Offline-first strategies in heterogeneous, distributed and virtualized infrastructures

Henry Cocos[a]
Department of Computer Science and Engineerung,
Frankfurt University of Applied Sciences,
Nibelungeplatz 1, Frankfurt D-60318, Germany
cocos@fb2.fra-uas.de

Abstract

This Ph.D project aims to investigate the outsourcing of cloud services to clients by including heterogeneous, client-side resources and cooperation between the cloud and the client in the execution of services. To this end, an offline-first strategy will be formulated, developed, and investigated as a counter-design to the cloud-first strategy, enabling the usability of services without permanent services and a permanent cloud connection. This pursues the goal of application autonomy with simultaneous cooperation with cloud services. The goal is to achieve autonomy and simultaneous cooperation of applications in this context. Methods of service migration are to be researched to enable the provision of services. The influence of migration on the quality of services will also be investigated. From this, methods for determining the quality of service provide the end user with information about the service provided. In summary, this should lead to an increase in the resilience of the use of services to increase the stability of services along the cloud-to-edge supply chain and enable transparent use for the end user.

1 RESEARCH PROBLEM

The Frankfurt University of Applied Sciences develops an innovative platform for the creation and use of virtual learning environments, the SKILL project (Strategic Competence Platform - Innovative Learning and Teaching) (Baun et al., 2024). This virtual teaching service is implemented to deliver virtual resources for university teaching activities. For this purpose, lecturers, students, and staff can create virtual environments to integrate these into their teaching. One of the concepts of the SKILL project involves the use of the service by students. The SKILL concept envisages that users use the virtual environments in the infrastructure provided. This means that the virtual machines are to be operated on server hardware and thus use the project's resources.

However, using the services requires the project's resources on the server side, as the end users of the services access the server hardware via a client. Since there are end users who themselves have devices with a large amount of resources, this could create opportunities and relieve the core service on the server side. A suitable method for relieving the load on cloud services is to outsource the service or parts of it to the client. Therefore, utilizing the clients' resources holds the potential for reducing load on the server side. Since the resources on the client side, which are not usually required elsewhere in parallel during the server-side service, can be used as an extension of the core service.

This concept can relieve the cloud service using the end user's resources. It also reduces latencies on the network side. Running the service locally can also benefit the latency when interacting with the service. One conceivable scenario is the use of the service via a long-distance data connection with low bandwidth, as can still be found in many rural areas today.

For example, according to a Federal Ministry for Digital and Transport (BMDV) survey, only 68% of households in rural areas have a 100 Mbit/s connection or better (BMD,). The figures for businesses are at a similar level. Low bandwidths on the network side are problematic when using an online cloud service. These problems can be found in cloud services beyond the SKILL project and in many scenarios in which a slow data connection impairs the use of a cloud service. It also affects many companies in rural locations, where the need for more broadband access over a considerable distance leads to low bandwidths on the network side. In such a case, low bandwidth becomes a problem when connecting a company to a cloud service.

An extension of the service using the client could

[a] https://orcid.org/0009-0001-7573-0361

counteract this problem, but this approach raises many open questions that need to be answered. For example, the clients are highly heterogeneous computer environments without uniform resources. The heterogeneity on the client side is determined, among other things, by resources such as CPU performance, memory size, and mass storage. This means that an end user can use the service with comparatively weak hardware, such as an older notebook, while another uses a desktop computer with a powerful CPU.

2 OUTLINE OF OBJECTIVES

One of the Ph.D. project's objectives is to research the outsourcing and cooperation of cloud services with the client, including the client-side resources. The aim is to investigate which services from the cloud could be outsourced to the client and when this procedure makes sense. For this purpose, an offline-first strategy shall be formulated, which, in contrast to the currently prevailing cloud-first strategy, makes services available in the local network and thus provides a network-independent provision of cloud services.

With this approach, end users should be able to use services in their immediate proximity without being dependent on the cloud. This is intended to give local services greater autonomy and resilience for local services, without sacrificing the advantages of a cloud environment. Defined criteria for the decision to outsource computing-intensive tasks are to be determined. Determining these metrics and criteria is a task for which there has yet to be a generally accepted method in current research.

Another objective is the investigation of Service migration (Rodrigues et al., 2021) for the distribution of resources. This method of distributing applications is concerned with placing services at the level of end devices and the cooperation between the cloud, end devices, and services, as well as the use of the services by the end user. The decisive factor here is which technical methods exist for local resource utilization and how these positively affect service use.

Generally, service migration can be achieved vertically (between the cloud and end devices) or horizontally (between end devices). The objectives here are, among other things, to research the influences on the service associated with the migration and the technical options for migrating services.

Another aim of the Ph.D. studies is to research the resilience of services (Eberz-Eder et al., 2021; Harchol et al., 2020) and the impact of service migration on the quality of service. For example, service migration can induce latency during use, negatively

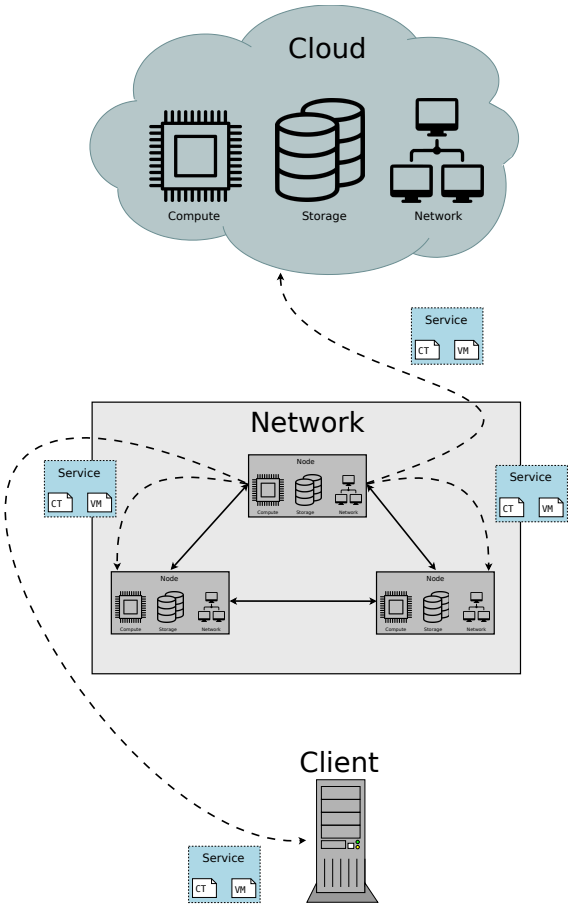

Figure 1: Concept of service migration

impacting the service's quality. Also, many factors for the decision to migrate need to be monitored, as the client's hardware resources (e.g., processor, main memory) differ from those of the server. Which quality characteristics (e.g., performance) are influenced by migration and how the migration affects the individual features must be investigated. An estimate of the migration's effect on service quality could be helpful in this context. The definition of quality presented here differs from the quality of service since, in the context of the project, the quality of the application's content is to be examined rather than the quality of the service in terms of connection metrics. These are content-related requirements for the quality of the service (e.g., correctness of the results depending on time constraints). The quality characteristics at the service's runtime would be helpful in such a context for decisions for or against the migration of a service. These and other criteria are to be researched during the Ph.D. project.

3 STATE OF THE ART

One significant focus lies on virtual machine migration and its effects on the quality of services (Choudhary et al., 2017). Service migration can only occur when the surrounding and isolating virtual machine is also migrated and moved from a source server offering the service to the terminal device, consuming the service. Reducing blackout times during migration of services can be achieved using live migration features in VMs. However, VMs are very heavy-weight, so migration over larger WAN links is only feasible for legacy services, which cannot switch to containerized environments (Winkelhofer, 2019). Container virtualization (also called operating system virtualization) is more suitable for operation on terminal devices since they have a smaller footprint regarding resource consumption. This isolation mechanism for application is very promising and widely used. However, it is only suitable for newer applications and micro-service arrangements.

Another critical point in this Ph.D. project is the network since service migration can only occur over a stable and efficient network link. One challenge is emulating a local network for the virtual machine or container applications. One possible solution widely used in the industry and science are overlay networks employed by Software Defined Networking (Wang et al., 2019). By using virtual overlay networks, applications can operate transparently with one another over a WAN link. The control flow and data flow are separated, making the administration of such overlay networks easier. However, establishing overlay networks over WAN to connect terminal devices to the servers operating the service is a considerable challenge that needs to be investigated in this Ph.D. project.

Most of the current research regarding service migration and resource sharing stems from the field of edge computing (Gedeon et al., 2019). This Ph.D. project follows a common principle but differentiates itself by spanning the focus to end devices with more resources (e.g., Laptops or client PCs). The focus of edge computing, on the other hand, lies heavily on using low-power and weak terminal devices. Outsourcing services and resource use require methods for distributing and providing components.

Edge computing can be seen as an extension of cloud computing, bringing services closer to the data producers and users, accelerating communication, and reducing latencies (Merkl and Cocos, 2020). The area of wireless sensor networks benefits from the edge computing paradigm (Cocos and Merkl, 2019). Although this project is investigating methods and technologies of edge computing, unlike the classic edge computing use case, this project focuses on the inclusion of clients. In this approach, the client should use its resources to extend the cloud service, and, at the same time, there should be the option of working cooperatively together with the cloud service.

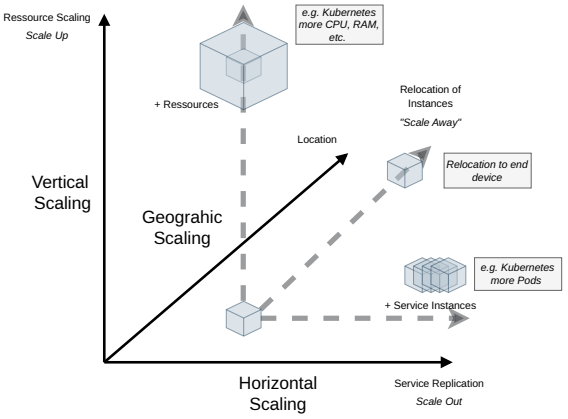

Figure 2: Dimensions of resource scaling

4 METHODOLOGY

Figure 1 schematically shows the migration between the individual layers, whereby the migration between the cloud/network and the client is vertical, and migration between devices at the same level (e.g., within the cloud) is referred to as horizontal migration.

The migration of the virtual resources between the different layers is taking place over the wide area network (WAN), which has a significantly lower bandwidth than the local area network (LAN). Therefore, the decision on the virtual resource's exact location and the migration's timing has to be made.

Using methods from edge and fog computing is beneficial in implementing suitable measures for migration and communication in said scenario. The connection and communication between virtual machines or containers place a high interest in the underlying networking technology. Overlay networks like VXLAN are a reasonable choice and, therefore, shall be employed in realizing the project.

Figure 2 presents the dimensions of scalability in virtual resources ranging from *vertical scaling* of resources and *horizontal scaling* of resources. These two dimensions are implemented by adding resources to the virtual resource (so-called upscaling) Alternatively, by replicating virtual resources (so-called out scaling). A third dimension not investigated in research is geographic scaling, which involves relocating services and their components closer to end de-

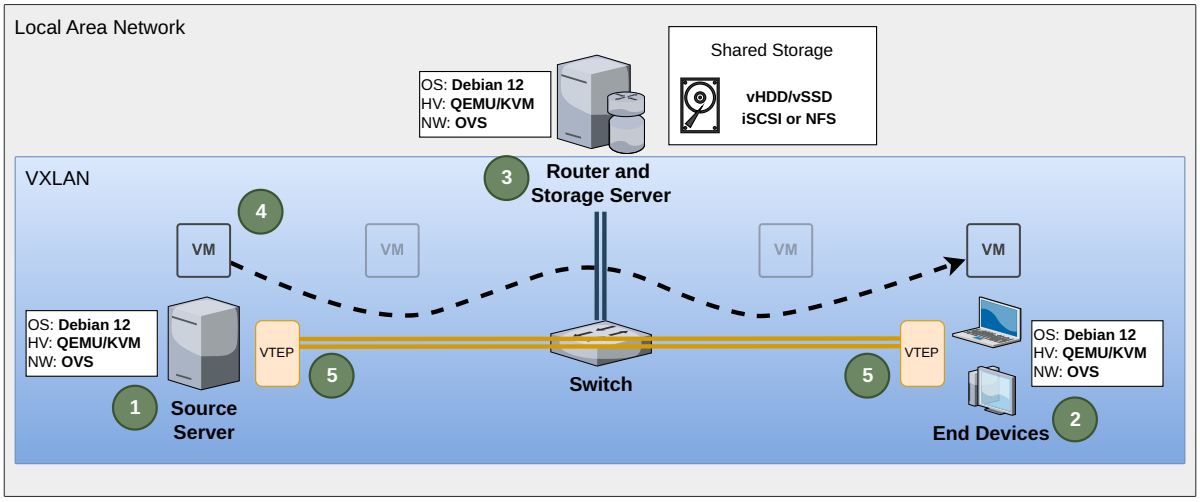

Figure 3: Experimental setup for WAN migration of services

vices. More precisely on the end device itself. This dimension is a focal point of research in this Ph.D. project.

5 EXPECTED OUTCOME

The expected outcome of this Ph.D. project is an increased resilience of cloud service offerings by leveraging the performance of terminal devices and an offline operation of services. The terminal devices become an enhancement of the cloud service, spanning its service offering to terminal devices, making services more accessible to end users. Another expected outcome is the reduced network latency of service offerings in conjunction with a reduced server load, making the operation of services more effective from a user perspective and increasing the reliability of the service offering. In summary, the Ph.D. project investigates ways to increase service offerings' autonomy and resilience by using vertical scalability and geographic scalability (see figure 2).

Investigating suitable technologies and methods for the seamless distribution of services and use of resources is another expected outcome of the Ph.D. project. In the end the findings of the research will be beneficial for the application of methods in the future, resulting in solid arguments for the embodiment of new standards and methods for the migration and communication of services in heterogeneous and distributed environments.

6 STAGE OF THE RESEARCH

The initial literature research phase has already been completed, resulting in a concept for implementing the Ph.D. project and the choice of technologies and methods. In the next step, initial experimentation with the technologies will give an insight into the feasibility of the technologies and methods.

An analysis of the migration of virtual units (VM, container) between different end devices (servers, laptops, etc.) concerning the influence of applications and different amounts of resources (vCPU, vRAM, etc.) on the behavior of the virtual environments during migration is set up as a first indicator for further research. This setup is highly relevant for legacy services and their behavior during migration and can be seen in figure 3.

Therefore, a first experiment is carried out to investigate the feasibility of live virtual machine migration over WAN using VXLAN overlay networks. This first experiment will give insights into the applicability of the setup and guidance on the direction of future research in the field.

ACKNOWLEDGEMENTS

I thank my supervisors, Prof. Dr. Martin Kappes and Prof. Dr. Christian Baun, for their support and guidance during my research.

REFERENCES

Bericht zum Breitbandatlas Teil 1: Ergebnisse (Stand Mitte 2021). Zugegriffen am: 31.05.2022.

Baun, C., Kappes, M., Cocos, H.-N., and Koch, M. M. (2024). Eine plattform zur erstellung und verwendung komplexer virtualisierter IT-Strukturen in lehre und forschung mit open source software. *Informatik-Spektrum*.

Choudhary, A., Govil, M. C., Singh, G., Awasthi, L. K., Pilli, E. S., and Kapil, D. (2017). A critical survey of live virtual machine migration techniques. *J. Cloud Comput.*, 6(1).

Cocos, H.-N. and Merkl, D. (2019). Decentralized Data Processing on the Edge – Accessing Wireless Sensor Networks with Edge Computing. In Weghorn, H., editor, *Proceedings of the 16th International Conference on Applied Computing 2019, Cagliari, Italy*, pages 265–269. IADIS, Nov. 2019.

Eberz-Eder, D., Kuntke, F., Schneider, W., and Reuter, C. (2021). Technologische Umsetzung des Resilient Smart Farming (RSF) durch den Einsatz von Edge Computing. In *41. GIL-Jahrestagung, Informations- und Kommunikationstechnologie in kritischen Zeiten*. Gesellschaft für Informatik e.V. Zugegriffen am: March 15, 2024.

Gedeon, J., Brandherm, F., Egert, R., Grube, T., and Mühlhäuser, M. (2019). What the Fog? Edge Computing Revisited: Promises, Applications and Future Challenges. *IEEE Access*, 7:152847–152878.

Harchol, Y., Mushtaq, A., Fang, V., McCauley, J., Panda, A., and Shenker, S. (2020). Making edge-computing resilient. In *Proceedings of the 11th ACM Symposium on Cloud Computing*, SoCC '20, page 253–266, New York, NY, USA. Association for Computing Machinery.

Merkl, D. and Cocos, H. (2020). Complex Event Processing on the Edge - Bringing Data Consolidation and Processing closer to Wireless Sensor Networks. In *2020 IEEE International Workshop on Metrology for Industry 4.0 IoT*, pages 395–400.

Rodrigues, T. K., Liu, J., and Kato, N. (2021). Application of cybertwin for offloading in mobile multiaccess edge computing for 6g networks. *IEEE Internet of Things Journal*, 8(22):16231–16242.

Wang, H., Li, Y., Zhang, Y., and Jin, D. (2019). Virtual machine migration planning in software-defined networks. *IEEE Transactions on Cloud Computing*, 7(4):1168–1182.

Winkelhofer, S. (2019). Konzeption und umsetzung von live-migration für docker-container.