FUTUREWEI

Whitepaper

A Look into The Future of Edge Data Center

To be removed for release

V 0.1	2/10/22 initial draft

Trademarks and Permissions

Notice

FUTUREWEI® TECHNOLOGIES, INC.

Boston Research Center

Address: 111 Speen Street, Suite 114
Framingham, MA 01701
United States of America

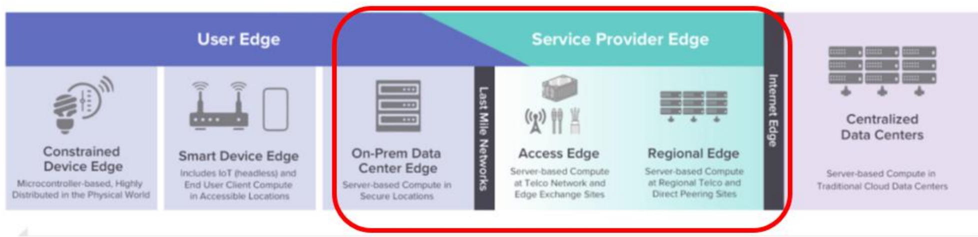
Website: http://www.futurewei.com/

Contents

1. Introduction 7
2. Evolution of The Edge 8
   - 2.1 The Edge Growth by Region and by Industry 8
   - 2.2 Data explosion in the AI Era 10
   - 2.3 New Opportunities at The Edge 11
   - 2.4 Edge Challenges 12
   - 2.5 Market analysis and prediction 13
3. Edge Solution for The Next Decade 14
   - 3.1 Intelligent edge based on GPU|DPU enabled HCI 14
   - 3.2 Software ecosystem at the edge 15
4. Recommended Strategies 15
   - 4.1 A future-proof edge data center solution with intelligent hardware 15
   - 4.2 Cloud-native integration 16
   - 4.3 Move upward in the software stack 17
   - 4.4 More flexible business model 18
5. References 20

摘要

当前谁在用边缘数据中心，用来做什么：目前边缘数据中心的常见应用对数据的使用模式主要有三大类：**数据服务**（例如电信），**数据采入**类（例如 IoT 和 ADAS），**实时数据处理**（例如高端医疗和制造）。按照行业划分，当前在边缘部署中占比较高的行业有：**企业 IT，汽车，电信，医疗，零售，制造业，智慧城市和智能电网**等基础设施类。在边缘生态中有较大份额是基于微型终端如手机和 IoT 设备，这些对于 IT 设备商不构成直接机会，不在本文讨论范围。但是他们带来的数据会增加边缘设备需求，设备商的机会点集中在下图红框的区域，包含 ROBO，区域中心，接入中心等等场景。



边缘数据中心的增长点：未来几年边缘数据应用的趋势主要体现在**数据量的高速增长和数据处理的复杂度不断提升**。过去两年的疫情给**医疗行业**带来了巨大增长，边缘数据中心在医疗行业的高速增长不会随着疫情停止，值得持续关注。边缘应用高速增长的板块还有**自动驾驶，制造业，企业 IT 服务，智慧城市**。此外 **Web 3.0 和元宇宙**类应用也值得关注。相比前面的一些相对成熟板块，这两类应用有更高复杂性，对边缘会带来新挑战和机遇，尤其在**生态和计算**侧。

什么样的架构最能满足我们预测的趋势：在边缘数据中心里，运算，存储，网络的三元核心依然成立，并且依然是以计算为核心。已知的重点需求依然有效：**统一高效的管理平台**，存算合一的**模块化部署**。在此基础上，由于高成长领域例如 **AI/ML** 类应用的快速演变，软件应用的生命周期越来越快，对基础架构提出更多挑战。**整合云原生容器化平台**，**验证支持成熟及新锐软件生态**将大大增强产品竞争力。硬件方面应增加 **GPU/DPU 选项**避免被某些机会点屏蔽。传统 HCI 的架构能够降低运维复杂度，减少硬件维度，从而帮助客户整体降低 TCO，是个很好的基座。在软件生态方面必须加强，形成**软硬一体**的方案。

怎样基于现有产品构建更有竞争力的解决方案：以现有 HCI 方案为基础，在软件，硬件，商务三方面加强，提供类云的用户体验和更高的性价比。软件上整合云原生平台，内嵌集成应用模板支持一键部署各类成熟生态体系和开源及商用软件。对于最流行的应用，推出 App 专用**超融合一体机**，例如 Spark 大数据一体机。硬件上提供 GPU/DPU 选项。商务模式方面可以学习 HPE®的 GreenLake 风格的 PayU 模式进一步帮助客户降低 TCO 提高竞争力。制定商务模式很简单，但是能形成差异竞争力的核心技术是**用量监测分析工具**。

(参考) Linux Foundation 边缘计算 2021 年度报告摘要

• 尽管受到 COVID-19 的影响，但新的边缘基础设施和应用部署仍在增长。与去年的报告相比，十分之七的地区预测有所增加。

• 需要巨大的基础设施投资来支持不断增长的设备和基础设施边缘需求。我们估计，在 2019 年至 2028 年期间，边缘计算累计资本支出高达 **8000 亿美元**。(去年估计 7000 亿)

• 作为云计算的自然延伸，边缘云结构越来越被视为"第四次工业革命"的关键推动因素，其中**物联网，全球共享经济和零边际成本制造**提供了前所未有的巨大机遇。

• 随着边缘部署的扩展和新的**芯片选项**的出现，架构决策成为资本支出和运营支出的重要因素。英特尔®，AMD®和 Nvidia®正处于激烈的竞争中，而 SmartNIC，FPGA 和 DPU 等大大提高人工智能（AI）和机器学习（ML）算法的性价比。

• SD-WAN（软件定义的广域网）和 SASE（安全访问服务边缘）代表了边缘应用的关键虚拟网络技术。**5G 当前没有王牌应用**，并不是主流边缘应用的要求。

• 来自主流公有云服务商的混合云和多云解决方案正在将云体验扩展到偏远地区，并承诺提供一致的应用程序和运营体验。

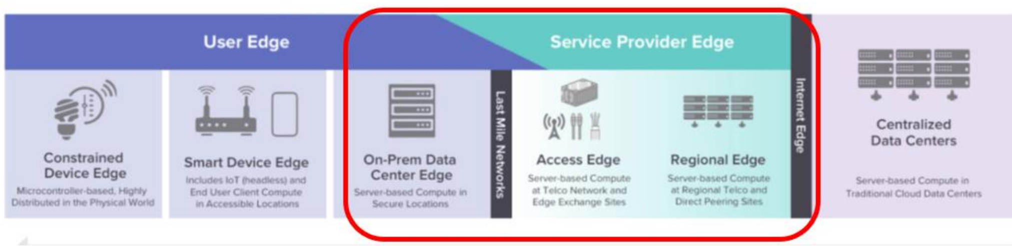
• 现成的边缘应用程序和市场正在变得实用，这要归功于 Google® 于 2020 年 12 月推出的**生态系统**等举措，该生态系统由来自 30 多家独立软件供应商 （ISV） 的 200 多个应用程序组成。IBM®和亚马逊®的类似努力，以及"一体化"定价模式，表明了云式消费如何影响边缘。

• **开源项目**使组织能够加速边缘应用程序的采用和部署。这些社区还促进了整个行业的标准化，加快了创新步伐，同时降低了供应商锁定的风险。

• 保守预测，基础设施边缘部署的全球 IT 电力消耗将从 2019 年的 1 吉瓦增加到 2028 年的 40 吉瓦以上，**复合年增长率（CAGR）为 40%**。据预测，到 2028 年，全球 37%的边缘计算基础设施将用于与移动和住宅消费者相关的用例，其余 **63%支持垂直市场的应用程序**，**如医疗保健，制造，能源，物流，智能城市，零售和运输**。

Executive Summary

Who is currently using edge data centers: There are three main types of data usage patterns for common applications in edge data centers: **data services** (such as telecommunications) and **data acquisition** (such as IoT and ADAS), **real-time data processing** (e.g., high-end medical, manufacture). The industries that currently account for the highest proportion of edge deployment are **enterprise IT, automotive, telecommunications, medical, retail, manufacturing, smart cities, and smart grids**. Opportunity for equipment vendors is concentrated in the areas within the red box in the following figure (LF edge taxonomy), typically including: ROBO, edge DC, regional DC, access DC.



Growth point at the edge: The epidemic in the past two years has brought great challenges and opportunities to the **medical industry**, and the rapid growth of edge data centers in the medical industry will not stop with the epidemic, which deserves continuous attention. Edge applications are also growing rapidly in **autonomous driving, manufacturing, enterprise IT services, and smart cities**. In addition, Web 3.0 and metaverse class applications are also worth paying attention to. Compared with some of the more mature sectors in the past, these two types of applications have higher complexity and bring new challenges and opportunities to the edge, especially on the **ecological and computational** side.

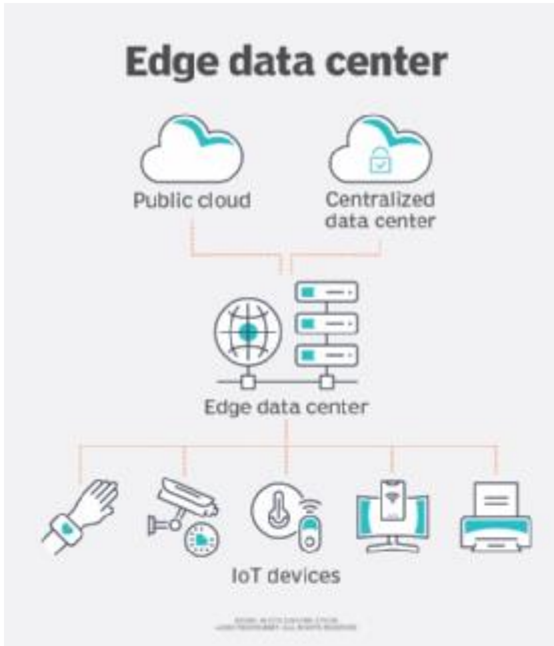
Best architecture for the predicted trend: In the edge data center, the three-dimensional core of computing, storage, and networking is still established, and it is still computing as the core. The known key requirements remain valid: **a unified and efficient management platform**, a **modular deployment** that integrates computing and storage. Life cycle of software applications is getting faster and faster, posing more challenges to infrastructure. **Integrating cloud-native containerized** platforms to **verify and support mature and cutting-edge software ecosystems** will greatly enhance product competitiveness.

Competitive solution based on existing products: On top of current HCI products, integrate cloud-native platform with certified app store on software side. Add GPU/DPU options on the hardware side. Certify popular applications such as Spark, Elastic Search, MongoDB®, ethereum network dApp framework, etc. For the most popular applications, consider introducing a pre-integrated one-click deploy solution.

1. Introduction

To understand the edge data center business better, we first need to understand why companies are building edge data center, and what do they run on them.

This techtarget article [1] gave a great explanation of what is an edge data enter, and some typical applications. As shown in the diagram below, edge data center sits between cloud or traditional data center and end user / devices.



Key benefits that can offered by edge data center includes:

- **Optimized network topology** with better latency and user experience, also reduce internet traffic, scale out beyond central data center space limit.
- **Flexible deployment** (size, power, location, modular) and easy to scale out.

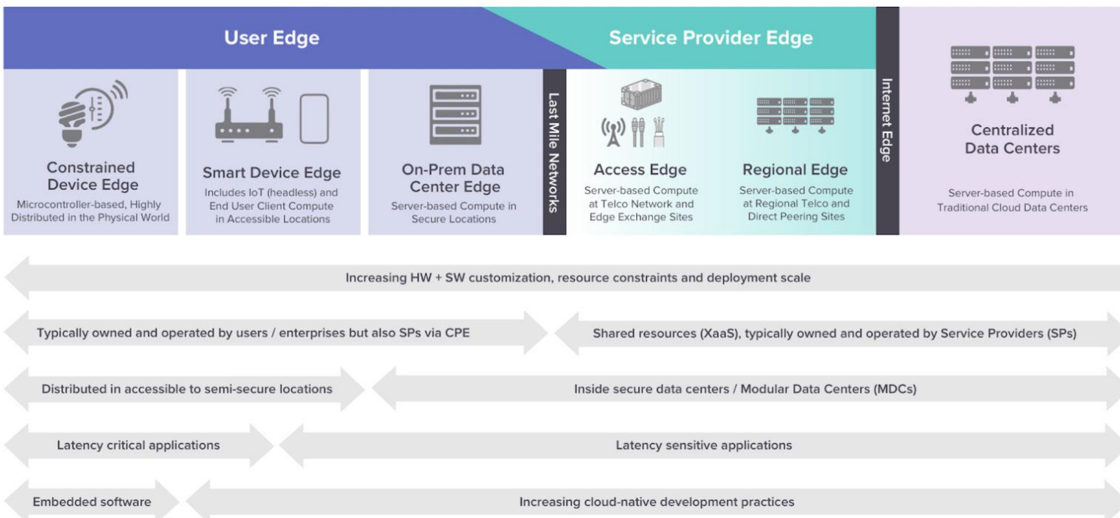
Typical applications include:

- **Telco**. With cell-tower edge data centers, telecom companies can get better proximity to end users by connecting mobile phones and wireless sensors.
- **IoT**. Edge data centers can be used for processing data generated by IoT devices, it extends the total data processing capability, reduce amount of data transferred, and reduce latency.
- **Healthcare**. Some medical equipment, such as those used for robotic surgeries, would require extremely low latency and network consistency, of which, edge data centers can provide.
- **Autonomous vehicles**. Like IoT devices, we can consider autonomous vehicles a very large and complex sensor.
- **Smart factories**. Edge data centers can be used for machine predictive maintenance, as well as predictive quality management. It can also be used for efficiency regarding robotics used within inventory management.

With these characteristics in mind, we will analyze current industry trend and try to predict the direction for future edge data center evolution.

2. Evolution of The Edge

The Linux Foundation's LF Edge has created a formal taxonomy that has received many accolades and is getting widespread adoption. The LF Edge taxonomy visualizes edge computing through the continuum of physical infrastructure that comprises the internet, from centralized data centers to devices.

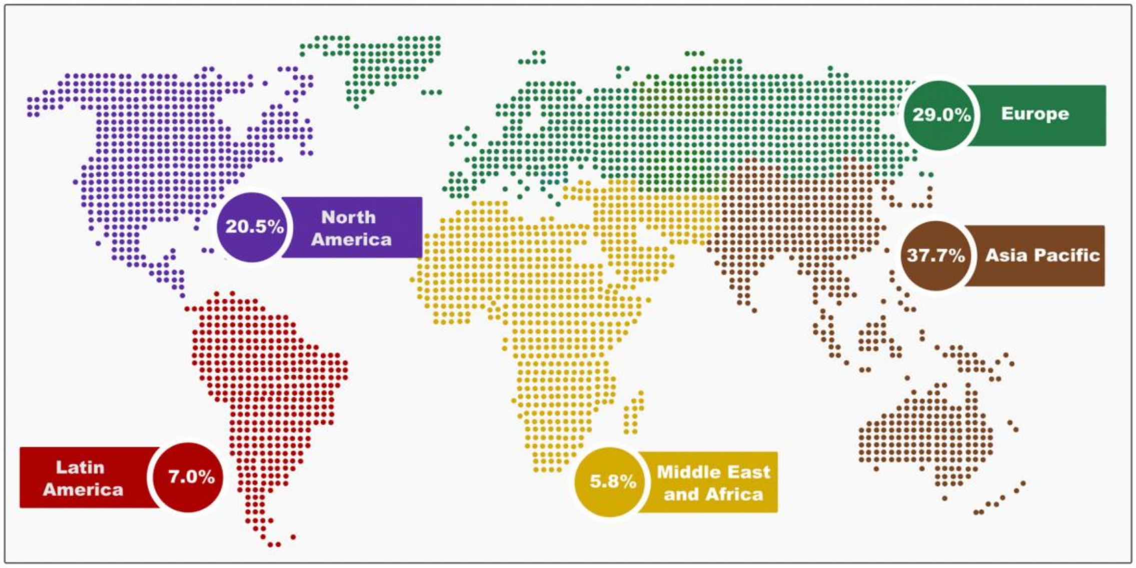


The edge infrastructure includes compute, storage and network, and compute has always been the core focus for edge, and we do not see this changing in any foreseeable future. However, the rapid growth of data at the edge will bring massive growth and demands to all key sectors.

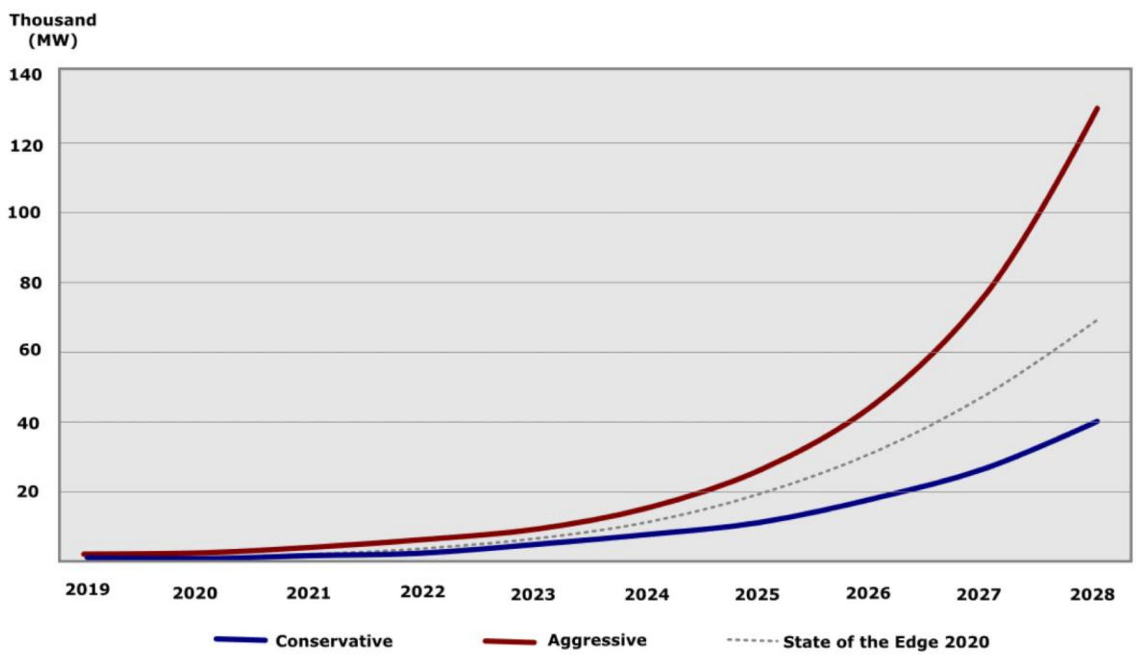
According to "State of the Edge Report 2021" [2], as we enter a hyper-connected era, we are seeing increasing HW + SW customization, resource constraint, and deployment scale at the edge.

2.1 The Edge Growth by Region and by Industry

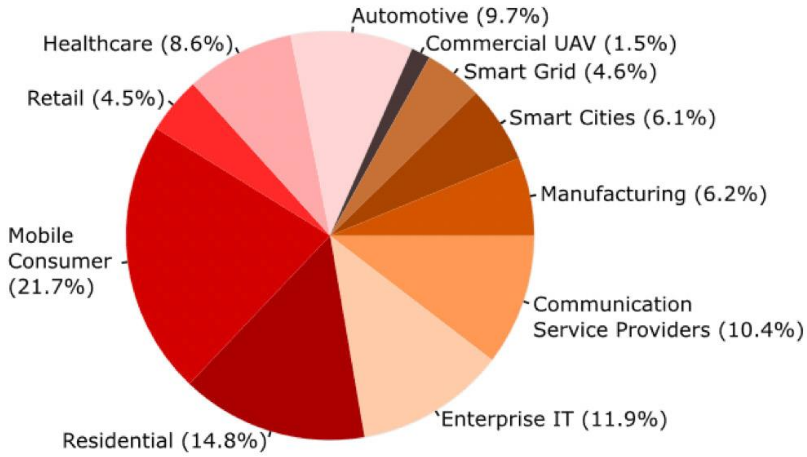
The diagram below shows "Regional Distribution of Edge Deployments" on the global scale.



Expected growth of total edge deployment is exponential with higher than traditional Data Center AGAR:



The chart below shows current break down of edge deployment scale by industry:



Cloud Service Providers (CSP) have driven the early deployments of edge, as they cloudify their networks. In the short to medium term, infrastructure edge demand for Enterprise IT will still be driven by cloud service use cases that are complemented and enhanced with edge computing capabilities. However, it is predicted that in the long-term, Infrastructure Edge demand will be driven by "edge native" use cases such as closed loop enterprise IT functions using AR and VR.

The pandemic has brought specific uses-cases into sharp focus, such as those relating to remote healthcare and assisted living. By 2028, it is forecast that 8.6% of the global Infrastructure Edge will support healthcare use cases. This is a significant expansion from the 6.8% forecast in 2020. We will look at some highlighted growth areas in this chapter.

2.2 Data explosion in the AI Era

IDC® predicts 175ZB data by 2025 [3]. Among these data, 90ZB are predicted to be produced by IoT devices, and an estimated 30% of the 175ZB data will be consumed real-time, which is mostly happening at the edge.

The goal of edge computing is to process data and services as close to the end user as possible. It's an architecture that allows the compute and content delivery process to happen within 10 milliseconds or less of the user. The trend driving the edge computing model is the increased use of consumer mobile devices, especially consumption of video and virtual reality content and the growth of sensors as part of the Internet of Things.

Revisit DCF 2021 predictions [4] told us: As predicted, the whales were extremely busy in edge computing in 2021. Amazon Web Services is adding 30 new Local Zones around the world to extend its edge computing network, along with new ways to connect to the "Internet of Billions of Things." AWS also demonstrated how it can use its edge infrastructure to target specific vertical markets by creating a Private Local Zone for capital markets inside the NASDAQ® data center in Carteret, New Jersey, which will allow NASDAQ to move its exchanges onto the AWS Cloud.

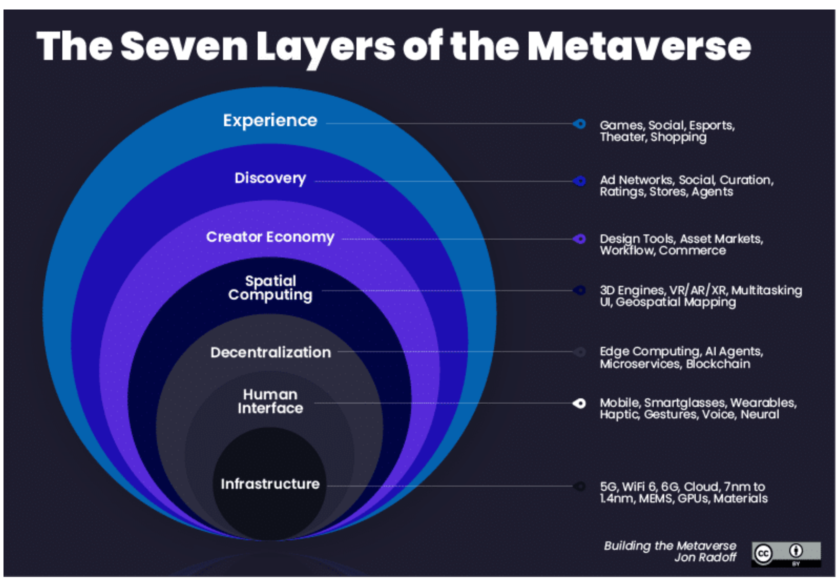
2.3 New Opportunities at The Edge

Latest hot trends in IT industry such as web 3.0 and metaverse all expect large amount of data coming from the edge, and heavily rely on edge's data processing capability, HCI style elastic deployment at the edge would be very beneficial. However, the infrastructure needs to have the ability to adapt to a very dynamic environment and help **reduce the management overhead** throughout the application lifecycle.

The web 3.0 concept started as "Semantic Web" [5] as early as 2001 but has evolved well beyond that concept. Investopedia [6] compared the current web 3.0 concept to the web 2.0 and 1.0 and concluded some key characteristics, and the most important ones are:

- **Decentralization**: This is a core tenet of Web 3.0. With Web 3.0, the data generated by disparate and increasingly powerful computing resources, including mobile phones, desktops, appliances, vehicles, and sensors, will be sold by users through decentralized data networks, ensuring that users retain ownership control.
- **Trustless and permissionless**: Trustless means the network will allow participants to interact directly without going through a trusted intermediary, and permissionless means that anyone can participate without authorization from a governing body. Supporting technology can be blockchains or decentralized peer-to-peer networks, or a combination thereof—such decentralized apps are referred to as dApps.

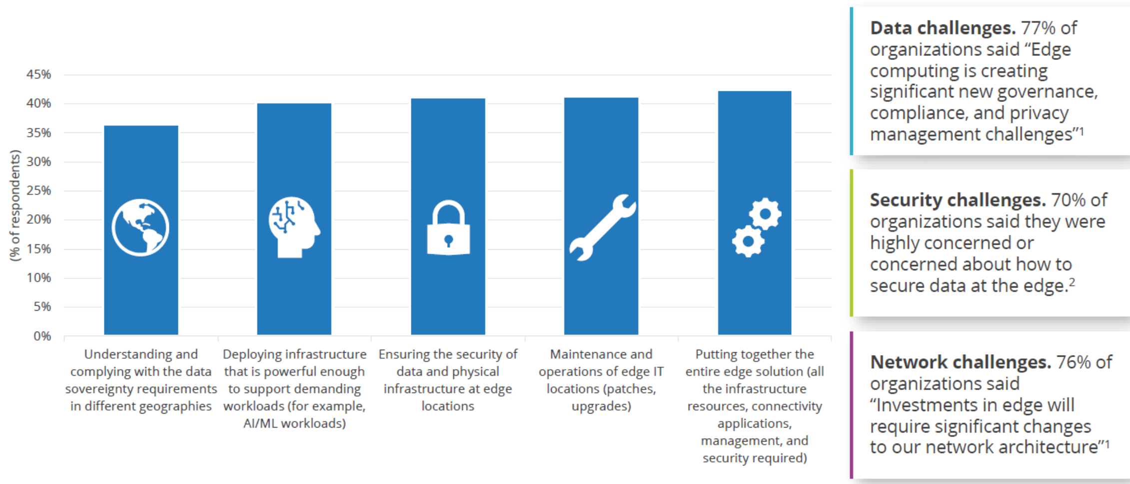
The diagram below shows core technologies used by Metaverse:



Some of the key technologies are shared with Web 3.0, such as block chain, mobile network etc., with additional components such as VR and AR to go beyond web content. Most of data around the Metaverse are predicted to be consumed at the edge at real-time. While there is no killer app created for Metaverse yet, vendors need to pay attention to the trend and be ready for handling the core technologies in this area.

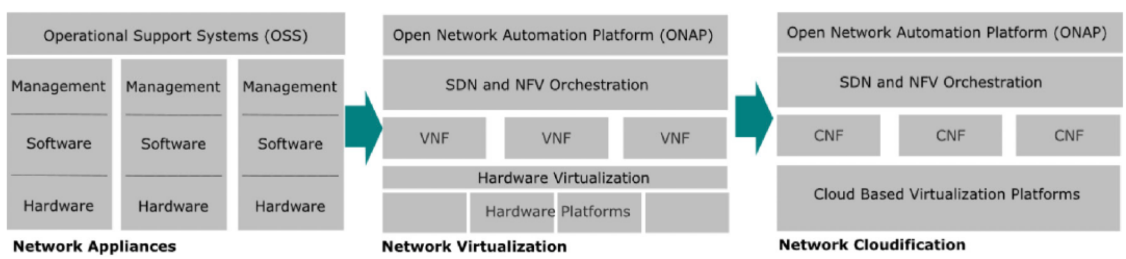
2.4 Edge Challenges

According to IDC “Future Enterprise Resiliency & Spending Survey” December 2021, the top 3 challenges faced by edge consumers are: data management, security, and network.



Graph source: Future Enterprise Resiliency & Spending Survey - Wave 10, December 2021, n = 764

Massive amount of data is making edge-to-center internet connection a major bottleneck in the edge-center-cloud infrastructure [7]. Historically, networks have been built using dedicated network appliances with vertically integrated hardware, software, and management functions. Over the past decades, network at the edge has gone through some major changes to be more virtual and software-defined, to deliver increased performance and agility.



Network is a main bottleneck in a center-edge-cloud layout, reducing data coming out from edge will be critical. This demands a rich matrix of data processing infrastructure at the edge, like what is already in the cloud. This requirement brings both application lifecycle management (ALM) challenge, and some potential security and compliance challenges. The ALM challenge can be largely helped by integrating common software infrastructure into our solution, to provide cloud-like 1-click deployment and unified management. Security is normally handled at the network and application level, not covered in the scope of this white paper.

Computational storage is frequently mentioned as a great way to meet such needs. With the added ability of simple data processing right at the storage node, edge efficiency could get a major boost, and achieve immediate boost of ROI for operators.

While the current trend in edge compute often involves tighter integration like HCI, hyperscale cloud providers are starting to investigate disaggregated architectures in Rack-Scale Architecture (RSA). The dHCI model start to show ROI advantage above certain scale.

With the addition of **NVMe over TCP** standards to the Linux kernel, disaggregation of compute, RAM and storage is becoming more compelling and reliable. Space and other constraints make these technologies especially interesting for edge environments. NVMe over TCP may become a very favorable feature in the coming years at the edge.

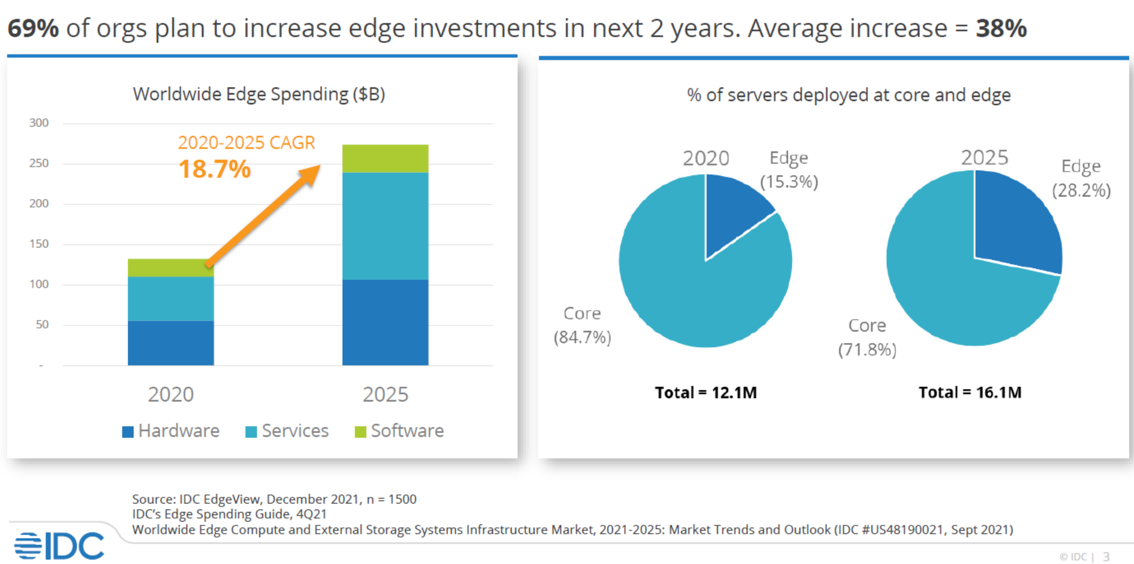
Each edge tier represents unique tradeoffs between scalability, reliability, latency, cost, security, and autonomy. In general, compute at the user edge reflects dedicated, operated resources relative to the users and processes they serve. Meanwhile, the Service Provider Edge and Public Cloud generally represent shared resources in much wider scope.

In most applications, User Edge workloads will run in concert with Service Provider Edge workloads. Workloads on the User Edge will be typically optimized for latency criticality, bandwidth savings, autonomy, safety, security, and privacy, whereas workloads on the Service Provider Edge will be optimized for scale. For example, an AI/ML model might be trained in a centralized cloud data center or on the Service Provider Edge but pushed down to the User Edge for execution.

The boundaries between edge tiers are not rigid. Certain technical and logistical limitations will always dictate where workloads are best run across the continuum based on any given context.

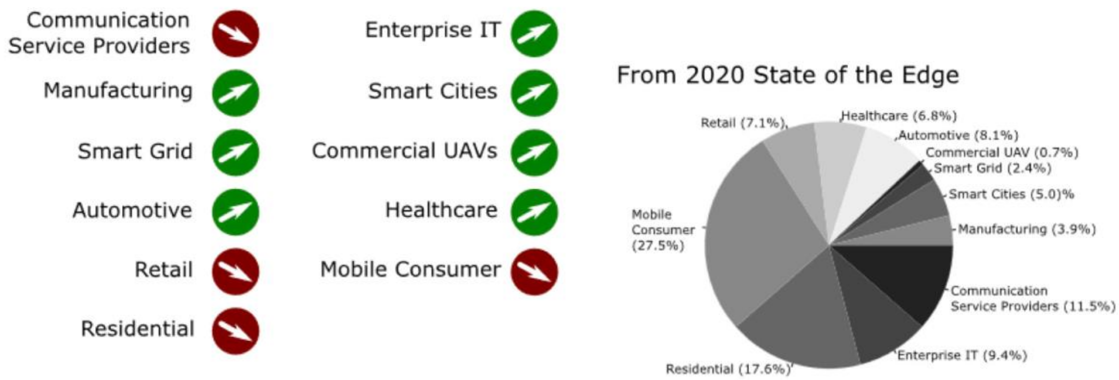
2.5 Market analysis and prediction

Analysts are very optimistic about edge growth [8]. In the latest IDC survey, they found edge spending is predicted to outpace the core by a wide margin.



In addition to the current key consumers of the edge such as healthcare and retail, bleeding edge new applications are starting to attract attention. For example, metaverse advocates believe that "*The future is here, and it's distributed: What the metaverse is and why it needs edge computing*". Also echoed by Web 3.0 advocates here in this article [9] – "*Web 3.0 is here. Don't Let Your Digital Product Fall behind*".

The chart below compares 2021 predictions with 2020 by industry, with the green arrows signifying industries that grows faster than average at the edge. Vendors should focus on opportunities in these areas.



While the newer applications such as “Commercial UAV” (Unmanned Aerial Vehicle), web 3.0, and metaverse are barely visible on the chart, they may grow at much faster rate.

3. Edge Solution for The Next Decade

ROI is the key factor when it comes to any infrastructure spending. Edge computing comprises combinations of systems that span a wide range of locations and conditions and support a diverse set of use cases. While one use case might demand high-powered GPUs for AI (Artificial Intelligence), another use case might demand low power consumption to lengthen battery life. The location of equipment, such as a micro edge data center or a wall-mounted industrial cabinet, places different constraints on the hardware. All these factors result in a wide range of edge hardware that will continue to diversify in 2021, and the core of edge solution should be computing. For this reason, **HCI and computational storage** are solid hardware foundation for edge solutions.

3.1 Intelligent edge based on GPU|DPU enabled HCI

Edge applications has been very messy and sophisticated in the past and will become even more so in the next a few years with the introduction of different hardware and software options. We observed more and more application of heterogeneous compute hardware matrix at the edge, such as GPU, DPU, FPGA in addition to traditional x86 and ARM CPUs. Some of these specialty hardware chipsets can offer magnitudes of higher efficiency in AI and ML compute workloads, dramatically changing the ROI landscape when introduced to the solution.

Due to the massive difference in certain workload, GPU|DPU will become a necessary for an edge solution to stay competitive. The complexity of setting up proper software environment for some of these chipsets also make it highly preferrable for vendors to preload systems with all necessary software infrastructure, ideally, even include a few popular applications that fully utilize the hardware and provide standard APIs for developers.

3.2 Software ecosystem at the edge

The importance of ecosystem supports all boils down to reducing cost. From administration to development, to daily operations. This is very similar to the concept behind HCI but pushes one step further up toward the user space applications. In a typical enterprise IT organization, a good portion of IT personnel resources are spent on bringing the hardware up and running with the targeted applications, the follow up monitoring, and troubleshooting when issue happens. Time spent in the application layer frequently much more than what is spent on setting up and operating the hardware.

When it comes to the edge, things get even more complicated. Unlike large core datacenter with many “standard” IT equipment, edge can have a very rich mixture of hardware and platforms to deal with. Tighter integration of software ecosystem from vendor will save enterprise IT a lot of effort, and lower knowledge requirements on the operator.

4. Recommended Strategies

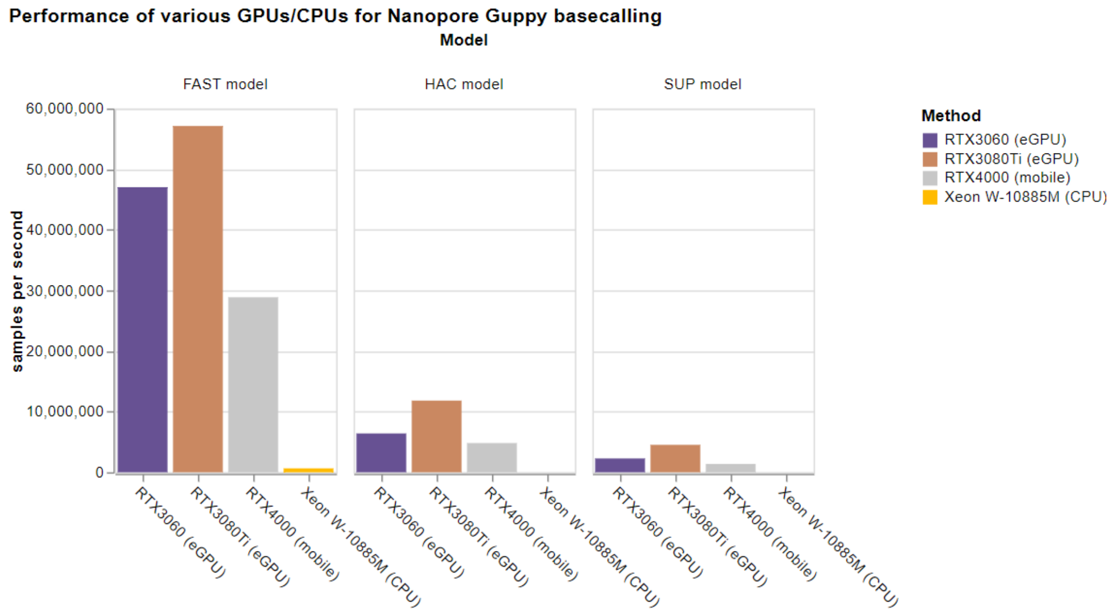
4.1 A future-proof edge data center solution with intelligent hardware

Many of the high growth industries at the edge are heavily relying on AI and Machine Learning. These compute workloads can be run extremely efficiently on GPU, DPU, and FPGA. Research by Xilinx® lab in 2019 [10] presented detailed comparison of energy efficiency of these different hardware architectures.

	CPU	GPU	FPGA
Input Processing	1	**1.79**×	1.41×
Image Arithmetic	1	**3.19**×	2.93×
Image Filters	1	3.17×	**3.89**×
Image Analysis	1	2.34×	**5.67**×
Geometric Transform	1	10.3×	**16.6**×
Features/ OF/ StereoBM	1	7.44×	**22.3**×

Ratios of Energy/Frame Reduction (Reference CPU)

In addition to the operational efficiency, GPU and FPGA also packs much higher computer power per cubic foot, and cost much less per TFLOPS compute power. The AI community consensus is that GPU can offer about 2 magnitude more of computing power density per $ compared to generic CPU, some reference can be easily found on the internet such as in this report [11].



One way to make a vendor's existing edge solution much more competitive for these kinds of workloads is to add the option of integrated GPU, DPU or FPGA. This will bring down both upfront cost and operational cost dramatically and make the solution a highly differentiated offering. On the other hand, if competitors are starting to offer such options earlier, it will become unbeatable in the bid. Based on existing product portfolio, **HCI is a better choice to add GPU support** compared to computational storage, due to its more generic hardware platform and chassis configuration.

4.2 Cloud-native integration

Adding cloud-native container capability to the edge data center not only bring all the benefits of container framework such as easier DevOps, more dynamic workload management etc., but also enables more seamless integration with public cloud. It also opens business opportunities to migrate some customers to on-premise set up for relatively stable and consistent workloads to achieve better ROI.

Amazon AWS has been offering a matrix of GPU assisted computing nodes for a while, however it is not cheap. This article from Lambdalabs® [12] compared on-premise GPU accelerated computing server with Amazon equivalent offerings and found that on-premise can save over 30% TCO over 3 years period.

	AWS (0%)	AWS (Partial)	AWS (100%)	Hyperplane
Upfront	$0	$126,729	$238,250	$109,008
Annual rental	$91,244	$42,241	$0	$0
Annual Co-Location Cost	$0	$0	$0	$15,000
Annual Admin Cost	$0	$0	$0	$10,000
Total Cost Over 3-years	$273,732	$253,444	$238,250	**$184,008**

Some other organizations gave even more aggressive on-premise GPU computing cost saving estimates (up to 84% savings), such as AIME [13]:

Cost for 1 Year

AWS p3.8xlarge Reserved Instance, All Upfront	**76.798,04 €**
AIME R400 4x RTX 2080TI, incl. electricity	**12.051,80 €** **You save 84,3%**
AIME R400 4x Tesla V100, incl. electricity	**41.226,40 €** **You save 46,3%**

Cloud-native integration make enormous amount of cloud-centric ecosystem instantly available to on-premise customers. There are many open-source and commercial Kubernetes based platforms available on the market, they can be great additions to the existing edge solutions.

This capability should be added to **both HCI and computational storage** offerings.

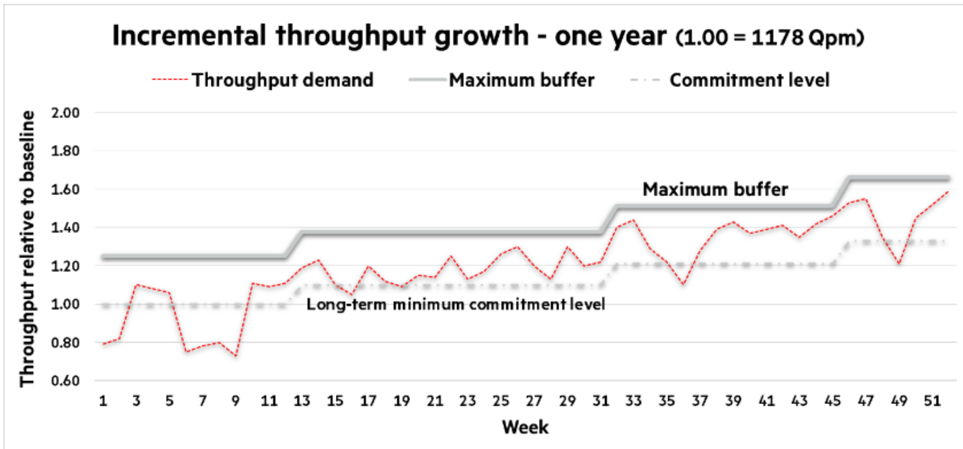
4.3 Move upward in the software stack

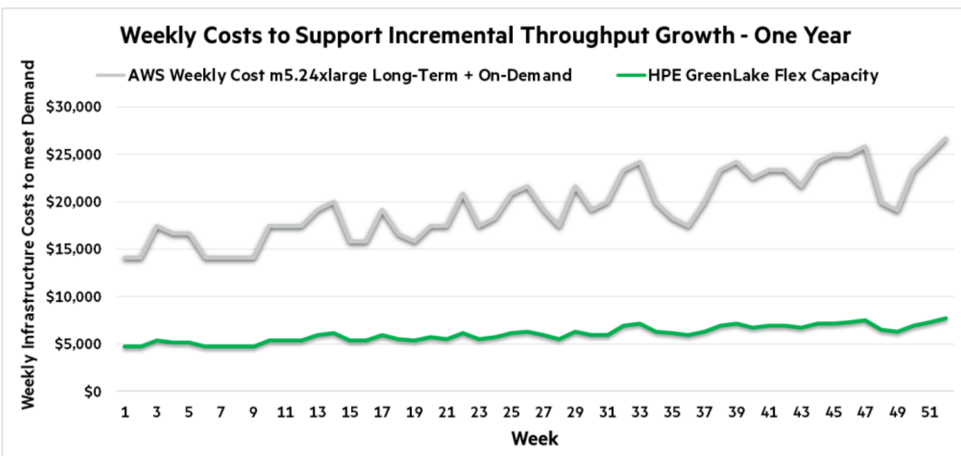
Open source and free software and applications have greatly changed Enterprise IT landscape in recent years. Many things that were only possible with proprietary hardware have been realized by all kinds of software defined approaches on commodity hardware. More choices are becoming available, and each have different balance between hardware, software, and IT personnel resources. This has forced traditional equipment vendor to gradually convert from a simple hardware provider role toward IT service enabler and service role. Frequently when an IT project has a clear expected result, ROI is fixed on the return side, and better number can only be achieved by reducing investment. Within the same IT budget constraint, CapEx spending can be justified by help reducing OpEx if a solution provides tight integration in the application layer to lower operational complexity.

Many cloud-native platforms are working toward this direction, such as DaoCloud's integrated app store that covers many popular applications as **"certified app" with 1-click deployment** support. This kind of integration is going to greatly reduce operational cost for customers and will greatly help move sales of the solution. **Both HCI and computational storage** products can benefit from such integration. Open-source community also have plenty great example of "1-click deploy" [14] of very useful features.

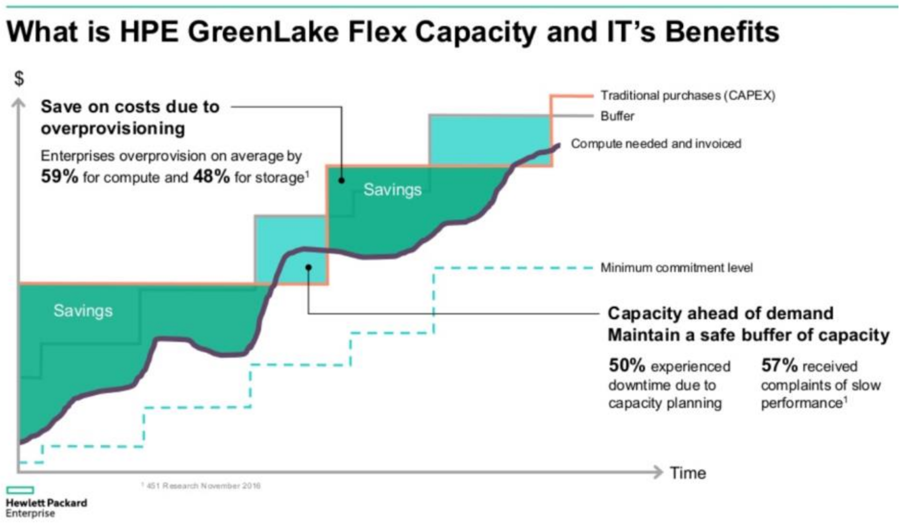
4.4 More flexible business model

The most important competitor for edge solution is public cloud offerings. HPE® GreenLake® style business model offers customer cloud-like flexibility without the hefty outbound network charge from typical public cloud vendors. The chart below shows how the GreenLake model can save cost for customers in a typical scenario [15]:





And here is how the GreenLake Flex Capacity compare to traditional IT purchase [16]:



The flexibility offered by business model innovation can greatly help reduce enterprise IT over-provisioning, and in turn reduce TCO and boost ROI. **Consumption analytics tool** is a key component in implementing such business models, HPE Cloud Cruiser is a good example.

5. References

[1] techtarget, [Online]. Available: https://www.techtarget.com/searchdatacenter/definition/edge-data-center.

[2] Linux Foundation, "State of the Edge 2021," [Online]. Available: https://stateoftheedge.com/reports/state-of-the-edge-report-2021/.

[3] IDC, "IDC: Expect 175 zettabytes of data worldwide by 2025," [Online]. Available: https://www.networkworld.com/article/3325397/idc-expect-175-zettabytes-of-data-worldwide-by-2025.html.

[4] datacenterfrontier, "Scorecard: Looking Back at DCF's 2021 Predictions," [Online]. Available: https://datacenterfrontier.com/scorecard-looking-back-at-dcfs-2021-predictions/.

[5] w3.org, [Online]. Available: https://www.w3.org/standards/semanticweb/.

[6] investopedia, [Online]. Available: https://www.investopedia.com/web-20-web-30-5208698.

[7] datacenterfrontier, "Edge to Core Networking is Critical for IoT Devices, AI and Automation," 10 1 2022. [Online]. Available: https://datacenterfrontier.com/edge-to-core-networking-is-critical-for-iot-devices-ai-and-automation/.

[8] edgeir, [Online]. Available: https://www.edgeir.com/the-future-is-here-and-its-distributed-what-the-metaverse-is-and-why-it-needs-edge-computing-20220127.

[9] emergeinteractive, [Online]. Available: https://www.emergeinteractive.com/insights/detail/web-3-0-trends-digital-product-strategy/.

[10] Xilinx, "Comparing Energy Efficiency of CPU, GPU and FPGA," [Online]. Available: https://arxiv.org/pdf/1906.11879.

[11] M. Benton, "GPU price / performance comparisons for Nanopore basecalling," 16 7 2021. [Online]. Available: https://hackmd.io/@Miles/BJc5hOkCu.

[12] lambdalabs, "Tesla V100 Server: On-Prem vs AWS Cost Comparison," [Online]. Available: https://lambdalabs.com/blog/8-v100-server-on-prem-vs-p3-instance-tco-analysis-cost-comparison/.

[13] AIME, "CLOUD VS. ON-PREMISE - Total Cost of Ownership Analysis," [Online]. Available: https://www.aime.info/en/blog/cloud-vs-on-premise-total-cost-of-ownership-analysis/.

[14] J. Kok, "One click deployment for measuring networking availability of your Kubernetes cluster," 2019. [Online]. Available: https://medium.com/@johankok/one-click-deployment-for-measuring-networking-availability-of-your-kubernetes-cluster-cab1243fb56f.

[15] HPE, "Comparing HPE on-premises infrastructure vs. Amazon Web Services (AWS): Pay-per-use consumption," [Online]. Available: https://community.hpe.com/t5/Servers-Systems-The-Right/Comparing-HPE-on-premises-infrastructure-vs-Amazon-Web-Services/ba-p/7019674#.YkH_rejMI2x.

[16] HPE, "HPE GreenLake and Cloud Technology Partners," [Online]. Available: https://www.slideshare.net/LizMastersLovelace/43019-hpe-greenlake-and-cloud-technology-partners-ctp.