

Ilari Räisänen

# **CLOUD REPATRIATION: FACTORS AND STRATEGIES FOR SUCCESSFULLY AND BENEFICIALLY LEAVING THE PUBLIC CLOUD**

Master's Thesis

Faculty of Information Technology and Communication Sciences

Examiners: University Lecturer Antti Sand

University Lecturer Timo Poranen

January 2025

## ABSTRACT

Ilari Räisänen: Cloud Repatriation: factors and strategies for successfully and beneficially leaving the public cloud  
Master's Thesis  
Tampere University  
Master's Programme in Information Technology  
January 2025

---

Cloud repatriation is the process of migrating away from public cloud. The subject of the migration can be data, workloads, or IT infrastructure either partly or completely. The target environment in cloud repatriation is either on-premises or a data center in a colocation center. This thesis addresses two research questions: what were the promised benefits of public cloud and when, if not always, are they realized, and what are the key factors and strategies that allow successful and beneficial cloud repatriation.

This thesis includes the review of existing literature and industry practices. It also inspects the common motivations for repatriation, including financial reasons, improved control, regulatory compliance and improved performance and latency for specific types of workloads. The analysis includes success factors such as understanding what makes a workload or data suitable for repatriation, using cloud-agnostic practices, and different migration strategies tailored for organizational needs. The thesis concludes that cloud repatriation is a nuanced process, needs thorough planning, and requires understanding the impact on both IT and business level.

Keywords: Cloud, public cloud, on-premises, cloud exit, cloud repatriation, migration, on-premises, infrastructure

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

# TIIVISTELMÄ

Ilari Räisänen: Cloud Repatriation: tekijät ja strategiat onnistuneeseen ja hyödylliseen julkisesta pilvestä poistumiseen

Diplomityö

Tampereen yliopisto

Tietotekniikan DI-ohjelma

Tammikuu 2025

---

Pilvipalvelun kotiuttaminen (cloud repatriation) on prosessi, jossa suoritetaan julkisesta pilvestä poistuminen. Siirron kohteena voi olla data, työkuormat (workloads) tai IT-infrastruktuuri joko osittain tai kokonaan. Kohdeympäristö pilvipalvelun kotiuttamisessa on joko organisaation omissa tiloissa sijaitseva infrastruktuuri (on-premises) tai konesali, joka sijaitsee yhteisomistuksessa olevassa datakeskuksessa (colocation center). Tämä diplomityö käsittelee kahta tutkimuskysymystä, mitkä olivat julkisen pilven luvattuja hyötyjä milloin ne toteutuvat, sekä mitkä ovat avaintekijät ja -strategiat, jotka mahdollistavat onnistuneen ja hyödyllisen pilvipalvelun kotiuttamisen.

Tässä diplomityössä tarkastellaan olemassa olevaa kirjallisuutta ja alan käytäntöjä. Lisäksi työ tarkastelee yleisiä kotiuttamisen syitä, kuten taloudellisia perusteita, tiukempaa kontrollia ympäristöstä, regulaatioiden noudattamista sekä suorituskyvyn ja viiveen parantamista tietyn tyyppisten työkuormien kohdalla. Analyysi sisältää onnistumistekijöitä, kuten sen ymmärtämistä, mitkä työkuormat tai data soveltuvat kotiuttamiseen, pilvipalveluista riippumattomien toimintatapojen käyttöönottoa ja erilaisten organisaatioiden tarpeisiin räätälöityjä strategioita. Diplomityön johtopäätös on, että pilvipalvelun kotiuttaminen on monisyinen prosessi, joka vaatii perusteellista suunnittelua ja edellyttää vaikutusten ymmärtämistä sekä IT- että liiketoimintatasolla.

Avainsanat: Pilvi, julkinen pilvi, pilvipalveluista poistuminen, pilvipalvelusta kotiuttaminen, migraatio, on-premises, infrastruktuuri

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck -ohjelmalla.

## **PREFACE**

I would like to thank all my fellow students who have made this phase of my life enjoyable, to say the least. I would also like to thank my family, my wonderful girlfriend, and other people close to me for their support, and even for their existence. I would also like to thank my employer, Evitec Solutions, for the opportunity to research an interesting topic during working hours.

Thus ends a journey of five and a half years, and I'm glad to move forward.

Tampereella, 11th January 2025

Ilari Räisänen

## USE OF AI IN THESIS

The AI tools used in my thesis and the purpose of their use has been described below:

**Name of the tool (and version):** ChatGPT (GPT-4o)

**Purpose of use and the part in which it was used:** I used ChatGPT to organize chapters and text for better flow, and to help me get inspiration when I was lacking it. I also used it to help me fix the syntax of table 4.1 in LaTeX. ChatGPT was used for these purposes throughout the whole thesis, except for chapters discussion and conclusion.

**Name of the tool (and version):** Concensus

**Purpose of use and the part in which it was used:** In the beginning of the writing process I also used Concensus, which is an academic search engine with AI features. It was used to find relevant research while not yet knowing all the keywords that could be useful.

I am aware that I am totally responsible for the entire content of the thesis, including the parts generated by AI, and accept the responsibility for any violations of the ethical standards of publications.

## CONTENTS

1.	Introduction . . . . .	1
2.	Background. . . . .	4
2.1	Public Cloud. . . . .	4
2.2	Private Cloud . . . . .	5
2.3	On-Premises . . . . .	5
2.4	Hybrid Cloud . . . . .	7
2.5	Cloud delivery model . . . . .	7
2.6	Virtualization and Containers . . . . .	9
2.7	Promises of Cloud . . . . .	12
2.8	Cloud repatriation: Migrating away from the public cloud . . . . .	13
3.	Benefits and Risks associated with the transition . . . . .	14
3.1	Costs . . . . .	14
3.1.1	Direct costs . . . . .	14
3.1.2	Indirect costs. . . . .	15
3.1.3	Opportunity costs . . . . .	15
3.2	Increased Control . . . . .	16
3.3	Effect on performance and latency . . . . .	16
3.4	Effect on operational Efficiency and Agility . . . . .	17
3.5	Effect on security and compliance . . . . .	18
3.6	Long Term Implications . . . . .	18
4.	Factors and strategies for repatriating from public cloud . . . . .	19
4.1	Motivations for the repatriation . . . . .	19
4.2	Repatriation scope . . . . .	20
4.3	Repatriation timeline . . . . .	21
4.4	Migration strategies . . . . .	23
4.5	Choosing Tools and Services . . . . .	25
4.6	Cost analysis . . . . .	29
4.7	Hardware and Infrastructure . . . . .	30
4.8	Other factors . . . . .	31
5.	Discussion . . . . .	33
6.	Conclusion . . . . .	36
	References. . . . .	37

## ABBREVIATIONS

On-prem	On-premises, hardware and software that is hosted on-site
SaaS	Software-as-a-Service
IaaS	Infrastructure-as-a-Service
PaaS	Platform-as-a-Service
CapEx	Capital Expenditure
OpEx	Operational Expenditure
CSP	Cloud Service Provider
AWS	Amazon Web Services
GCP	Google Cloud Platform
OCI	Oracle Cloud Infrastructure
RDS	Amazon's Relational Database Service
IaC	Infrastructure as Code

# 1. INTRODUCTION

Cloud computing has for a long time gained more traction, and enterprises are utilizing public cloud platforms for their cost-efficiency and flexibility. The move from on-premises-infrastructure - where an organization uses their own datacenters - to cloud, was driven by the promise of reduced costs, increased accessibility, superior service quality, improved efficiency, and heightened security. [48]

Before the cloud, computing was centralized. Large organizations like governments, educational institutes and major corporations owned and operated massive mainframes. Only largest organizations could afford their own mainframe, leading to the sharing of resources for smaller organizations. Over time, hardware got cheaper and even the smaller organizations could purchase their own on-premises infrastructure. [13]

In this century, organizations are moving back to using shared resources offered by public cloud service providers. With the launch of Amazon Web Services (AWS) in 2002, the IT landscape started to go through a pivotal phase. In 2006, Amazon Web Services released Elastic Compute 2 (EC2), which revolutionized the industry, as it offering scalable virtual computing that even smaller organizations could afford. This period marked a critical shift, as organizations realized the potential of renting instead of owning computing storage and power. For many organizations, this meant large potential savings from reducing the need for expensive on-premises infrastructure.

The critical factors leading to the widespread adoption of cloud computing included:

- Cost Savings from avoiding capital expenditure from hardware and operational costs.
- Scalability and flexibility based on demand for varying loads and growing companies.
- Focusing resources on core business.
- Better global reach without the need for internationally spread physical infrastructure.

Public cloud platforms like AWS have evolved to offer a diverse array of services, fulfilling the promises that led to the widespread adoption of cloud computing. However, research from 451 Research (S&P Global Inc.) shows that recently an increasing amount of or-

ganizations have been looking for tailored and flexible approaches to have their specific cost, performance and governance requirements met, instead of one-size-fits-all solution for their infrastructure [37]. As organizations' experience with cloud grows, they become better at finding the best "execution venue" for their workloads. This can be in the public cloud, on-premise or remote-infrastructure private cloud, or a hybrid-solution, which is a combination of the options. The majority of especially the large organizations don't rely on a single solution to cover their whole IT needs, but rather a combination. This means that many organizations that formerly had their entire IT in the public cloud, have needed to move some workloads and applications, for example, to other venues. This transition from the public cloud is called cloud repatriation, or alternatively cloud exit. [37]

Kiran Jewargi defines cloud repatriation as "the process of transferring IT workloads or applications from public clouds to onpremises or co-located data centers". [32] Another term often used with the same definition is cloud exit. However, especially the term cloud exit can often be misleading, as it implies a complete and full transition away from the cloud altogether. The shift can in fact be for only some workloads or applications, or might be only temporary. [37]

Successfully transitioning a workload, application or even a whole organization requires careful analysis of the benefits and risks, as well as creating and following a strategy. Knowing the key factors and strategies that enable successful and beneficial repatriation from public cloud is important to minimize the risks and to maximize the benefits for the organization performing the repatriation. This thesis's purpose is to identify these key factors and strategies, and find their effects on the potential risks and benefits. Different reasons to leave the cloud are explored, and the key factors and strategies of repatriation are associated to those reasons, for the purpose of providing a guide that helps to find the right approach to leaving the cloud based for each situation. Additionally, it is evaluated whether the original promises of cloud have indeed been realized.

This thesis relies on a literature review on existing academic literature, and the examination of various reports, surveys, practical guides and other resources on cloud repatriation. Exploring current situation of research on the topic helps to position the information from the examined resources within the existing academic context. This method provides a robust framework for creating insight into cloud repatriation. As this thesis is in the technical field, the practical findings are especially important.

The research questions this paper aims to answer are:

- **RQ1:** What were the promised benefits of public cloud and when, if not always, are they realized?
- **RQ2:** What are the key factors and strategies that allow successful and beneficial cloud repatriation?

The first part of the first research question is rather trivial to answer. The promised benefits of public cloud are found from virtually every cloud service providers websites. Finding out when they are realized is more difficult, as it requires looking at survey data from companies. Because the promises are made in comparisons to on-premises environments, evaluating them also requires the comparing between the models.

The second question will require a deep analysis of our findings. The main aspects considered in the analysis will be divided into IT level and company level.

On IT level, this paper examines:

- Cost considerations: what direct and indirect costs does the transition present?
- Performance issues
- Security concerns
- Data sovereignty and compliance

On the company level, this paper examines:

- Industry of the organization
- Size of the organization
- Geography
- Previous IT infrastructure models

Inspecting the effects on both IT level and on the more general company level provides a more meaningful picture of the transition's impact, while still keeping this paper focused on the technical field of study.

## 2. BACKGROUND

### 2.1 Public Cloud

Public cloud is a model for deploying and managing IT infrastructure. Public cloud is defined as computing services offered by a third party over the public internet [60]. The most common public cloud service providers are Amazon (Amazon Web Services, AWS) Microsoft (Azure), Alphabet (Google Cloud Platform, GCP). Large-scale public cloud providers, such as the forementioned, are often called hyper-scalers.

#### Hyper-scale providers

Hyper-scale providers offer the ability for the platform to scale massively with vast amounts of computing resources to meet the needs of the largest companies. Hyper-scale providers are also associated with a wide range of services and technologies that offer ease of automation and high efficiency.

Hyper scale providers offer:

- The ability for scaling massively with vast computing resources.
- The capacity for the needs of the largest organizations.
- A wide range of services and technologies to ease automation and provide high efficiency.

Providers like Microsoft Azure promise that public cloud can help companies save in costs from having to purchase, manage and maintain on-premises infrastructure [60]. Public clouds also utilize multitenancy, which means that multiple different customers use the same hardware, even though their virtualized environments are separated and the customers are not aware of each other. Multitenancy makes cloud computing more efficient, as resources don't have to be reserved for a single customer. This means less idle time for resources, which in turn means less resources needed for larger amount of customers. More efficient use of resources means that the cloud vendor can offer services for a lower cost, which is one of the reasons for the booming popularity of the cloud. [65]

## 2.2 Private Cloud

In their 2009 paper *Above the Clouds: A Berkeley View of Cloud Computing*, Armbrust and others define private cloud as internal datacenters of an organization that are not publically available [9]. Private cloud can be hosted on an on-premises datacenter or at a third-party service provider. The third party service provider usually either rents room for the customer's datacenters or rents and manages datacenters. Having a private cloud allows the organization to retain the ownership of the data and increase data protection and data privacy. As private clouds are focused on solving the problems of their customers, their solutions are smaller and more affordable. [57]

A private cloud can also be hosted by a cloud vendor as a virtual private cloud. A virtual private cloud offered by a cloud vendor is much like using their normal services, but without multitenancy. This means that the drawbacks and benefits are similar to using a public cloud. Compared to a regular private cloud, the customer will have little to no control over hardware and security protocols. A virtual private cloud is logically isolated from other tenants on the network level, but the physical hardware is shared. This means that isolation is achieved by virtualization. With a virtual private cloud comes with the benefits of public cloud, like rapid scalability, redundancy and lower up-front investments [61]. This means that it might be a good option for organizations that need to avoid multitenancy for regulatory compliance reasons.

## 2.3 On-Premises

In the context of computing, on-premises software is defined as local software that is installed and runs in the organization's own IT environment physically located on the site of the organization using it. This environment can be managed either by the organization itself or by another provider. The organization is responsible for maintaining the infrastructure [16]. Common aspects like scaling, redundancy, licenses, and hardware failures need to be thought of. This also gives the organization more control, which can be desired, especially with data and security - Among the organizations that moved away from the cloud surveyed by 451 Research, 36,2 % said that the primary reasoning behind moving from public cloud was information security concerns, and 22,6 % said it was data locality/data sovereignty issues [49]. The cost of on-premises usually peaks in the beginning, as upfront capital expenditure (CapEx) for purchasing hardware, licenses for software, and potentially finding and hiring staff for maintenance is high. A private cloud can be on-premises, if the environment is designed to offer cloud computing features. Often the terms private cloud and on-premises datacenters are used interchangeably [44]. In this thesis, a private cloud refers to data center and resources that can reside either on-premises of the organization or at a third party service provider's premises, such as a colocation center.

<b>Aspect</b>	<b>Public Cloud</b>	<b>Private Cloud</b>	<b>On-Premises</b>
<b>Location</b>	Hosted off-site by a third-party service provider.	Can be hosted on-site or off-site; infrastructure is dedicated to a single organization.	Physically located on-site, within the organization's own facilities.
<b>Resource Sharing</b>	Multi-tenancy: Resources are shared among multiple organizations	Resources are not shared; dedicated to one organization.	Resources are fully dedicated to the organization and managed internally.
<b>Management</b>	Managed by the cloud provider, reducing the IT burden on the organization.	Can be managed by the organization or a third party.	Managed by the organization, requiring in-house IT expertise.
<b>Control and Security</b>	Controlled by the provider, but with configurable security options. Compliance and security shared between client and vendor.	High level of control and security, especially if hosted on-premises.	Maximum control over the environment and data, supporting stringent compliance requirements.
<b>Cost</b>	Typically operates on a pay-as-you-go model, leading to more predictable operational expenses (OpEx).	Mix of capital expenditure (CapEx) (if on-premises) and OpEx, with potential for cost efficiency through resource optimization.	Upfront CapEx and ongoing operational expenses (OpEx) for maintenance and staff.

**Table 2.1.** Differences between Public Cloud, Private Cloud, and On-Premises.

As shown in Table 2.1, each model has its own benefits and drawbacks on aspects like location, resource sharing, management, control and security, and cost. While public cloud is always hosted on the service provider's premises, private cloud can be hosted either at the premises of the organization that uses it, or at off-site location such as a colocation center. On-premises datacenters are by definition hosted on the premises of the organization. As the hardware that is being used to host a public cloud is owned by the service provider, the organization that uses the public cloud doesn't carry as big of a burden on managing the infrastructure as in the case of private cloud or on-premises model. However, this also means that the organization has less control over resource sharing and on what hardware is used, for example. Costs of the public cloud model tend to be cheaper in capital expenditure.

Especially larger organizations with the required resources often opt for a combination of these models, which enables them to leverage both the control of private cloud or on-

premises model, and the flexibility and scalability of public cloud. This approach, which is known as hybrid cloud, allows organizations to run their workloads and store their data with optimal infrastructure.

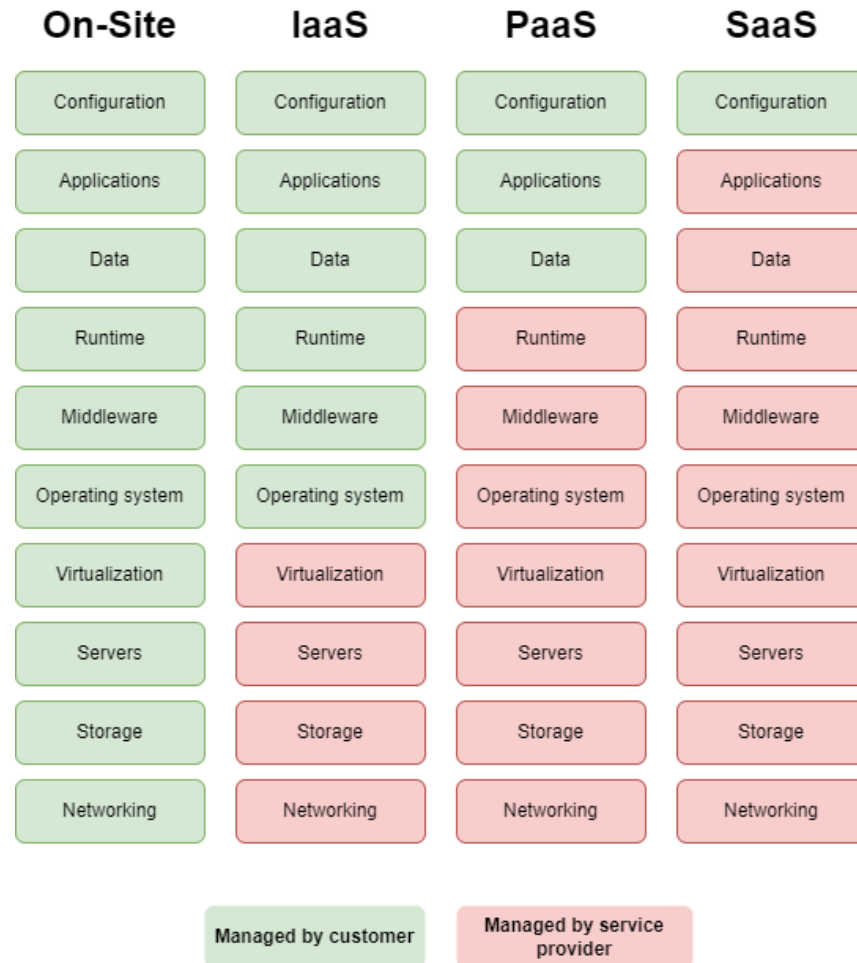
## **2.4 Hybrid Cloud**

Hybrid Cloud is the combination of private and public cloud infrastructures. It allows organizations to horizontally scale their infrastructure with public cloud for performance gain if necessary. It can also be used to lower the up-front costs of on-premises and private cloud infrastructure, while still gaining the benefits of private cloud [35]. Hybrid cloud can be a practical solution in situations where the organization has some requirements that only a private cloud can meet, such as restrictive data regulations, but can't invest heavily in private cloud or on-premises infrastructure, or finds public cloud more beneficial for their non-sensitive workloads and data. Another situation for hybrid cloud would be if the organization expects to grow at some point, and needs the possibility for scaling rapidly, and they already have large investments in on-premises infrastructure.

According to the Flexera 2023 State of the Cloud report [33], hybrid cloud is the approach for a majority of organizations. Multi-cloud, which means using multiple different clouds, was the strategy for 87 % of organizations. As a subset of multi-cloud, 72 % of the organizations reported utilizing hybrid cloud. [33] When repatriating from public cloud, the goal is often hybrid cloud [42].

## **2.5 Cloud delivery model**

Cloud delivery model essentially describes the level of abstraction. The three most commonly recognized models are Software-as-a-Service (SaaS), Platform-as-a-Service (PaaS) and Infrastructure-as-a-Service (IaaS). Each of these can be deployed in both public and private cloud environments. Public SaaS is when a vendor hosts a software in a publicly accessible cloud. Private SaaS is when an organization hosts software in its own private cloud on-site, restricted to the organization's internal network. Public PaaS includes development platforms that are accessible over the internet. Private PaaS would be a platform such as Red Hat OpenShift in a private cloud. Similarly, public IaaS offers virtualized resources over the public internet, and private IaaS mimics similar provisioning of resources while staying on the organization's private cloud. [28]



**Figure 2.1.** Cloud service models' levels of responsibility compared to on-site [28]

As shown in the Figure 2.1, SaaS is the most comprehensive of the service models. The organization is responsible for only the configuration of the software, the service provider handles the rest. Some of the most known SaaS products are Outlook, Slack and Netflix. SaaS means an entire application that is owned, deployed and managed by the service provider. The software is also hosted by the service provider, and accessed by the customer over the internet with a web browser. There is no need for installation of the software on the customer's devices, and group access is convenient. SaaS-based applications are very widely adopted, as for example email services accessed through the web browser are counted into the category. SaaS is also convenient for organizations with less resources to maintain software, and are often priced so that alternatives are rarely viable, if cost is the only concern. SaaS does provide less control and might not comply with some regulations in highly regulated industries.

PaaS is in between SaaS and IaaS, in terms of control and responsibility. The organization is responsible for configuration, applications that they host on the platform, and the data that is stored and used by the applications. PaaS provides the customer with an on-demand environment for developing, testing, delivering and maintaining software. PaaS

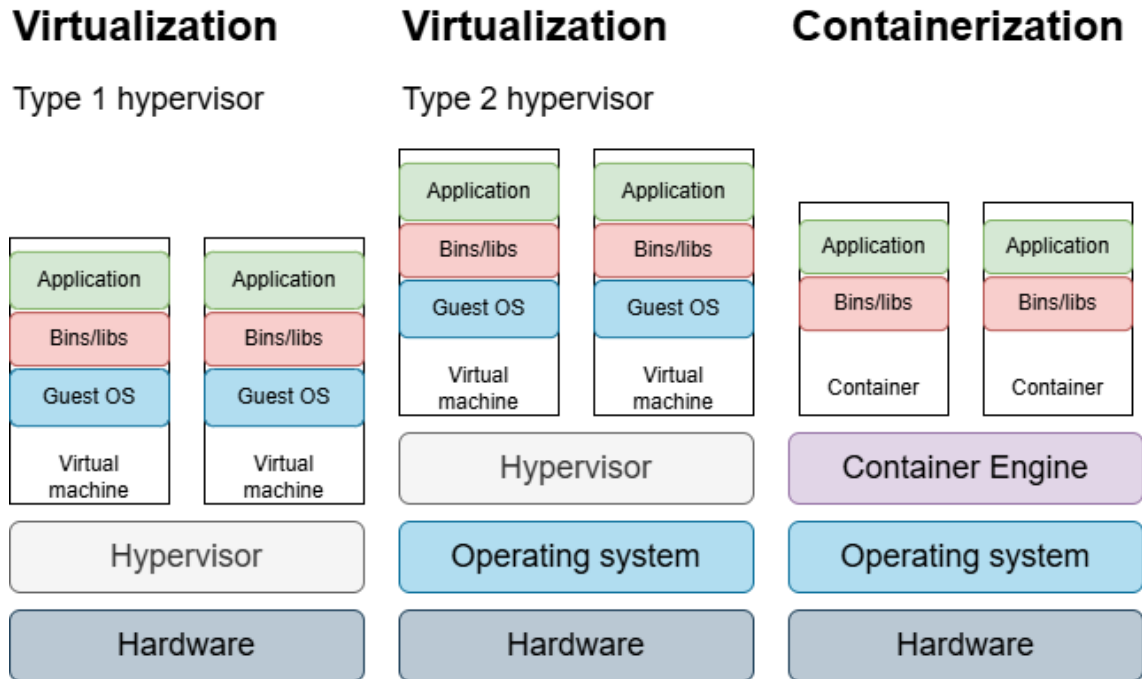
removes the need of setting up and managing the infrastructure like network, servers and storage. Some examples of PaaS are Google App Engine and Heroku. Compared to SaaS, the customer has control over data and applications that they deploy on the platform. PaaS is a good option for developers and teams without access to infrastructure expertise, or not enough staff to manage the complexity of handling both infrastructure and development, as it allows them to focus on building and deploying applications without worrying about infrastructure. Using PaaS also allows them to lower the amount of code they need to write, as PaaS providers offer many built-in components such as directory services to ease the process.

IaaS is the least abstracted cloud computing service model, and therefore closest to having an on-premises datacenter. The customer is responsible of everything outside of servers, storage, networking and virtualization. The customer chooses what infrastructure they need and rents it in a pay-as-you-go manner. The maintenance of the physical infrastructure is done by the provider. This gives the customer most control out of the three models. Compared to an on-premises solution, IaaS has lower overhead and maintenance costs. IaaS is a good option for developers and teams that have a greater need for control and the necessary expertise. Popular IaaS options include Amazon Web Services (AWS) and Google Compute Engine (GCE). [28]

## **2.6 Virtualization and Containers**

Central technologies in the realm of cloud computing, virtualization and containerization are used to pursue benefits like standardization, elasticity, scalability, efficient use of their resources, and portability, fault containment and security through isolation. Virtualization has existed since an IBM project in 1964, so it is not a new technology. In virtualization, multiple isolated operating systems run on top of a hypervisor layer. [47]

Virtualization allows maximizing the use of physical hardware. It enables allocating an application a part of the resources of a server, instead of having to purchase and set up a separate physical machine for every application. Managing virtual machines is also easier compared to physical machines. Manual setup, that is required for physical machine, is error-prone and requires a lot of work and effort. In the case of virtual machines, the work can be made more reliable and effortless by automating IT management workflows. The required services can be installed through automated deployment and configuration tools for each new instance, and spinning up new machines takes far less time, as does retiring a virtual machine that is no longer needed, to conserve resources. [67]



**Figure 2.2.** Difference between containerisation and virtualisation. [67]

Hypervisors are the coordination layer in virtualization, and they act as an interface between the virtual machines and hardware, or between virtual machines and the operating system. As shown in figure 2.2, there are two types of hypervisors: Type 1 hypervisors, also called "bare-metal" hypervisors, are running straight on the hardware of the machine, without any operating system in between. Virtual machines are run on top of the hypervisors, and they contain the operating system, applications and binaries. Type 2 hypervisors, also known as "hosted hypervisors" run on top of the operating system.

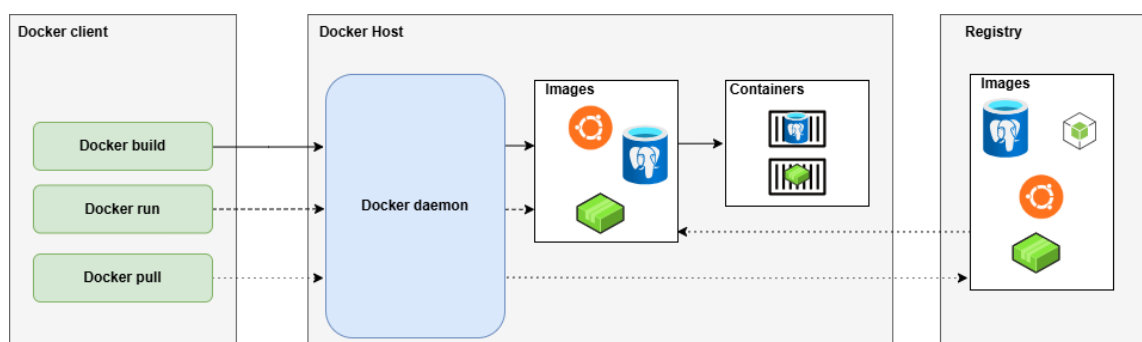
[67] Type 1 hypervisors are better for bigger, more resource-intensive workloads, as it offers better performance and stability, with the drawback of being more complex to deploy and manage. Type 2 hypervisors are better for development, being easy to install and manage, as well as being able to run on any hardware that the underlying OS can run on. The drawback is lower stability and security, as problems with the OS can carry over to the hypervisor and virtual machines.

As virtual machines run on a hypervisor, they gain some security benefits, like the rest of the virtual machines possibly being unaffected and having a higher chance of successfully disinfecting with an operating system rollback, compared to non-virtualized physical systems. Though, virtual machines are not invulnerable. A compromised hypervisor can lead to potentially every compromising every virtual machine running on the hypervisor. [67] In simpler situations, virtualization can bring unnecessary overhead. Containerization is a more recent technology, that requires less setup and maintaining, though it has some core differences to virtualization.

Containers make developing and delivering distributed applications simpler. They allow the user to very quickly boot up an isolated environment with an application and its dependencies, ensuring the application runs reliably between different host systems. While virtualization runs virtual machines on a hypervisor layer, with the virtual machines having their own operating system, containers share the host operating system's kernel, which is demonstrated in figure 2.2. This allows the fast boot up times and potentially smaller resource overhead, compared to conventional virtualization. [59]

As the containers share the kernel of the host operating system, there are some limits to containerization. For example, containerized Linux application needs Linux as the host operating system. This is not a problem with virtual machines, as they have their own operating systems. This limitation can be solved by running the container using a virtual machine's operating system. Recently, it is also possible to run Linux containers on Windows 10 and 11 without the need for virtualization, by using WSL2, which is a Linux kernel built by Microsoft. [69]

Images are pre-configured and standardized templates, and in the context of containers, they are used for running a containerized environment. They contain everything that a software like Docker needs to produce an user space instance, in this context a container. In the image there can be included things like operating systems, middleware, libraries and application code. Container images, however, are more lightweight, as they contain only the minimum required binaries and dependencies. Having an easily distributable, lightweight images is what allows containerization to be portable and scalable. [21] [59] In the context of cloud repatriation, containers are central in migrating applications and workloads to on-premises. Having cloud-agnostic images allows easier migration to on-prem or even to other vendors. Container images are not inherently cloud-agnostic. For example platform-specific proprietary software, such as Amazon Machine Images in Amazon EC2, can lead to vendor lock-in. [5] This can be avoided by following practices and following guidelines that promote interoperability, such as the Open Container Initiative Image Specification. [3] In the context of this cloud repatriation, building images that follow such guidelines ensures less friction when migrating out of public cloud.



**Figure 2.3.** Architecture of Docker [64]

Docker, released in 2013, is a platform that allows users to develop, ship and run applications using containers. Figure 2.3 shows how Docker utilizes client-server architecture. Docker host contains the Docker daemon, which builds, runs and distributes the images. The Docker client is the method the user usually interacts with Docker. The commands like "docker build" that the client receives will be communicated to the daemon as Docker API requests. A Docker registry stores images. There is a public registry called Docker Hub that can be used by anyone, but it is not uncommon to have a private registry. When the daemon receives the command to pull an image, it looks for the image in the registry that has been configured for it. Docker Desktop is an application that contains the Docker daemon and the Docker client, as well as other tools such as Docker Compose. [64]

Especially larger applications often require multiple containers that are able to communicate together. These groups are known as clusters. Automatically managing clusters is called container orchestration, with Kubernetes being the most popular orchestration tool at the time of writing. Managing multiple different containers possibly even across different machines is very difficult manually. Container orchestration tools like Kubernetes are used to manage and automate the lifecycle of containerized applications through service discovery, provisioning and deployment, load balancing and scaling, allocating resources between containers, and monitoring health and performance. [62] Many cloud vendors offer their own container orchestration systems, such as Amazon Elastic Container Service. While these tools offer ease of use on their corresponding platforms, they also contribute to vendor lock in. As most cloud vendors also offer support for Kubernetes, in terms of vendor lock in and future migrations, it is a more future-proof choice.

## 2.7 Promises of Cloud

With as much as 90 % of organizations using public cloud, according to Flexera 2024 State of the Cloud Report, opting for an on-premises environment is a rare and difficult choice. [1] While some organization have opted to move partly away from the public cloud, most have found the promised benefits of the cloud beneficial enough to move to the public cloud and stay there.

1. **Scalability and Flexibility** One of the key benefits is it's scalability and flexibility. Rapid scaling to new levels of demand means that less infrastructure is sitting idle, depreciating in value. It also eliminates the need for as large up-front costs as purchasing hardware and estate for maintaining it. [58]
2. **Cost-effectiveness** Cloud computing is cost-effective, especially at a smaller scale. It moves expenses from capital to operational, allows pay-as-you-go pricing, and in many cases benefits the customer with lower prices through economies of scale. [58]
3. **Faster time to market and testing of ideas** Spinning up instances and deleting

them in seconds allows rapid deployment and faster development and testing of ideas. This cultivates innovation, as testing new ideas isn't limited by slow procurement and existing hardware. [4]

4. **Disaster recovery and backup** Ensuring data security and business continuity is important for a large amount of organizations. Cloud platforms offer redundancy to store data in multiple locations to mitigate risks of data loss and down time through data center accidents. Automated backups, encryption and snapshots are other methods of increasing data resilience. [58]
5. **Security** Cloud computing can improve security, as large and reputable cloud service providers utilize advanced security solutions and employ top security expertise. These tools may not be available especially for smaller organizations, if they didn't use public cloud service providers. Usually cloud providers have a shared security model, which means that the provider is responsible of the security of the cloud, but the user is responsible of security in the cloud. This means, that the security outside of hardware and the rest of the infrastructure that runs the cloud services is still much dependent on the expertise of the customer. Most public cloud providers however, like AWS, offers tools like automated audit management and guidance through documentation and partner companies specialized in security for the user to secure their content. [51] [53]

The benefit depends on the deployed workload, the chosen cloud service provider and implementation strategy. Some legacy applications might need heavy restructuring to be able to scale up and work efficiently in a cloud environment. Cost effectiveness may diminish and turn more expensive in the long run, and savings might not be achievable with certain types of workloads, such as workloads with intense but stable computing [17]. Additionally, all of the promises require proper management from the user to be realized.

## 2.8 Cloud repatriation: Migrating away from the public cloud

451 Research's (S&P Global Inc.) survey data also shows that in the recent years there has been growing movement away from public cloud: The research organization surveyed over 600 data center/colocation respondents in it's Voice of the Enterprise: Data Centers 2021 survey. This survey indicated that during the past year as of answering the poll, 48 % of the responends stated that they had migrated an application or a workload from the public hyper-scale cloud to elsewhere. [37] The motivation for leaving the cloud varies from budget to security reasons.

### 3. BENEFITS AND RISKS ASSOCIATED WITH THE TRANSITION

Moving from the public cloud to an on-premises environment is most often not a simple task, and involves many risks, while the benefits might be hard to materialize. While assessing a potential repatriation, being aware and understanding the risks and benefits is crucial. Having a clear picture of them is also important for developing a strategy that minimizes the risks and maximizes the benefits, or whether the benefits are attainable in the first place.

#### 3.1 Costs

Transitioning from public cloud provider to another venue comes with direct and indirect costs. Understanding the key components that make up the costs is crucial when creating a strategy for the transition, and even more so, if the main motivation for the transition is the cost.

##### 3.1.1 Direct costs

Direct costs are costs that can be traced to their source, like hardware. Direct costs are often easier to predict than indirect costs. Direct costs include [22]:

1. **Hardware:** Acquiring hardware will create a large cost up-front. Purchasing or leasing servers, networking equipment, storage and other necessary hardware. From people who migrated away from hyperscale public cloud, majority moved workloads to their own datacenters [37], making hardware a major cost for most migrators.
2. **Software licenses:** Acquiring licenses will also create a large up-front cost. Operating systems, virtualization tools and other software that is necessary to run and manage an environment for the organizations workloads.
3. **Facility costs:** Running and managing datacenters requires space and equipment. Rooms, electricity, cooling (and other right conditions such as cleanliness), fire security and security against attacks.
4. **Migration services and tools:** Migrating requires specific expertise, and the orga-

nization might need to hire professional services if they don't possess the expertise themselves. This can also include migration software, training cost, and data migration fees from the origin public cloud platform.

5. **Network:** Upgrading the current network infrastructure to meet the new requirements of increased loads, increased security needs and possibly higher connectivity and performance standards.

These direct costs make up the most immediate and visible portion of the costs associated with the repatriation. Less tangible factors also impact the overall financial effect, creating indirect costs.

### 3.1.2 Indirect costs

Indirect costs are costs are not identified with a clear, direct single objective [31], and therefore can't be easily traced to their source. This can be things like software architecture changes and lost productivity.

1. **Software architecture changes:** The applications might require significant rework or refactoring, especially if they were made to be native for the cloud platform that the organization is migrating away from.
2. **Downtime:** Downtime can have an impact on customer satisfaction and business revenue.
3. **Training and hiring:** Existing staff needs to be trained or new staff hired to fill the expertise requirements for maintaining software and hardware. This also includes support.
4. **Compliance and Security:** Ensuring compliance with regulation and standards involves the need for security expertise and auditing.
5. **Maintenance:** Upgrading and maintaining hardware and software will create ongoing costs.

Even though these indirect costs are not simple to calculate, their impact can be substantial. This list is not comprehensive.

### 3.1.3 Opportunity costs

Opportunity costs refer to the cost difference between the highest valued use of a resource and the alternative use of the resource. In other words, it's the missed benefit of choosing the most valuable use for a resource, like the time of employees. [25] In the context of cloud repatriation, opportunity costs can have direct and indirect effects. Choosing to focus on migrating away from public cloud will divert the focus of the organi-

zation, which can have an impact on competitiveness on core business activities and on innovation, increasing time to market for new products. It can also hinder initiatives that create growth like research and development and market expansion. In terms of human capital, IT employees shifting their focus from strategic projects to operational work such as network and hardware maintenance will not only have an impact on the core business activity, but also possibly job satisfaction and talent retention. This is only true if the organization does not already possess the required infrastructure and personnel, and can also be counteracted by hiring new employees specifically for these operational tasks.

### **3.2 Increased Control**

Transitioning to a private cloud or on-premises solution offers more control to the organization over their IT environment. In a private cloud, the organization has exclusive control over the infrastructure, including storage, servers and network. The infrastructure can be customized to the organization's needs, as all the resources are isolated and controlled by the organization. [44] This can be important for an organization that prefers the freedom for their own architectural decisions, picking the hardware and software that they want to use. Such increased control can also enable an organization to meet regulatory compliance requirements in a way that might not be possible with public cloud. This can be valuable in highly regulated industries, such as healthcare and finance.

### **3.3 Effect on performance and latency**

Making any general conclusions about performance between public cloud and on-premises is redundant, as it has many variables that affect the end result, such as the underlying hardware of the on-premises datacenter and specific cloud platform and how well optimized the on-premises environment is. Comparing results from studies reveals though, that an on-premises data center can provide less latency and sometimes better performance in specific workloads [24] [55]. However, a suboptimal on-premises environment will have worse performance and latency when compared to some public cloud options [36], especially with specific types of workloads [55] .

Lower latency is at least partly a result from shorter distance from end users to the servers. The difference in performance exists mostly in the cheaper cloud tiers. The performance gap gets much narrower with more expensive cloud computing options, but the price can get also higher than having an on-premises environment. Though, it is possible to get a cheaper option in the cloud with medium-tier options with a negligible difference in performance, at least when comparing to a small scale on-premises data center. [24] Especially in the case of larger organizations and high-performance computing, it seems that an on-premises data center is often more cost effective [17]. Exploring the difference

in performance and latency with cost effectiveness in mind outside of high-performance computing using larger on-premises data centers would require more studies.

### **3.4 Effect on operational Efficiency and Agility**

Operational efficiency can be defined as the ratio of outputs, such as application functionality for end users, to inputs, such as the amount of costs associated with the functionality or computing power used to deliver said functionality. This means that operational efficiency can be improved by either reducing the input or increasing the output more than the input. [12] To make more use of the resources they have, organizations often want to optimize their efficiency. For example a movie theater will always have baseline amount of staff and the same amount of electricity used to run the projector and other electronics, but the used seats that earn the movie theater money through sold tickets will not stay the same. The excess resources wasted are called capacity waste. It would be more efficient for the movie theater to run a smaller projector in a smaller room and less staff, when they sell less tickets. This same principle applies to computing. Operational efficiency is improved via reducing waste. The waste can be [12]:

- Overproduction: Producing more output than can be used.
- Inefficient use of resources: Using inputs more than is necessary to produce a unit of output.
- Costs of poor quality.
- Creation of functionality of no or little value

Moving to on-premises from public cloud gives an organization more control over allocating resources, potentially reducing various forms of waste. On-premises environments can be tuned especially for stable and predictable workloads. However, the impact on agility must be weighted against the potential benefits of efficiency.

Public cloud platforms inherently enable rapid scaling and on-demand provisioning, which makes it easier to test, deploy and iterate. On-premises infrastructure, by contrast, typically requires acquiring and configuring hardware and software to scale and provision for new workloads. This reduces the ability to react to shifts in market or demand spikes, diminishing the agility that the public cloud provides. However, using modern infrastructure technologies mentioned earlier, such as containers, orchestration tools and virtualization, some of the challenges with agility can be mitigated, as they enable more dynamic management and scaling of applications even without the elasticity offered by the public cloud. [63] To mitigate the lack of redundancy and slow hardware and software scaling without investing into excess infrastructure, it is possible to use a colocation center or even data centers as a service. Colocation centers allow customers to spread out their equipment geographically, which grants redundancy. Colocation also is a way to improve

scaling, when compared to on-premises datacenters. Some colocation centers also offer integrations to public cloud services to meet sudden spikes in demand. [34]

### **3.5 Effect on security and compliance**

Security and compliance are a common motivator for organizations to repatriate workloads from the public cloud, as 451 Research 2022 survey shows [49]. While cloud service providers have very high standards for security, multi-tenancy and shared responsibility model may conflict with compliance with government regulations and industry standards. On-premises implementation allows organizations to directly oversee and govern the security measures, and also using a colocation center gives much more oversight in comparison to public cloud. [66] These advantages do come with the expectation of in-house security expertise and enabling processes like auditing and regular patching. Compliance does not inherently mean security [38], and with a bad implementation, it might be worse than being in the cloud.

### **3.6 Long Term Implications**

Deciding to repatriate a workload or the whole IT infrastructure is not a short term, one-time event. In the short term, there is likely infrastructure investments and process adjustments, depending on the existing on-premises infrastructure that the organization has and the scale of the repatriation. If the repatriation is being executed with the goal of cutting costs, the financial gain will most likely be in the long term. Long term implications extend beyond costs and infrastructure adjustments, though. As the organization relies less on external cloud services, the personnel skill sets might change. The organization's IT might need to learn more about handling parts that were previously abstracted or managed by the cloud vendor, such as virtualization, container orchestration and security protocols. With more limited resources, the organizational culture might become more restrictive on resource planning and scaling. Repatriation might even lead to the quitting of expertise, as they might feel like they want to keep working with public cloud platforms. Leaving these platforms to a data center may also make adapting emerging technologies slower and more difficult.

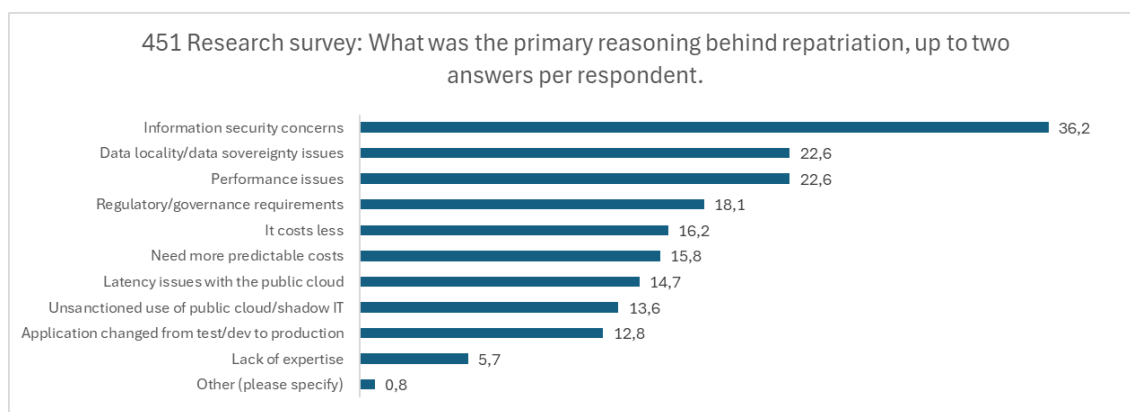
## **4. FACTORS AND STRATEGIES FOR REPATRIATING FROM PUBLIC CLOUD**

The reason for migrating from public cloud depends heavily on the organizational context and the parts of infrastructure that are being migrated. The range of migration can vary from the whole organization's IT infrastructure to a singular workload. Migration strategy will be constructed based on the organization's needs and goals.

### **4.1 Motivations for the repatriation**

When transitioning away public cloud services, it is imperative to evaluate at least the core considerations associated with the shift.

- What are the reasons for leaving - if it is because of budget reasons, careful cost estimation of managing infrastructure is required. As public cloud platforms not only provide computing resources, but also heavy abstraction to make managing those resources simpler and thus require less human resources on the customer side.
- What are the risks for your organization? Potential data breaches or large initial costs can ruin an organizations finances.
- What regulational requirements does the organization have to meet, and will there be changes in the regulation?
- What does this transition implications does the repatriation have on the organizations business?



**Figure 4.1.** Primary reasoning behind repatriation, up to two answers per respondent [49].

Figure 4.1 shows the the primary reasons behind repatriation according 451 Research's 2022 survey targeted to IT decision makers from organizations that had gone through a repatriation. The number represents percentage of respondents that chose each option. The sum of the rows adds up to over 100 %, as respondents were allowed up to two answers. According to the survey, the most common primary reason for leaving was information security concerns. Second most common was reason was shared between data locality/data sovereignty issues and performance issues. Totaling up "it costs less" and "need more predictable costs", costs were a primary reason for over 30 % of the respondents. [49]

As can be seen in Figure 4.1, the reasons to repatriate vary quite a lot between organizations. The strategy, scope and timeline for leaving should be built around the motivations for leaving, so that they are successfully realized.

## 4.2 Repatriation scope

As mentioned earlier in the thesis, the scope of the repatriation can differ vastly. Often it is not necessary or beneficial to migrate everything the cloud. The approach will be very different whether the organization needs to migrate only critical data, all the workloads, some of the workloads, or the whole IT infrastructure.

To repatriate the whole IT infrastructure would require very specific needs from the organization, and is rarely a viable option. More commonly, organizations move only a part of their workloads and data. Workloads that are usually suited for repatriation are:

- **Mission-critical workloads:** Some workloads have requirements such as real-time processing, low latency and high availability. Having such applications hosted on-premises allows organizations to optimize performance for example through hardware optimization. It also allows the organization to be more in control in the case of an outage.

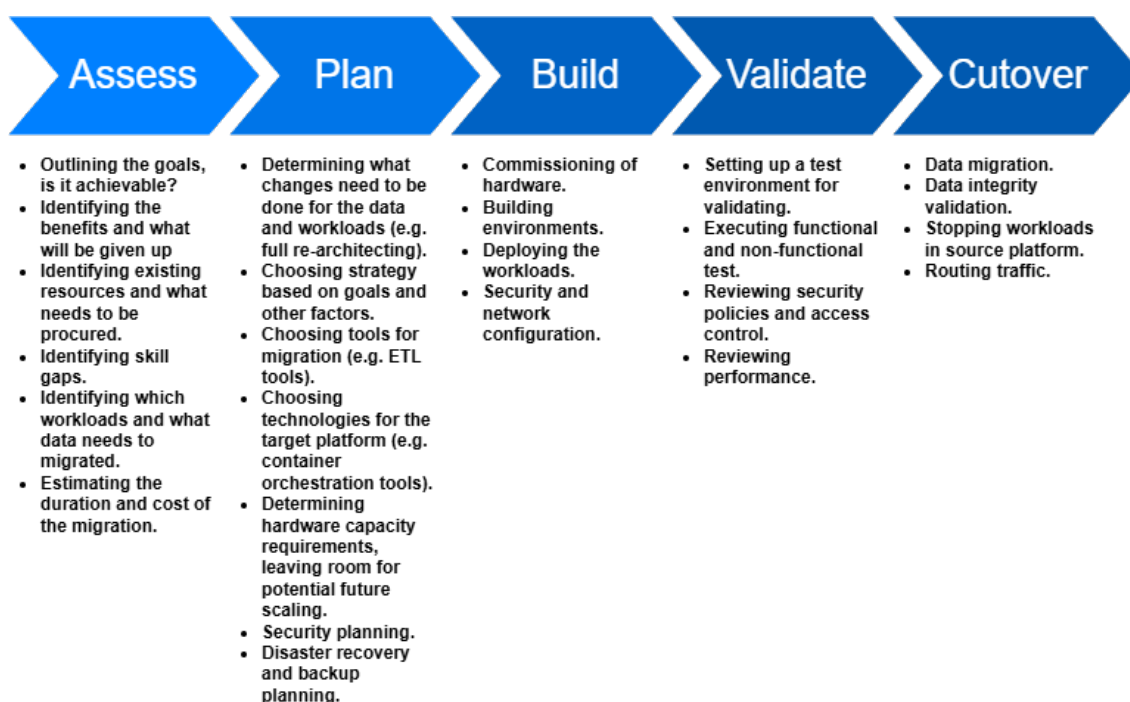
- **Other workloads that require low latency:** Hosting workloads physically closer to the end user allows lower latency than is possible with the public cloud, by the least through the limitation of the speed of light.
- **Sensitive workloads:** Some workloads handle data that has compliance requirements, such as keeping the data in a specific geographic location, which can be satisfied more easily with on-premises.
- **Stable, predictable and intensive workloads:** It may be more cost effective to run workloads that require significant storage and computing resources, as long as the usage is stable and predictable[19] .

If only data needs to be migrated, the focus shifts on the extraction, transformation and loading (ETL) processes. It is critical to ensure data integrity, security and consistency, and to keep in mind data format compatibility, compliance and possible downtime. In cases where the data being migrated resides in a single cloud platform, the platform's own tools, such as AWS Data Migration Services [6], are worth considering. If the data that needs repatriating spans multiple sources, tools like Fivetran allow consolidating data into a single destination [10]. Using a tool outside of the big cloud service providers also reduces the risk of vendor lock-in.

Flexera's 2024 State of the Cloud Report shows, that the data that is often stored on-premises includes consumer data, corporate financial data and other sensitive data [1]. However, it is common that even during the migration, the scope is not completely clear. There might be shadow IT, or during planning something like important networking was overlooked, which is why it is necessary to be prepared for changes to the scope during the migration. [50]

### 4.3 Repatriation timeline

Careful planning of a timeline is crucial for a successful transition to on-premises, as mistakes in the beginning will accumulate downstream, snowballing into more costly and time-consuming to handle. During planning, the organization lays down a timeline with milestones for the transition. It is important to consider the current state of the organizations deployments and data in the cloud environment, and to form an idea of the goals for the transition. [27]



**Figure 4.2.** Timeline example for repatriation [14].

Repatriation timeline is much like a cloud migration timeline. In Figure 4.2, the timeline is cut into 5 distinct steps. It starts with assessing readiness for migration, which leads to preparing for the migration by developing a migration plan, building the environment and deploying the workloads, then validating the results, and finally ends with cutover phase. [14]

Multiple factors have an effect on the timeline of the migration. Large data volume, strict data security, slow transfer speed, lacking user preparation and system incompatibility can complicate and hinder the process, and should be accounted for in the planning of the timeline. [15]

If the migration has a fixed deadline for completion, the choice of approach should reflect that. In a very time-sensitive migration, it might be beneficial to utilize the big bang strategy, and migrate first, optimize later. [15] [50] Big bang approach presents the most risks. In many cases, different assets have different time constraints. In these cases, a phased migration might be a better, less risky approach. If the time constraint is less strict, and especially if the goal of the migration is to gain efficiency, it would be logical to optimize for the target environment first.

## 4.4 Migration strategies

Migration strategy is central to a successful migration from public cloud to on-premises or other private infrastructure. The strategy must be planned based on the organizations goals, constraints, resources, and the scope of the migration. These are some strategies that can be utilized.

### Phased migration

Phased migration is an incremental approach, where workloads and data are migrated gradually. This is beneficial when real-time validation is required, disruptions to operations need to be minimized, and when workloads can be migrated separately without causing much more effort. This approach is ideal for complex migrations, and calls for careful coordination and prioritizing workloads and data. [7]

### Big Bang Migration

In this approach, everything is migrated at once. This method can be fast and efficient, but it requires careful planning, is error-prone and therefore may cause significant downtime. Simple, smaller environments that can tolerate downtime are most suited for big bang migration. [7]

### Hybrid migration

Hybrid migration is a combination of both phased and big bang migration. This approach is suitable when the migration has workloads and data that require different needs, such as some workloads that tolerate no downtime, and other workloads that are simple and less critical. [7]

### 7 R's of migration

The 7 R's of migration are AWS's strategies for migrating applications. The strategies are *retire*, *retain*, *rehost*, *relocate*, *repurchase*, *replatform* and *refactor or re-architect*. [2] While these strategies were made with migrating to cloud in mind, they prove useful in cloud repatriation as well.

### Retire

Retiring means the decommissioning of a workload or an application. This is often the approach for applications that offer no business value if retained in the cloud or repatriated to on-premises. There are some workloads that are useful in the cloud, but not

on-premises. The application might also bring some security risks or have bad performance. AWS suggests to consider retiring if there has been no inbound connection to the application in 90 days. [2]

### **Retain**

Retaining means keeping the application in the source environment. This strategy is useful when the organization is not ready to migrate the application, for example if it needs to wait for some dependencies to migrate first. If an application provides business value only in the cloud, retaining is the clear approach. Retaining temporarily is also beneficial, if migrating would disrupt important development being done on the application. [2]

### **Rehost**

Rehosting, also known as lift and shift, means migrating an application without making any changes. This is an easy strategy that allows migrating quickly from a single or multiple sources without needing to consider compatibility or performance. In the long term, this is often not an efficient strategy, as changing platforms often requires optimization to reach efficiency. Rehosting can be a good approach, when the migration must be done quickly. Optimizing and re-architecting can be done after migrating. [2]

### **Relocate**

Relocating means migrating entire servers from a platform to another to migrate the applications. It is much like rehosting, but instead of moving the application without modifications, it means moving whole servers to their cloud equivalents. [2] Like rehosting, relocating is a quick and easy strategy. This can be a good approach when the application spans multiple servers, and there is no time or resources to refactor the application and re-architecture infrastructure.

### **Repurchase**

Repurchasing, also known as drop and shop, means purchasing another product to replace the functionality that the old application provided. Repurchasing is beneficial when the purchased product gives more business value than the current application, such as in cases where migrating, refactoring and maintaining the application would cost in the long run than purchasing the same functionality from a vendor. It is common to move from a self built, outdated application to a third party equivalent, or moving from a traditional license to SaaS. [2]

## Replatform

Replatforming also known as lift and reshape, means migrating and optimizing the application to the target platform for efficiency, reduced costs or to utilize tools that exist in the target platform. This approach takes more effort than rehosting and relocating, and therefore is best suited for situations where the organization has enough time and resources to make optimizations, but not enough for complete refactoring or re-architecting. [2]

## Refactor or re-architect

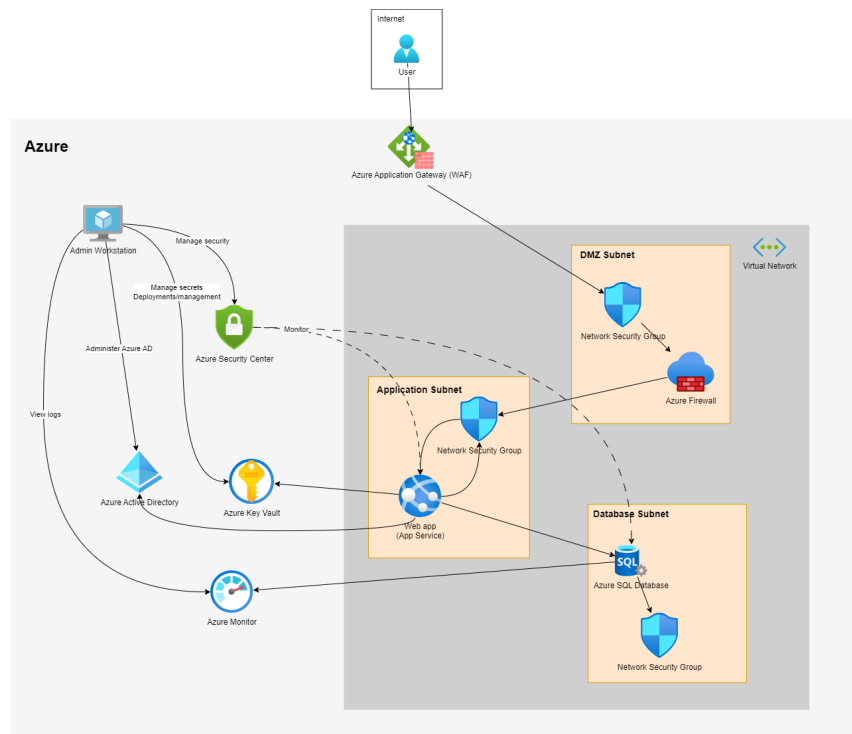
This approach includes modifying the applications architecture to benefit from the target platform's features [2]. If the application is designed benefit from cloud-native tools on the source platform, there are different ways to approach the refactoring:

- Refactor in a way that the application can utilize existing design patterns, but using the target platform's equivalent tools and functionalities.
- Refactor in a way that utilizes different tools and functionalities offered by the target platform.
- Refactor by removing the cloud-native optimizations completely, making the application cloud-agnostic, and then migrate.

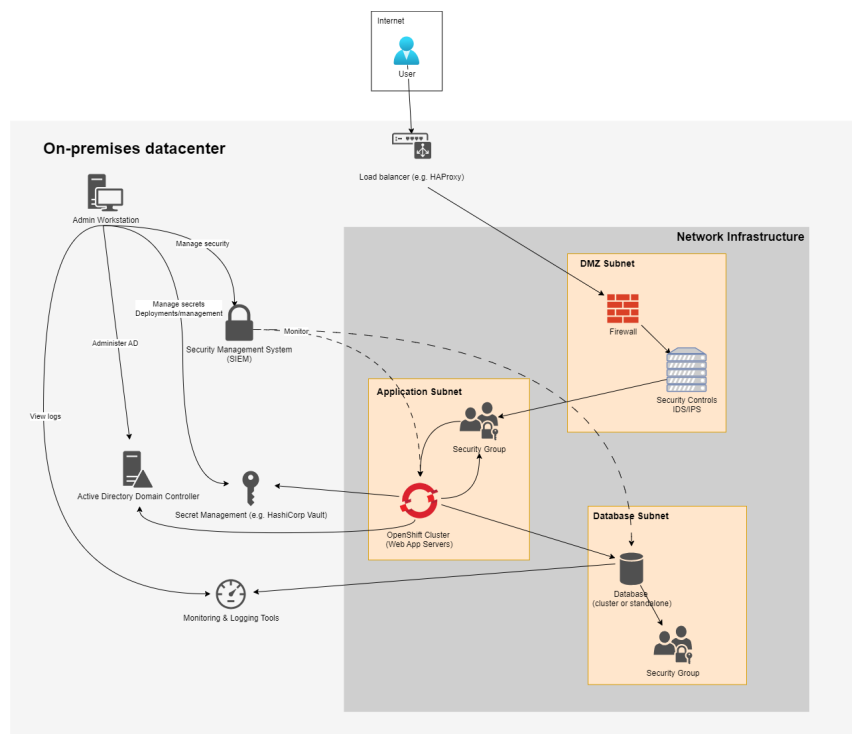
The best choice depends on the goals. If efficiency is important, the third option is most likely not a good approach.

## 4.5 Choosing Tools and Services

Transitioning from to on-premises does not mean that the transitioner needs to give up all the tools usually associated with cloud. Using tools like Docker for containerization, Kubernetes for orchestration, and an IaC tool like Terraform for automated infrastructure management will reduce the gap between traditional on-premises and cloud. Below, in Figure 4.3, there is an example of a simple web application architecture in Azure, and a similar architecture with minimal modifications using tools available for on-premises environments.



**Figure 4.3.** Example of an Azure Architecture for a web application.



**Figure 4.4.** Architecture example of an on-premise version of the Azure Architecture in Figure 4.3.

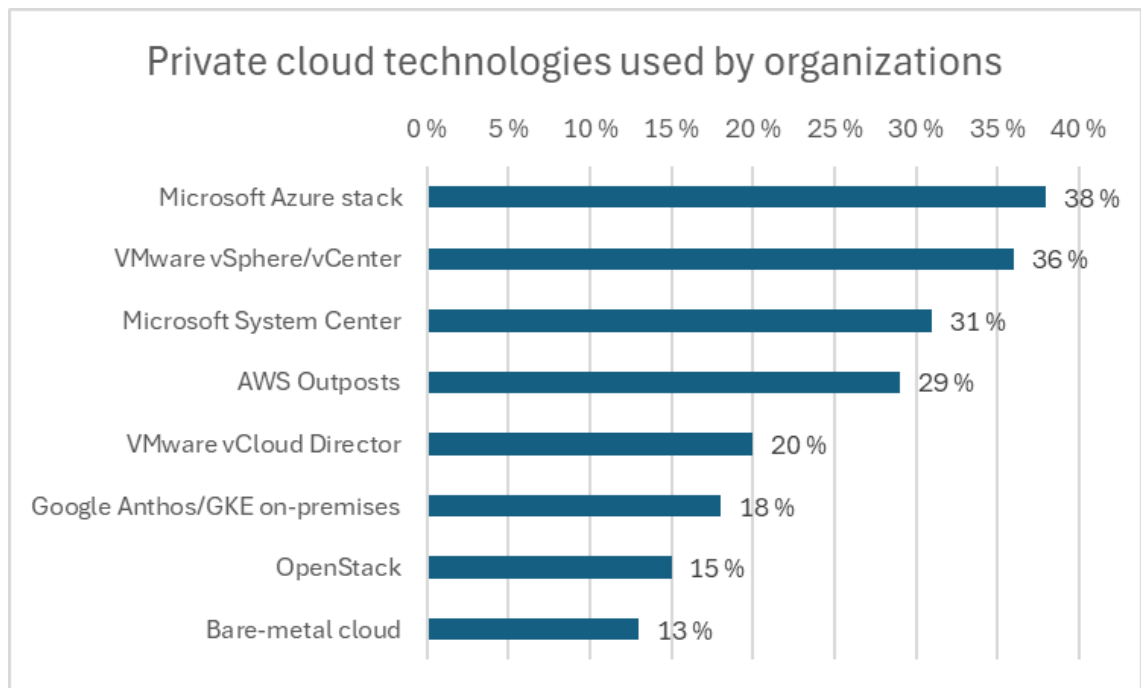
In the architectural examples (Figures 4.3 and 4.4), Azure's vendor-specific tools are replaced with on-premises equivalents. They show how similar architectures between public cloud and on-premises can be with the right tools. Azure native services such as App Services and Azure SQL Database are replaced with OpenShift-orchestrated, containerized workloads and on-premises databases. This demonstrates that it is possible to substitute cloud-based services with self-managed options.

Many public cloud platforms even offer their proprietary tools for a more convenient repatriation experience, but they do usually come with a cost. Many of the tools have platform agnostic counterparts, that are often also free and open source, and don't cause any vendor lock-in. In Table 4.1 below, there are some options for repatriation.

Source Platform	Source Technology to Replace	Repatriation Target Technology	Remarks
Azure	Azure VMs (IaaS)	1. Rehost to on-prem (on-premises) using Azure Stack Hub 2. Migrate to on-prem hypervisor (e.g. VMware/Hyper-V)	Azure Stack Hub makes migration easier with familiarity but has costs [43]. On-prem has no-cost options, and tools for easy V2V conversion [54].
	Azure App Service (PaaS)	1. Re-platform to on-prem application servers 2. Containerize and run on Kubernetes/OpenShift	Transitioning from PaaS may require re-architecture. Using containers maintains scalability and portability.
	Azure Functions (Serverless)	1. Adopt on-prem serverless frameworks (Knative, OpenFaaS) 2. Containerize functions for Kubernetes	Runtime management responsibility. Integrated tools like monitoring may need replacing as well.
	Azure SQL Database (Managed DB)	1. On-prem RDBMS (e.g. SQL Server) 2. On-prem DBaaS platforms	Similar DB engine makes migration easier. Use backup/restore or replication tools to minimize downtime.
AWS	EC2 Instances (IaaS)	1. Rehost to on-prem EC2 VMs 2. Bare-metal for performance	Lift and shift EC2 workloads with minimum changes, or V2V to run i.e. VMWare [54].
	EKS (Managed Kubernetes)	1. EKS Anywhere 2. Kubernetes/OpenShift	EKS Anywhere allows using familiar AWS tools on-premises. Kubernetes/OpenShift minimize vendor dependency and give more control.
	RDS (Managed DB)	1. On-prem self-managed databases (e.g. MySQL, PostgreSQL) 2. Managed DB on-premises, like Amazon RDS on AWS Outpost	AWS DMS for smoother transition. Managed DBs like RDS on Amazon Outpost requires less effort, is not free [6].
GCP	GCE (Compute Engine VMs)	1. Rehost to on-prem VMs 2. Deploy on-prem VMware using GCP VMware Engine	Exporting images from GCE is straightforward [23]. VMware Engine makes it easier to migrate VMware virtual machines to on-prem, but costs and adds vendor dependency [68].
	GKE (Managed Kubernetes)	1. Deploy GKE on-prem 2. Use Kubernetes/OpenShift on-prem	GKE on-prem allows using familiar tools [52]. Kubernetes/OpenShift minimize vendor dependency and give more control.
	Cloud SQL	1. On-prem self-managed databases (e.g. MySQL, PostgreSQL) 2. Containerized databases	Backup/restore for data migration. Containerized DBs are more portable, scalable, generally more efficient use of resources and make testing easier.
OCI (Oracle)	Oracle Cloud VMs	1. Rehost to on-prem VMs 2. Bare-metal servers	Oracle tools simplify migrations. Verify licensing compliance. Bare-metal is a good option for performance-critical workloads.
	Oracle Autonomous Database	1. On-prem Oracle DB (RAC, Exadata) 2. Another on-prem RDBMS	Migrating from Autonomous DB may require re-architecting. Exadata improves performance and consolidation.
	Oracle Kubernetes Engine (OKE)	1. Upstream Kubernetes 2. Oracle Linux Cloud Native Environment	Oracle supports on-prem Kubernetes ecosystems [39]. Adjust CI/CD pipelines for seamless integration.
IBM Cloud	IBM Cloud VMs/Databases	1. Re-platform to on-prem VMs 2. Use on-prem DB2 or IBM database solutions	IBM migration tools (e.g. Lift [29]) ease migration. Power Systems integration for legacy workloads.
	IBM Cloud Kubernetes Service (IKS)	1. On-prem Kubernetes 2. Use OpenShift or IBM Cloud Pak	IBM Cloud Pak offers consistency and supports multicloud management [30]. OpenShift maintains native Kubernetes compatibility.
Multi-cloud	Various containerized workloads	1. Consolidate to on-prem Kubernetes/OpenShift 2. Use cloud-agnostic IaC (e.g. Terraform, Ansible)	Standard Open Container Initiative images simplify migration [3]. Cloud-agnostic IaC ensures consistent infrastructure provisioning and management.
	Cloud-native serverless workloads	1. On-prem serverless (Knative, OpenFaaS) 2. Event-driven containers	Serverless on-prem requires event-driven frameworks or containerized workloads. OpenFaaS is easier to adopt but costs for full access [40]. Knative can be complex for simple workloads. [46] Evaluate performance impact.
	Managed Databases (RDS-like services)	1. Self-managed on-prem databases 2. On-prem DBaaS solutions	Tools like PostgreSQL offer consolidating multiple databases to one, and from multiple sources [18]. Backup and monitoring are important after consolidating and migration.

**Table 4.1.** Repatriation choices, target technologies, and considerations for various cloud platforms and services. Edited from source [14].

Table 4.1 offers comparisons of repatriation options for some of the most common cloud platforms. It can be used as a practical guide to evaluate options based on the source platform. Each row on the column "Repatriation Target Technology" has two different options for target technologies. For many of the rows, one of these options is a tool offered by the source cloud platform. Using such tool allows hosting the workload on-premises with minimal impact in user experience, but ties it to the cloud vendor. Therefore, there is also cloud-agnostic options to reduce vendor lock-in. There are many more tools and technologies than listed in the table. This table can help when choosing between the most popular technologies, but a more careful exploration of options may be beneficial.



**Figure 4.5.** Private cloud technologies used by organizations according to Flexera's State of the Cloud Report survey 2024 (n=753) [1].

Flexera's 2024 State of the Cloud Report surveyed organizations on the tools they use for their private clouds. While popularity should not be a sole reason for choosing technologies, it does give an indication of which tools are likely to be up to standard.

## 4.6 Cost analysis

Moving to on-premises can lead to large up-front costs if the organization does not possess existing necessities like hardware and software. The costs can be more predictable going forward, compared to cloud. In section 3.1 we already covered what constitutes these costs. It is important to also know how to analyze these costs. The current sum of the cloud costs should be summed up, while being certain that the expenses are optimized for an accurate comparison. This should include every environment that is being

migrated from. At least the recurring costs of the services that are currently utilized, the licensing fees, data ingress and egress fees and the operational costs like costs related to personnel and tools for managing the cloud should be covered when calculating the sum.

To estimate the costs of an on-premises option, at least the costs of ownership, including staffing, and hardware and software need to be covered. There is a possibility of unexpected costs, such as hardware breaking or licenced software no longer fulfilling the needs and needing replacement, it is important to include enough room for them in the final calculation.

Cloud cost calculators like Holori [26] allow users to input compute, storage and network requirements and choose a provider to get a range of possible costs for a cloud setup. Total cost of ownership (TCO) calculators provide the same service, but for on-premises workloads, for example Azure offers such service [56]. However, it is imporant to consider the motivations of a cloud service provider when using their tool for calculating the cost of not using their service. Using multiple different tools and tools that are not provided by a cloud service provider will give a more accurate picture of the total cost of ownership.

After the assessment of current cloud provider costs and proposed on-premises costs, making a thorough cost-benefit analysis is crucial. A cost-benefit analysis will tell the organization about the potential return of investment and the optimal setup. [45] An organization might find, for example, that moving some workloads to on-premises is the most cost-effective route, while keeping the rest in the cloud. A cost-benefit analysis will also help prioritize the right workloads to migrate, which helps with creating the most beneficial timeline for the transition.

## 4.7 Hardware and Infrastructure

Proper sizing of servers, network infrastructure and storage is critical. Underestimating the necessary hardware will lead to performance bottlenecks and failure, whereas overestimating will lead to unnecessarily increased costs. This unnecessary capacity is called capacity waste - the capacity above what is necessary to serve the actual user demand with acceptable quality and business risk. [12] In the cloud, provisioning more resources to facilitate load changes is quick. On-premises environments require planning and purchases much ahead to respond to an increased need for capacity. In his 2018 article Eric Bauer divides capacity into four categories: Core capacity, variance capacity, reserve capacity and excess capacity [12].

- **Core capacity:** the capacity needed to serve the average or mean demand.
- **Variance capacity:** the capacity needed to cover statistical peaks in demand.
- **Reserve capacity:** the spare capacity over variance and core capacity that is used

to mitigate service quality impact. Includes capacity for failures, called *redundancy*, and *capacity for lead time demand*, which is used to cover the rising demand in between when additional capacity was decided to be added and when new capacity is available to serve demand. This means, for example if lead time for increasing capacity is 5 minutes, the lead time reserve capacity should cover 10 minutes of growth. Long lead times require larger lead time reserve capacity. Reserve capacity should also cover other contingencies, such as sudden, unforecasted spikes in demand.

Some organizations will already possess some or all the required resources for migrating and running a unit like a workload on-premises. Especially many of the larger organizations have the data centers and expertise for such smaller scale migrations.

## 4.8 Other factors

### Security

According to 451 Research's (S&P Group Inc.) results, the most common reason for repatriation was security concerns. When considering security as one of the reasons for repatriation, it is crucial to make sure the on-premise environment actually ends up more secure than the public cloud. While not having control over infrastructure's security may seem risky, cloud providers have put in place rigorous security operations. As an example, according to Statista, Azure is responsible for the data of 1 billion users worldwide, understandably making security vital for their success. [8] While public cloud service providers can leverage economies of scale to invest in security measures, organizations using their own data centers need to implement security themselves. The organization should consider, whether they have or can acquire the staff and resources for maintaining:

- Physical security: Depends on the environment, can include measures such as fire protection and guards.
- Access control: Includes enforcing and maintaining permissions and restrictions for users, groups systems.
- Network protection: Includes implementing and maintaining tools such as firewalls and intrusion detection systems.
- Data encryption: Includes using encryption algorithms and proper key maintenance.
- Incident detection and response: Includes tools such as monitoring systems, processes like escalation procedures and training.
- Redundancy: Backup systems and failover mechanisms protect continuity of operations from disasters.

- Compliance: Training and understanding of relevant regulations, as well as auditing for them.

Big cloud service providers like Amazon Web Services had multi-layered security, consisting of security guards, fencing and intrusion detection technology. [41] This protection covers the malicious aspect of physical security. Internal aspects, such as accidental data breaches by an employee, are mitigated by granting access rights on the principle of least privilege. Employees are granted a time-bound access to a specified layer of the data center and are restricted from other areas. Requests are thoroughly reviewed, access is monitored, and access logs are audited. [20]

When repatriating to a data center, these responsibilities fall closer to the organization. If the data center is owned by and hosted at the organization and the data that is being handled is sensitive, measures like security guards and intrusion detection need to be implemented. However, if the organization goes with this approach, it is very likely that they already possess physical security against intrusions. That leaves protection against internal and environmental threats. Thorough access management can help cover internal threats, intentional and unintentional. Lack of internal security can lead to data breaches. Environmental threats include the threats that arise from environmental events like floods, fires, and blackouts. Redundancy through insulated datacenters helps prevent catastrophic downtime. Ideally, datacenters would be geographically separated, but this can be costly to implement. A cheaper option would be failover to the public cloud, if the organization's regulations allow it.

### **Data migration**

Microsoft defines data migration as "the process of selecting, preparing, extracting, and transforming data and permanently transferring it from one computer storage system to another." [11] Data migration is usually one of the most demanding and crucial parts of migrating to a new environment. Data assets can be spread out in different locations and states, making collecting and organizing the important assets a complex task. Data conversion, meaning transforming the format of the data, is also often required. If the data is sensitive, migrating it securely is going to be priority. If the data is not sensitive, the migrating organization must only ensure the transfer is seamless, which can be addressed by using tools that check the integrity of the data. [amin\_opportunities\_2021]

## 5. DISCUSSION

This thesis identified some of the factors and strategies that are essential for a successful and beneficial cloud repatriation, and explored the common motivations for such transition. By evaluating these motivations, potential benefits and risks, it was concluded that repatriating requires careful planning tailored for the specific organizational needs, such as workload and data characteristics, compliance requirements and cost factors.

Repatriation appears to be an option mostly for large and stable organizations. This seems sensible, as it requires enough initial capital, and larger organizations can benefit from economies of scale when building on-premises infrastructure. On-premises infrastructure is also a long-term investment, and cost benefits seem to start to accumulate on a longer timeframe. Therefore, it appears that public cloud is usually a more suitable option for smaller organizations.

Repatriation rarely means completely leaving public cloud, but instead repatriating workloads and other parts of IT infrastructure that usually fits more into on-premises environments. It was shown that this includes workloads with intense but stable computing, mission-critical workloads, and workloads handling sensitive data. Dynamic workloads with fluctuating demand appear to be more fit for public cloud.

### **Hybrid cloud as predominant outcome**

As some workloads fit on-premises better than others, it would make sense that the result of repatriation would most often be hybrid cloud, where the organization utilizes both public cloud and on-premises data centers for their distinct benefits. This was confirmed by survey data, as most large organizations have a hybrid cloud approach. It is possible that this movement continues as organizations start to explore options outside of public cloud, and as more tools and technologies make it easier to integrate on-premises environments into the existing public cloud infrastructure, while also making on-premises environments provide many of the benefits of public cloud.

Hybrid cloud being more widely adopted might indicate that cloud technology is no longer seen as the only choice towards infrastructure modernization, but instead as a tool that fills in its role in the same way any other tool does.

## **Risk management**

Repatriation is difficult to execute and involves a lot of risks. These risks might be hard to mitigate or to even notice without experience. When planning large and complex repatriations, it might be beneficial to start with smaller, non-critical workloads to gain experience. This can even show whether it is indeed beneficial for the organization to even repatriate.

Mitigating the potential risks is an important part of repatriation. This involves choosing the right strategies and creating a realistic timeline. For complex repatriation, a phased approach is going to mitigate the risk of larger mistakes. This approach can also be used for critical workloads, so that the risk of downtime is minimized. Big bang approach should be used only for workloads that are non-critical.

Double billing was identified as a challenge in repatriation, meaning the cost of maintaining both on-premises and public cloud environments during the process. If the repatriation takes longer than expected, the cost of double billing will likely become a very large part of the financial impact of the repatriation. When it is known that there will be double billing, having the right strategy executing it efficiently becomes especially important, so that the overlap duration is minimized.

Another dimension worth considering when assessing cloud repatriation is the erosion of expertise in the management of on-premises environments. With increasing adoption of public cloud platforms, the workforce is also oriented towards specific platforms and cloud-native practices. Thus, the lack of skilled personnel available is a risk to be considered. As an alternative option, training existing employees is an option, but it requires time, resources, adds complexity and poses its own risks.

## **Technical complexity**

The complexity of the repatriation depends on multiple factors, such as the source platform and what options are available for the target platform. If the source platform is PaaS, repatriating to an on-premises environment with very different capabilities is going to be difficult and can require complete re-architecting of the repatriated workload.

Choosing the correct technologies for the target platform can help to make the target platform more similar and ensure a more seamless transition. Table 4.1 can help with the choices. It is also beneficial to explore alternatives outside of the table, as new technologies emerge rapidly.

## **Evaluation of findings and research work**

It is important to evaluate the possible sources of errors and the reliability of the results of this thesis and acknowledge the potential sources of errors. First, a large portion of

the data used in this thesis comes from publicly shared stories of repatriation. This might introduce selection bias, as the successful stories might be more likely to be publicized.

Second, as the landscape of cloud technology evolves at a rapid pace, and new stories on cloud repatriation were being published throughout the process of writing this thesis. Thus, the findings might become outdated quickly. New technologies or regulations can change the feasibility of repatriation of a workload, for example.

Third, the data focuses on large organizations. Because of this, making conclusions for smaller organizations that can't reach economies of scale based on the findings is difficult.

Finally, a large portion of the findings relies on secondary data from industry surveys. While those provided good real-world data, they lack controlling for variables, such as what was the pricing of public cloud services at the time, or if the organizations had access to exceptionally cheap hardware.

There is also considerable interpretation, which can introduce researcher bias. Certain factors, such as costs, that were more interesting or relevant to me, were weighted more heavily in the analysis. This might give a skewed view of the impact of repatriation. These factors should be considered when reading this thesis.

## 6. CONCLUSION

The findings in this thesis show that cloud repatriation is a difficult and complex but still achievable process, when the approach is correct. The correct approach is very context-based, depending on multiple factors such as the goals of the repatriation. While the potential benefits are substantial, the risks highlight the importance of thorough planning.

Key findings demonstrate that to be able to move flexibly between public cloud and on-premises environments and avoid vendor lock-in, it is important to not use vendor specific tools and technologies and instead adopt their cloud-agnostic counterparts. Furthermore, it is important to have an exit strategy. This thesis also demonstrates that repatriation has varying effects on performance, operational efficiency and security, often positive effects on latency and compliance, usually positive effects on control and usually negative effects on agility. These effects are dependent on multiple factors, such as the organizations expertise and other relevant resources that they already possess, and the structure of the workloads being repatriated.

Ultimately, the findings emphasize the benefits of staying cloud-agnostic, and showcases repatriation as a viable option for specific workloads and especially larger organizations with the required resources. It is important to consider cloud repatriation as a nuanced process, instead of black-and-white choice between public cloud and on-premises.

### **Future directions**

Currently, not much research exists on cloud repatriation. As organizations look for optimal environments for their workloads, there will hopefully be more case studies on which workloads are the best candidates for repatriation. This would allow for a more thorough definition of what and when to repatriate, and what kind of benefits to expect. While cloud adoption is still on the rise, hybrid cloud appears to gain more ground, offering interesting opportunities for research. Alternatively, cloud vendors might react to the repatriation trends and add new services to their offering, changing the context entirely.

## REFERENCES

- [1] *2024 State of the Cloud Report | Flexera*. en. Tech. rep. Flexera, 2024. URL: [https://info.flexera.com/CM-REPORT-State-of-the-Cloud?lead\\_source=Organic%20Search](https://info.flexera.com/CM-REPORT-State-of-the-Cloud?lead_source=Organic%20Search) (visited on 09/06/2024).
- [2] *About the migration strategies - AWS Prescriptive Guidance*. URL: <https://docs.aws.amazon.com/prescriptive-guidance/latest/large-migration-guide/migration-strategies.html> (visited on 12/14/2024).
- [3] *About the Open Container Initiative - Open Container Initiative*. URL: <https://opencontainers.org/about/overview/> (visited on 12/06/2024).
- [4] *Advantages Of Cloud Computing*. en-US. Publisher: Google Cloud. URL: <https://cloud.google.com/learn/advantages-of-cloud-computing> (visited on 10/11/2024).
- [5] *Amazon Machine Images in Amazon EC2 - Amazon Elastic Compute Cloud*. URL: <https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/AMIs.html> (visited on 12/06/2024).
- [6] *Amazon RDS on Outposts Pricing | On-Premises Managed Database | Amazon Web Services*. en-US. URL: <https://aws.amazon.com/rds/outposts/pricing/> (visited on 12/08/2024).
- [7] Mustyala Anirudh. "Seamless Data Migration: Best Practices for Transitioning to Cloud Environments and Kubernetes Pods". en. In: *International Research Journal of Modernization in Engineering Technology and Science* 03.04 (Aug. 2024), pp. 2433–2444. ISSN: 25825208. DOI: 10.56726/IRJMETS8638. (Visited on 12/19/2024).
- [8] Liv Appleton. *How Secure is Microsoft Azure?* en-GB. July 2024. URL: <https://netcentrix.com/news/how-secure-is-microsoft-azure/> (visited on 11/01/2024).
- [9] Michael Armbrust, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy H. Katz, Andy Konwinski, Gunho Lee, David A. Patterson, Ariel S. Rabkin, Ion Stoica, and Matei A. Zaharia. *Above the Clouds: A Berkeley View of Cloud Computing*. Tech. rep. UCB/EECS-2009-28. 2009. URL: <https://www2.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.pdf> (visited on 09/06/2024).
- [10] *Automated data movement platform | Fivetran*. en. Publisher: Fivetran. URL: <https://www.fivetran.com/data-movement> (visited on 12/18/2024).
- [11] Azure. *What is Data Migration? | Microsoft Azure*. Publication Title: What is Data Migration? | Microsoft Azure. URL: <https://azure.microsoft.com/en-us/resources/cloud-computing-dictionary/what-is-data-migration>.

- [12] Eric Bauer. “Improving Operational Efficiency of Applications via Cloud Computing”. In: *IEEE Cloud Computing* 5.1 (2018), pp. 12–19. DOI: 10.1109/MCC.2018.011791710. (Visited on 09/06/2024).
- [13] Catherine Beaton. *The History of Cloud Migration - Pulsion Technology*. Mar. 2023. URL: <https://www.pulsion.co.uk/blog/the-history-of-cloud-migration> (visited on 09/06/2024).
- [14] Wipro Tech Blogs. *Cloud repatriation — Practical scenarios and solutions*. en. Sept. 2024. URL: <https://wiprotechblogs.medium.com/cloud-repatriation-practical-scenarios-and-solutions-6b3b9a5285f0> (visited on 12/08/2024).
- [15] Shelley Bougnague. *Cloud Migration Timeline: How Long Does the Process Take?* en. URL: <https://www.cloudficient.com/blog/cloud-migration-timeline-how-long-does-the-process-take> (visited on 12/18/2024).
- [16] Cegal. *On-Premises*. Publisher: Cegal. URL: <https://www.cegal.com/en/dictionary/onpremises> (visited on 09/06/2024).
- [17] Alan Chalker, Curtis W. Hillegas, Alan Sill, Sharon Broude Geva, and Craig A. Stewart. “Cloud and on-premises data center usage, expenditures, and approaches to return on investment: A survey of academic research computing organizations”. en. In: *Practice and Experience in Advanced Research Computing*. Portland OR USA: ACM, July 2020, pp. 26–33. ISBN: 978-1-4503-6689-2. DOI: 10.1145/3311790.3396642. URL: <https://dl.acm.org/doi/10.1145/3311790.3396642> (visited on 12/19/2024).
- [18] *Chapter 29. Logical Replication*. en. Nov. 2024. URL: <https://www.postgresql.org/docs/17/logical-replication.html> (visited on 12/13/2024).
- [19] *Cloud vs. on-premises datacenters: How to choose for your workload*. en. URL: <https://www.redhat.com/en/blog/cloud-vs-on-premises> (visited on 12/17/2024).
- [20] *Data Centers - Our Controls*. en-US. URL: <https://aws.amazon.com/compliance/data-center/controls/> (visited on 11/01/2024).
- [21] *Docker: Lightweight Linux Containers for Consistent Development and Deployment | Linux Journal*. URL: <https://www.linuxjournal.com/content/docker-lightweight-linux-containers-consistent-development-and-deployment> (visited on 12/06/2024).
- [22] Adam Earls. *What’s involved in cloud repatriation cost calculation? | TechTarget*. en. Publisher: TechTarget. URL: <https://www.techtarget.com/searchdatacenter/feature/Whats-involved-in-cloud-repatriation-cost-calculation> (visited on 12/19/2024).
- [23] *gcloud compute images export | Google Cloud CLI Documentation*. en. URL: <https://cloud.google.com/sdk/gcloud/reference/compute/images/export> (visited on 12/08/2024).
- [24] Robert Györödi, Marius Iulian Pavel, Cornelia Györödi, and Doina Zmaranda. “Performance of OnPrem Versus Azure SQL Server: A Case Study”. In: *IEEE Access* 7 (2019). Conference Name: IEEE Access, pp. 15894–15902. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2019.2893333. (Visited on 12/19/2024).

- [25] David R. Henderson. *Opportunity Cost - Econlib*. Publication Title: Econlib. June 2018. URL: <https://www.econlib.org/library/Enc/OpportunityCost.html> (visited on 09/06/2024).
- [26] *Holori - The Best Cloud Calculator*. en-US. Publisher: Holori. URL: <https://holori.com/cloud-calculator/> (visited on 09/13/2024).
- [27] *How to Migrate from Cloud to On-Premise, and Why | Erbis Blog*. URL: <https://erbis.com/blog/how-to-migrate-from-cloud-to-on-premise-and-why/> (visited on 09/13/2024).
- [28] *IaaS vs. PaaS vs. SaaS*. en. URL: <https://www.redhat.com/en/topics/cloud-computing/iaas-vs-paas-vs-saas> (visited on 10/13/2024).
- [29] *IBM Lift*. en. Publisher: IBM. May 2024. URL: <https://www.ibm.com/products/lift> (visited on 12/13/2024).
- [30] *IBM Support*. en-US. Publisher: IBM. URL: <https://www.ibm.com/mysupport/s/topic/0TO0z0000006uqxGAA/www.ibm.com/mysupport/s/topic/0to0z0000006uqxgaa/cloud-pak-for-multicloud-management> (visited on 12/13/2024).
- [31] *Indirect Cost: Definition and Example | National Institutes of Health*. Publisher: National Institutes of Health. URL: <https://oamp.od.nih.gov/division-of-financial-advisory-services/indirect-cost-branch/indirect-cost-submission/indirect-cost-definition-and-example> (visited on 12/19/2024).
- [32] Kiran Jewargi. "Public Cloud to Cloud Repatriation Trend". en. In: *Scholars Journal of Engineering and Technology* 11.1 (Jan. 2023), pp. 1–3. ISSN: 23479523, 2321435X. DOI: 10.36347/sjet.2023.v11i01.001. URL: [https://saspublishers.com/media/articles/SJET\\_111\\_1-3\\_FT.pdf](https://saspublishers.com/media/articles/SJET_111_1-3_FT.pdf) (visited on 10/11/2024).
- [33] Tanner Luxner. *Cloud computing Stats: Flexera 2023 State of the Cloud Report*. en-US. Apr. 2023. URL: <https://www.flexera.com/blog/finops/cloud-computing-trends-flexera-2023-state-of-the-cloud-report/> (visited on 12/14/2024).
- [34] Kit MacLean. *Understanding Scalability In Colocation*. en-us. May 2024. URL: <https://www.databank.com/resources/blogs/understanding-scalability-in-colocation/> (visited on 12/06/2024).
- [35] Yaser Mansouri, Victor Prokhorenko, and M. A. Babar. "An Automated Implementation of Hybrid Cloud for Performance Evaluation of Distributed Databases". In: *Journal of Network and Computer Applications* 167 (2020), pp. 2–25. DOI: 10.1016/j.jnca.2020.102740. (Visited on 09/06/2024).
- [36] Jaymie Rae Medina and Jennalyn Mindoro. *On-premise File Server Vs Cloud Storage With Incident Management: a Comparative Study*. en. SSRN Scholarly Paper. Rochester, NY, Nov. 2023. DOI: 10.2139/ssrn.4645107. URL: <https://papers.ssrn.com/abstract=4645107> (visited on 12/19/2024).
- [37] Melanie Posey, Dan Thompson, Pedro Schweizer. *Cloud repatriation: What it is, what it isn't, and why it's not going away*. Sept. 2021. URL: <https://www.spglobal.com>

- com/marketintelligence/en/documents/cloud-repatriation-what-it-is-what-it.pdf (visited on 09/06/2024).
- [38] *Mind the Cybersecurity Compliance Gap | Tripwire*. en. Publisher: Tripwire. URL: <https://www.tripwire.com/resources/guides/mind-the-cybersecurity-compliance-gap> (visited on 12/06/2024).
- [39] *OCI Kubernetes Engine*. en. URL: <https://www.oracle.com/cloud/cloud-native/kubernetes-engine/> (visited on 12/13/2024).
- [40] *OpenFaaS Plans & Pricing*. en. URL: <https://www.openfaas.com/pricing/> (visited on 12/13/2024).
- [41] *Perimeter Layer - Amazon Web Services (AWS)*. en-US. URL: <https://aws.amazon.com/compliance/data-center/perimeter-layer/> (visited on 11/01/2024).
- [42] Mary Pratt. *The great repatriation? IT leaders reset cloud strategies to optimize value*. en-US. URL: <https://www.cio.com/article/2520890/the-great-repatriation-it-leaders-reset-cloud-strategies-to-optimize-value.html> (visited on 12/14/2024).
- [43] *Pricing - Azure Stack Hub | Microsoft Azure*. en. URL: <https://azure.microsoft.com/en-us/pricing/details/azure-stack/hub/> (visited on 12/08/2024).
- [44] *Private Cloud vs. Public Cloud: What is the difference? | VMware Glossary*. en. Publisher: VMWare. URL: <https://www.vmware.com/topics/private-cloud-vs-public-cloud> (visited on 12/14/2024).
- [45] Adarsh Rai. *How to Estimate Your Cloud Migration Costs in 3 Steps?* en-US. Section: Cloud Migration. Aug. 2023. URL: <https://blog.economize.cloud/estimate-cloud-migration-costs/> (visited on 09/13/2024).
- [46] Ran Ribenzaft. *Serverless Open-Source Frameworks: OpenFaaS, Knative, & more*. en-US. Apr. 2020. URL: <https://www.cncf.io/blog/2020/04/13/serverless-open-source-frameworks-openfaas-knative-more/> (visited on 12/13/2024).
- [47] Fernando Rodríguez-Haro, Felix Freitag, Leandro Navarro, Efraín Hernández-sánchez, Nicandro Farías-Mendoza, Juan Antonio Guerrero-Ibáñez, and Apolinar González-Potes. "A summary of virtualization techniques". In: *Procedia Technology*. The 2012 Iberoamerican Conference on Electronics Engineering and Computer Science 3 (Jan. 2012), pp. 267–272. ISSN: 2212-0173. DOI: 10.1016/j.protcy.2012.03.029. (Visited on 11/29/2024).
- [48] Maria Rutkin. *7 Reasons Why You Should Move from On-Premise to Cloud Computing - IBIS Technology*. Publication Title: IBIS Technology. June 2017. URL: <https://ibistechnology.com/why-you-should-move-from-on-premise-to-cloud-computing> (visited on 06/06/2024).
- [49] Pedro Schweizer. *Cloud repatriation: the who, the where, the why - 451 Alliance*. Dec. 2022. URL: <https://blog.451alliance.com/cloud-repatriation-the-who-the-where-the-why/>.

- [50] *Scope, strategy, and timeline - AWS Prescriptive Guidance*. URL: <https://docs.aws.amazon.com/prescriptive-guidance/latest/strategy-large-scale-migrations/scope-strategy-time.html> (visited on 12/18/2024).
- [51] *Security and compliance - Overview of Amazon Web Services*. URL: <https://docs.aws.amazon.com/whitepapers/latest/aws-overview/security-and-compliance.html> (visited on 10/11/2024).
- [52] *Set up Google Distributed Cloud (GKE Enterprise on-premises)*. en. URL: <https://cloud.google.com/kubernetes-engine/enterprise/docs/setup/on-premises> (visited on 12/08/2024).
- [53] *Shared Responsibility Model - Amazon Web Services (AWS)*. en-US. URL: <https://aws.amazon.com/compliance/shared-responsibility-model/> (visited on 10/11/2024).
- [54] *StarWind V2V Converter Help : Using StarWind V2V Converter with Microsoft Azure*. URL: <https://www.starwindsoftware.com/v2v-help/UsingStarWindV2VConverterwithMicrosoftAzure.html> (visited on 12/08/2024).
- [55] Venkata Tadi. "Performance and Scalability in Data Warehousing: Comparing Snowflake's Cloud-Native Architecture with Traditional On- Premises Solutions Under Varying Workloads". en. In: *European Journal of Advances in Engineering and Technology* 9.5 (2022), pp. 127–139. ISSN: 2394 - 658X. DOI: 10.5281/zenodo.13319605. (Visited on 12/19/2024).
- [56] *Total Cost of Ownership (TCO) Calculator | Microsoft Azure*. en. URL: <https://azure.microsoft.com/en-us/pricing/tco/calculator/> (visited on 09/13/2024).
- [57] Vassil Vassilev, Karim Ouazzane, Viktor Sowinski-Mydlarz, Herbert Maosa, Sabin Nakarmi, Martin Hristev, and Sorin Radu. "Network Security Analytics on the Cloud: Public vs. Private Case". In: *2023 13th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*. Noida, India: IEEE, 2023, pp. 151–156. DOI: 10.1109/Confluence56041.2023.10048889. (Visited on 06/06/2024).
- [58] Prathamesh M Vichare, Poonam P Sawant, and Waman R Parulekar. "The Future of Cloud Computing: Benefits and Challenges". en. In: *International Journal of Progressive Research in Engineering Management and Science (Ijprems)* 04.05 (2024), pp. 2233–2237. DOI: 10.4236/ijcns.2023.164004. (Visited on 10/11/2024).
- [59] *What is a Container? | Docker*. en-US. URL: <https://www.docker.com/resources/what-container/> (visited on 11/30/2024).
- [60] *What is a Public Cloud - Definition | Microsoft Azure*. URL: <https://azure.microsoft.com/en-us/resources/cloud-computing-dictionary/what-is-a-public-cloud> (visited on 09/06/2024).
- [61] *What is a virtual private cloud (VPC)?* en-us. Publisher: Hewlett Packard Enterprise. URL: <https://www.cloudflare.com/learning/cloud/what-is-a-virtual-private-cloud/> (visited on 10/12/2024).
- [62] *What Is container orchestration*. en-US. URL: <https://cloud.google.com/discover/what-is-container-orchestration> (visited on 12/06/2024).

- [63] *What is Data Center Colocation? | Glossary.* en-FI. URL: <https://www.hpe.com/fi/en/what-is/data-center-colocation.html> (visited on 12/06/2024).
- [64] *What is Docker?* en. 200. URL: <https://docs.docker.com/get-started/docker-overview/> (visited on 11/30/2024).
- [65] *What is multitenancy? | Multitenant architecture.* en-us. Publisher: Hewlett Packard Enterprise. URL: <https://www.cloudflare.com/learning/cloud/what-is-multitenancy/> (visited on 10/12/2024).
- [66] *What is On-Premises vs. Cloud? | Glossary.* en-FI. URL: <https://www.hpe.com/fi/en/what-is/on-premises-vs-cloud.html> (visited on 12/06/2024).
- [67] *What Is Virtualization? | IBM.* en. Publisher: IBM. Mar. 2023. URL: <https://www.ibm.com/topics/virtualization> (visited on 11/29/2024).
- [68] *Workload VM migration | Google Cloud VMware Engine Documentation.* en. URL: <https://cloud.google.com/vmware-engine/docs/concepts-migration-options> (visited on 12/08/2024).
- [69] *WSL.* en. Publisher: Docker. 100. URL: <https://docs.docker.com/desktop/features/wsl/> (visited on 11/30/2024).