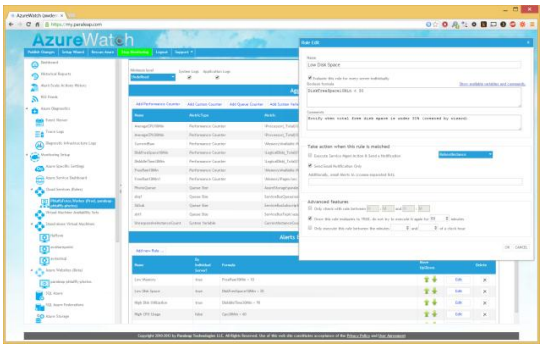
Source: [Author] - via my. paraleap.com

Annex 27: AzureWatch event logs view



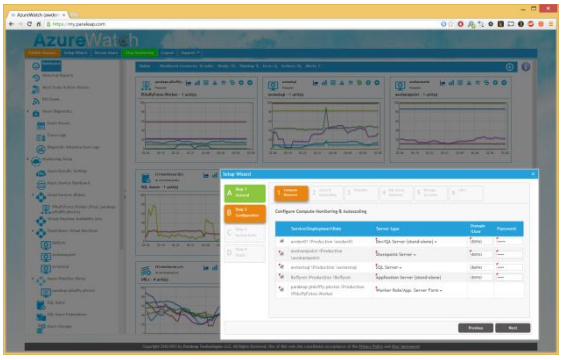
Source: [Author] - via my. paraleap.com

Annex 28: AzureWatch Monitoring - Configuration



Source : Author - via my. paraleap.com

Annex 29:AzureWatch Setup Wizard



Source: [Author] - via my. paraleap.com