
Read the Docs Template Documentation

Release 1.0

Read the Docs

Feb 02, 2020

Introduction to Cloud Technologies

1	Introduction to Cloud Computing, Containers, Microservices and DevOps	1
2	What is UCO!	3
3	UCO Technology Stack	5
4	Why Unicorn?	7
5	Platform Usage Video	9
6	Online Documentation Contents	11
6.1	Cloud Computing Fundamentals	11
6.2	Adoption of Cloud Computing paradigm	19
6.3	Software engineering, DevOps	28
6.4	Overprovisioning Cost KPI	35
6.5	Getting started with UCO Dashboard	36
6.6	Managing Cloud Resources	43
6.7	Managing Components and Applications	47
6.8	Monitoring Application and Scale	60
6.9	Managing and Using Volumes	63
6.10	Advanced Monitoring Configuration - UNICORN Monitoring Service	64
6.11	Advanced Security Configuration - Unicorn Perimeter Security (IDS configuration)	70
6.12	Advanced Monitoring Configuration - Unicorn Analytic Service	73
6.13	Advanced Scaling Configuration - Decision-Making & Auto-Scaling service	77

Introduction to Cloud Computing, Containers, Microservices and DevOps

Cloud Computing has reached virtually all areas of society and its impact on service development, production, provision and consumption is manifold and far-reaching. It lowers innovation barriers and thereby impacts industry, small and large businesses, governments and society, and offers significant benefits for everyone.

According to [Gartner Inc](#) , while cloud computing has been an application deployment and infrastructure management paradigm for many years now, the cloud market is still expanding, reaching an impressive \$200bn milestone projection for 2016 with an increasing growth rate of 16%. In this digital economy, Small and Medium Enterprises (SMEs) and today's Startups are migrating core services and products of their business to the cloud. Recent studies shows that in 2015 more than 37% of SMEs have embraced the cloud to run parts of their business, while projections show that by 2020 this number will grow and reach 80%. However, properly preparing for tomorrow's cloud challenges is crucial if one wants to unleash the full potential of the technology.

Below is a set of resources in the form of dissemination and scientific articles, implementation examples, blog entries, videos, tutorials and courses and at different levels of difficulty. The purpose of collecting this information is to give potential users of the UNICORN platform quick access to useful and high quality resources on different related topics, namely, general information on Cloud Computing, information on the challenges one must face when deciding to adopt this technology and, finally, aspects related to the agile processes of software development for Cloud Computing.

Moreover, you can read some useful information about analytic services, decision making, auto-scaling and monitoring using the UNICORN platform.

CHAPTER 2

What is UCO!

A framework that allows the design and deployment of secure and elastic by design cloud applications and services.

Unified DevOps Tool



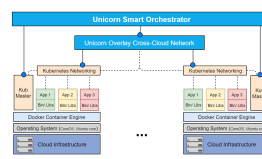
Offer a single tool for application **development**, **deployment**, and **management** during the whole application lifecycle

Monitoring & Resource Adaptation



Unicorn **elasticity library** supports apps to elastically **(de-)allocate resources** and provides **real-time monitoring** and **analytics**

True Multi-Cloud Deployments



Unicorn supports **transparent** and **automated multi-cloud deployments** for services to span across cloud zones and geographical regions

Privacy & Security Adoption



Unicorn **security** and **privacy design libraries** prevent **data breaches** and **ensure** customer **privacy**

CHAPTER 3

UCO Technology Stack

Developed on top of popular and open-source frameworks including Kubernetes, Docker, CoreOS to support multi-cloud application runtime management .. image:: 5UsageGuide/assets/arch1.png

CHAPTER 4

Why Unicorn?

- All Unicorn apps are packaged and enabled by Docker Runtime Engine to create and isolate the containerized execution environment. * Docker Runtime Engine and Docker Compose are tools sufficient for small deployments, but limited to a single host.
- Kubernetes can support the orchestration of large-scale distributed containerized deployments spanning across multiple hosts. * However Kubernetes has limitations in regard on the provisioning and deprovisioning of infrastructure resources and the auto-scaling. * Also Kubernetes cannot support cross-cloud deployments.
- Underlying containerized environment based on CoreOS which enables fast boot times and secure-out-of-the Docker runtime. * Enhanced by security service to filter network traffic and apply privacy preserving ruling.
- Unicorn Smart Orchestrator is suitable for Highly Available (HA) host management. * Taps into auto-scaling offered by cloud offerings to estimate and assess app elasticity behavior and scaling effects. * Low-cost and self-adaptive monitoring to reduce network traffic propagation.
- Unicorn Smart Orchestrator enables deployments across multiple cloud sites. * Cross-cloud network overlay is provided.
- Compatibility with Docker Compose is preserved as an extension of Docker Compose is used to describe, configure and deploy multi-container applications using YAML syntax.

CHAPTER 5

Platform Usage Video

This video will help you get to started (english subtitles are included):

The platform usage video is also available with english voiceover:

6.1 Cloud Computing Fundamentals

Selecting relevant resources in the field of Cloud Computing is on the one hand simple given the large amount of information available on the Internet, but at the same time can be a challenge if you really want to keep the best of the best. Below are the links to different very interesting resources with different levels of complexity so that the reader can adapt the reading to the knowledge they have of the different tools.

6.1.1 Magazine articles

Title: A Cloud You Can Trust

Source: <https://spectrum.ieee.org/computing/hardware/the-cloud-is-the-computer>

Author: Christian Cachin and Matthias Schunterterm

Date: November 2011

Summary: This past April, Amazon's Elastic Compute Cloud service crashed during a system upgrade, knocking customers' websites off-line for anywhere from several hours to several days. That same month, hackers broke into the Sony PlayStation Network, exposing the personal information of 77 million people around the world. And in June a software glitch at cloud-storage provider Dropbox temporarily allowed visitors to log in to any of its 25 million customers' accounts using any password—or none at all. As a company blogger drily noted: "This should never have happened."

Target audience: Startups/SMEs

Difficulty: Medium

Title: Pros and Cons of Cloud Computing Technology

Source: <https://www.ijsr.net/archive/v5i7/ART2016314.pdf>

Author: Sandeep Mukherji and Shashwat Srivastava

Date: July de 2016

Summary: Cloud computing technology today is the most enchanting technology that has become something for every business users or plans to move to. There are different types of cloud beneficial for the different users of the World, for the simple public cloud is most suitable, for the corporate world private cloud raise its hand and for the other users who are not completely corporate nor public hybrid cloud is chosen, all these three categories make all the uses complete for the world. Cloud computing is a boon for the business solution that every enterprise wants to adopt it. But having lots of pros in cloud computing, there are some cons also. This Paper is trying to put some light on those areas.

Target audience: Startups/SMEs

Difficulty: Medium

Title: The Cloud Is The Computer

Source: <https://spectrum.ieee.org/computing/hardware/the-cloud-is-the-computer>

Author: Paul McFedries

Date: Aug 2018

Summary: In the past few years we've seen Sun's slogan morph from perplexing to prophetic. As we do more and more online, we see that the network—that is, the Internet—is now an extension of our computers, to say the least. Particularly with wireless technologies, we see that a big chunk of our computing lives now sits out there in that haze of data and connections known as the cloud . In fact, we're on the verge of cloud computing , in which not just our data but even our software resides within the cloud, and we access everything not only through our PCs but also cloud-friendly devices, such as smart phones, PDAs, computing appliances, gaming consoles, even cars.

Target audience: Startups/SMEs

Difficulty: Medium

Title: Escape From the Data Center: The Promise of Peer-to-Peer Cloud Computing

Source: <https://spectrum.ieee.org/computing/networks/escape-from-the-data-center-the-promise-of-peertopeer-cloud-computing>

Author: Ozalp Babaoglu and Moreno Marzolla

Date: Sep 2014

Summary: Today, cloud computing takes place in giant server farms owned by the likes of Amazon, Google, or Microsoft—but it doesn't have to. Not long ago, any start-up hoping to create the next big thing on the Internet had to invest sizable amounts of money in computing hardware, network connectivity, real estate to house the equipment, and technical personnel to keep everything working 24/7. The inevitable delays in getting all this funded, designed, and set up could easily erase any competitive edge the company might have had at the outset.

Target audience: Startups/SMEs

Difficulty: Medium

Title: A Comprehensive Guide to Secure Cloud Computing. ISBN:0470589876 9780470589878

Source: <https://www.amazon.es/Cloud-Security-Comprehensive-Secure-Computing/dp/0470589876>

Author: Ronald L. Krutz Russell Dean Vines

Date: Sep 2014

Summary: Well-known security experts decipher the most challenging aspect of cloud computing-security Cloud computing allows for both large and small organizations to have the opportunity to use Internet-based services so that they can reduce start-up costs, lower capital expenditures, use services on a pay-as you-use basis, access applications only as needed, and quickly reduce or increase capacities. However, these benefits are accompanied by a myriad of security issues, and this valuable book tackles the most common security challenges that cloud computing faces. The authors offer you years of unparalleled expertise and knowledge as they discuss the extremely challenging topics of data ownership, privacy protections, data mobility, quality of service and service levels, bandwidth costs, data protection, and support. As the most current and complete guide to helping you find your way through a maze of security minefields, this book is mandatory reading if you are involved in any aspect of cloud computing. Coverage Includes: Cloud Computing Fundamentals Cloud Computing Architecture Cloud Computing Software Security Fundamentals Cloud Computing Risks Issues Cloud Computing Security Challenges Cloud Computing Security Architecture Cloud Computing Life Cycle Issues Useful Next Steps and Approaches).

Target audience: Startups/SMEs

Difficulty: Medium

6.1.2 Scientific articles

Title: Security Problems of Platform-as-a-Service (PaaS) Clouds and Practical Solutions to the Problems

Source: <https://www.computer.org/csdl/proceedings/srds/2012/2397/00/4784a463.pdf>

Author: Mehmet Tahir Sandıkkaya and Ali Emre Harmancı

Date: Noviembre 2012

Summary: Cloud computing is a promising approach for the efficient use of computational resources. It delivers computing as a service rather than a product for a fraction of the cost. However, security concerns prevent many individuals and organizations from using clouds despite its cost effectiveness. Resolving security problems of clouds may alleviate concerns and increase cloud usage; in consequence, it may decrease overall costs spent for the computational devices and infrastructures. This paper particularly focuses on the Platform-as-a-Service (PaaS) clouds. Security of PaaS clouds is considered from multiple perspectives including access control, privacy and service continuity while protecting both the service provider and the user. Security problems of PaaS clouds are explored and classified. Countermeasures are proposed and discussed. The achieved solutions are intended to be the rationales for future PaaS designs and implementations.

Target audience: Startups/SMEs

Difficulty: Medium

Title: Performance of Cloud Based Solutions. The Impact of Public Cloud, Private Cloud and Hybrid Cloud

Source: <https://www.grin.com/document/294956?lang=es>

Author: Mehmet Tahir Sandıkkaya and Ali Emre Harmancı

Date: Noviembre 2014

Summary: The cloud paradigm introduces a change in visualization of system and data owned by an enterprise. Further, the on service-based sharing of resources such as storage, hardware and applications which are delivered with cloud computing in a total different way has facilitated coherence of the resources

and economies of scale through its pay-per-use business model. It is no longer a collection of devices on a physical location and run a particular software program with all the needed data and resources present at a physical location but instead is a system which is geographically distributed with consideration on both application and data. But even when the development of distributed cloud architectures and services are all dealing with the same issues of scalability, elasticity over demand, broad network access, usage measurement, security aspects such as authorization and authentication, and many other concepts related to multitenant services in order to serve a high number of concurrent users over the internet, is it the main goal for companies to find the right solution for their requirements. The right solution can be a public, a private or a hybrid cloud and although the issues are very similar in any of these solutions, it depends further on the degree of potency of one or some issues which are related to different kind of industries and organisations to evaluate the right cloud based approach for a particular company. This means we have to agree that the evolution of a new paradigm requires adaptation in usage patterns and associated functional areas to fully benefit from the paradigm shift.

Target audience: Startups/SMEs

Difficulty: Medium

Title: Security Issues: Public vs Private vs Hybrid Cloud Computing

Source: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.245.1453&rep=rep1&type=pdf>

Author: R.Balasubramanian and M.Aramudhan

Date: Octobre 2012

Summary:Cloud computing appears as a new paradigm and its main objective is to provide secure, quick, convenient data storage and net computing service. Even though cloud computing effectively reduces the cost and maintenance of IT industry security issues plays a very important role. More and more IT companies are shifting to cloud based service like Private, Public and Hybrid cloud computing. But at the same time they are concerned about security issues. In this paper much attention is given to Public, Private and Hybrid cloud computing issues., as more business today utilize cloud services and architectures more threats and concerns arise. An analysis of comparative benefits of different styles of cloud computing by using SPSS is also discussed here.

Target audience: Startups/SMEs

Difficulty: Medium

Title: Survey on Microservice Architecture - Security, Privacy and Standardization on Cloud Computing Environment

Source: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.245.1453&rep=rep1&type=pdf>

Author: Washington Henrique Carvalho Almeida, Luciano de Aguiar Monteiro, Raphael Rodrigues Hazin, Anderson Cavalcanti de Lima and Felipe Silva Ferraz

Date: Octobre 2017

Summary:Microservices have been adopted as a natural solution for the replacement of monolithic systems. Some technologies and standards have been adopted for the development of microservices in the cloud environment; API and REST have been adopted on a large scale for their implementation. The purpose of the present work is to carry out a bibliographic survey on the microservice architecture focusing mainly on security, privacy and standardization aspects on cloud computing environments. This paper presents a bundle of elements that must be considered for the construction of solutions based on microservices.

Target audience: Startups/SMEs

Difficulty: Medium

Title: SMicroservices Architecture based Cloudware Deployment Platform for Service Computing

Source: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7473049>

Author: Dong Guo, Wei Wang , Guosun Zeng, Zerong Wei

Date: Octobre 2017

Summary: With the rising of Cloud computing, evolution have occurred not only in datacenter, but also in software development, deployment, maintain and usage. How to build cloud platform for traditional software, and how to deliver cloud service to users are central research fields which will have a huge impact. In recent years, the development of microservice and container technology make software paradigm evolve towards Cloudware in cloud environment. Cloudware, which is based on service and supported by cloud platform, is an important method to cloudalize traditional software. It is also a significant way for software development, deployment, maintenance and usage in future cloud environment. Furthermore, it creates a completely new thought for software in cloud platform. In this paper, we proposed a new Cloudware PaaS platform based on microservice architecture and light weighted container technology. We can directly deploy traditional software which provides services to users by browser in this platform without any modification. By utilizing the microservice architecture, this platform has the characteristics of scalability, auto-deployment, disaster recovery and elastic configuration.

Target audience: Startups/SMEs

Difficulty: Medium

6.1.3 Tutorials

Title: Cloud Computing Services Models - IaaS PaaS SaaS Explained

Source: <https://youtu.be/36zducUX16w>

Author: Ecourse Review

Date: April 2017

Summary: The Three Delivery Models: Cloud computing provides different services based on three delivery configurations. When they are arranged in a pyramid structure, they are in the order of SaaS, PaaS, and IaaS. The Three Services:

SaaS - Software as a Service

This service provides on-demand pay per use of the application software for users and is independent of a platform. You do not have to install software on your computer, unlike a license paid program. Cloud runs a single occurrence of the software, making it available for multiple end-users allowing the service to be cheap. All the computing resources that are responsible for delivering SaaS are totally managed by the vendor. The service is accessible through a web browser or lightweight client applications.

End customers use SaaS regularly. The most popular SaaS providers offer the following products and services:

Google Ecosystem including Gmail, Google Docs, Google Drive, Microsoft Office 365, and Salesforce.

PaaS - Platform as a Service

This service is mostly a development environment that is made up of a programming language execution environment, an operating system, web server, and database. It provides an environment where users can build, compile, and run their program without worrying about an hidden infrastructure. You manage the data and application resources. All

the other resources are managed by the vendor. This is the realm for developers. PaaS providers offer the following products and services:

Amazon Web services, Elastic Beanstalk, Google App Engine, Windows Azure, Heroku, and Force.com

IaaS - Infrastructure as a Service

This service provides the architecture and infrastructure. It provides all computing resources but in a virtual environment so multiple users can have access. The resources include data storage, virtualization, servers, and networking. Most vendors are responsible for managing them. If you use this service, you are responsible for handling other resources including applications, data, runtime, and middleware. This is mostly for SysAdmins. IaaS providers offer the following products and services:

Target audience: Startups/SMEs

Difficulty: Medium

Title: Introduction to Cloud | Cloud Computing Tutorial for Beginners

Source: <https://youtu.be/usYySG1nbfl>

Author: Edureka!

Date: April 2017

Summary: This Edureka video on “Introduction To Cloud” will introduce you to basics of cloud computing and talk about different types of Cloud provides and its Service models. Following is the list of content covered in this session:

1. What is Cloud?
2. Uses of Cloud
3. Service Models
4. Deployment Models
5. Cloud Providers
6. Cloud Demo - AWS, Google Cloud, Azure

Target audience: Startups/SMEs

Difficulty: Medium

Title: How It Works: Cloud Microservices

Source: <https://youtu.be/Pvbr5d2mIZs>

Author: IBM Think Academy

Date: October 2019

Summary: Microservices are an important piece of the new approach to cloud — many tiny pieces, in fact. But how do they all work together? Find out in this video from Think Academy. For more information on IBM Cloud, please visit: <http://www.ibm.com/cloud>

Target audience: Startups/SMEs

Difficulty: Medium

Title: Difference Between APIs, Services and Microservices?

Source: <https://youtu.be/qGFRbOq4fmQ>

Author: IBM Cloud

Date: April 2017

Summary: What's the difference between APIs, services, and microservices? Watch now and learn more here: ibm.co/2o1paz1

Target audience: Startups/SMEs

Difficulty: Medium

Title: Learn All About Microservices — Microservices Architecture Example:

Source: <https://dzone.com/articles/microservices-tutorial-learn-all-about-microservic>

Author: DZone

Date: May 2018

Summary: This microservices tutorial is the third part in this microservices blog series. I hope that you have read my previous blog, What is Microservices, which explains the architecture, compares microservices with monoliths and SOA, and explores when to use microservices with the help of use cases.

Target audience: Startups/SMEs

Difficulty: Medium

6.1.4 Courses

Title: Introduction cloud computing

Source: <https://www.edx.org/es/course/introduction-cloud-computing-microsoft-cloud200x>

Author: Microsoft

Date: 2018

Summary: Cloud computing, or “the cloud”, has gone from a leading trend in IT to mainstream consciousness and wide adoption.

This self-paced course introduces cloud computing concepts where you'll explore the basics of cloud services and cloud deployment models.

You'll become acquainted with commonly used industry terms, typical business scenarios and applications for the cloud, and benefits and limitations inherent in the new paradigm that is the cloud.

This course will help prepare you for more advanced courses in Windows Server-based cloud and datacenter administration.

Target audience: Startups/SMEs

Difficulty: Medium

6.1.5 Blogs

Title: Blog Cloud Computing David Linthicum

Source: <https://www.infoworld.com/blog/cloud-computing/>

Author: David Linthicum, Cloud Computing

Date: 2015

Summary: David Linthicum, the CTO and founder of Blue Mountain Labs, is widely recognized as a thought leader in the cloud computing industry – and with good reason. He travels to deliver keynote speeches on cloud, has contributed to or authored 13 books, and writes the Cloud Computing blog. With provocative titles such as “Shocker: Government agency drafts sensible cloud computing strategy,” “Wozniak is wrong about cloud computing” and “Wake up, IT: Even CFOs see value in the cloud,” Linthicum isn’t afraid to state his opinion – and with his extensive cloud background, everyone should take note. Check out Linthicum’s monthly column on SearchCloudComputing.com, in which he discusses topics that range from the cloud API wars to cloud portability and interoperability struggles.

Target audience: Startups/SMEs

Difficulty: Medium

Title: Cloud Tech

Source: <https://www.cloudcomputing-news.net>

Author: Cloud Tech

Date: 2016

Summary: CloudTech is a leading blog and news site that is dedicated to cloud computing strategy and technology. With authors including IBM’s Sebastian Krause, Clouconomics author Joe Weinman, and Ian Moysse from the Cloud Industry Forum, CloudTech has hundreds of blogs about numerous cloud-related topics and reaches over 320,000 cloud computing professionals. A recent post How the Financial Services Industry Is Slowly Waking Up to Cloud Computing by Rahul Singh of HCL Technologies provides an interesting analysis of how banks can overcome the barriers to cloud migration.

Target audience: Startups/SMEs

Difficulty: Medium

Title: The best page with the 25 best blogs in the world of cloud computing

Source: <https://www.cloudendure.com/blog/top-25-must-read-cloud-computing-blogs/>

Author: Cloudendure

Date: 2018

Summary: In this page you will find the 25 most relevant blogs in the world about cloud computing.

Target audience: Startups/SMEs

Difficulty: Medium

Title: Top five must-read cloud computing blogs

Source: <https://www.cloudendure.com/blog/top-25-must-read-cloud-computing-blogs/>

Author: Werner Vogels, All Things Distributed, Randy Bias, Cloudscaling ,Adrian Cockcroft, Adrian Cockcroft's, Lori MacVittie, Two Different Socks, David Linthicum, Cloud Computing

Date: 2015

Summary: Werner Vogels, All Things Distributed: These top bloggers are in no particular order, but kicking off the list with the CTO of what's arguably cloud computing's king vendor, Amazon Web Services, should come as no surprise. Werner Vogels' blog, All Things Distributed, is not strictly about cloud computing, but it does give insight into AWS' strategy and vision, which should be of interest to any cloud pro. With some personality, regular updates and blogs jam-packed with information, Vogels proves he's one to beat in the blogging arena. Randy Bias, Cloudscaling blog: Randy Bias has a lengthy resume. Featured in publications as prestigious as Forbes, The Economist and The Wall Street Journal, Bias has a lot to say about cloud computing. This founding CEO and CTO of Cloudscaling also regularly contributes, along with other cloud experts and pioneers, to Cloudscaling's blog. With posts on Amazon's secret to success, the open source cloud market and the rise of DevOps, Bias focuses on what he calls "the profound disruption" cloud computing causes. If anyone can help IT navigate this disrupted landscape, Bias is at the top of the list. Adrian Cockcroft, Adrian Cockcroft's blog: Though simply titled, Adrian Cockcroft's blog has been covering the complex world of cloud architectures and performance tools – as well as his interest in cars and music – since 2004. He may not update the blog as often as Bias or Vogels, but his monthly (or so) insights are always worth a read. As a cloud architect at Netflix, Adrian Cockcroft is keen on discussing how his company handles the cloud computing evolution and how that compares to other industry leaders and trends, tackling subjects like Platform as a Service's (PaaS) place in the market and the use of popular AWS products. Lori MacVittie, Two Different Socks: Recognized as one of the Top Women in Cloud by CloudNOW, a nonprofit consortium featuring female leaders in cloud computing, Lori MacVittie is already well-known as a cloud leader. MacVittie's blog, Two Different Socks, delves into cloud performance struggles, DevOps and other vital issues in the cloud industry, with the help of charts, infographics and fun images. As the senior technical marketing manager at F5 Networks, MacVittie still had time to co-author a book called The Cloud Security Rules. David Linthicum, Cloud Computing: David Linthicum, the CTO and founder of Blue Mountain Labs, is widely recognized as a thought leader in the cloud computing industry – and with good reason. He travels to deliver keynote speeches on cloud, has contributed to or authored 13 books, and writes the Cloud Computing blog. With provocative titles such as "Shocker: Government agency drafts sensible cloud computing strategy," "Wozniak is wrong about cloud computing" and "Wake up, IT: Even CFOs see value in the cloud," Linthicum isn't afraid to state his opinion – and with his extensive cloud background, everyone should take note. Check out Linthicum's monthly column on SearchCloudComputing.com, in which he discusses topics that range from the cloud API wars to cloud portability and interoperability struggles.

Target audience: Startups/SMEs

Difficulty: Medium

6.2 Adoption of Cloud Computing paradigm

Potential adopters of Cloud Computing technology perceive the adoption decision as very challenging. Cloud Computing adoption should be driven by change management, competence, maturity, among others. Cloud Computing is still in its stage of emergence, and there is still a lack of both knowledge and empirical evidence about which issues are the most significant for Cloud Computing adoption decisions. The following resources have been selected to give some light to these issues.

6.2.1 Magazine articles

Title: What is Docker and why is it so darn popular?

Source: <https://www.zdnet.com/article/what-is-docker-and-why-is-it-so-darn-popular/>

Author: Steven J. Vaughan-Nichols

Date: Nov 2016

Summary: Docker is hotter than hot because it makes it possible to get far more apps running on the same old servers and it also makes it very easy to package and ship programs. Here's what you need to know about it.

Target audience: Startups/SMEs

Difficulty: Medium

Title: A Beginner's Guide to Kubernetes:

Source: <https://medium.com/containermind/a-beginners-guide-to-kubernetes-7e8ca56420b6>

Author: Imesh Gunaratne

Date: Aug 2017

Summary: Kubernetes has now become the de facto standard for deploying containerized applications at scale in private, public and hybrid cloud environments. The largest public cloud platforms AWS, Google Cloud, Azure, IBM Cloud and Oracle Cloud now provide managed services for Kubernetes. A few years back RedHat completely replaced their OpenShift implementation with Kubernetes and collaborated with the Kubernetes community for implementing the next generation container platform. Mesosphere incorporated key features of Kubernetes such as container grouping, overlay networking, layer 4 routing, secrets, etc into their container platform DC/OS soon after Kubernetes got popular. DC/OS also integrated Kubernetes as a container orchestrator alongside Marathon. Pivotal recently introduced Pivotal Container Service (PKS) based on Kubernetes for deploying third-party services on Pivotal Cloud Foundry and as of today there are many other organizations and technology providers adapting it at a rapid phase.

Target audience: Startups/SMEs

Difficulty: Medium

Title: What is Kubernetes? Container orchestration explained

Source: <https://www.infoworld.com/article/3268073/kubernetes/what-is-kubernetes-container-orchestration-explained.html>

Author: Serdar Yegulalp

Date: Aug 2017

Summary: The rise of containers has reshaped the way people think about developing, deploying, and maintaining software. Drawing on the native isolation capabilities of modern operating systems, containers support VM-like separation of concerns, but with far less overhead and far greater flexibility of deployment than hypervisor-based virtual machines.

Target audience: Startups/SMEs

Difficulty: Medium

6.2.2 Scientific articles

Title: Model-Driven Management of Docker Containers

Source: <https://ieeexplore.ieee.org/document/7820337/>

Author: Fawaz Paraiso ; Stéphanie Challita ; Yahya Al-Dhuraibi ; Philippe Merle

Date: Aug 2017

Summary: With the emergence of Docker, it becomes easier to encapsulate applications and their dependencies into lightweight Linux containers and make them available to the world by deploying them in the cloud. Compared to hypervisor-based virtualization approaches, the use of containers provides faster start-ups times and reduces the consumption of computer resources. However, Docker lacks of deployability verification tool for containers at design time. Currently, the only way to be sure that the designed containers will execute well is to test them in a running system. If errors occur, a correction is made but this operation can be repeated several times before the deployment becomes operational. Docker does not provide a solution to increase or decrease the size of container resources in demand. Besides the deployment of containers, Docker lacks of synchronization between the designed containers and those deployed. Moreover, container management with Docker is done at low level, and therefore requires users to focus on low level system issues. In this paper we focus on these issues related to the management of Docker containers. In particular, we propose an approach for modeling Docker containers. We provide tooling to ensure the deployability and the management of Docker containers. We illustrate our proposal using an event processing application and show how our solution provides a significantly better compromise between performance and development costs than the basic Docker container solution.

Target audience: Startups/SMEs

Difficulty: Medium

Title: Docker Reference Architecture: Designing Scalable, Portable Docker Container Networks

Source: <https://success.docker.com/article/networking>

Author: Docker

Date: Aug 2018

Summary: Docker containers wrap a piece of software in a complete filesystem that contains everything needed to run: code, runtime, system tools, system libraries – anything that can be installed on a server. This guarantees that the software will always run the same, regardless of its environment. By default, containers isolate applications from one another and the underlying infrastructure, while providing an added layer of protection for the application.

What if the applications need to communicate with each other, the host, or an external network? How do you design a network to allow for proper connectivity while maintaining application portability, service discovery, load balancing, security, performance, and scalability? This document addresses these network design challenges as well as the tools available and common deployment patterns. It does not specify or recommend physical network design but provides options for how to design Docker networks while considering the constraints of the application and the physical network.

Target audience: Startups/SMEs

Difficulty: Medium

Title: Docker Commands—The Ultimate Cheat Sheet

Source: <https://hackernoon.com/docker-commands-the-ultimate-cheat-sheet-994ac78e2888>

Author: Nick Parsons

Date: Aug 2017

Summary: If you don't already know, Docker is an open-source platform for building distributed software using "containerization," which packages applications together with their environments to make them more portable and easier to deploy. Thanks to its power and productivity, Docker has become an incredibly popular technology for software development teams. However, this same power can sometimes make it tricky for new users to jump into the Docker ecosystem, or even for experienced users to remember the right command. Fortunately, with the right learning tools you'll have a much easier time getting started with Docker. This article will be your one-stop shop for Docker, going over some of the best practices and must-know commands that any user should know.

Target audience: Startups/SMEs

Difficulty: Medium

Title: Features Eclipse Che

Source: <https://www.eclipse.org/che/features/>

Author: Eclipse Che

Date: October 2017

Summary: Eclipse Che is a developer workspace server and cloud IDE built for teams and organizations

Target audience: Startups/SMEs

Difficulty: Medium

Title: Understanding cloud computing adoption issues: A Delphi study approach

Source: <https://www.sciencedirect.com/science/article/pii/S016412121630036X>

Author: Science Direct

Date: Aug 2016

Summary: This research paper reports on a Delphi study focusing on the most important issues enterprises are confronted with when making cloud computing (CC) adoption decisions. We had 34 experts from different domain backgrounds participated in a Delphi panel. The panelists were IT and CC specialists representing a heterogeneous group of clients, providers and academics, divided into three subpanels. The Delphi procedure comprised three stages: brainstorming, narrowing down and ranking. The panelists identified 55 issues of concerns in the first stage, which were analyzed and grouped into 10 categories: security, strategy, legal and ethical, IT governance, migration, culture, business, awareness, availability and impact. The top 18 issues for each subpanel were ranked, and a moderate intrapanel consensus was obtained. Additionally, 16 follow-up interviews were conducted with the experts to get a deeper understanding of the issues and why certain issues were more significant than others. The findings indicate that security, strategy and legal and ethical issues are the most important. The discussion resulted in highlighting certain inhibitors and drivers for CC adoption into a framework. The paper is concluded with key recommendations with focus on change management, competence and maturity to inform decision-makers in CC adoption decisions.

Target audience: Startups/SMEs

Difficulty: Medium

6.2.3 Tutorials

Title: 47 advanced tutorials for mastering Kubernetes:

Source: <https://techbeacon.com/top-tutorials-mastering-kubernetes>

Author: Tech Beacon

Date: Oct 2017

Summary: Best tutorials for mastering Kubernetes

Target audience: Startups/SMEs

Difficulty: Medium

Title: Top Tutorials To Learn Kubernetes

Source: <https://medium.com/quick-code/top-tutorials-to-learn-kubernetes-e9507e76d9a4>

Author: Medium Blog

Date: Nov 2017

Summary: Learn Kubernetes

Target audience: Startups/SMEs

Difficulty: Medium

Title: Kubernetes cluster architecture:

Source: <https://www.coursera.org/lecture/deploy-micro-kube-ibm-cloud/kubernetes-cluster-architecture-9DY7Z>

Author: Coursera, IBM

Date: April 2017

Summary: First Kubernetes cluster architecture

Target audience: Startups/SMEs

Difficulty: Medium

Title: Create first images in docker

Source: <https://docs.docker.com/develop/develop-images/baseimages/>

Author: Docker

Date: Aug 2016

Summary: Create images in docker

Target audience: Startups/SMEs

Difficulty: Medium

Title: Docker Series—Building your first image:

Source: <https://medium.com/pintail-labs/docker-series-building-your-first-image-8a6f051ae637>

Author: Docker

Date: Nov 2016

Summary: The Docker image we are going to build is the image used to start the pintail-whoami container we used in the article Docker Series—Starting your first container.

Target audience: Startups/SMEs

Difficulty: Medium

Title: Network Configuration

Source: <https://docs.docker.com/v1.5/articles/networking/>

Author: Docker

Date: Aug 2016

Summary: When Docker starts, it creates a virtual interface named docker0 on the host machine. It randomly chooses an address and subnet from the private range defined by RFC 1918 that are not in use on the host machine, and assigns it to docker0. Docker made the choice 172.17.42.1/16 when I started it a few minutes ago, for example — a 16-bit netmask providing 65,534 addresses for the host machine and its containers. The MAC address is generated using the IP address allocated to the container to avoid ARP collisions, using a range from 02:42:ac:11:00:00 to 02:42:ac:11:ff:ff...

Target audience: Startups/SMEs

Difficulty: Medium

6.2.4 Courses

Title: DOCKER FUNDAMENTALS + ENTERPRISE DEVELOPERS COURSE BUNDLE

Source: <https://training.docker.com/instructor-led-training/docker-fundamentals-enterprise-developers>

Author: Docker

Date: 2017

Summary: As the follow-on to the Docker Fundamentals course, Docker for Enterprise Operations is a role-based course designed for an organization's Development and DevOps teams to accelerate their Docker journey in the enterprise. The course covers in-depth core advanced features of Docker EE and best practices to apply these features at scale with enterprise workloads. It is highly recommended to complete the Docker Fundamentals course as a pre-requisite.

Target audience: Startups/SMEs

Difficulty: Medium

Title: General docker courses

Source: <https://training.docker.com/instructor-led-training>

Author: Docker

Date: 2017

Summary: Docker Fundamentals: (2 days, \$1495) - <https://success.docker.com/training/courses/docker-fundamentals>

This is THE introductory Docker course to give your team the best foundation for enterprise grade Docker use-cases.

Docker for Enterprise Developers: (2 days, \$1995) - <https://success.docker.com/training/courses/docker-for-enterprise-developers>

As the follow-on to our Docker Fundamentals course, Docker for Enterprise Developers is a role-based course designed for an organization's Development and DevOps teams to accelerate their Docker journey in the enterprise.

Docker for Enterprise Operations:(2 Days, \$1995) - <https://success.docker.com/trainingcourses/docker-for-enterprise-operations>

This course is the second level of Docker's core curriculum for the enterprise and is focused on the Docker Operator role in administration of Docker Enterprise Edition Advanced.

Docker Troubleshooting & Support : (2 Days, \$1995) - <https://success.docker.com/training/courses/docker-support-troubleshooting>

The Docker Troubleshooting & Support course is a role-based course designed for an organization's support teams to troubleshoot the variety of issues that arise in their Docker journey.

Docker Security: (1 Day, \$995) - <https://success.docker.com/training/courses/docker-security>

This is the Docker Security course for your entire organization. Get everyone "on-the-same page" and working together to secure your Dockerized environment. This hands-on workshop style course will give your team an overview of important security features and best practices to protect your containerized services.

Target audience: Startups/SMEs

Difficulty: Medium

6.2.5 Blogs

Title: 5 TIPS TO LEARN DOCKER IN 2018

Source: <https://blog.docker.com/2018/01/5-tips-learn-docker-2018/>

Author: Victor Coisne

Date: January, 2018

Summary: 5 TIPS TO LEARN DOCKER IN 2018

Target audience: Startups/SMEs

Difficulty: Medium

Title: Play With Docker: the Docker Playground and Training site

Source: <https://training.play-with-docker.com>

Author: Victor Coisne

Date: 2018

Summary: Play with Docker (PWD) is a Docker playground and training site which allows users to run Docker commands in a matter of seconds. It gives the experience of having a free Linux Virtual Machine in browser, where you can build and run Docker containers and even create clusters. Check out this video from DockerCon 2017 to learn more about this project. The training site is composed of a large set of Docker labs and quizzes from beginner to advanced level available for both Developers and IT pros at training.play-with-docker.com.

Target audience: Startups/SMEs

Difficulty: Medium

Title: DockerCon 2018

Source: <https://2018.dockercon.com>

Author: DockerCon

Date: 2018

Summary: In case you missed it, DockerCon 2018 will take place at Moscone Center, San Francisco, CA on June 13-15, 2018. DockerCon is where the container community comes to learn, belong, and collaborate. Attendees are a mix of beginner, intermediate, and advanced users who are all looking to level up their skills and go home inspired. With a 2 full days of training, more than 100 sessions, free workshops and hands-on labs and the wealth of experience brought by each attendee, DockerCon is the place to be if you're looking to learn Docker in 2018.

Target audience: Startups/SMEs

Difficulty: Medium

Title: Docker Meetups

Source: https://www.meetup.com/es-ES/Docker-Online-Meetup/?chapter_analytics_code=UA-48368587-1

Author: Docker Meetups

Date: 2018

Summary: Look at our Docker Meetup Chapters page to see if there is a Docker user group in your city. With more than 200 local chapters in 81 countries, you should be able to find one near you! Attending local Docker meetups are an excellent way to learn Docker. The community leaders who run the user group often schedule Docker 101 talks and hands-on training for newcomers!

Can't find a chapter near you? Join the Docker Online meetup group to attend meetups remotely!

Target audience: Startups/SMEs

Difficulty: Medium

Title: Docker Captains

Source: <https://www.docker.com/community/captains>

Author: Docker Captains

Date: 2018

Summary: Captains are Docker experts that are leaders in their communities, organizations or ecosystems. As Docker advocates, they are committed to sharing their knowledge and do so every chance they get! Captains are advisors, ambassadors, coders, contributors, creators, tool builders, speakers, mentors, maintainers and super users and are required to be active stewards of Docker in order to remain in the program.

Follow all of the Captains on twitter. Also check out the Captains GitHub repo to see what projects they have been working on. Docker Captains are eager to bring their technical expertise to new audiences both offline and online around the world – don't hesitate to reach out to them via the social links on their Captain profile pages. You can filter the captains by location, expertise, and more.

Target audience: Startups/SMEs

Difficulty: Medium

Title: Training and Certification

Source: <https://europe-2017.dockercon.com>

Author: DockerCon

Date: 2018

Summary: The new Docker Certified Associate (DCA) certification, launching at DockerCon Europe on October 16, 2017, serves as a foundational benchmark for real-world container technology expertise with Docker Enterprise Edition. In today's job market, container technology skills are highly sought after and this certification sets the bar for well-qualified professionals. The professionals that earn the certification will set themselves apart as uniquely qualified to run enterprise workloads at scale with Docker Enterprise Edition and be able to display the certification logo on resumes and social media profiles.

Target audience: Startups/SMEs

Difficulty: Medium

Title: Kubernetes Oficial Blog

Source: <https://kubernetes.io/blog/>

Author: Kubernetes

Date: 2017

Summary: Kubernetes Blog

Target audience: Startups/SMEs

Difficulty: Medium

Title: Blog Official Eclipse Che

Source: <https://che.eclipse.org/what-is-eclipse-che-815d110f64e5>

Author: Eclipse Che

Date: 2016

Summary: Eclipse Che Blog

Target audience: Startups/SMEs

Difficulty: Medium

6.3 Software engineering, DevOps

DevOps and the cloud are closely related. The vast majority of cloud development projects use DevOps, and the list will get longer and longer. The benefits of using DevOps with cloud projects are also being better defined. The following resources provide relevant information on the methods of software development process related to DevOps and particularly for the case of Cloud Computing.

6.3.1 Magazine articles

Title: Dos and don'ts: 9 effective best practices for DevOps in the cloud

Source: <https://techbeacon.com/dos-donts-9-effective-best-practices-devops-cloud>

Author: David Linthicum

Date: May 2016

Summary: DevOps and cloud computing are joined at the hip. Why? DevOps is about streamlining development so user requirements can quickly make it into application production, while the cloud offers automated provisioning and scaling to accommodate application changes.

Target audience: Startups/SMEs

Difficulty: Medium

Title: The Relationship Between The Cloud And DevOps

Source: <https://www.forbes.com/sites/forbestechcouncil/2017/07/21/the-relationship-between-the-cloud-and-devops/#4207c8822957>

Author: Aater Suleman

Date: 2017

Summary: Most companies understand that if they want to increase their competitiveness in today's swiftly changing world, they can't ignore digital transformation. DevOps and cloud computing are oft-touted as vital ways companies can achieve this needed transformation, though the relationship between the two is often confusing, as DevOps is about process and process improvement whereas cloud computing is about technology and services. Not mutually exclusive, it's important to understand how the cloud and DevOps work together to help businesses achieve their transformation goals.

Target audience: Startups/SMEs

Difficulty: Medium

6.3.2 Scientific articles

Title: Key Factors Impacting Cloud Computing Adoption

Source: DOI: 10.1109/MC.2013.362 IEEE Computer (Volume: 46, Issue: 10, October 2013)

Author: Lorraine Morgan and Kieran Conboy

Date: October 2013

Summary: Findings from multiple cloud provider case studies identify 9 key factors and 15 subfactors affecting the adoption of cloud technology. These range from technological issues to broader organizational and environmental issues.

Target audience: Startups/SMEs wanting to know the technological, organizational, and environmental factors that impact cloud services adoption.

Difficulty: Low

Title: Identification of SME Readiness to Implement Cloud Computing

Source: DOI: 10.1109/ICCCSN.2012.6215757 / 2012 International Conference on Cloud Computing and Social Networking (ICCCSN)

Author: Kridanto Surendro & Adiska Fardani

Date: June 2012

Summary: Cloud Computing allows the use of information technology based on the on-demand utility. This technology can provide benefits to small and medium enterprises with limited capital, human resources, and access to marketing network. A survey conducted on SMEs in the district of Coblong Bandung to dig up the IT needs and analyze their readiness to adopt cloud computing technologies. The survey results stated that SMEs' respondents are more suitable to implement Software as a Service with public cloud deployment method. SMEs are ready to implement this technology, but requires appropriate training and role models that can be used as an example because their technology adoption characteristics that are late majority.

Target audience: Startups/SMEs

Difficulty: Medium

Title: Big Data Drives Cloud Adoption in Enterprise

Source: DOI: 10.1109/MIC.2013.63 IEEE Internet Computing (Volume: 17, Issue: 4, July-Aug. 2013)

Author: Huan Liu (Accenture Technology Labs)

Date: June 2013

Summary: The need to store, process, and analyze large amounts of data is finally driving enterprise customers to adopt cloud computing at scale. Understanding the economic drivers behind enterprise customers is key to designing next generation cloud services.

Target audience: Startups/SMEs wanting to know important drivers to cloud adoption beyond cost reduction.

Difficulty: Low

Title: Elements of Cloud Adoption

Source: DOI: 10.1109/MCC.2014.7 IEEE Cloud Computing (Volume: 1, Issue: 1, May 2014)

Author: Samee U. Khan (North Dakota State University)

Date: May 2014

Summary: Sharing experiences in transitioning from traditional computing paradigms to the cloud can provide a blueprint for organizations to gauge the depth and breadth of cloud-enabled technologies.

Target audience: Startups/SMEs.

Difficulty: Low

6.3.3 Tutorials

Title: What is DevOps? | DevOps Training - DevOps Introduction & Tools | DevOps Tutorial

Source: <https://youtu.be/3EyT1i0wYUY>

Author: edureka!

Date: 2016

Summary: This DevOps tutorial takes you through what is DevOps all about and basic concepts of DevOps and DevOps Tools. This DevOps tutorial is ideal for beginners to get started with DevOps. Check our complete DevOps playlist here: <http://goo.gl/O2vo13>

DevOps Tutorial Blog Series: <https://goo.gl/P0zAf>

Target audience: Startups/SMEs

Difficulty: Medium

Title: Cloud Computing Adoption: Management Challenges and Opportunities

Source: https://www.youtube.com/watch?v=gR_psSPqAnU

Author: Julian Kudritzki and Andrew Reichman

Date: Sept 2016

Summary: Industry data from Uptime Institute and 451 Research evidence a rapid rate of cloud computing adoption for enterprise IT departments. Organizations weigh cloud benefits and risks, and also evaluate how cloud will impact their existing and future data center infrastructure investment. In this video, Uptime Institute COO Julian Kudritzki and Andrew Reichman, Research Director at 451 Research discuss the various aspects of how much risk, and how much reward, is on the table for companies considering a cloud transition.

Target audience: Startups/SMEs

Difficulty: Medium

Title: On-Premises to Cloud: AWS Migration in 5 Super Easy Steps

Source: <https://serverguy.com/cloud/aws-migration/>

Author: Sahil Chugh

Date: May 2018

Summary: Five steps describing how to migrate applications and data to the AWS Cloud and detailing several strategies to do that. Wonderful article with lot of detailed information and examples.

Target audience: Startups/SMEs

Difficulty: Medium

Title: On-Premises to Cloud: AWS Migration in 5 Super Easy Steps

Source: <https://github.com/google/go-cloud/tree/master/samples/tutorial>

Author: @zombiezen and @vangent, users of GitHub

Date: Jul 2018

Summary: With the premise that the best way to understand Go Cloud is to write some code and use it, the authors details how to build a command line application that uploads files to blob storage on both AWS and GCP. Blob storage stores binary data under a string key, and is one of the most frequently used cloud services.

Target audience: Startups/SMEs

Difficulty: Low

Title: Google Cloud tutorial: Get started with Google Cloud

Source: <https://www.infoworld.com/article/3267929/cloud-computing/google-cloud-tutorial-get-started-with-google-cloud.html>

Author: Peter Wayner

Date: Apr 2018

Summary: From virtual machines and Kubernetes clusters to serverless functions and machine learning APIs — how to navigate Google's cloud riches.

Target audience: Startups/SMEs

Difficulty: Low

Title: Guide to planning your cloud adoption strategy

Source: <https://docs.newrelic.com/docs/using-new-relic/welcome-new-relic/plan-your-cloud-adoption-strategy/guide-planning-your-cloud-adoption-strategy>

Author: New Relic Inc.

Date: Apr 2018

Summary: Follow these guides as your roadmap through each phase of your cloud adoption journey: from planning your migration, to migrating your applications to the cloud, and, finally, to running your applications in the cloud successfully.

Target audience: Startups/SMEs

Difficulty: Low

6.3.4 Courses

Title: DOCKER FUNDAMENTALS + ENTERPRISE DEVELOPERS COURSE BUNDLE

Source: <https://training.docker.com/instructor-led-training/docker-fundamentals-enterprise-developers>

Author: Docker

Date: 2017

Summary: As the follow-on to the Docker Fundamentals course, Docker for Enterprise Operations is a role-based course designed for an organization's Development and DevOps teams to accelerate their Docker journey in the enterprise. The course covers in-depth core advanced features of Docker EE and best practices to apply these features at scale with enterprise workloads. It is highly recommended to complete the Docker Fundamentals course as a pre-requisite.

Target audience: Startups/SMEs

Difficulty: Medium

Title: General docker courses

Source: <https://training.docker.com/instructor-led-training>

Author: Docker

Date: 2017

Summary: Docker Fundamentals: (2 days, \$1495) - <https://success.docker.com/training/courses/docker-fundamentals>

This is THE introductory Docker course to give your team the best foundation for enterprise grade Docker use-cases.

Docker for Enterprise Developers: (2 days, \$1995) - <https://success.docker.com/training/courses/docker-for-enterprise-developers>

As the follow-on to our Docker Fundamentals course, Docker for Enterprise Developers is a role-based course designed for an organization's Development and DevOps teams to accelerate their Docker journey in the enterprise.

Docker for Enterprise Operations:(2 Days, \$1995) - <https://success.docker.com/trainingcourses/docker-for-enterprise-operations>

This course is the second level of Docker's core curriculum for the enterprise and is focused on the Docker Operator role in administration of Docker Enterprise Edition Advanced.

Docker Troubleshooting & Support : (2 Days, \$1995) - <https://success.docker.com/training/courses/docker-support-troubleshooting>

The Docker Troubleshooting & Support course is a role-based course designed for an organization's support teams to troubleshoot the variety of issues that arise in their Docker journey.

Docker Security: (1 Day, \$995) - <https://success.docker.com/training/courses/docker-security>

This is the Docker Security course for your entire organization. Get everyone "on-the-same page" and working together to secure your Dockerized environment. This hands-on workshop style course will give your team an overview of important security features and best practices to protect your containerized services.

Target audience: Startups/SMEs

Difficulty: Medium

6.3.5 Blogs

Title: 4 things you need to know about managing complex hybrid clouds

Source: <https://techbeacon.com/4-things-you-need-know-about-managing-complex-hybrid-clouds>

Author: David Linthicum

Date: 23 January de 2018

Summary: You need to know about managing complex hybrid clouds

Target audience: Startups/SMEs

Dificultad: Medium

Title: How to choose: Open-source vs.commercial cloud management tools

Source: <https://techbeacon.com/how-choose-open-source-vs-commercial-cloud-management-tools>

Author: David Linthicum

Date: 5 December de 2017

Summary: With so many open-source and commercial tools available for cloud management, how do you, as an IT operations management professional, decide which is the best fit for your team's needs? Here's a quick rundown of the general strengths and weaknesses of open-source versus commercial tool options, and when it makes sense to have one, the other, or a mix.

You need a plan for where free and open-source tools fit into your overall cloud management strategy, and where you should consider commercial options to complete the picture and meet your overall needs.

Target audience: Startups/SMEs

Dificultad: Low

Title: Infoworld — Cloud Computing

Source: <https://www.infoworld.com/blog/cloud-computing/>

Author: David Linthicum

Date: Sept 2018

Summary: Infoworld's outstanding cloud computing blog is written by David Linthicum, a consultant at Cloud Technology Partners and a sought-after industry expert and thought leader. His Infoworld blog is exclusively devoted to cloud computing and updated frequently. David's recent blog 'Featuritis' Could Lead You to the Wrong Cloud recommends that enterprises concentrate on strategy and not features when making cloud service choices.

Target audience: Startups/SMEs

Difficulty: Low

Title: All Things Distributed

Source: <https://www.allthingsdistributed.com/>

Author: Werner Vogels

Date: Sept 2018

Summary: All Things Distributed is written by the world-famous Amazon CTO Werner Vogels. His blog is a must-read for anyone who uses AWS. He publishes sophisticated posts about specific AWS services and keeps his readers up-to-date on the latest AWS news. Recent blog posts include: Accelerating Data: Faster and More Scalable ElastiCache for Redis and New Ways to Discover and Use Alexa Skills.

Target audience: Startups/SMEs

Difficulty: Low

Title: Compare the Cloud

Source: <https://www.comparethecloud.net/>

Author: Several authors

Date: Sept 2018

Summary: All Things Distributed is written by the world-famous Amazon CTO Werner Vogels. His blog is a must-read for anyone who uses AWS. He publishes sophisticated posts about specific AWS services and keeps his readers up-to-date on the latest AWS news. Recent blog posts include: Accelerating Data: Faster and More Scalable ElastiCache for Redis and New Ways to Discover and Use Alexa Skills.

Target audience: Startups/SMEs

Difficulty: Medium

Title: Google Cloud Platform Blog

Source: <https://cloudplatform.googleblog.com/>

Author: Google and Cloud experts

Date: Sept 2018

Summary: Google Cloud Platform's blog contains hundreds of articles written by Google cloud experts, and actually dates back to 2008. This vast blog discusses the products, customers, and technical features of Google's cloud solution, while articles can range from product blurbs to extremely detailed technical explanations. A recent post, Evaluating Cloud SQL Second Generation for Your Mobile Game, describes how Google's structured query language can be applied to the special needs of game development.

Target audience: Startups/SMEs

Difficulty: Low

Title: Prudent Cloud

Source: <http://www.prudentcloud.com/>

Author: Subraya Mallya

Date: Sept 2018

Summary: The story behind why this blog was started helps to explain why it is a great blog for any cloud tech enthusiast and why Subraya is in zsah's top 50 Cloud Bloggers. Subraya writes 'Back when I was at

a large company, I had the opportunity to interact with a large set of customers. During my interactions, it became apparent that there was a wide gap in the way we (technologists) thought about how technology gets used and the way our customers thought technology should work. So, I started PrudentCloud to share my thoughts and ideas on how these gaps can be bridged.

Target audience: Startups/SMEs

Difficulty: Low

Title: ClearSky

Source: <https://www.clearskydata.com/blog/author/ellen-rubin>

Author: Ellen Rubin

Date: Sept 2018

Summary: Ellen's experience as an entrepreneur is what gives this 'business and the cloud' blog an edge. She writes from a place of experience and as she has a proven track record in leading strategy, market positioning and go-to-market for fast-growing companies, you can trust her words. Ellen writes blogs that are everyday user-friendly, making this blog a great destination for a start-up or a business at the start of their cloud journey.

Target audience: Startups/SMEs

Difficulty: Low

Title: Right Scale

Source: <https://www.rightscale.com/blog/users/kim-weins>

Author: Kim Weins

Date: Sept 2018

Summary: Kim's blog is another blog targeted at every day users. She has the right balance of technical and commercial knowledge making this blog a well-balanced finished piece. She is currently the Vice President of Marketing for RightScale but obtained a B.S. in engineering from Duke University. Her writing often has a financial focus; comparing pricing and questioning spend.

Target audience: Startups/SMEs

Difficulty: Low

6.4 Overprovisioning Cost KPI

Overprovisioning Cost is the cost for reserving additional resources per unit of time to satisfy unrealised demand. To demonstrate the importance of the Overprovisioning Cost KPI let us consider an example of an e-commerce company that uses cloud infrastructure to offer its services.

For this, the company requests from its cloud provider to reserve additional idle resources to satisfy unforeseen demand in order to keep the application's performance at the desired levels. These idle resources cause unnecessary cost which can be eliminated using the elasticity features of the UNICORN platform.

How to measure?

To measure the over-provisioning, one has to know for a given period of time the actual resources needed (red line) and also the available resource capacity (black line). The area between these lines gives the over-provisioning cost.

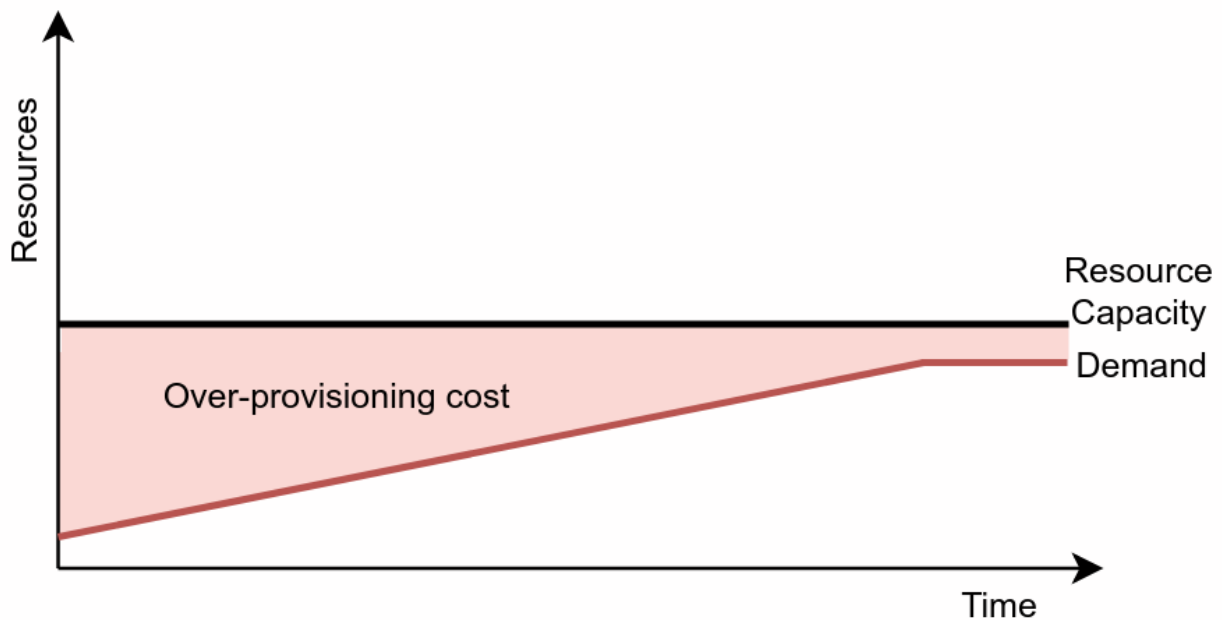


Fig. 1: Over-provisioning cost

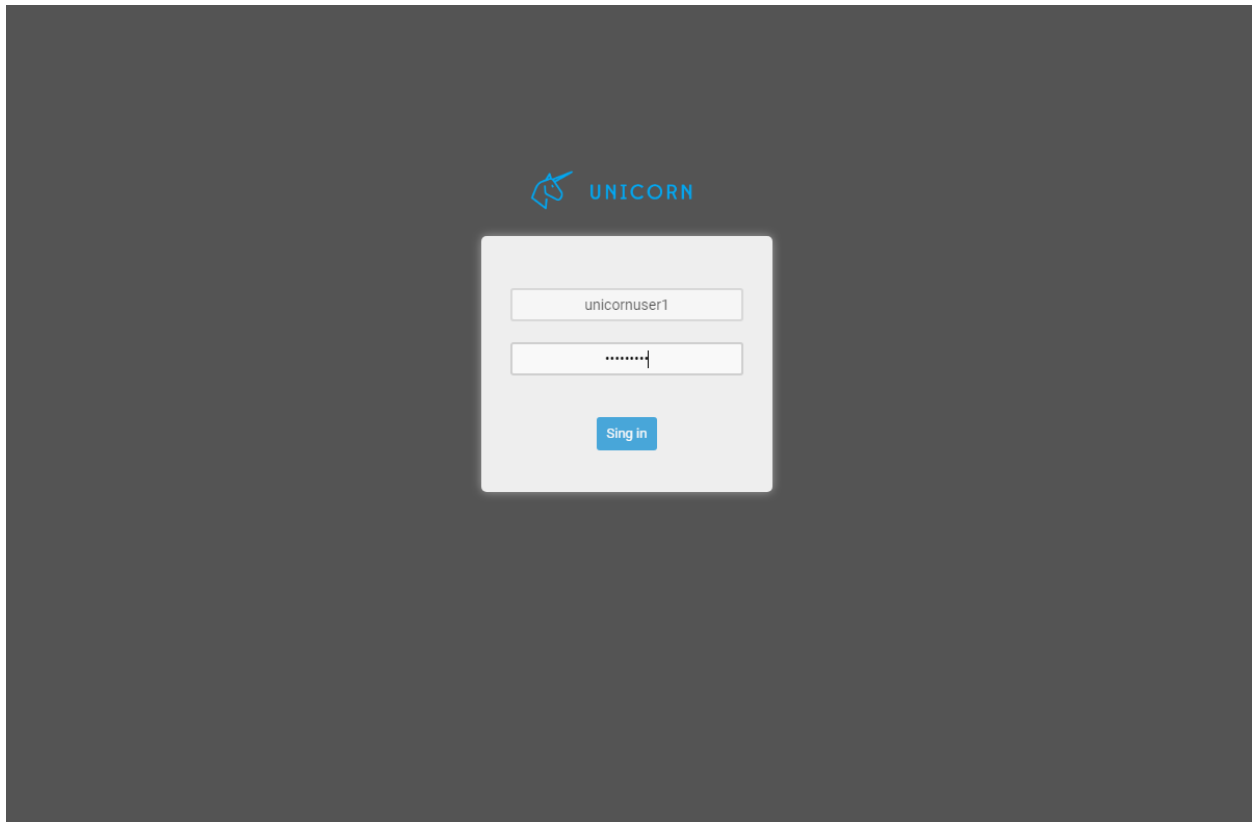
6.5 Getting started with UCO Dashboard

Here you can find instructions for the basic functionalities of UCO Platform.

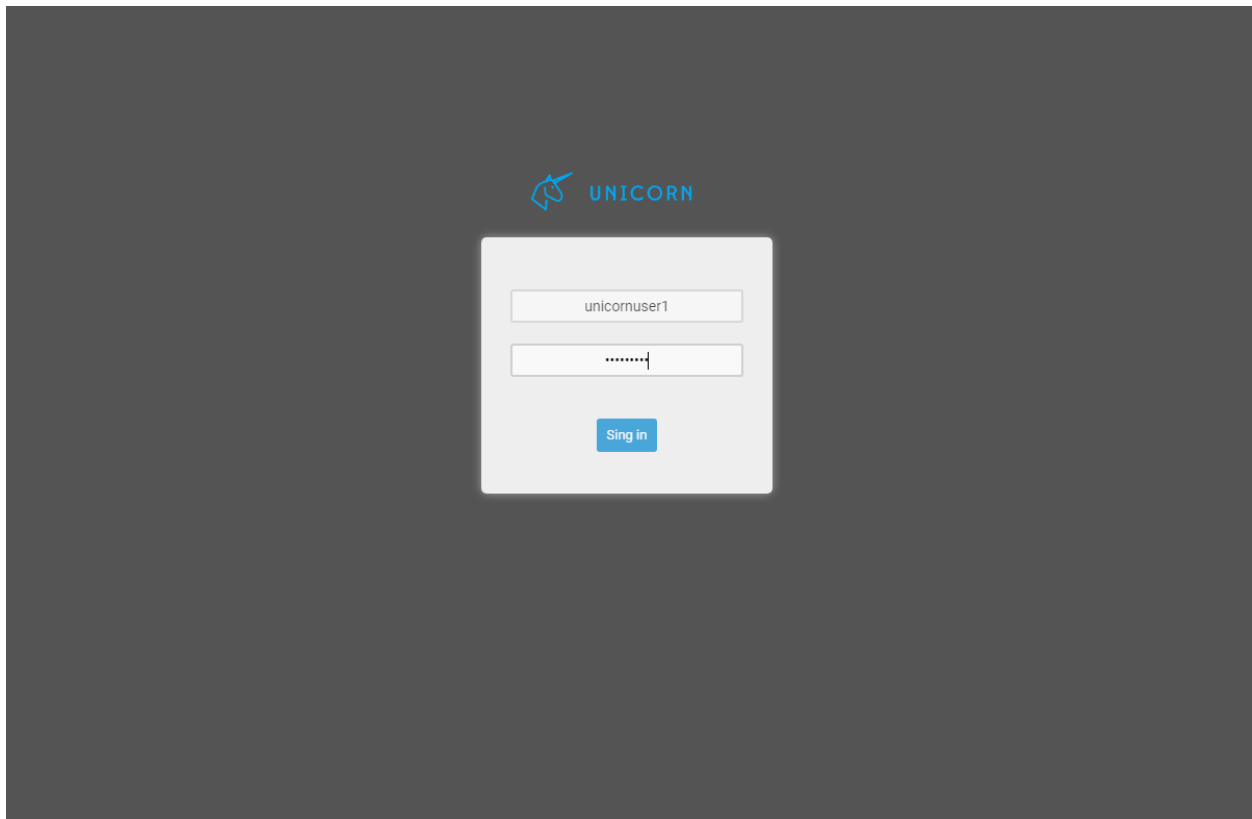
6.5.1 Register and Login

Login

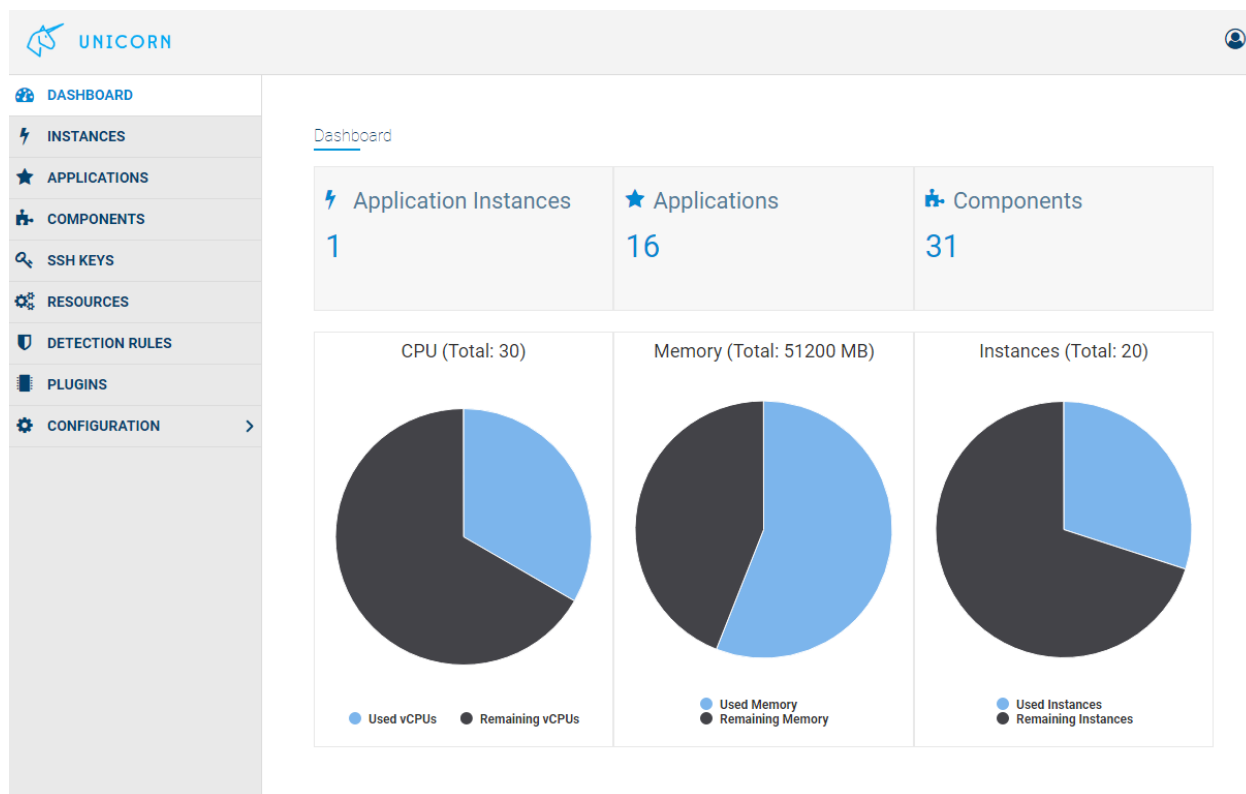
- Click the <Log in> button .



- Provide your login credentials and click the <SIGN IN> button.

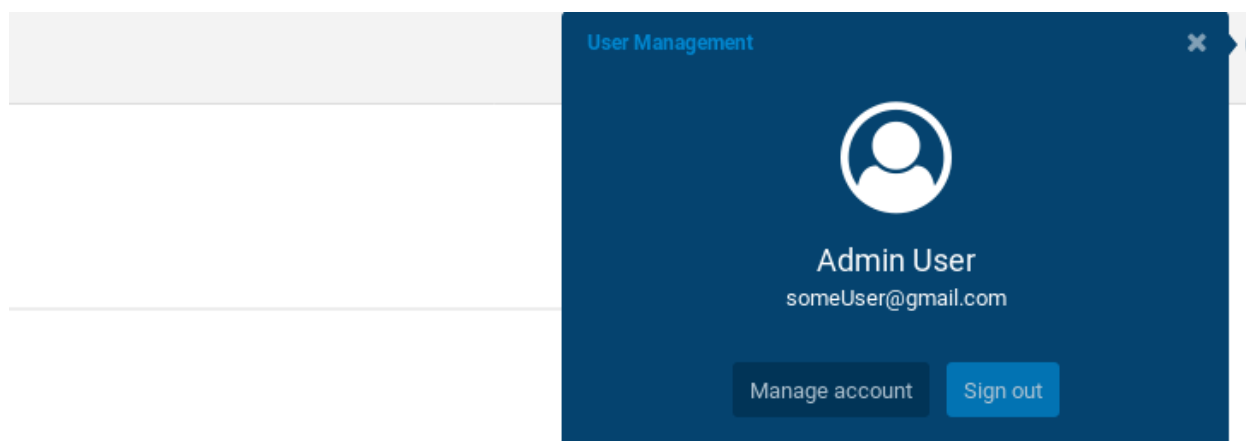


- Upon successful authentication the following screen will be presented.

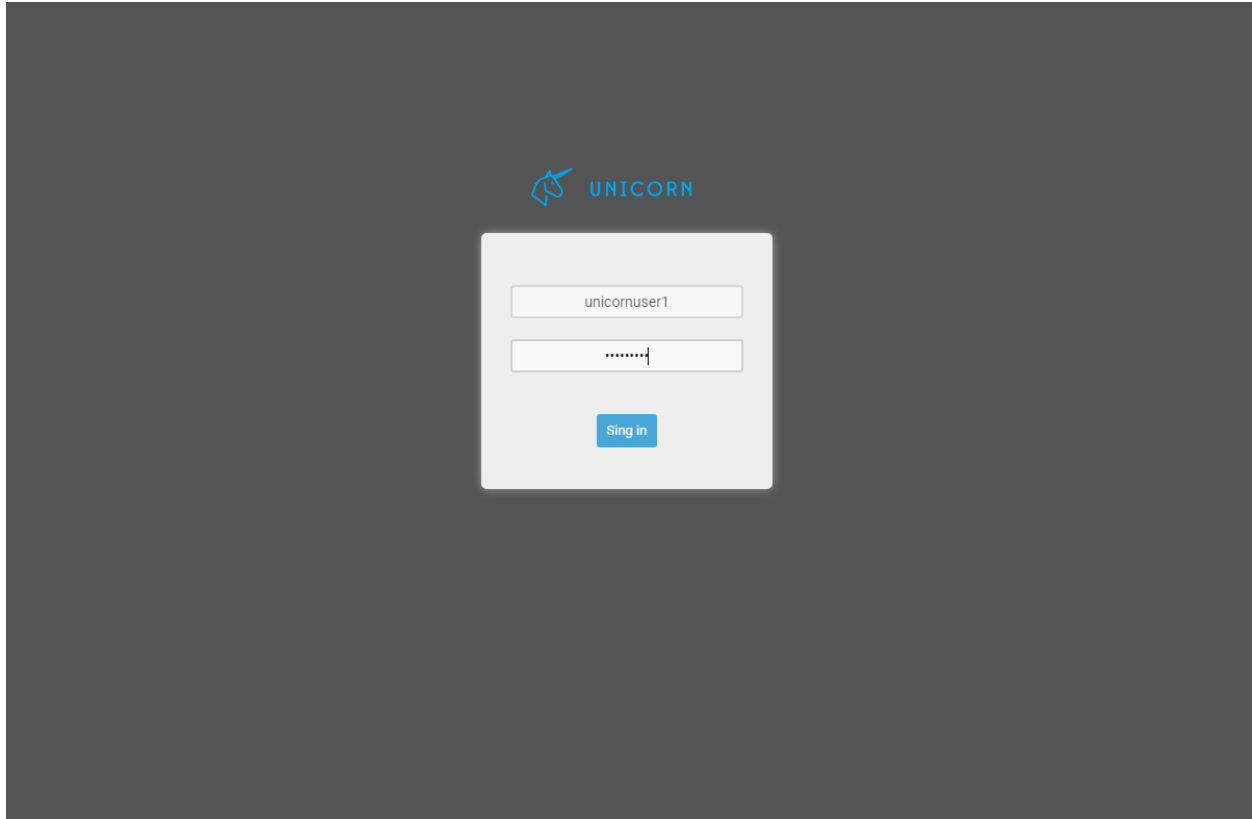


Logout

- In order to perform logout click the <Log-out > Button.



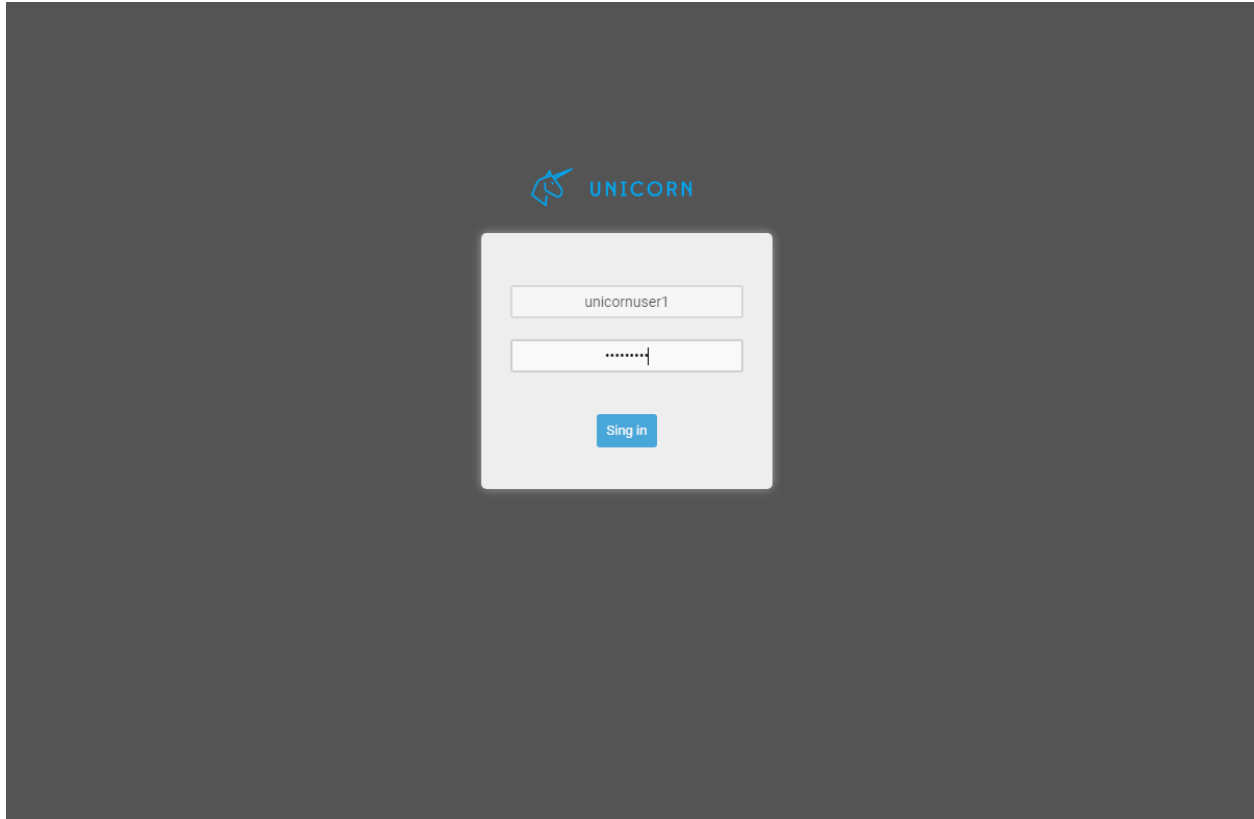
- Upon successful logout the following screen will be presented.



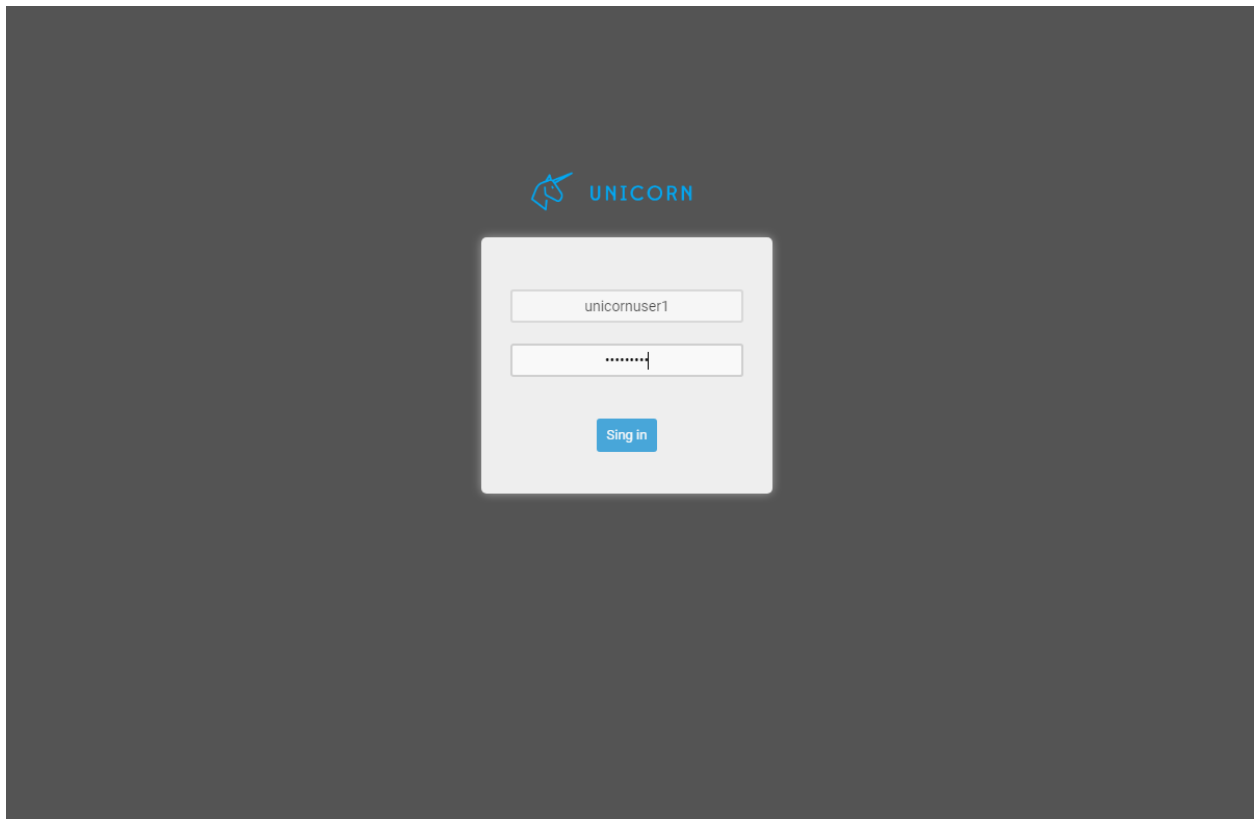
Create account

Note: Registration is currently offered only by contacting us

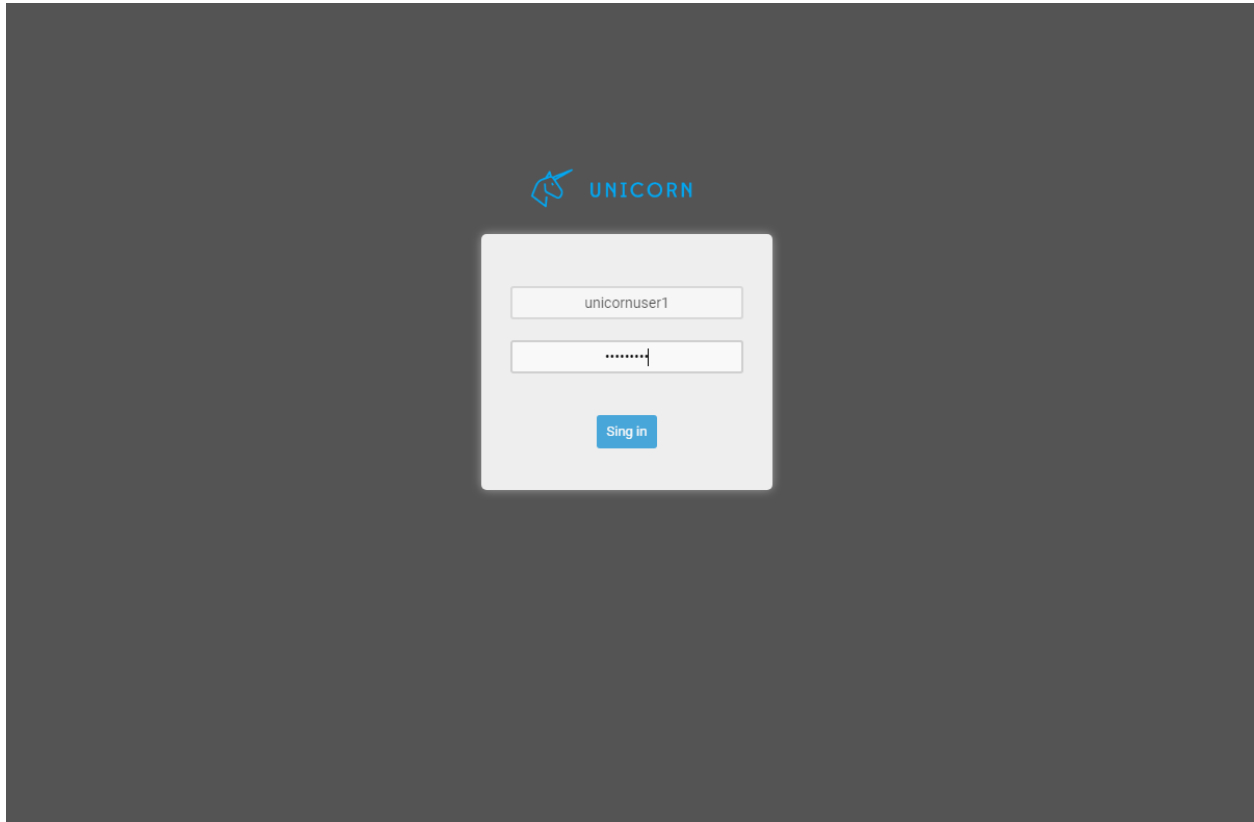
- Visit the Unicorn Platform [Unicorn Platform Webpage](#)
- When a user access the Unicorn the following screen is shown:



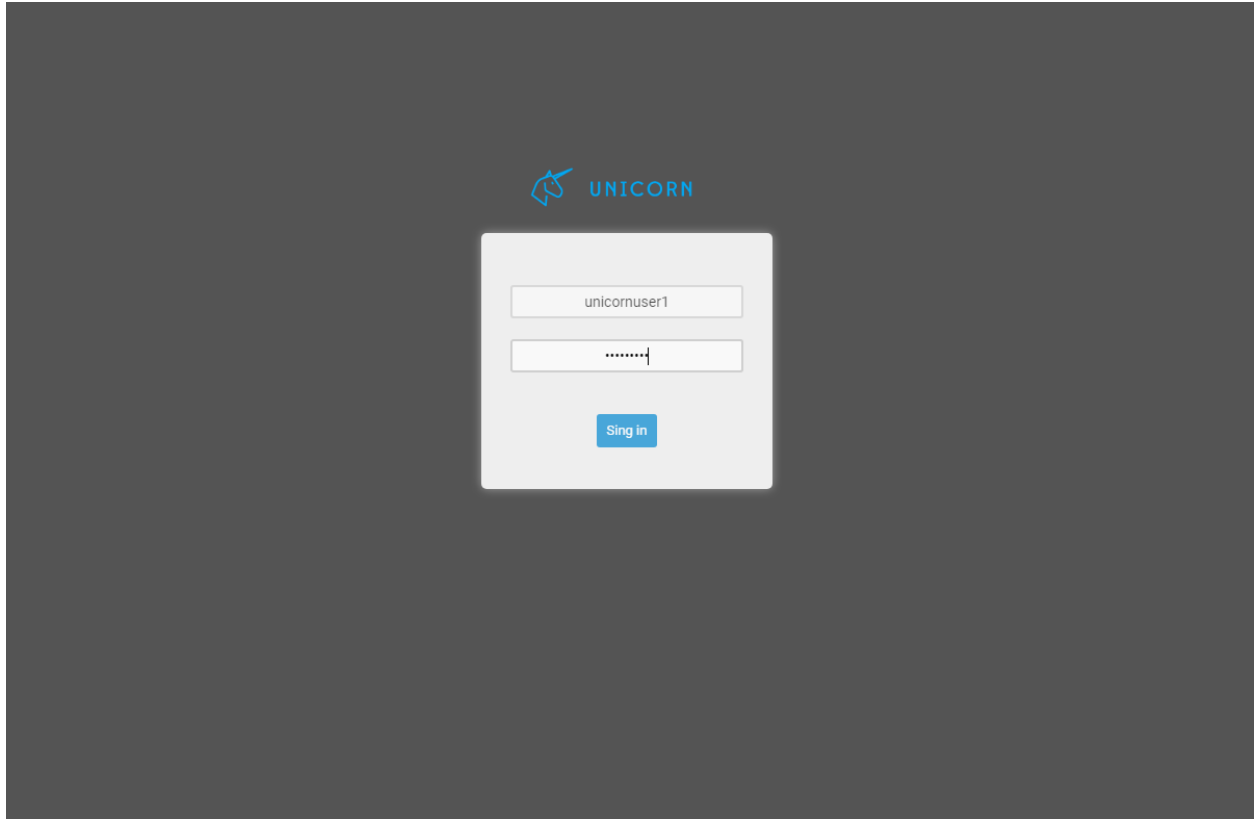
- Click on login button.



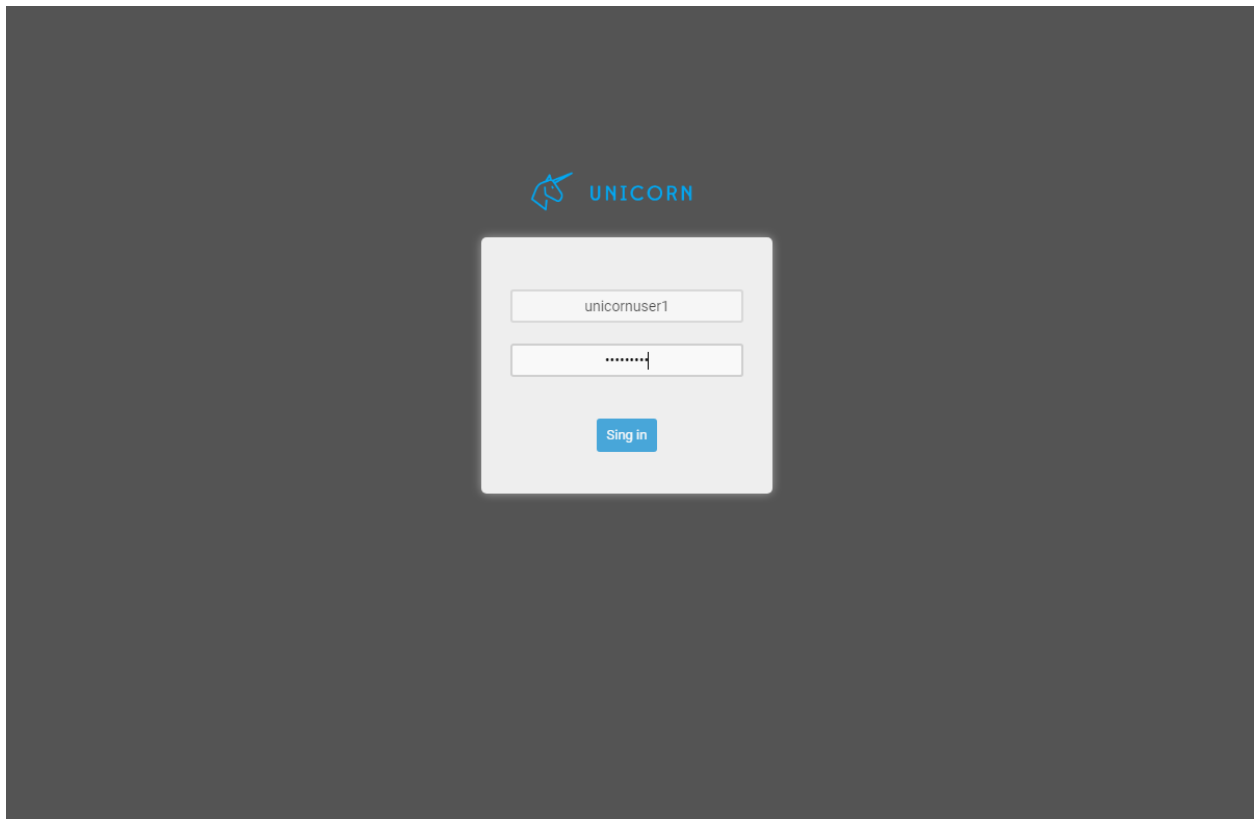
- Click on “Create Account” button.



- Provide account information and click on “CREATE ACCOUNT” button.



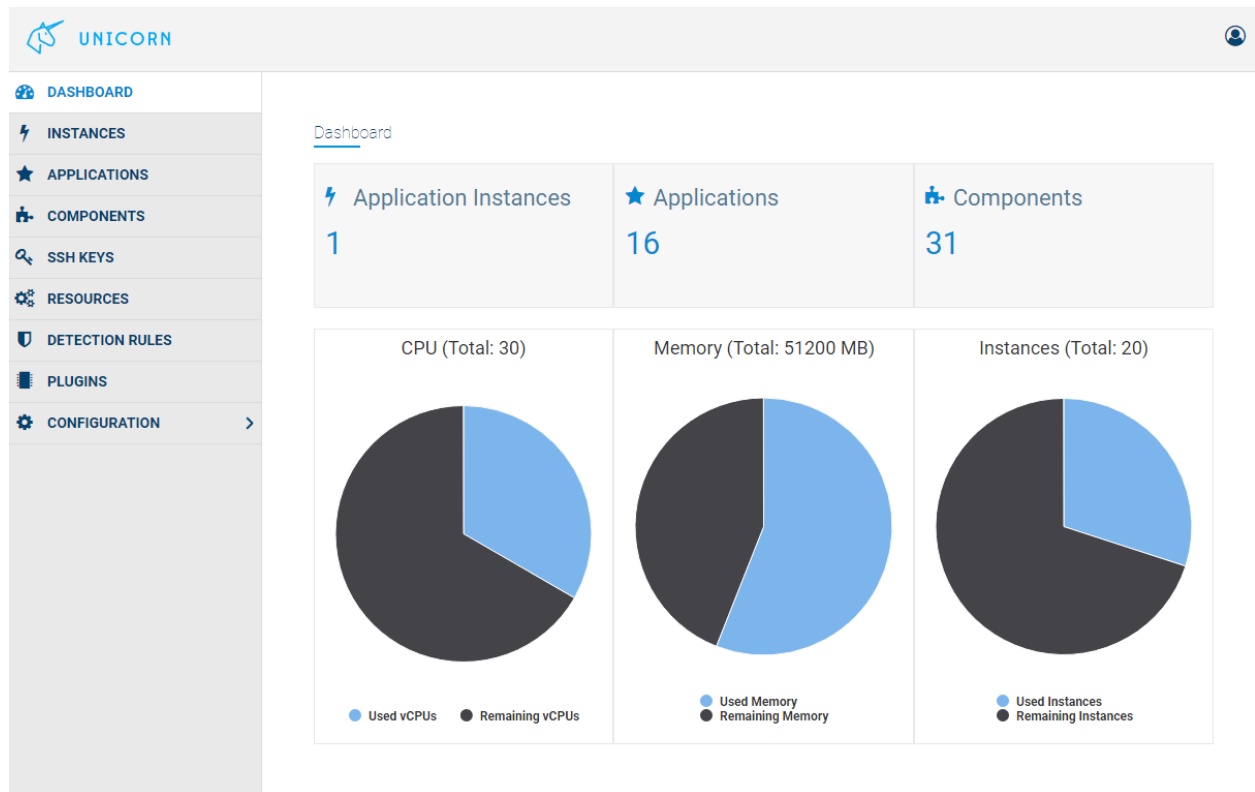
- Go to your email and click on Unicorn account information link.



- You have successfully Created an New Account.

6.5.2 Dashboard Main View

- Through the dashboard, an overview of the components, applications and deployed application instance is provided.
- In UNICORN, applications are described as microservices composed by smaller components.
- Also, an overview of the available and used, aggregated cloud resources of the user is provided.



6.5.3 Platform Usage Video

This screencast can also help you get started:

6.6 Managing Cloud Resources

6.6.1 Manage Cloud Resources

- Each user of UNICORN shall add cloud resources in compatible cloud providers in order to allow the deployment of an application
- OpenStack, Amazon AWS and Google Cloud are supported
- Appropriate forms for of the each cloud providers are available;
 - OpenStack

UNICORN

DASHBOARD

INSTANCES

APPLICATIONS

COMPONENTS

SSH KEYS

RESOURCES

DETECTION RULES

PLUGINS

CONFIGURATION

Resources | Create

Resources Management

Name *

Type name

Image ID

Type Image ID

Nexus IP

Type Nexus IP

Endpoint *

Type endpoint

Password *

Type password

Project *

Type domain

Resource Type *

OpenStack

Network ID

Type the Network ID

Proxy

Type proxy url or IP address

Username *

Type username

Domain *

Type domain

Regions

Select an existing Region

Select...

Or add a new one

Name



Name

+

-

Select this resource as default

- Amazon AWS



DASHBOARD

INSTANCES

APPLICATIONS

COMPONENTS

SSH KEYS

RESOURCES

DETECTION RULES

PLUGINS

CONFIGURATION

Resources > Create

Resources | Create

Resources Management

Name *

Type name

Resource Type *

Amazon WS

Image ID

Type Image ID

Network ID

Type the Network ID

Nexus IP

Type Nexus IP

Public Key *

Paste your public key

Private Key *

Paste your private key

Regions

Select an existing Region

Select...

Or add a new one

Name

Name

+

-

☐ Select this resource as default

Save

© 2018 UNICORN | [About](#) | [Terms of Service](#) | [Privacy](#)

- Google Cloud

UNICORN

DASHBOARD

INSTANCES

APPLICATIONS

COMPONENTS

SSH KEYS

RESOURCES

DETECTION RULES

PLUGINS

CONFIGURATION

Resources > Create

Resources | Create

Resources Management

Name *

Type name

Resource Type *

Google Cloud Compute

Image ID

Type Image ID

Network ID

Type the Network ID

Nexus IP

Type Nexus IP

Username *

Type username

Private Key *

Paste your private key

Regions

Select an existing Region

Select...

Or add a new one

Name

Name

+

-

Select this resource as default

Save

© 2018 UNICORN | [About](#) | [Terms of Service](#) | [Privacy](#)

6.6.2 Manage Your Keys

You can add keys for your account with the appropriate form.

Name	Default ^	Date Created	
myLaptop	✓	27/11/2018 - 14:05	⋮

6.7 Managing Components and Applications

6.7.1 Manage Componets

- In UNICORN, applications are described as **microservices** composed by smaller **components**.
- A list of the existing components is provided.

UNICORN

COMPONENTS

Components Management

Create new

Name
Search by Name

Filter Reset

Identifier	Name ▾	Visibility	Date Created
uMRzm4UKry	APIGatewayCC	Public	27/11/2018 - 15:11
h8b6w4dDzf	BookService	Public	27/11/2018 - 14:33
sVNNwekfwE	CCFrontend	Public	27/11/2018 - 15:41
8ybwnk5ftq	CCRatings	Public	27/11/2018 - 15:39
DP7gyR5R7z	CCReviews	Private	27/11/2018 - 15:40

<< < > >>

- Also, the user can create new components.
- For each component the name, the architecture and the way that the specific component scales have to be defined.
- Then the docker container image that is used for this component has to be defined.
 - Custom Docker registries are supported.

[Component](#) > [Edit](#)[Components](#) | [Edit](#)

Components Management

The screenshot shows a web form for configuring a component. It includes fields for Name, Architecture, Elasticity Profile, Distribution Parameters (Registry and Image), and Minimum Execution Requirements (vCPUs, RAM, Storage, Hypervisor Type, and a GPU-Enabled checkbox).

Name *
WordPress

Architecture *
x86

Elasticity Profile *
Horizontal

Distribution Parameters

Docker Registry *
nexus:39580

Docker Image *
wordpress:4

Minimum Execution Requirements

vCPUs *
1

RAM (MB) *
2048

Storage (GB) *
20

Hypervisor Type *
ESXI

☐ GPU-Enabled

- Execution Requirements are defined, and custom health checks can be added to ensure that the service is deployed properly..

The screenshot shows a form for configuring a health check. It includes fields for HTTP, Command, and Time Interval.

Health Check

HTTP *
http://localhost:80

Command *
e.g. mysql -u username -p`Password` status

Time Interval *
10s

- The service of the component must use environmental variables that can be configured by the user.
- **As an example we have a WordPress component;**
 - Environmental variables as WORDPRESS_DB_USER and WORDPRESS_DB_PASSWORD can be used for the configuration of the component, and default values can be added
 - Adding @ and the component name (e.g: @MariaDB) means that the WordPress component will dynamically get the IP that the corresponding component will get once deployed.

Environmental variables

Key	Value	
WORDPRESS_DB_PASSWORD	wordpress	+ -
WORDPRESS_DB_HOST	@MariaDB	-
WORDPRESS_DB_USER	wordpress	-

- **The interfaces exposed and the interfaces required by the service have to be defined.**
 - A user can select one of the existing interfaces, like the TCP access through port 80, or define a new interface.
 - For the definition of the required interface, an existing exposed interface of another component has to be selected.

Exposed Interfaces

Select an existing Interface

× httpAccessInterface (Port: 80, Type: ACCESS, Protocol: TCP)
 ×

Or add a new one

Name	Port	Interface Type	Transmission Protocol	
<input type="text" value="Name"/>	<input type="text" value="Port"/>	<input type="radio"/> Core <input checked="" type="radio"/> Access	<input type="radio"/> TCP <input checked="" type="radio"/> TCP/UDP <input type="radio"/> UDP	+ -

Required Interfaces

Label	Select an Interface	
<input type="text" value="Type a label for the required interface"/>	sqlInterface (Port: 3306, Type: ▼)	+ -

- Additional details like volumes, devices and labels for the component can be defined.

Plugins

Select an existing Plugin

Select...

Volumes

Name

/var/path

+

-

Devices

Key

e.g GPU

Value

100

+

-

Labels

Select an existing Label

Content Management

Or add a new one

Name

Name

+

-

☒ (Public) If this option is checked, anyone could see this component

Save

6.7.2 Manage Applications

- After the needed components have been defined, the user can proceed with the definition of the application.
- The application will be created through the help of a visual graph editor and then will be available for deployment.

UNICORN

DASHBOARD
INSTANCES
APPLICATIONS
COMPONENTS
SSH KEYS
RESOURCES
DETECTION RULES
PLUGINS
CONFIGURATION

Applications
Application Management
Create new

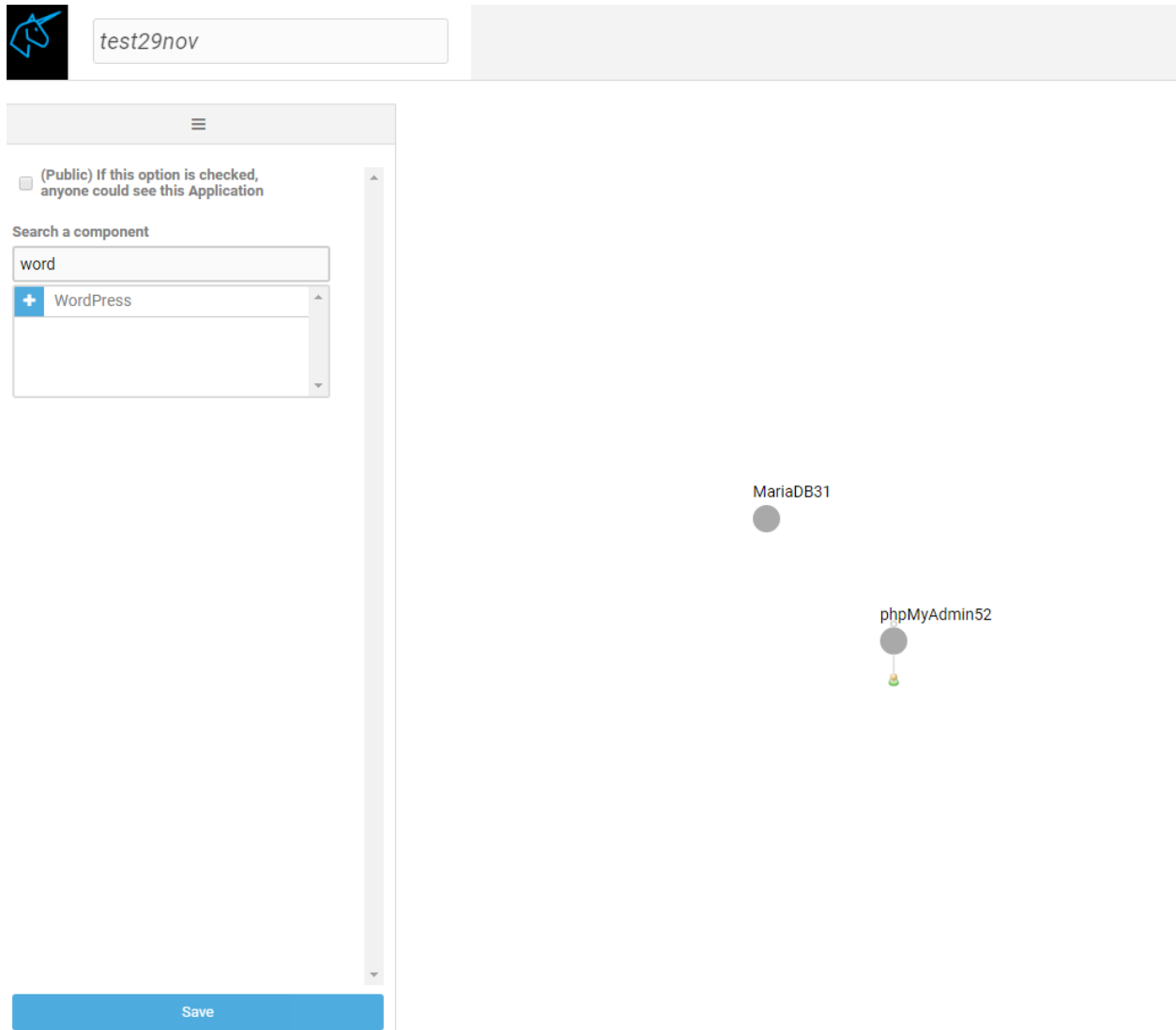
Name
Search by Name

Filter
Reset

Identifier	Name	Visibility	Date Created
9FDXMvmDvK	AudioDetector	Public	27/11/2018 - 14:05
VC8V345VQp	AudioVideoDetector	Public	27/11/2018 - 14:05
Ku6YbdATtx	ccapp	Public	27/11/2018 - 16:03
qmd2J1hesT	CMS	Public	28/11/2018 - 17:22
uGW41GFZVu	CMSApp	Public	27/11/2018 - 14:05

<<
<
>
>>


- At the visual editor, the application components are presented as the nodes of a graph, and the connection between the nodes is describing the interfaces between the services.



- Through the left side panel, the components can be retrieved and added to the editor.

The screenshot displays a web application management interface. At the top, there is a header bar with a logo on the left and a search bar containing the text "test29nov". Below the header, the interface is divided into two main sections. On the left, a sidebar contains a menu icon and a search bar with the text "word". Below the search bar, a panel titled "Generic info:" displays the ID "J86QsMcJeT" and the Name "WordPress". Below this, a section titled "WordPress63" includes edit and delete icons. Further down, a section titled "Required Interfaces:" lists "1. Interface: sqlInterface" and includes a search bar for candidate components. Below that, a section titled "Exposed Interfaces:" lists "1. Interface: httpAccessInterface". At the bottom of the sidebar is a blue "Save" button. On the right, a large canvas displays a diagram with three nodes: "WordPress63" (a blue circle with a green plug icon), "MariaDB31" (a grey circle), and "phpMyAdmin52" (a grey circle with a green plug icon). Lines connect "WordPress63" to "MariaDB31" and "WordPress63" to "phpMyAdmin52".

- By selecting the required interface and dragging it to another node, the connection between the interfaces of the components can be done.



test29nov

☰



☐ (Public) If this option is checked, anyone could see this Application

Search a component

Generic info:

ID: J86QsMcJeT

Name: WordPress

WordPress63  

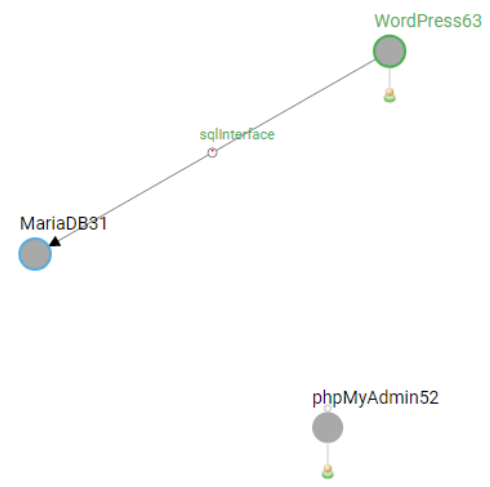
Required Interfaces:

1. Interface: sqlInterface

Exposed Interfaces:

1. Interface: httpAccessInterface

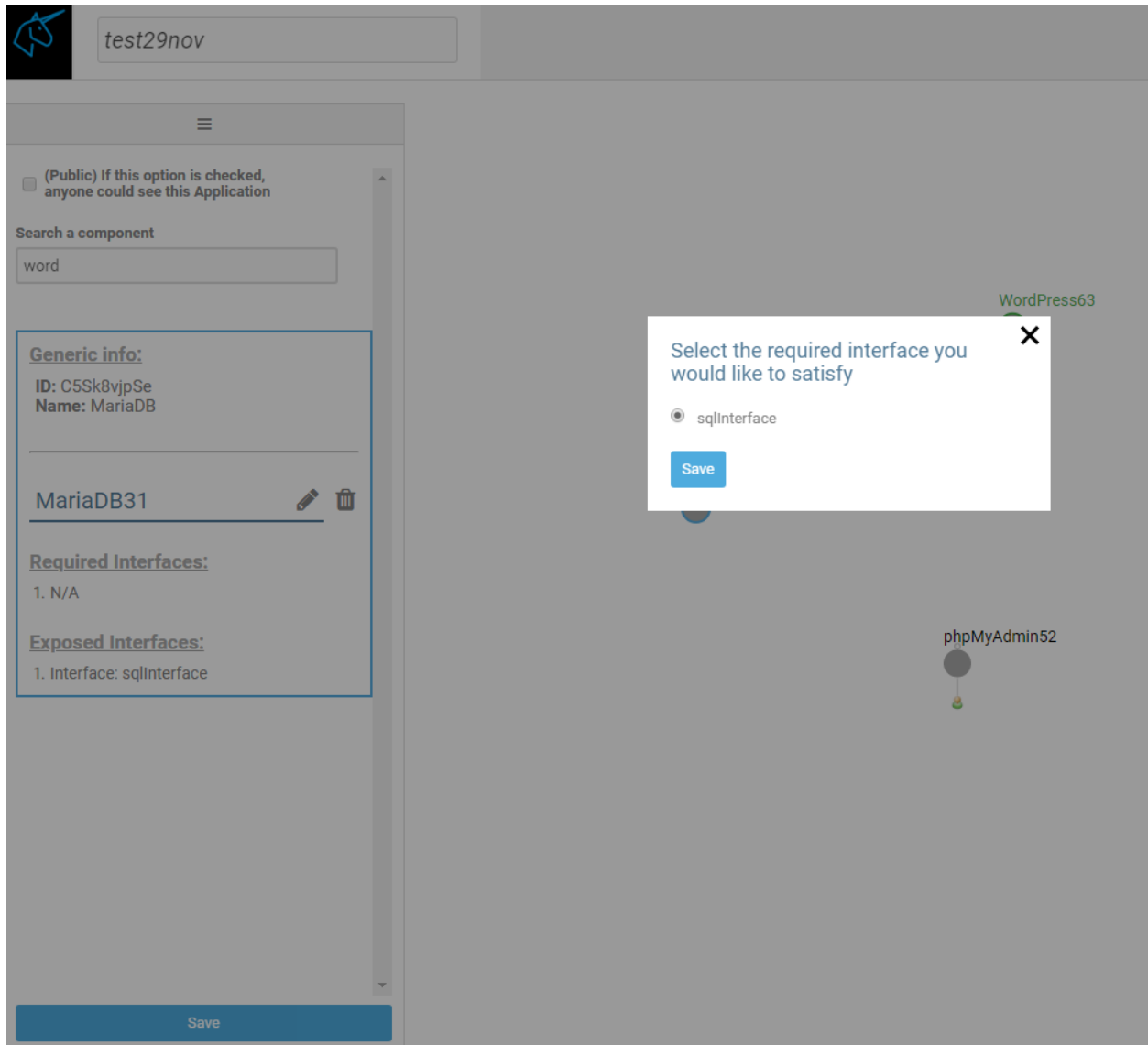
Save



```

graph TD
    MariaDB31((MariaDB31)) -- sqlInterface --> WordPress63((WordPress63))
    phpMyAdmin52((phpMyAdmin52)) -- httpAccessInterface --> WordPress63
        
```

- This procedure is followed until all required interfaces have been connected in order to save a valid application graph.



The screenshot shows the Unicorn application management interface. On the left, a sidebar contains a search bar with the text "word" and a "Generic info" section. The "Generic info" section displays the ID "J86QsMcJeT" and the Name "WordPress". Below this, the "WordPress63" component is listed with edit and delete icons. The "Required Interfaces" section shows "1. Interface: sqlInterface", and the "Exposed Interfaces" section shows "1. Interface: httpAccessInterface". A "Save" button is at the bottom of the sidebar. The main area on the right displays a dependency graph with three nodes: "WordPress63" (green), "MariaDB31" (blue), and "phpMyAdmin52" (grey). A solid line labeled "sqlInterface" connects "WordPress63" to "MariaDB31". A dashed line connects "MariaDB31" to "phpMyAdmin52".

6.7.3 Manage Applications instances

- Now the application can be instantiated and deployed to the cloud that user desires.

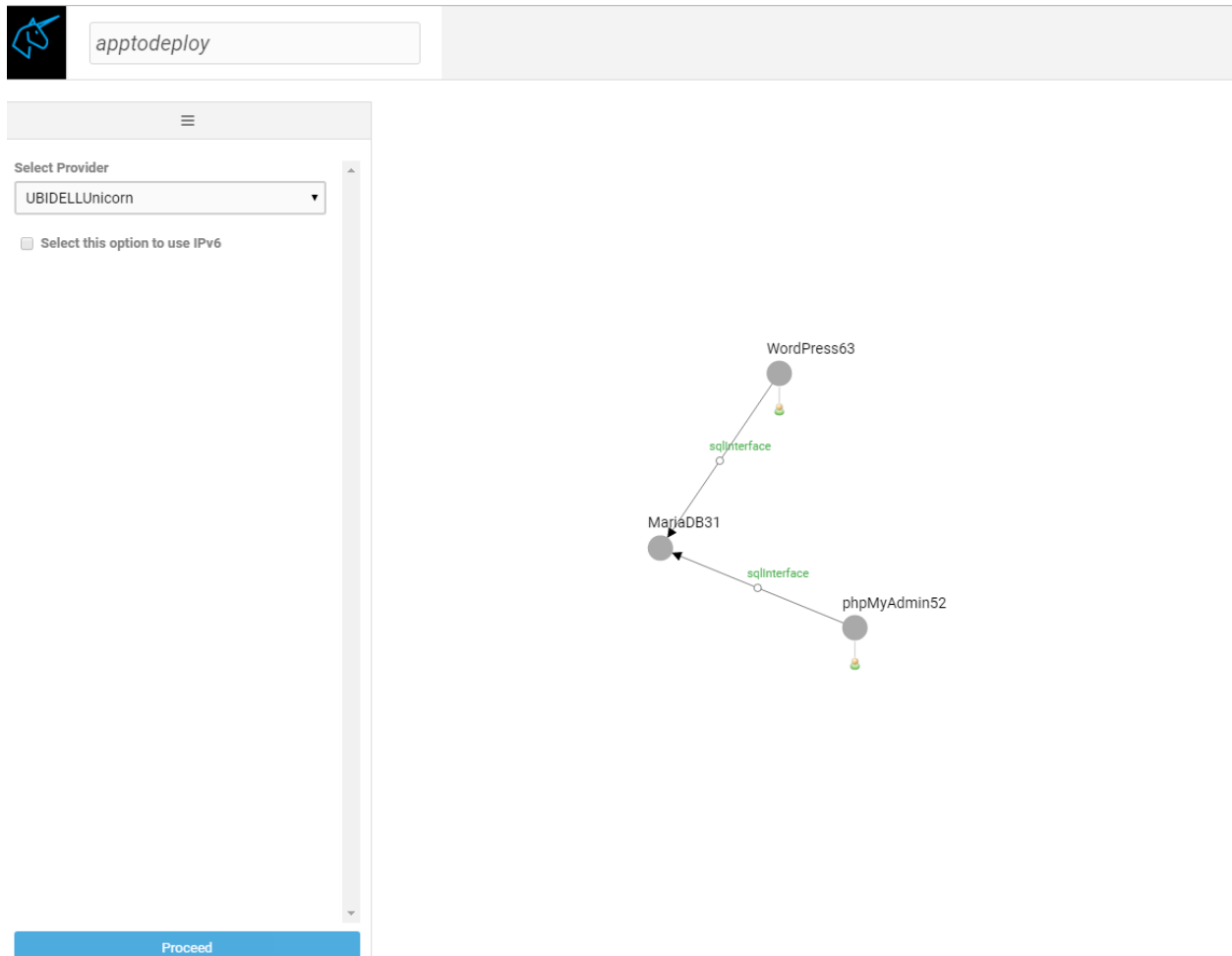
The screenshot shows the Unicorn Applications Management page. The left sidebar contains a navigation menu with options: DASHBOARD, INSTANCES, APPLICATIONS (selected), COMPONENTS, SSH KEYS, RESOURCES, DETECTION RULES, PLUGINS, and CONFIGURATION. The main area is titled "Applications" and "Application Management". It features a search bar labeled "Name" with the placeholder "Search by Name", and "Filter" and "Reset" buttons. Below this is a table with the following data:

Identifier	Name	Visibility	Date Created	
V8GasiYW6S	test29nov	Private	29/11/2018 - 15:37	Create Instance Delete
dDebKhWuaa	TwoNodeAppArm	Public	27/11/2018 - 14:05	

At the bottom of the table are navigation controls: "<<", "<", ">", ">>". A "Create new" button is located in the top right corner of the main area.

- By pressing "Proceed" the deployment starts.

- However, the user can also configure the application components before deployment.



- Prior to the deployment the user can activate the embedded Intrusion Prevention Mechanism (IPS) and Intrusion Detection Mechanism (IDS), by selecting the checkbox.
- The minimum and maximum amount of workers per node are defined in order to specify the scalability profile of the application.

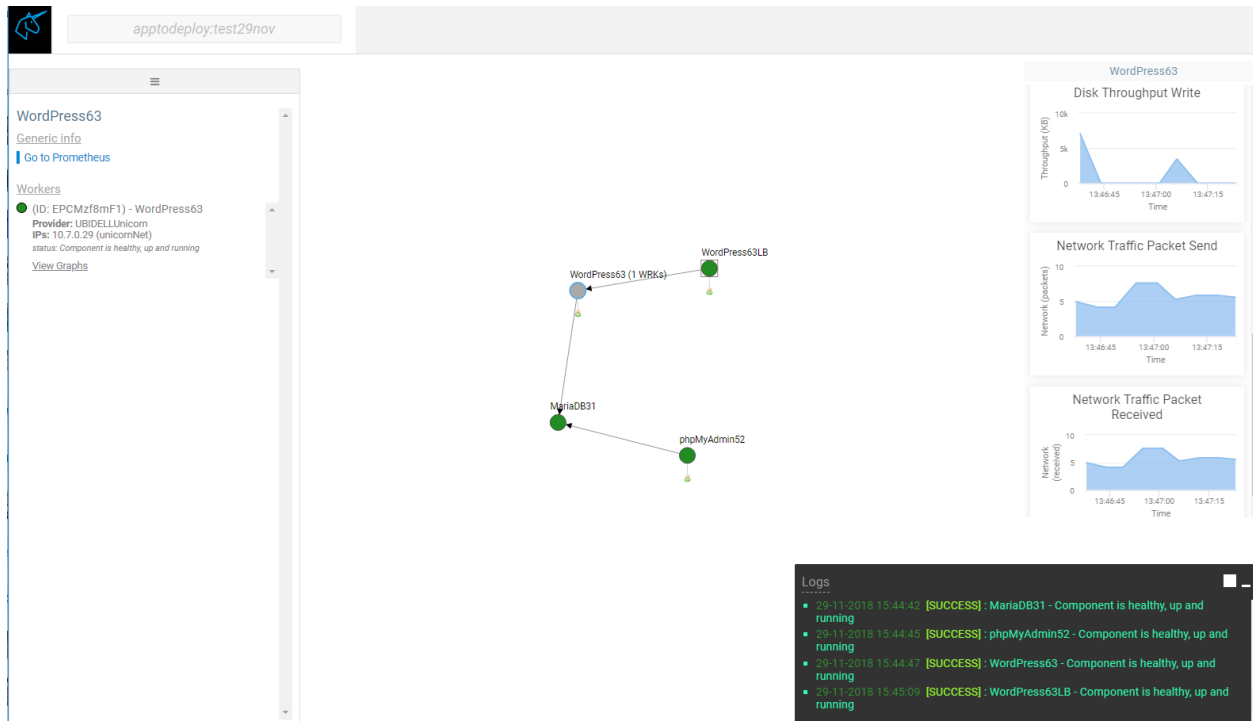
The screenshot shows the 'Configure "WordPress63" Component' dialog box. The title bar includes the component name and ID: EPCMzf8mF1. The dialog is divided into several sections: 'Select SSH Key' with a dropdown set to 'myLaptop'; 'Select Provider' with a dropdown set to 'UBIDELLUnicom'; 'Select Region *' with a dropdown set to 'eu-east'; 'Prevention Rules' with a checkbox for 'Activate IPS'; 'Detection Rules' with a checkbox for 'Activate IDS' and a 'Select Rulesets' dropdown; 'Workers' with input fields for 'Minimum Workers' (1) and 'Maximum Workers' (1); 'Flavor' with input fields for 'vCPUs *' (1), 'RAM *' (2048), and 'Storage *' (20); and a 'Health Check' section at the bottom.

- The environmental variables of the application can also be configured prior to the deployment.

This screenshot shows the 'Health Check' configuration section. It includes a 'HTTP' field with the value 'http://localhost:80', a 'Command' field with the value 'e.g. mysql -u username -p'Password' status', and a 'Time Interval' field set to '10s'. Below this is the 'Overridable Parameters' section with a 'Command' field. The 'Interfaces' section contains a table with 'Name' and 'Port' columns, showing 'httpAccessInterface' on port '80'. The 'Environmental Variables' section has a table with 'Key' and 'Value' columns, listing 'WORDPRESS_DB_PASSV' (wordpress), 'WORDPRESS_DB_HOST' (@MariaDB31), and 'WORDPRESS_DB_USER' (wordpress). The 'Plugins' section at the bottom shows a list of plugins with checkboxes for selection, including 'proc/proc/softirqs', 'proc/proc/net/softnet_stat', 'apps', 'statsd/stats', 'proc/proc/net/snmp6', 'proc/proc/net/sockstat', 'proc/proc/net/sockstat6', 'proc/sys/kernel/mm/ksm', 'proc/proc/net/snmp', 'charts.d', 'proc/proc/vmstat', and 'proc/proc/pc'.

- Deployment procedure needs few minutes to finish. The user is constantly informed by viewing the logs aggregated from a

- The total deployment time depends on the cloud infrastructure selected, as the spawning of new VMs might take more time in some IaaS.
- Total time is also affected by the network delays between the cloud infrastructure and the docker registry that is used to fetch the components container image.



- When deployment finishes all nodes turn green
 - On the instance list the application is shown as “DEPLOYED”.
- Monitoring metrics are presented for each one of the application nodes

UNICORN

Application Instances

Application Instances Management

Name: Search by Name Status: -- Select -- Application: -- Select --

Filter Reset

Identifier	Name	Application Name	Quotas	Status	Date Created
Qxf3hVmlJZ	apptodeploy	test29nov	vCPUs: 4 Memory: 8192 MB Storage: 80 GB	WAITING_ORCHESTRATOR	29/11/2018 - 15: View Graph
904s5gnzgn	testide	CMS	vCPUs: 4 Memory: 8192 MB Storage: 80 GB	DEPLOYED	29/11/2018 - 09:32

6.8 Monitoring Application and Scale

6.8.1 Monitoring Applications instances

- When deployment finishes all nodes turn green
 - On the instance list the application is shown as “DEPLOYED”.
- Monitoring metrics are presented for each one of the application nodes

Application Instances Management

Name: Search by Name Status: -- Select -- Application: -- Select --

Filter Reset

Identifier	Name	Application Name	Quotas	Status	Date Created
Qxf3hVmlJZ	apptodeploy	test29nov	vCPUs: 4 Memory: 8192 MB Storage: 80 GB	WAITING_ORCHESTRATOR	29/11/2018 - 15:00
904s5gnzgn	testde	CMS	vCPUs: 4 Memory: 8192 MB Storage: 80 GB	DEPLOYED	29/11/2018 - 09:32

6.8.2 Application Elasticity

- From the application instance list, the user must select the “Elasticity Policies” option for the deployed application, in order to configure how the application scales.
- By selecting the appropriate function, user can to aggregate the monitoring results in various ways.
- For the monitored parameter we select the metric and it’s dimension from appropriate lists.
- An operand shall be added to the policy and the threshold that the policy shall confirm to.
- The period field is used to set the size of the time window that the metric values are collected and aggregated for the policy enforcement.

The screenshot shows the UNICORN configuration interface. On the left is a sidebar with navigation links: DASHBOARD, INSTANCES, APPLICATIONS, COMPONENTS, SSH KEYS, RESOURCES, DETECTION RULES, PLUGINS, and CONFIGURATION (which is highlighted). The main area displays a configuration form for a policy named 'policy12'.

Policy Configuration:

- Name:** policy12
- Expression:**
 - Select Function: Average
 - Select Component: WordPress63
 - Select Metric: _cpu_cpu_percentage_average (percentage)
 - Select Dimension: All dimensions
 - Select Operand: >
 - Threshold: 80
- Period (Seconds):** 20
- Inertia Time (Minutes):** 1

Actions:

- Action 1:**
 - Select Component: MariaDB31
 - Type: Info
 - Message: Scaling action needed.
 - Buttons: + (blue), - (red)
- Action 2:**
 - Select Component: WordPress63
 - Type: Scale Out
 - Workers: 2
 - Buttons: - (red)

A 'Save' button is located at the bottom left of the configuration area.

- On the scaling action we can select the component to scale in or scale out, and the number of workers to scale.
 - After a scaling action is performed, some time is needed for having the component workers deployed. For this reason we should ensure that for this period we don't fire additional scaling actions.
 - This is done through the "Inertia" field that is used to define the time in minutes that after a scaling action is done, no further action is performed.
- Multiple scaling actions can be added.
- The policy can be saved and will be enforced to the application within few seconds.

The screenshot shows the UNICORN configuration interface. On the left is a sidebar with navigation links: DASHBOARD, INSTANCES, APPLICATIONS, COMPONENTS, SSH KEYS, RESOURCES, DETECTION RULES, PLUGINS, and CONFIGURATION (which is highlighted). The main area displays a configuration form for a scaling rule named 'policy12'.

Expression Section:

- Name ***: policy12
- Select Function**: Average
- Select Component**: WordPress63
- Select Metric**: _cpu_cpu_percentage_average (percentage)
- Select Dimension**: All dimensions
- Select Operand**: >
- Threshold**: 80

Period (Seconds) *: 20

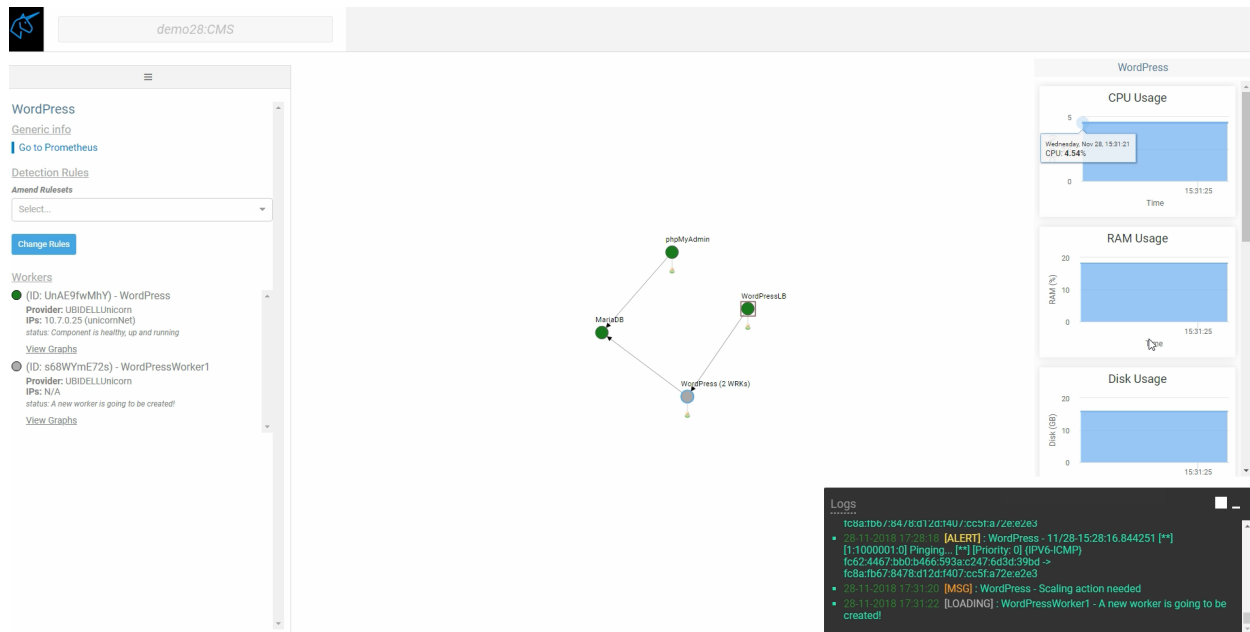
Inertia Time (Minutes) *: 1

Actions Section:

- Action 1:**
 - Select Component**: MariaDB31
 - Type**: Info
 - Message**: Scaling action needed.
- Action 2:**
 - Select Component**: WordPress63
 - Type**: Scale Out
 - Workers**: 2

At the bottom of the form is a 'Save' button.

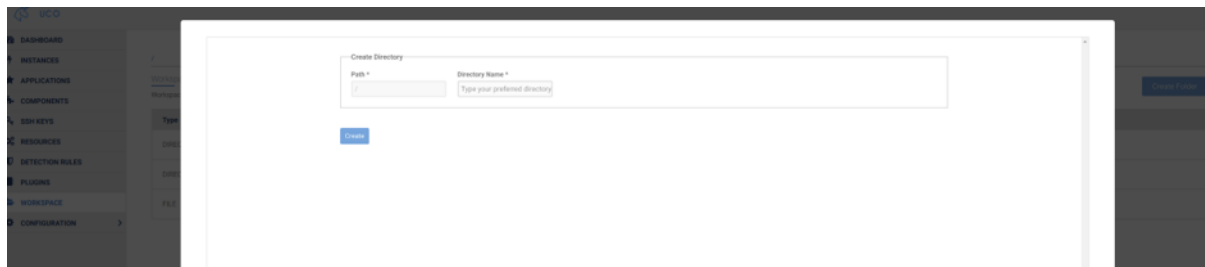
- In this example we had initially only one worker of the WordPress component.
- **But due to the scaling rule, an additional node has been created.**
 - A load balancer had been already deployed from the initial deployment since we had defined that this component might need multiple workers.
 - The scaling action is visible to the user through the logs and the number on workers in the “WordPress” node in the graphs.



6.9 Managing and Using Volumes

6.9.1 Unicorn Volumes and Workspace

On you workspace, create a proper path for your containers



On you workspace and on the proper path, upload yours files



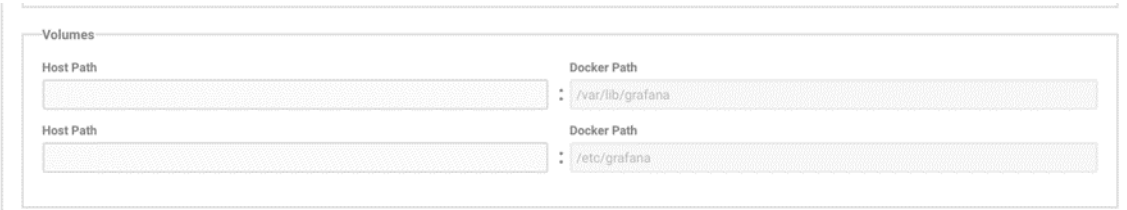
6.9.2 Unicorn Volumes and Application

On component registration or edit, you can add your container volume paths



The screenshot shows a 'Volumes' section with two entries. Each entry has a 'Name' field and a corresponding '+' or '-' button. The first entry has the name '/var/lib/grafana' and a '+' button. The second entry has the name '/etc/grafana' and a '-' button.

At last, on create instance and on component Parameterization, set the host paths using the yours workspace path (use only directory paths and NO file paths)



The screenshot shows a 'Volumes' section with two entries. Each entry has a 'Host Path' field and a 'Docker Path' field. The first entry has an empty 'Host Path' field and a 'Docker Path' field containing '/var/lib/grafana'. The second entry has an empty 'Host Path' field and a 'Docker Path' field containing '/etc/grafana'.

6.10 Advanced Monitoring Configuration - UNICORN Monitoring Service

6.10.1 Overview

Unicorn Monitoring Service is an open source monitoring system that follows an agent-based architecture that embraces the producer-consumer communication paradigm. The Unicorn Monitoring Service runs in a non-intrusive and transparent manner to any underlying cloud as neither the metric collection process nor metric distribution and storage are dependent to the underlying platform APIs and communication mechanisms. In turn, the Monitoring Service takes into consideration the rapid changes that occur due to the enforcement of elastic actions to the underlying infrastructure and the application topology.

6.10.2 Features

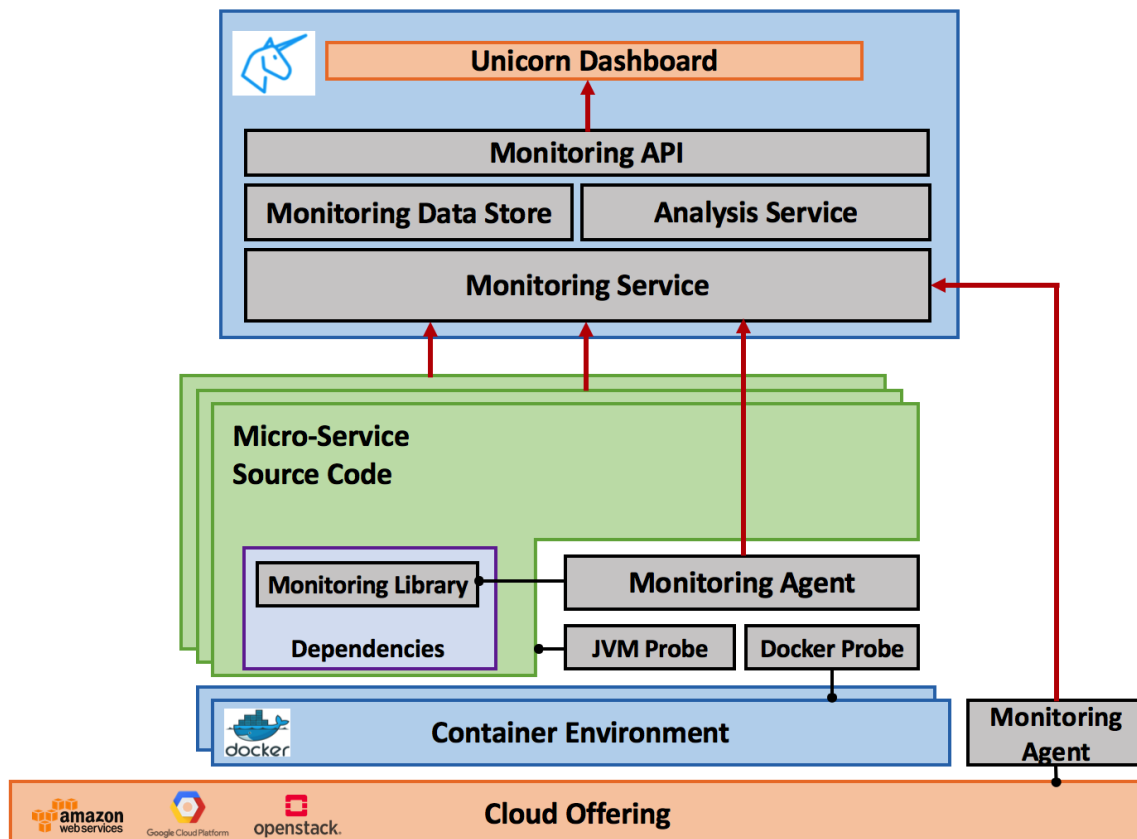
- Monitoring cloud and container level utilization
- Monitoring cloud application behavior and performance
- Metric collector development toolkit
- Access to historical and real-time monitoring metric data
- Runtime monitoring topology adaptation acknowledgement
- Monitoring rule language for metric composition, aggregation and grouping
- Code annotations

6.10.3 Components

- **Monitoring Agents:** lightweight entities deployable on any cloud element to be monitored responsible for coordinating and managing the metric collection process on the respective cloud element (e.g., container, VM), which includes aggregation and dissemination of monitoring data to the Monitoring Service over a secure control plane.

- **Monitoring Probes:** the actual metric collectors that adhere to a common metric collection interface with specific Monitoring Probe implementations gathering metrics from the underlying infrastructure.
- **Monitoring Library:** the source code annotation design library supporting application instrumentation for Unicorn compliant cloud applications.
- **Monitoring Service:** the entity easing the management of the monitoring infrastructure by providing scalable and multi-tenant monitoring alongside the Unicorn platform.
- **Monitoring Data Store:** a distributed and scalable data store with a high-performance indexing scheme for storing and extracting monitoring updates.
- **Monitoring REST API:** the entity responsible for managing and authorizing access to monitoring data stored in the Monitoring Data Store.

6.10.4 Architecture



6.10.5 How to enable Unicorn Monitoring

Warning: Currently the Unicorn Monitoring can be enabled for JAVA Maven Projects

To enable the Unicorn Monitoring in your application you have to follow the simple steps below

- **Step 1:** Add the Catascopia Monitoring dependency to your `.pom` file

```

1 <dependency>
2   <groupId>eu.unicornH2020.catascopia</groupId>
3   <artifactId>Catascopia-Agent</artifactId>
4   <version>0.0.3-SNAPSHOT</version>
5 </dependency>

```

- **Step 2:** Enable Catascopia Monitoring in your application either by code or annotation

An example SpringBoot DemoApplication should look like this:

```

1 @SpringBootApplication
2 public class DemoApplication {
3     public static void main(String[] args) throws Exception {
4         SpringApplication.run(DemoApplication.class, args);
5         Catascopia agent = new Catascopia();
6         agent.startMonitoring();
7     }
8 }

```

or

```

1 @SpringBootApplication
2 public class DemoApplication {
3
4     @UnicornCatascopiaMonitoring
5     public static void main(String[] args) throws Exception {
6         SpringApplication.run(DemoApplication.class, args);
7     }
8
9 }

```

At this point default monitoring is enabled and you will see monitoring data printed to your console.

- **Step 3:** For configuration, add a `catascopia.properties` file preferably under `/src/main/resources` and customize monitoring process

Table 1: Configurable Parameters

Parameter Name	Default Value	Description
app.id	myapp	Application Id
agent.id	myservice	Agent Id
agent.logging	false	Enable logging
agent.aggregation	false	Enable metric aggregation with window set to service connector
service.connector	PrintRawStreamConnector	Other options: PrintAggregatedStreamConnector (print aggregated stream), NetdataConnector
probe.config	local	There are three different types of probes for Catascopia: local: these are the probes embedded in the probe library part of Catascopia Monitoring and basically monitor the underlying JVM, dependency:these probes are not part of Catascopia Monitoring but are added to the classpath through mvn dependencies, remote:these probes are added from remote sources

Example of `catascopia.properties`

```

1  ### Catascopia application monitoring config file###
2
3  #application id
4  app.id: myapp
5
6  #agent id
7  agent.id: myservice
8
9  #enable logging
10 agent.logging: true
11
12 #enable metric aggregation with window set to connector dissemination periodicity
13 agent.aggregation:true
14
15 #service connector configuration
16 service.connector: NetdataConnector
17 service.connector.ip: 127.0.0.1
18 service.connector.port: 8125
19 service.connector.rate: 10000
20
21 #catascopia probes configuration
22 probes.config:custom-probe,dependency:eu.unicornH2020.catascopia.probe,100

```

6.10.6 Unicorn Catascopia Monitoring Probes

Monitoring Probes are dynamically pluggable to Monitoring Agents via the Agent's probe loader which embraces the class reflection paradigm to dynamically link, configure and instantiate Monitoring Probes at runtime in immutable execution environment.

In the Unicorn Monitoring Probe Repository there exists a number of publically available Monitoring Probes that can be used by users. To date, the Unicorn Monitoring Probe Repository hosts a JVM, J2EE and Docker Probe.

6.10.7 How to enable probes

- **Step 1** : Add Catascopia Monitoring to your application as demonstrated above.
- **Step 2** : Add the Maven dependency for the probe to your project's `.pom` file

```

1  <!-- example utilization of spring boot probe for catascopia monitoring -->
2  <dependency>
3    <groupId>eu.unicornH2020.catascopia</groupId>
4    <artifactId>SpringBootProbe</artifactId>
5    <version>0.0.1-SNAPSHOT</version>
6  </dependency>

```

- **Step 3**: Create a `probe.properties` file to alter the default probe configuration.

Table 2: Configurable Parameters for probe.properties

Parameter Name	Default Value	Description
service.endpoint	http://localhost	Service Endpoint
service.port	8080	Service Port
service.headers	(empty)	if multiple headers will be appended then they must be delimited by “;” e.g., X-MY_CUSTOM_API_KEY:1234;X-ANOTHER_HEADER:3845fgd85930dkf

- **Step 4:** In your application booter “import” the `CatascopiaMetricFilter` and `CatascopiaMetricProvider` and start `Catascopia`

```

1 @SpringBootApplication
2 @Import({CatascopiaMetricFilter.class, CatascopiaMetricProvider.class})
3 public class DemoApplication {
4
5     @UnicornCatascopiaMonitoring
6     public static void main(String[] args) throws Exception {
7         SpringApplication.run(DemoApplication.class, args);
8     }
9
10 }
```

6.10.8 How to develop custom probes

Nonetheless, Developers are free to create their own Monitoring Probes and Metrics, by adhering to the properties defined in the Monitoring Probe API which provides a common API interface and abstractions hiding the complexity of the underlying Probe functionality

The `ExampleProbe` below includes the definition of two `SimpleMetric`'s, denoted as `Metric1` and `Metric2`, that periodically report random integer and double values respectively. In this code snippet we also observe that for a user to develop a Monitoring Probe, she must only provide default values for the Probe periodicity and a name, a short description of the offered functionality and a concrete implementation of the `collect()` method which, as denoted by the name, defines how metric values are updated.

```

1 package eu.unicornH2020.catascopia.probe;
2
3
4 import java.util.Random;
5
6 import eu.unicornH2020.catascopia.probe.Probe;
7 import eu.unicornH2020.catascopia.probe.exceptions.CatascopiaMetricValueException;
8 import eu.unicornH2020.catascopia.probe.metricLibrary.CounterMetric;
9 import eu.unicornH2020.catascopia.probe.metricLibrary.SimpleMetric;
10 import eu.unicornH2020.catascopia.probe.metricLibrary.TimerMetric;
11
12 public class ExampleProbe extends Probe {
13
14     private static final int DEFAULT_SAMPLING_PERIOD = 10000;
15     private static final String DEFAULT_PROBE_NAME = "ExampleProbe";
16     private SimpleMetric<Double> metric1;
17     private SimpleMetric<Integer> metric2;
18     private CounterMetric metric4;
19 }
```

(continues on next page)

(continued from previous page)

```

20     public ExampleProbe(String name, long period) {
21         super(name, period);
22
23         this.metric1 = new SimpleMetric<Double>("m1", "%", "an example of a_
↪random percentage", true, 0.0, 100.0);
24         this.metric2 = new SimpleMetric<Integer>("m2", "", "an example of a_
↪random integer", true, 0, 10);
25
26         this.metric4 = new CounterMetric("m4", "an example of a counter metric",
↪ 0, 10, 1);
27
28         this.addMetricToProbe(this.metric1);
29         this.addMetricToProbe(this.metric2);
30         this.addMetricToProbe(this.metric4);
31     }
32
33     public ExampleProbe() {
34         this(DEFAULT_PROBE_NAME, DEFAULT_SAMPLING_PERIOD);
35     }
36
37     @Override
38     public String getDescription() {
39         return "An exemplary probe showcasing the offered functionality";
40     }
41
42     @Override
43     public void collect() throws CatascopiaMetricValueException {
44         TimerMetric tmetric = new TimerMetric("tmetric", "an example of a timer_
↪metric", this.getPeriod());
45         this.addMetricToProbe(tmetric);
46         tmetric.start();
47
48         Random r = new Random();
49
50         this.metric1.setValue(r.nextDouble() * 100);
51         this.metric2.setValue(r.nextInt(10));
52
53         try {
54             this.metric4.increment();
55         }
56         catch (CatascopiaMetricValueException e) {
57             e.printStackTrace();
58             this.metric4.reset();
59         }
60
61         try {
62             Thread.sleep(r.nextInt(1000));
63         }
64         catch (InterruptedException e) {
65             e.printStackTrace();
66         }
67
68         tmetric.finished();
69
70         // System.out.println(this.toJSON());
71     }
72

```

(continues on next page)

(continued from previous page)

```

73     public static void main(String[] args){
74         ExampleProbe p = new ExampleProbe();
75         p.activate();
76     }
77 }

```

6.10.9 Available Metrics Libraries

Table 3: Metrics Libraries

Metric Class	Description
CounterMetric	Emits a cumulative metric that represents a monotonically increasing counter
SimpleMetric	Emits a value for a referenced metric periodically.
TimerMetric	Emits the time consumed for the completion of a referenced task (e.g., API call).

6.11 Advanced Security Configuration - Unicorn Perimeter Security (IDS configuration)

6.11.1 Creation of a Ruleset

Unicorn Dashboard allows users to create their own Rulesets. The Detection Rules link opens a page with all the already available Rulesets (icmp, mysql, testDashboard), as is shown in the first Figure below. The Create new button leads the user to a new page where a new Ruleset can be created. What is needed is a name for the identification of the new set of rules and then the rules themselves. Each rule signature is written in the Rule Name box and the plus button adds the rule to the Ruleset. The second figure depicts the creation of a new such Ruleset named testVM_ruleset. A rule that raises an alert every time a packet comes from the public IP of the testVM is already included and a new rule raising an alert for every packet with the private IP of the testVM is about to be added. The third Figure shows again the already available Rulesets with the addition of the testVM_ruleset.

The screenshot shows the Unicorn Dashboard interface. On the left is a sidebar with navigation links: DASHBOARD, INSTANCES, APPLICATIONS, COMPONENTS, SSH KEYS, RESOURCES, DETECTION RULES (highlighted), PLUGINS, and CONFIGURATION. The main content area is titled 'Intrusion Detection Rulesets' and 'Intrusion Detection Rulesets Management'. It features a search bar, 'Filter' and 'Reset' buttons, and a table of existing rulesets.

Name	Date Created
icmp	10/12/2018 - 10:18
mysql	10/12/2018 - 10:18
testDashboard	18/02/2019 - 15:09

At the bottom of the page, there is a footer with the text: © 2019 UNICORN | [About](#) | [Terms of Service](#) | [Privacy](#)

The top screenshot shows the 'Intrusion Detection Rulesets | Edit' page. The left sidebar contains navigation links: DASHBOARD, INSTANCES, APPLICATIONS, COMPONENTS, SSH KEYS, RESOURCES, DETECTION RULES (selected), PLUGINS, and CONFIGURATION. The main content area has a breadcrumb 'Intrusion Detection Rulesets > Edit' and a title 'Intrusion Detection Rulesets Management'. A form for editing a ruleset is shown with the name 'testVM_ruleset'. Under 'Alerting Rules', there are two rules defined:

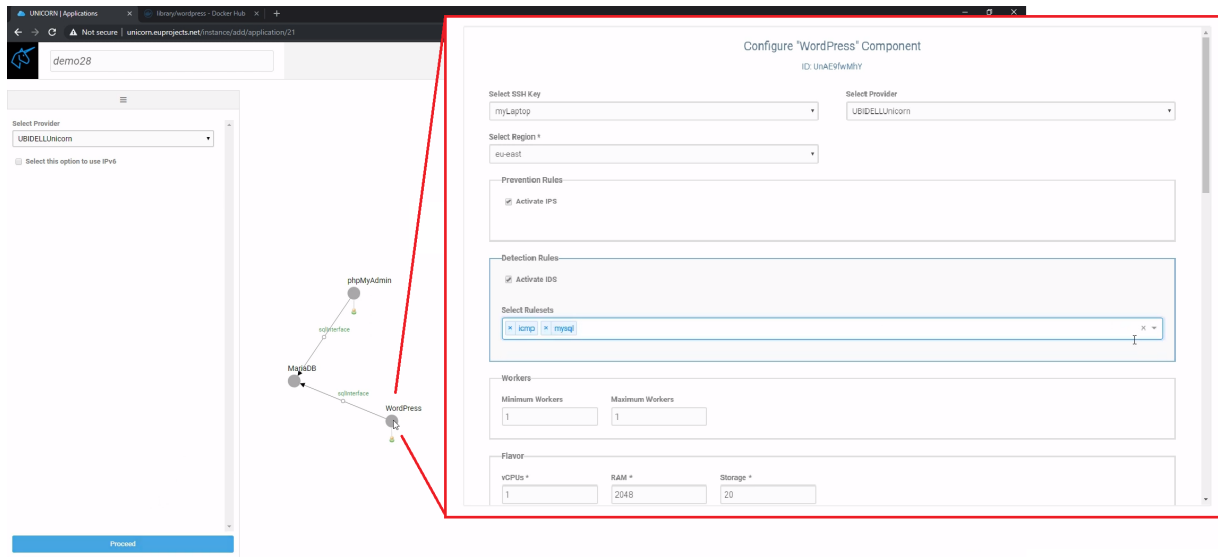
- Rule 1: alert tcp 10.142.0.4 any -> \$HOME_NET any (msg:"packet from 10.142.0.4"; GID:1; sid:10000007; rev:001;)
- Rule 2: alert tcp 34.73.160.110 any -> \$HOME_NET any (msg:"packet from 34.73.160.110"; GID:1; sid:10000006; rev:001;)

The bottom screenshot shows the 'Intrusion Detection Rulesets' overview page. It includes a 'Create new' button and a search bar. Below the search bar is a table of existing rulesets:

Name	Date Created
icmp	10/12/2018 - 10:18
mysql	10/12/2018 - 10:18
testDashboard	18/02/2019 - 15:09
testVM_ruleset	27/02/2019 - 10:59

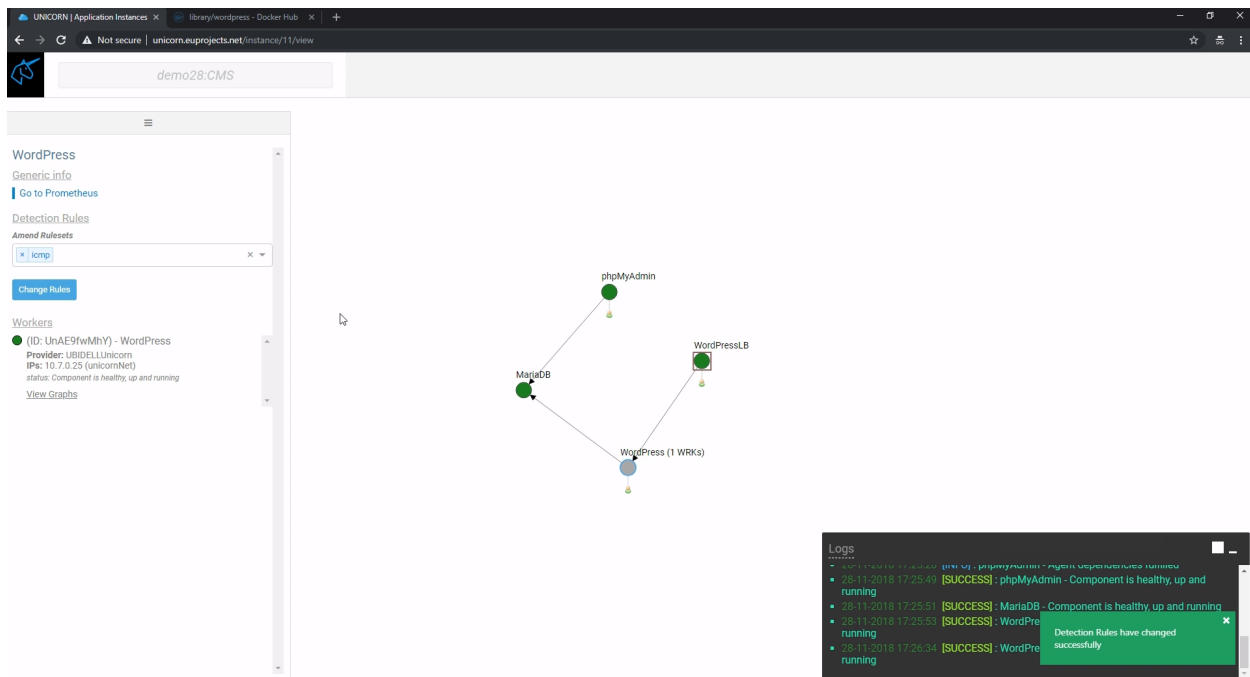
6.11.2 Configuration of an application component

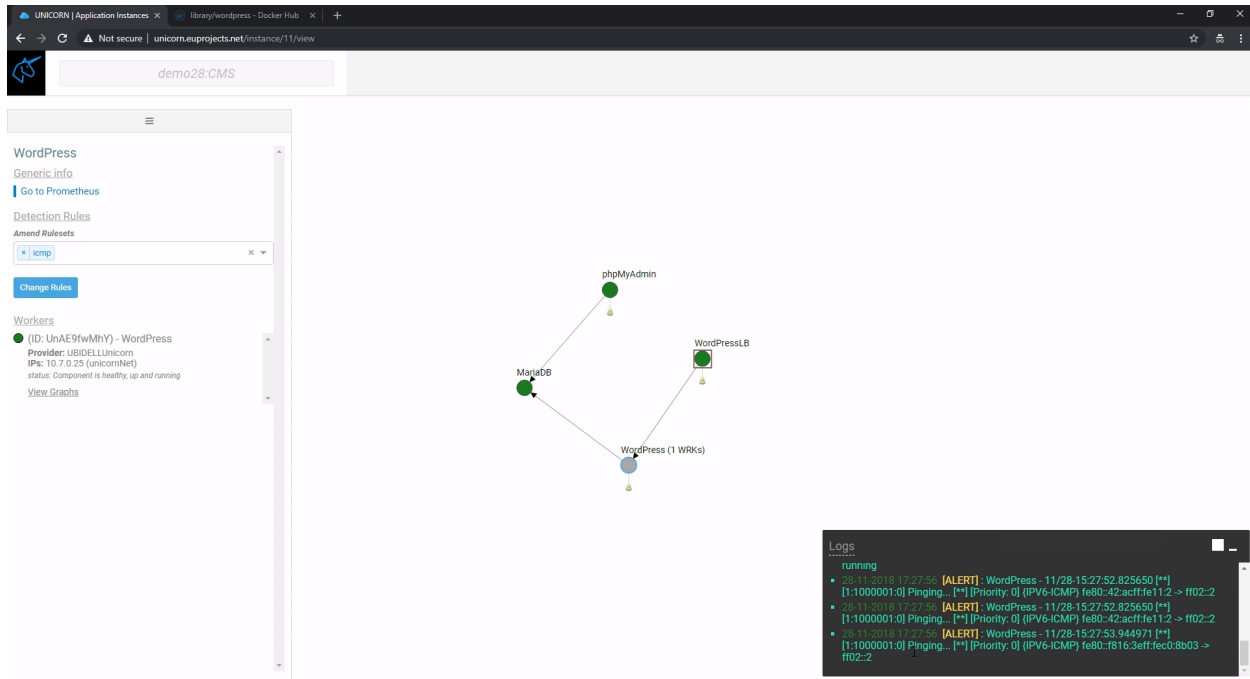
During the process of application initialization, a Unicorn user is able to configure any of the participating components of the application by clicking on the corresponding node of the service graph. The forth Figure depicts a service with three components (nodes). An overflow window appears when we click on the component that interests us, which allows for the activation and deactivation of the impended intrusion prevention and detection mechanisms (IPS, IDS) by selecting the appropriate check box. In the Figure both IPS and IDS are activated. IDS corresponds to a Snort instance that is containerized and deployed in the application execution environment. Such a Snort instance can be configured with different rule categories, which include a number of Snort rules, aka Rulesets. The process of creating a Snort Ruleset has already been described above. The user can choose, through the Unicorn dashboard, any of the Snort Rulesets from a dropdown menu. In this dropdown menu except from the Rulesets created by the users, there are predefined fine-grained Rulesets with rules that cover narrow vulnerability exploitation attempts, and broader Rulesets organized in two level. A detailed description of the offered Snort Rulesets is given in deliverable D4.2. As the forth Figure depicts, two Rulesets have been selected for Snort to use, icmp and mysql.



6.11.3 Set and unset Rulesets during application runtime

Provided that the IDS is already enabled during the application initialization, the user can set and unset Rulesets even when the application is running as can be seen in the fifth Figure. A Ruleset that raises alerts every time an ICMP packet reaches the node is selected in the fifth Figure. The green information box at the down right corner of the screen confirms the user's selection. As a result, when PING requests reach the node alerts are created by Snort and are shown in the Logs window (sixth Figure). Multiple Rulesets can be active at the same time.





6.12 Advanced Monitoring Configuration - Unicorn Analytic Service

6.12.1 Overview

Unicorn Analytic service allows the user to construct *insights*. An insight is a high-level analytic metric that is composed from raw metrics exposed by an application. The following is an example of an insight which calculates the cpu user utilization of every instance of a service called *service-streaming*

```
cpu_user_utilization =
  COMPUTE
    ARITHMETIC_MEAN( service-streaming:cpu_user, 60 SECONDS)
  EVERY 30 SECONDS
```

More specifically, the above insight specifies that every 30 seconds the system will calculate the last minute's average utilization of all containers running the *service-streaming*.

6.12.2 Features

- Easy to define valuable, simple or complex analytic expressions
- Translate the user-defined expressions to distributed execution engine language (*Spark Streaming*)
- Feeds the generated results to **elasticity policies** service

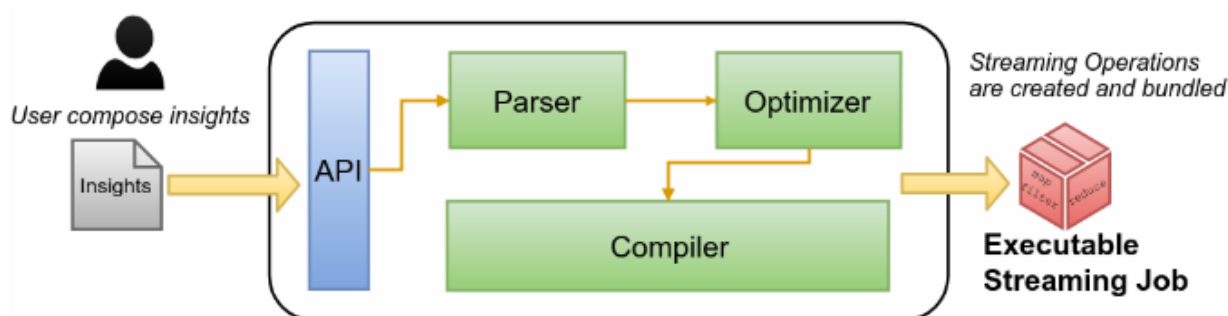
6.12.3 Components

- **Parser:** Parses the raw insights and generates the abstract model
- **Compiler:** Takes the abstract model and generates the low-level commands and optimizations

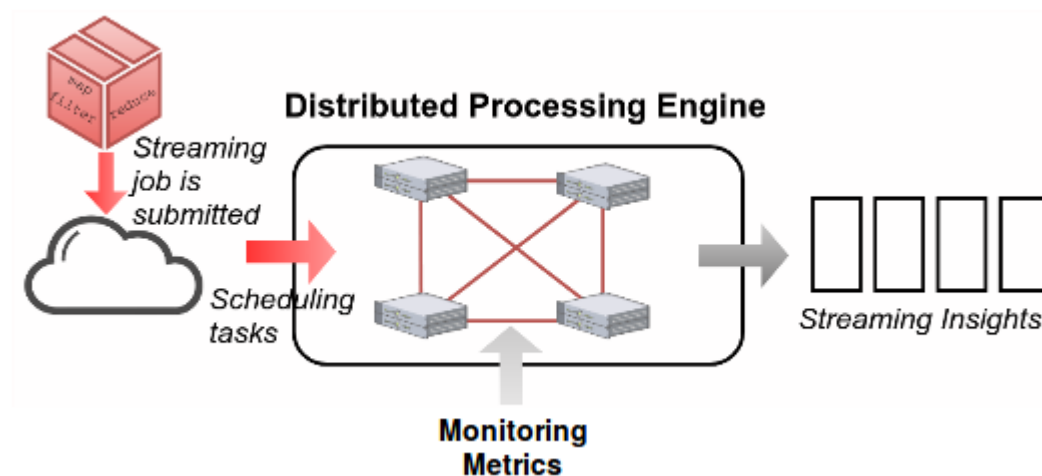
- **Manager:** Coordinates the compilation phase and submits the generated artifact to distributed streaming engine
- **Underline Streaming Engine:** Is responsible for real-time execution, in the first release we support only Spark

6.12.4 Architecture

Analytic Service supports users in composing analytic queries that are automatically translated and mapped to streaming operations suitable for running on distributed processing engines. This aids both advanced and inexperienced users to abstract and rigorously express complex analytics operations over streaming data, along with query constraints such as sample size and upper error-bounds for query execution to output approximate and in time answers. Thus, Analytic Service adopts a declarative programming paradigm, allowing users to describe analytic insights through a simple and powerful query modeling language.



Previous image depicts a high-level and abstract overview of the Analytic Service compiler. Users submit ad-hoc queries following the declarative query model and the system compiles these queries into low-level streaming commands. After that the system, automatically, submits the executable artifact to the underline distributed engine, as the following picture shows.



At the first version of analytic service, we integrated only the spark as underline distributed engine.

6.12.5 Constructing Insights

An insight is composed of 3 parts. The **COMPUTE** part, allows the user to compose simple or complex analytic expressions. The simple expressions can be either a window operation or an accumulated operation. The

windowed operation takes as input metric streams from a time period and performs an aggregation. For example: `ARITHMETIC_MEAN(service:metric, 30 SECONDS)`. On the other hand, an accumulated operation takes as input only the metric streams of interest and computes the results based on previous results. For example: `RUNNING_MEAN(service:metric)`

The **EVERY** part, specifies how often the calculations should occur (e.g., `EVERY 30 SECONDS`). Finally, there is an optional part **WITH**, which allows the user to specify different optimizations.

6.12.6 Available Operations

Windowed Operations

Windowed operations are used for aggregating values of interest in a time period in order to produce a summary statistic. Currently supported operations are:

- `ARITHMETIC_MEAN`
- `SUM`
- `COUNT`
- `MIN`
- `MAX`
- `SDEV`
- `VARIANCE`
- `GEOMETRIC_MEAN`
- `MODE`
- `MEDIAN`
- `PERCENTILE[p]`
- `TOP_K [k]`

Accumulated Operations

In contrast to windowed operations, these operations accumulate previous results in order to calculate the next result. Currently supported operations are:

- `RUNNING_MEAN`
- `RUNNING_SDEV`
- `RUNNING_MAX`
- `RUNNING_MIN`
- `EWMA` (Exponential Weighted Moving Average)

6.12.7 Insight Examples

Next, we will present some examples for useful insights from raw monitoring metrics.

EXAMPE 1

A useful insight that many companies need to monitor and take decisions on that is cpu utilization. So the following insight returns the average CPU utilization of a service for 30 seconds every 10 seconds.

```
cpu_utilization = COPMUTE (
    ARITHMETIC_MEAN( service:cpu_user, 30 SECONDS )
    + ARITHMETIC_MEAN( service:cpu_sys, 30 SECONDS )
) EVERY 10 SECONDS ;
```

EXAMPE 2

The free space in ram can be crucial for some applications, thus, the following expression gives us the RAM Average Usage for 10 minutes every 30 seconds.

```
ram_usage_per_service = COPMUTE
    ARITHMETIC_MEAN( service:ram , 10 MINUTES )
EVERY 30 SECONDS ;
```

EXAMPE 3

Next we present an insight for maximum number of HTTP Requests per Second for 10 minutes which computes every 30 seconds grouped by region. For devops and developers, who works on web-based applications, the peak of traffic in a specific region can be a critical factor.

```
http_requests_per_seconds_by_region = COPMUTE
    MAX( service:requests_per_seconds , 10 MINUTES ) BY Region
EVERY 30 SECONDS ;
```

EXAMPE 4

With the following example we can determine the difference between two serial time-windows (30 seconds) for user's cpu usage. This insight will be computed every 10 seconds.

```
cpu_usage_diff = COPMUTE (
    ARITHMETIC_MEAN( service:cpu_user, 30 SECONDS )
    - ARITHMETIC_MEAN( service:cpu_user, 30 SECONDS, 30 SECONDS )
) EVERY 10 SECONDS ;
```

6.12.8 Insight Optimizations

The optimizations allow users to:

1. Prioritize query execution over other queries so that when there is a high load influx, high priority queries are not delayed.
2. Enforce query execution over a sample of the available measurements to output immediate results. For the latter, users can denote the exact sample size as a percentage from the available measurements the sampling technique must obey when constructing the sample.

Insight Prioritization

In the following example, the user defines two queries: `cpu_utilization` and `ram_usage_per_service`. The `cpu_utilization` has salience 5 and `ram_usage_per_service` has salience 1. This means that, in high load influx, the `cpu_utilization` will be executed even 5 times more than `ram_usage_per_service`, thus the `cpu_utilization` will be more updated. If the query engine is stable, both queries will be executed as usual.

```
cpu_utilization = COPMUTE (
    ARITHMETIC_MEAN( service:cpu_user, 30 SECONDS )
    + ARITHMETIC_MEAN( service:cpu_sys, 30 SECONDS )
) EVERY 10 SECONDS WITH SALIENCE 5;

ram_usage_per_service = COPMUTE
    ARITHMETIC_MEAN( service:ram, 10 MINUTES )
EVERY 30 SECONDS WITH SALIENCE 1;
```

Sampling

In this example, the user wants to have a 25% sample of the query, so the system drops out the 25% from the input data.

```
cpu_utilization_sample = COPMUTE (
    ARITHMETIC_MEAN( service:cpu_user, 30 SECONDS )
    + ARITHMETIC_MEAN( service:cpu_sys, 30 SECONDS )
) EVERY 10 SECONDS WITH SAMPLE 0.25;
```

Complex Optimizations

Finally, we can have a compination of two these optimizations. Next, we present an insight which `SAMPLE 25%` and `SALIENCE 3`.

```
COPMUTE (
    ARITHMETIC_MEAN( service:cpu_user, 30 SECONDS )
    + ARITHMETIC_MEAN( service:cpu_sys, 30 SECONDS )
) EVERY 10 SECONDS WITH SAMPLE 0.25 AND SALIENCE 3;
```

6.12.9 Reference

You can find a more detailed description of our system in the following paper:

“StreamSight: A Query-Driven Framework for Streaming Analytics in Edge Computing”, Z. Georgiou, M. Symeonides, D. Trihinas, G. Pallis, M. D. Dikaiakos, 11th IEEE/ACM International Conference on Utility and Cloud Computing (UCC 2018), Zurich, Switzerland, Dec 2018.

6.13 Advanced Scaling Configuration - Decision-Making & Auto-Scaling service

6.13.1 Overview

The Decision Making & Auto-Scaling service allows the user to specify optimization strategies for adapting a cloud application on multiple cloud providers, based on cost, quality and performance preferences. This is achieved by

defining **elasticity policies** for adapting the cloud application both at the design-time of an application and during its runtime execution. There are two ways to define elasticity policies for the Unicorn application: i) via *High-level policies* that specify an optimization strategy per service; and ii) via *Low-level policies* that follow an *IF-THEN-ACTION* approach, where a scaling action is triggered when a set of conditions is satisfied. For this, the Decision Making & Auto-Scaling service continuously monitor these conditions at regular time intervals through application and infrastructure high-level analytic insights.

6.13.2 Features

- Define and Manage Elasticity policies for cost, quality and performance optimization of a UNICORN-enabled application
- Autonomous Runtime Monitor & Enforcement of Elasticity policies
- Resource-Aware & Transparent Multi-Cloud Elasticity Control
- Continuous assessment of elasticity policies and adaptation

6.13.3 Components

- **Elasticity Manager:** Allows the user to create, modify and remove elasticity policies
- **Elasticity Controller:** Enforces the scaling rules of a cloud application
- **Analysis Service:** Access to real-time analytic metric data
- **Monitoring Service:** Access to historical monitoring metric data
- **Resource Manager:** Retrieve elasticity capabilities and propose scaling decisions

6.13.4 Architecture

6.13.5 How to use?

The elasticity policies can be defined both at the design time of the application through the UNICORN docker-compose file, and during runtime via the service graph of the application.

The following example shows a low-level scale out policy for the **streaming_svc** service:

```
scale_out_streaming_svc =  
  WHEN  
    average_requests_5m > 100  
    AND  
    average_cpu_5m    > 80  
  THEN  
    SCALE OUT ( 1 service_streaming WITH 30 SECONDS COOLDOWN)
```

The first part (**WHEN**) contains two conditions. These conditions specify that the `average_requests_5m` must be greater than 100, while also the `average_cpu_5m` must exceed 80(%) in order to trigger the scaling action. The action (specified in **THEN** part), designates that 1 more `service_streaming` service should be provisioned. Note that the `WITH 30 SECONDS COOLDOWN` is a configurable time period which is used to give time to the system to provision/de-provision new resources and absorb any changes, in order to prevent false scaling alerts.

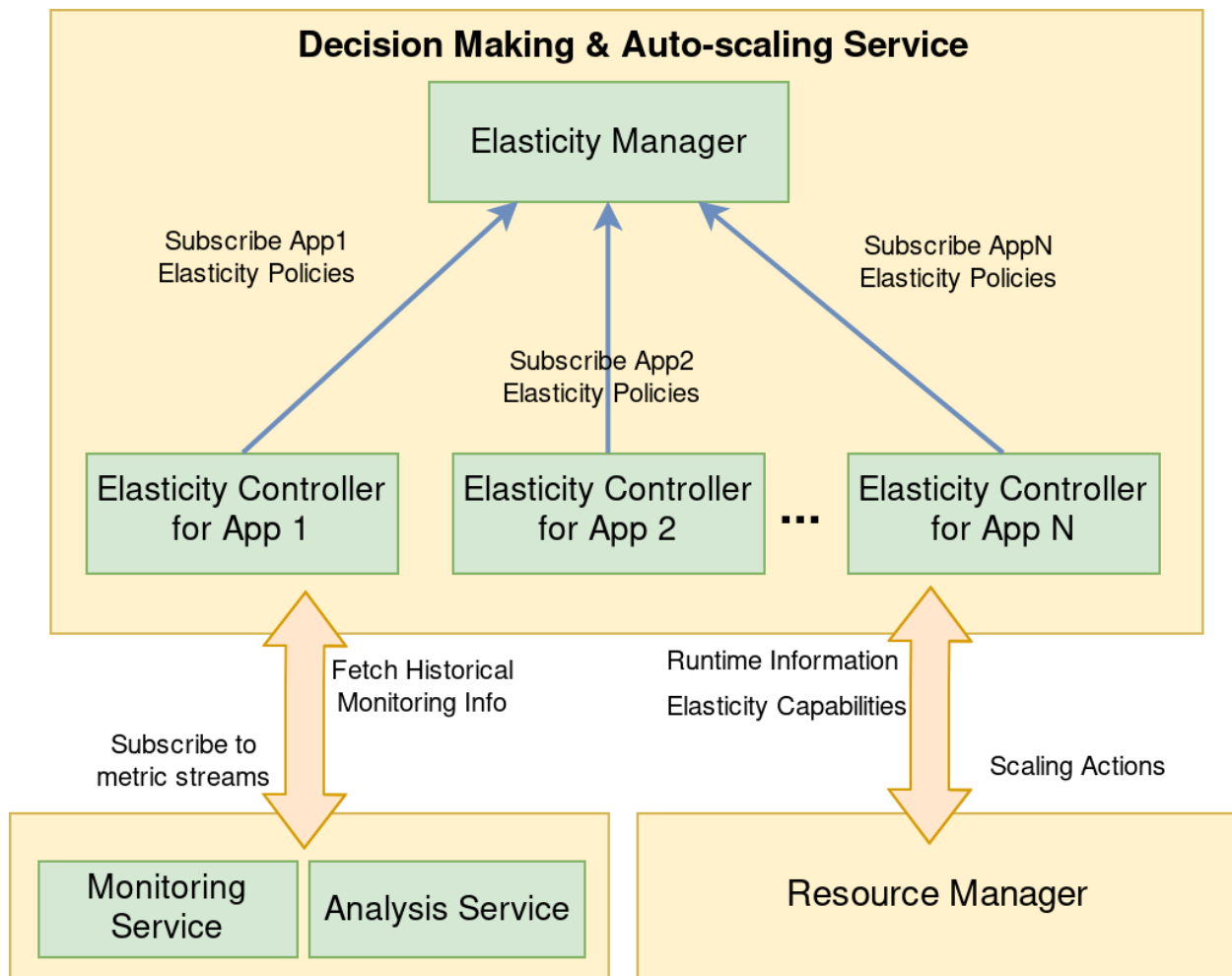


Fig. 2: Decision-Making & Auto-Scaling Service Reference Architecture

6.13.6 Elasticity Policy

An elasticity policy can be constructed in two different ways. The first way is the *high-level policy* which allows the user to specify for a given service of the UNICORN Service Graph, a high-level optimization strategy. Currently, the language supports three optimization strategies per service:

- Cost optimization
- Availability optimization
- Balance between cost and availability

Note that the default strategy is enabled when users do not specify the *Awareness* construct.

```
elasticity_streaming_svc =  
  SET SERVICE streaming_svc AWARE ON COST USING avg_cpu_streaming_utilization
```

The above segment, shows an example of a high-level elasticity policy. The policy specifies that the service *streaming_svc* from the UNICORN Service graph should be aware on cost using the average CPU utilization of that service. It is important to mention that the analytic insight *avg_cpu_streaming_utilization* must have the following properties:

- Indicate the current workload of a service.
- Its value should be decreased when applying a scaling out action, i.e., add more service instances.
- Its value should be increased when applying a scaling in action, i.e., remove service instances.

The second way to define elasticity policies is to construct a low-level policy. This feature is recommended for advanced users as they can express policies with a higher degree of detail. This policy is composed from a set of conditions and a scaling action. Each condition contains an expression of an *analytic insight* (left hand-side) and a number (right hand-side) which are operated by a binary operation (e.g., <, >, ==).

Enablers

The Decision-Making and Auto-Scaling service offers to users the ability to activate various optimization modules, namely *Enablers*. Currently, UNICORN supports the following two enablers:

- UNICORN Predictor
- UNICORN Decision Timeframe Sensitivity

UNICORN Predictor

This enabler predicts the values of the analytic insights specified in the *ElasticityTrigger* construct. The segment below shows how the enabler is activated. The horizon parameter specifies how far to predict the values (5 minutes), the confidence specifies the maximum acceptable error (95%), and the history parameter denotes how far in the past historic points are considered (2 weeks).

```
predicted_scale_out =  
  WHEN avg_cpu_streaming_utilization > 80  
  ENABLE (UNICORN_PREDICTOR[horizon=300, confidence=0.95, history=10080])  
  PERFORM SCALE OUT ( 1 streaming_svc WITH 5 MINUTES COOLDOWN)
```

UNICORN Decision Timeframe Sensitivity

This enabler can be activated as shown in the segment below, to enable the dynamic change of the decision timeframe. The Decision timeframe is the time period of the aggregation function used in the analytic insight

(*avg_cpu_streaming_utilization*). It uses a confidence value, with higher values resulting to larger time periods. The benefit of this approach is that the user doesn't need to manually find and set the aggregation period of the metric streams used for scaling decisions.

```
sensitivity_scale_out =  
    WHEN avg_cpu_streaming_utilization > 80  
    ENABLE (UNICORN_SENSITIVITY[confidence=0.95])  
    PERFORM SCALE OUT ( 1 streaming_svc WITH 5 MINUTES COOLDOWN)
```

Available Actions

Currently two actions are supported for horizontal scalability. The `SCALE OUT` action, which provisions a new instance of a service and the `SCALE IN` action, that de-provisions an existing instance of a service.

Elasticity Grammar

The Table below presents the Elasticity language grammar rules in EBNF syntax.

ElasticityPolicy	::=	<ElasticityPolicyID> “.” (<HighLevelPolicy> <LowLevelPolicy> “WITH PRIORITY” <Priority>
ElasticityPolicyID	::=	<String>
HighLevelPolicy	::=	“SET” <Service> [<Awareness>] “USING” <InsightID>
Service	::=	<GraphID> “.” <GraphInstanceID> “.” <ServiceID>
Awareness	::=	“AWARE ON” <Strategy>
Strategy	::=	“COST” “AVAILABILITY”
LowLevelPolicy	::=	“WHEN” <ElasticityTrigger> [<Enablers>] “PERFORM” <ElasticityAction>
ElasticityTrigger	::=	<ElasticityCondition> (“AND” <ElasticityCondition>)*
ElasticityCondition	::=	<InsightID> RelOp <Number>)
Enablers	::=	“ENABLE” “(” <Enabler> “(” <Enabler>)* “)”
Enabler	::=	<EnablerName> “[” <Parameters>”]
EnablerName	::=	<String>
Parameters	::=	<KeyValue> (“,” <KeyValue>)
KeyValue	::=	<String> “=” <String>
ElasticityAction	::=	<ReplicationAction> <InformationAction>
ReplicationAction	::=	(“SCALE OUT” “SCALE IN”) <PlacementConfig>
PlacementConfig	::=	“(” <PositiveInt> <Resource> <Cooldown> “)”
PositiveInt	::=	[1-9]([0-9])*
Resource	::=	<Service> <Service> IN <Cluster>
Cluster	::=	<String>
Cooldown	::=	<PositiveInt> <TimeUnit> “COOLDOWN”
TimeUnit	::=	“MILLISECONDS” “SECONDS” “MINUTES” “HOURS”