EDGE COMPUTING IN CLOUD COMPUTING

Sungho Kim [1] and ByunghoonKim [2]

[1] Department of Computer Science, Korea University, Seoul, Korea
[2] South High School, Torrance, California, USA

ABSTRACT

The cloud computing environment offers significant flexibility and access to computing re- sources at a reduced cost. This technology is rapidly transforming the landscape of e-services across various fields. In this paper, we examine cloud computing services and applications, high- lighting examples of services offered by leading Cloud Service Providers (CSPs) such as Google, Microsoft, Amazon, HP, and Salesforce. We also showcase innovative cloud applications in areas like e-learning, Enterprise Resource Planning (ERP), and e-governance. This study aims to help individuals and organizations recognize how cloud computing can deliver customized, reliable, and cost-effective solutions across a diverse range of applications.

1. INTRODUCTION

Over the past decade, cloud computing has garnered significant attention from individuals and organi- zations across various fields. This innovative environment offers high flexibility and access to computing resources at multiple levels of abstraction, all at a reduced cost.Cloud Service Providers (CSPs)—such as Google, Microsoft, and Amazon—offer cloud computing resources and services on a leasing basis, allowing customers to utilize these services as needed through specific business models. General cloud services in areas like business, education, and governance are available online and accessible via web browsers, with data and software hosted on cloud servers in data centers.

Cloud services are typically categorized into three main models, often referred to as cloud service models, each representing a distinct level of service within a Service-Oriented Architecture (SOA):

1.1. Software as a Service (SaaS)

In the SaaS model, CSPs manage and operate application soft- ware, operating systems, and computing infrastructure. Users interact with SaaS through a web-based application interface, where complete software applications are provided over the Internet, accessible through devices like laptops, tablets, and smartphones. Unlike traditional software, SaaS eliminates the need for customers to purchase licenses, install, update, maintain, or manage the software on their own devices. SaaS also offers advantages like multi-tenancy, configurability, and scalability. Popular SaaS providers include Zoho, Google Apps, and Salesforce.com.

Key Advantages

- Multi-tenancy: A single instance of software serves multiple customers.
- Configurability: Users can customize settings to fit specific needs.
- Scalability: Software resources can scale up or down based on demand.

David C. Wyld et al. (Eds): Signal, CDKP, AdNLP, MLT, SAI – 2025
pp. 133-145, 2025. CS & IT - CSCP 2025 DOI: 10.5121/csit.2025.150910

Examples Of Saas Providers

- Zoho (Business Applications)
- Google Apps (ProductivityTools)
- Salesforce.com (CRM Solutions)

Additional Insight

According to Gartner (2023), the global SaaS market is expected to grow to $232 billion by 2024, driven by increased demand in healthcare, finance, and education sectors.

1.2. Platform as a Service (PaaS)

In PaaS, CSPs manage and maintain both the system software (like the operating system) and computing resources. Customers are responsible for managing and running the application software on these virtual resources, though they have limited or no control over the underlying operating system and hardware. Unlike SaaS, which provides ready-to-use applications, PaaS allows users to design, develop, test and manage applications directly on the cloud, offering full control over the application lifecycle. PaaS also facilitates collaboration among team members across different locations. Examples of PaaS providers include Windows Azure, Google App Engine and Aptana Cloud.

Key Advantages:

- Allows users to design, develop, test and deploy applications directly in the cloud.
- Provides tools to support team collaboration across locations.
- Reduces time-to-market for applications.

Examples of PaaS Providers

- Windows Azure (Microsoft Cloud Platform)
- Google App Engine (Web Application Hosting)
- Aptana Cloud (Development Tools)

Additional Insight:

McKinsey (2022) reports that PaaS platforms are revolutionizing software development, enabling companies to reduce infrastructure management time and focus on delivering innovative solutions.

1.3. Infrastructure as a Service (IaaS)

In the IaaS model, CSPs offer virtualized resources like network bandwidth, storage, memory, and processing power. Customers are responsible for managing the operating system and application software on these virtual resources. IaaS uses virtualization to transform physical resources into logical resources that can be dynamically allocated and released based on customer needs. Well-known IaaS providers include Dropbox, Amazon EC2 and Akamai.

Key Advantages:

- Scalability: Resources can be dynamically allocated based on demand.
- Cost Savings: Reduces capital expenditure on physical hardware.

- Flexibility: Customers have complete control over the operating system and applications.

Examples of IaaS Providers:

- Dropbox (Storage Solutions)
- Amazon EC2 (Elastic Compute Cloud)
- Akamai (Cloud Delivery and Infrastructure)

Additional Insight:

According to Statista (2023), the IaaS market surpassed $150 billion globally, with major industries like online gaming, media streaming, and enterprise IT using IaaS to achieve high performance and cost optimization

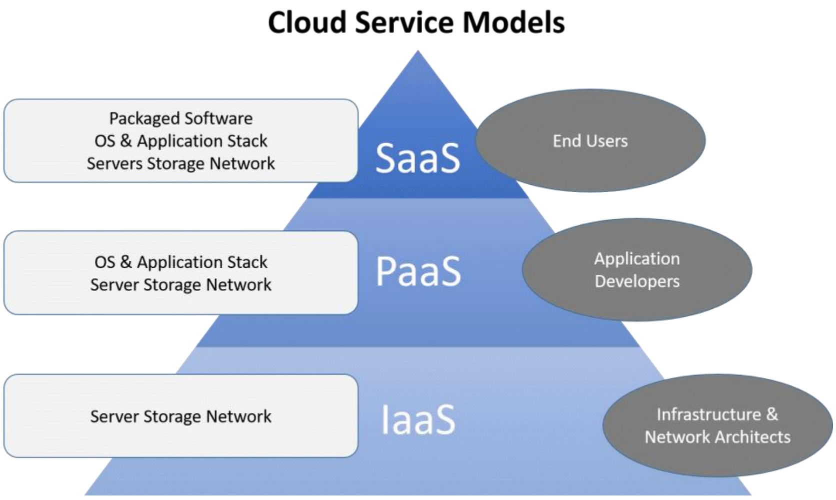

Figure 1: Services provided in cloud computing environment (uniprint,n.d.)

These service models illustrate the different levels of responsibility and control shared between CSPs and customers, as shown in Table 1.

	Application Software	Operating System	Virtual resources /HW
Saas	CSP	CSP	CSP
PaaS	CUSTOMER	CSP	CSP
IaaS	CUSTOMER	CUSTOMER	CSP

Table 1: Resource assignment in cloud service models

The cloud services described above can be provided to cloud customers by CSPs through different applications. In this paper, we explore cloud computing services and applications, we give examples for cloud services provided by the most common CSPs such as Google, Microsoft, Amazon, HP and Sales force and we present innovative applications for cloud computing in e-government, e-learning and Enterprise Resource Planning (ERP). The objective of our study is to help individuals and organizations understand how cloud computing can provide them with customized, reliable and cost-effective services in a wide variety of applications services are public including Amazon EC2, Google App Engine and Salesforce.com.

Private Cloud

A private cloud's computing resources are used solely by one organization, which may manage it internally or via a CSP. This setup is generally more secure than public clouds since it is accessed only by trusted users within the organization. Community and hybrid clouds, in contrast, offer options that lie between public and private models. (Cloudflare.com, n.d.)

Community Clouds

These clouds function similarly to private clouds but are shared among multiple organizations with common missions, policies and security requirements. A common example is educational clouds, which enable universities and research institutions to collaborate on education.

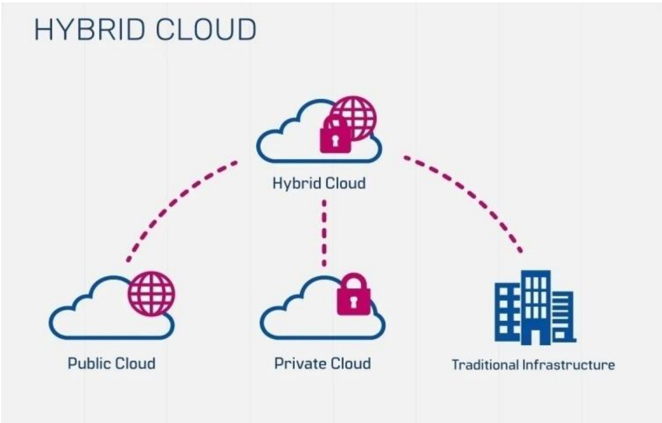

Figure 2: Community Cloud (Ramya Mohana krishnan, 2021)

Hybrid Clouds

A hybrid cloud combines elements of public, private, or community clouds into a single, integrated infrastructure. Each component retains its distinct identity but is interconnected through standardized technology, allowing for seamless management as a unified system. Hybrid clouds allow organizations to maximize resources by handling critical tasks within the private component while outsourcing less sensitive operations to the public component. (Chris Murphy, 2024)

2. ANALYSIS OF CLOUD COMPUTING AND EDGE COMPUTING

Cloud computing systems are classified as public cloud, private cloud, community cloud and hybrid cloud. These classes are known as deployment models and they describe the scope of services offered on the cloud to the customers.

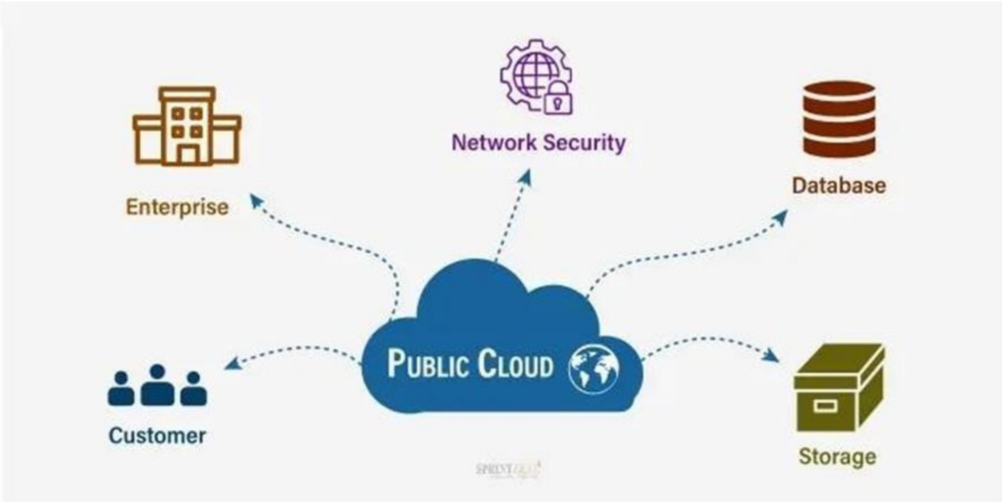

Figure 3: Public Cloud (cloud now tech, 2022)

Public Cloud

Public clouds provide infrastructure and cloud services accessible to the general public via the Internet. Owned and managed by a CSP, these services are typically offered on a pay per use basis.

However, because public cloud users are generally considered untrusted, security and privacy are major concerns with this cloud model. Many well-known cloud systems exhibit characteristics that make them ideal for future IT applications and services. (geeks for geeks, 2024)

The National Institute of Standards and Technology (NIST) has outlined five key characteristics essential to cloud computing systems, which are summarized as follows:

On-demand self-service: Cloud services, such as web applications, servers, processing power, storage, and networking, can be automatically allocated as consumers need them, eliminating the need for manual intervention.

Broad Network Access (mobility): Users can access cloud resources via the internet from virtually anywhere, anytime, and on various devices like smartphones, laptops, or PDAs.

Resource Pooling: Cloud environments combine physical and virtual resources, which are shared but not bound to specific locations. Customers generally lack control over or knowledge of where resources are housed.

Rapid Elasticity: Cloud resources are flexible and can be scaled up or down quickly according to demand. This elasticity creates the impression of unlimited resources, available in any quantity whenever needed.

Measured Services: Cloud providers monitor, manage, and optimize resources under a pay-as-you-go model. Consumers access these resources similar to utilities like electricity or water, paying only for what they use.

Additional characteristics of cloud computing include:

Multitenancy: The cloud serves multiple users at once, allowing shared resources across network, host, and application levels while maintaining isolation for each user's unique virtual setup.

Scalability: Cloud infrastructure is highly scalable; providers can expand capacity with minimal adjustments to hardware and software.

Reliability: High reliability is achieved through multiple backup sites, making the cloud an ideal solution for critical operations and disaster recovery.

Economies of Scale: Cloud infrastructures are designed for maximum scale to reduce costs, often by situating facilities in areas with lower real estate and power costs.

Cost Effectiveness: Consumers can rent exactly the computing resources and IT services they need, avoiding large investments in complex infrastructure and reducing costs for both organizations and individuals.

Customization: Cloud environments are adaptable, allowing users to configure infrastructure and applications to meet specific needs.

Efficient Resource Utilization: Resources are allocated only as needed, maximizing efficiency in their use.

Maintainability: Cloud service providers (CSPs) alleviate the software and hardware maintenance responsibilities for users.

Collaboration: Platform as a Service (PaaS) supports collaborative work among users within an organization or across multiple organizations.

Virtualization: The cloud isolates physical resources virtually, allowing users to focus on their applications without managing underlying hardware.

Green Technology: Cloud computing optimizes resource sharing among users, reducing the need for large, power-intensive resources.

High Performance: Cloud infrastructures offer high-performance environments with extensive storage and powerful computing capabilities.

Google AdWords and AdSense: These are advertising platforms used to promote and monetize content.

Cloud Computing Services

Examples of cloud services provided by leading CSPs include:

Windows Azure App fabric: This provides infrastructure support for applications that operate in the cloud or within an organization.

Windows Azure Marketplace: An online market place where users can buy and sell Software applications and data.

Google Cloud Services: Google offers a variety of integrated applications and services, making it a popular CSP choice for businesses. By handling software development and maintenance, Google saves users significant time and cost. Key services include:

Gmail: An email service with 25GB of storage, low spam levels, and mobile accessibility. Gmail also includes an integrated chat feature that saves conversations as emails.

Google Docs: Enables users to create and store spreadsheets, word documents, and presentations on cloud servers. Files can be accessed online anytime, promoting collaboration among geographically dispersed team members. Security is ensured through advanced encryption accessible only to authorized users.

Google Analytics: A tool for monitoring website traffic.

Amazon Web Services (AWS): AWS provides a scalable, cost-efficient cloud platform for businesses of all sizes. With AWS, companies can flexibly scale infrastructure to meet their needs and only pay for resources used. AWS prioritizes data security, utilizing certifications like ISO 27001, FISMA moderate, HIPAA, and SAS 70 Type II. Key services offered include:

Amazon Elastic Compute Cloud (Amazon EC2): Provides configurable cloud-based computing resources.

Amazon Simple Storage Service (Amazon S3): A secure, scalable storage solution for data of any size, accessible online.

Amazon Virtual Private Cloud (Amazon VPC): Connects an organization's on-premises infrastructure to the AWS cloud via a Virtual Private Network (VPN).

Picasa: A cloud-based tool for show casing and uploading product images.

What is Edge Computing?

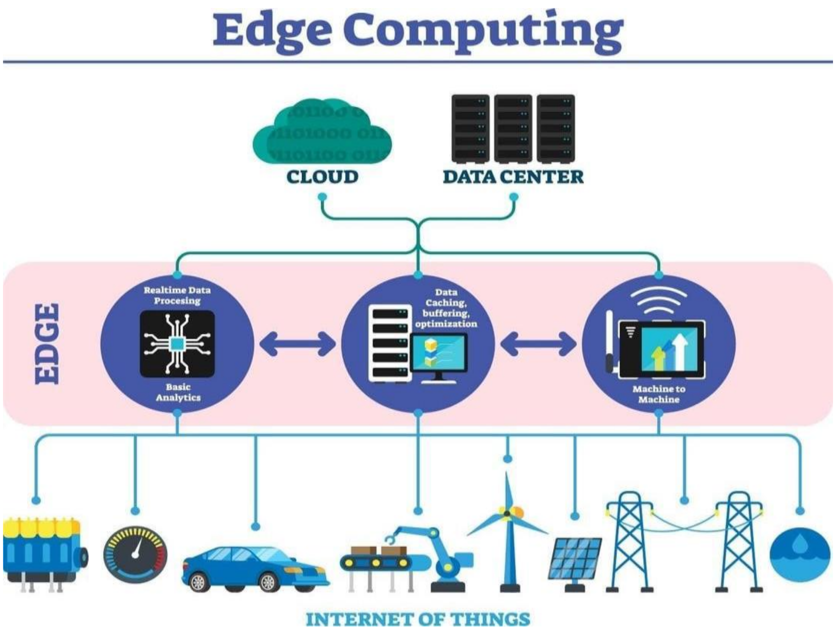

Figure 4: Edge Computing (Anushka mahajan 901, 2023)

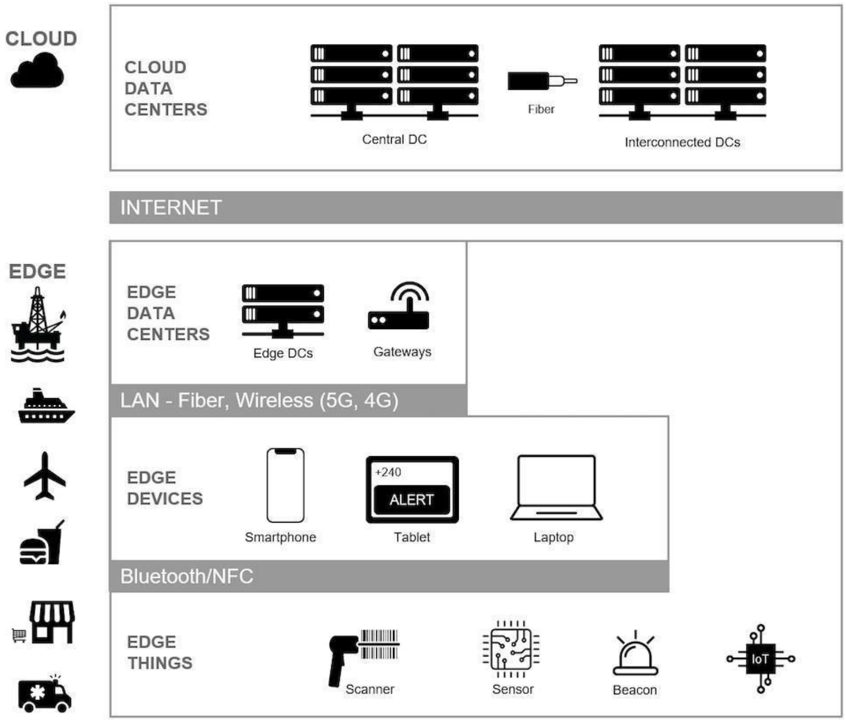

Figure 5: Architecture of Cloud and Edge Computing Ecosystem (Prabakar,2022)

Edge computing refers to a distributed computing paradigm that processes data at or near the source of its generation, rather than relying on a centralized data center or cloud. The aim is to reduce latency, enhance speed, and improve performance by bringing computation and data storage closer to the end-user. By minimizing the need for data to travel long distances, edge computing facilitates real-time processing, which is especially critical in applications requiring immediate responses.

Why Is It Useful?

Edge computing is essential in scenarios where:

1. Low Latency is critical, such a sin autonomous vehicles or industrial automation.
2. Bandwidth Optimization is necessary, a sin smart cities where sensors generate vast amounts of data.
3. Data Sovereignty must be maintained, a sin health care settings where patient data must remain localized for regulatory reasons.

Real-Life Example: Smart Traffic Management Systems

One of the most impactful uses of edge computing is in smart traffic management. Cities equipped with IoT-enabled cameras and sensors can process video and traffic data locally using edge devices. These systems can:

- Identify traffic congestion in real-time.
- Adjust traffic light patterns dynamically.
- Alert drivers about accidents or hazards ahead.

For instance, a system implemented in Barcelona leverages edge computing to analyze traffic patterns and manage vehicular flow efficiently. This reduces commute times and lowers carbon emissions, enhancing urban sustainability.

Image 1 : Los Angeles Department of Transportation (LADOT). (n.d.).

The Los Angeles Department of Transportation (LADOT) utilizes the Automated Traffic Surveil- lance and Control (ATSAC) system to manage the city's traffic flow. ATSAC employs a network of sensors and cameras to monitor real-time traffic conditions and adjust signal timings accordingly. This system enhances traffic efficiency and reduces congestion across Los Angeles.

Edge computing plays a vital role in the efficiency of ATSAC. The system processes data locally at or near its sources, such as traffic sensors and cameras placed strategically across the city. This local processing reduces latency, enabling real-time analysis of traffic conditions and immediate adjustments to signal timings. For example, during peak traffic hours or in response to unexpected congestion caused by accidents or construction, the system can dynamically recalibrate signals to improve traffic flow and minimize delays. This real-time adaptability is a hallmark of edge computing's utility in urban infrastructure.

A key feature of ATSAC is its integration with Los Angeles's broader traffic management framework, which includes public transportation systems, emergency services, and parking management. This integration ensures that the system not only optimizes vehicular traffic but also facilitates smoother coordination across various transportation modalities. The system's local edge processing capabilities ensure that even if connectivity to a central cloud is disrupted, critical functions like traffic signal adjustments can continue uninterrupted, ensuring reliability in dynamic urban conditions.

In addition to improving traffic efficiency, ATSAC significantly contributes to environmental sustainability. By reducing vehicle idling and congestion, the system lowers fuel consumption and decreases greenhouse gas emissions. Moreover, the data generated and processed by ATSAC provides valuable insights for urbanplanners, enabling data-driven decisions for infrastructure development and policy-making.

Future developments in ATSAC are likely to further integrate advanced edge computing technologies. For instance, pairing ATSAC with vehicle-to-infrastructure(V2I) communication could allow connected vehicles to receive real-time traffic updates and optimize their routes accordingly. Similarly, incorporating artificial intelligence could enhance the system's predictive capabilities, enabling it to anticipate congestion patterns and take preemptive measures.

The ATSAC system exemplifies how edge computing can revolutionize urban mobility. By processing data at the edge, LADOT ensures swift responses to dynamic traffic conditions, enhances system reliability, and contributes to sustainable urban development. This makes it a model for other cities looking to adopt edge-enabled traffic management solutions to address the growing challenges of urban congestion and mobility.

On the other hand, Barcelona's smart traffic management system exemplifies the transformative potential of edge computing and IoT technologies. It not only improves the quality of life for residents but also serves as a global model for sustainable urban mobility solutions. Barcelona's smart traffic management system is a corner stone of its broader smart city initiatives, designed to improve urban efficiency and sustainability. By deploying a network of IoT-enabled cameras and sensors across critical intersections, the city has created an infrastructure capable of collecting real-time data on traffic flow, vehicle density, and environmental conditions. Leveraging edge computing, the system processes this data near its source, minimizing latency and enabling rapid decision-making. This is particularly vital for real-time adjustments to traffic signal timings, ensuring smoother vehicle flow and reducing congestion.

One of the standout features of the system is its ability to dynamically manage traffic signals based on current conditions. This reduces idling times, shortens commute durations, and enhances over all mobility with in the city. Furthermore, the system detects accidents or hazards promptly, allowing for immediate rerouting and swift emergency responses. By integrating these capabilities with Barcelona's broader smart city platform, which includes public transportation data and parking systems, the city offers a cohesive and efficient urban experience.

The benefits of this system extend beyond convenience. It significantly lowers carbon emissions by reducing vehicle idling and fuel consumption, aligning with Barcelona's environmental goals. Additionally, the wealth of data generated by the system provides city planners with actionable insights for future infrastructure improvements and policy decisions.

Despite its success, the system faces challenges such as scalability and data privacy concerns. To address these, Barcelona has adopted modular designs for its infrastructure, ensuring the system can expand as needed. It has also implemented robust encryption and compliance measures to protect collected data under European GDPR regulations. Looking ahead, Barcelona plans to integrate artificial intelligence for predictive traffic analytics, enhance vehicle-to-infrastructure communication, and support sustainability initiatives such as low-emission zones and alternative transportation methods.

Barcelona's smart traffic management system exemplifies the transformative potential of edge computing and IoT technologies. It not only improves the quality of life for residents but also serves as a global model for sustainable urban mobility solutions.

Both Barcelona and Los Angeles demonstrate how smart traffic management systems can effectively utilize real-time data to enhance urban mobility and reduce congestion. Edge computing's role in these scenarios not only improves system efficiency but also ensures reliability, even when the connection to a central cloud server is weak or lost.

The Rise of Edge Computing: Key Insights and Research Highlights

Edge computing is one of the fast-developing landscapes of cloud computing that solves some very critical challenges and opens up a lot of new opportunities. Following are some of the key trends, statistics, and research areas to identify the rising importance of the sector.

Key Statistics of Edge Cloud Computing

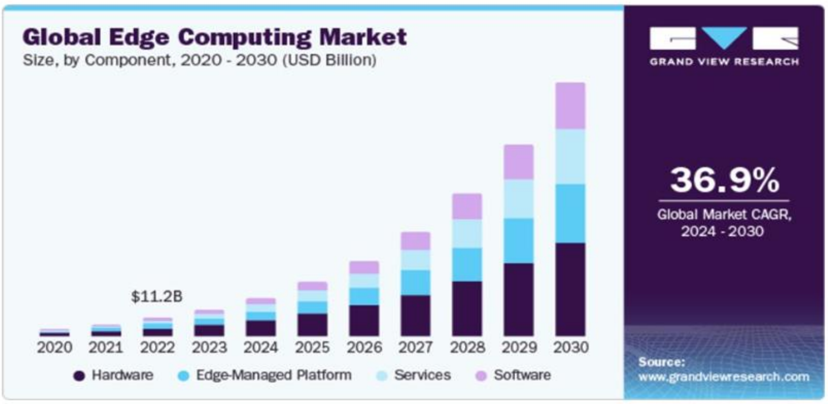

Image 2: Global Edge Computing Market (Grand View Research, 2024)

The volume of data generated globally is anticipated to reach a staggering 180 zetta bytes by2025. Notably, 70% of this data, largely produced by IoT devices, will be processed at the edge of networks rather than in centralized data centers, highlighting the pivotal role of edge computing in managing this immense data influx.

The market for edge computing is experiencing remarkable growth, with its valuation projected to expand from $11.2 billion in 2024 to $43.4 billion by 2027, reflecting a compound annual growth rate (CAGR) of 36.9%through 2030. This rapid expansion underscores the increasing adoption of edge computing solutions across various industries.

The proliferation of IoT devices further fuels the demand for edge computing. By2025,the number of IoT-connected devices is expected to exceed 41.4 billion. This surge in smart devices necessitates localized data processing at the edge to meet the challenges of managing vast data volumes efficiently and effectively.

The rollout of 5G networks is revolutionizing the adoption of edge computing. Industry forecasts indicate that by 2025, 75% of organizations will implement 5G-enabled edge solutions. This integration promises to deliver faster and more efficient systems, transforming industries with enhanced connectivity and low-latency performance.

Emerging Research Areas

Emerging research areas in edge computing reflect its transformative potential across diverse domains and the continuous innovation required to address technological challenges. One key area of focus is the exploration of novel architectures tailored for specific applications. Researchers are delving into how edge computing can enhance smart cities, precision agriculture, and advanced health care systems by enabling efficient, localized data processing and decision making.

The development of distributed systems is another significant research focus. These systems aim to address scalability and reliability issues inherent in centralized computing models. By decentralizing processes and distributing resources closer to the data source, researchers are creating frameworks that are more adaptable and resilient to dynamic operational demands.

The integration of artificial intelligence (AI) with edge computing represents a groundbreaking advancement in technology. This combination enables real-time data processing and decision-making with greater efficiency and accuracy. AI algorithms deployed at the edge allow for rapid analysis of data streams, facilitating applications such as predictive maintenance, autonomous systems, and enhanced user experiences.

Security and privacy remain paramount concerns in edge computing, prompting researchers to design robust measures to safeguard sensitive data and systems. These efforts focus on developing advanced encryption methods, secure communication protocols, and resilient system architectures to address vulnerabilities inherent in decentralized environments. As edge computing expands, ensuring the security and privacy of distributed systems becomes increasingly critical, particularly in sectors such as healthcare and finance, where sensitive data is frequently processed. Researchers are exploring innovative techniques, such as zero-trust architectures and blockchain, to fortify the security of edge networks while maintaining efficiency.

The integration of emerging network technologies like 5G and 6G further enhances the potential of edge computing. These networks enable faster data transmission and lower latency, making edge solutions more viable for applications requiring real-time responsiveness, such as autonomous vehicles and industrialautomation.The seamless collaboration between edge computing and advanced network infrastructure is a key area of ongoing research, driving improvements in system performance and scalability.

Finally, edge computing’s synergy with fog and cloud computing models is an essential focus for optimizing resource allocation. By leveraging the strengths of each model, researchers are working to create hybrid frameworks that maximize efficiency and minimize costs. This collaborative approach ensures that computing resources are utilized effectively while meeting the diverse demands of modern applications, from smart agriculture to personalized medicine. As these research areas evolve, edge computing continues to reshape industries, offering solutions that are faster, smarter, and more secure.

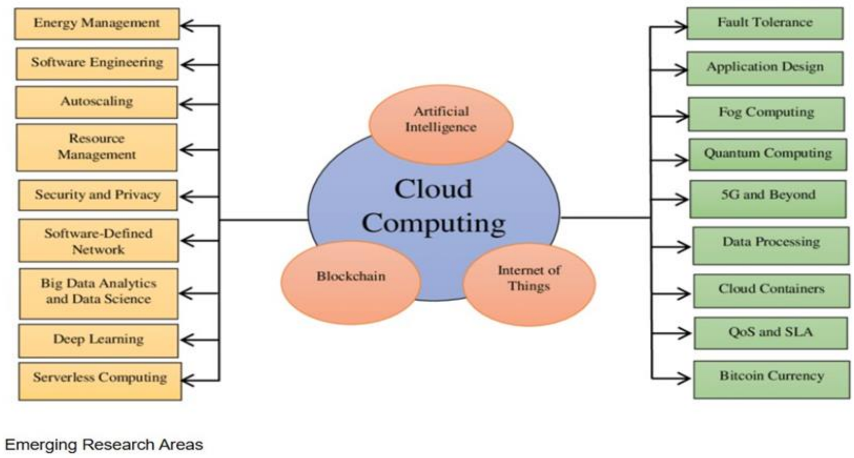

Image 3: Emerging Research Areas (Prabakaretal., 2022)

Edge computing is not an extension of cloud computing but a pivoting technology that shapes the future of data processing. The role it plays in the world, which is getting smarter day by day, in enabling faster, smarter, and more reliable systems, will continue to grow.

REFERENCES

[1] Anushka mahajan 901. (2023, July 20). Serverless computing and edge computing integration: Bridging the gap for real-time applications. Medium. https://medium.com

[2] Armbrust, M., Fox, A., Griffith, R., Joseph,A.D., Katz, R.H., Konwinski, A.,Lee, G., Patterson, D., Rabkin, A., Stoica, I., & Zaharia, M. (2010). A view of cloud computing. Communications of the ACM, 53(4), 50–58. https://doi.org/10.xxxx

[3] Buyya, R., Yeo, C. S., Venugopal, S., Broberg, J., & Brandic, I. (2009).Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. Future Generation Computer Systems, 25(6), 599–616. https://doi.org/10.xxxx

[4] Chris Murphy. (2024, February29). What is hybrid cloud? Oracle. https://www.oracle.com/cloud/hybrid-cloud/what-is-hybrid-cloud

[5] Cloudflare.com. (n.d.). What is a private cloud? Cloudflare. https://www.cloudflare.com/learning/cloud/what-is-a-private-cloud

[6] Gartner.(2019). 75% of enterprise-generated data to be created and processed at the edge by 2022.Seagate. https://www.seagate.com

[7] Gartner.(2023). Forecast analysis: Cloud computing worldwide. Gartner. https://www.gartner.com

[8] Geeks for Geeks. (2024, May 14). What is public cloud? Geeks for Geeks. https://www.geeksforgeeks.org/what-is-public-cloud

[9] International Data Corporation (IDC).(2018,December14).Worldwidedatatoreach175zettabytesby2025.

[10] Rack Solutions. https://www.racksolutions.com

[11] Intel.(2021).5Genabling the next wave of edge computing applications. Intel Corporation. https://www.intel.com Journal of Emerging Trends in Computing and Information Sciences. (2012,July6). Retrieved from CIS Journal. http://www.cisjournal.org

[12] Los Angeles Department of Transportation (LADOT). (n.d.). Los Angeles Department of Transportation. City of Los Angeles. https://ladot.lacity.gov

[13] Markets and Markets. (2020). Edge computing market worth $43.4 billion by 2027. https://www.marketsandmarkets.com

[14] McKinsey & Company.(2022).Cloud's trillion-dollar prize is up for grabs. McKinsey. https://www.mckinsey.com

[15] Microsoft Azure. (2023). What is Platform as a Service (PaaS)? Microsoft Azure. https://azure.microsoft.comPrajwal. (2023, May 10). What is public cloud? Sprint zeal. https://www.sprintzeal.com/blog/what-is-public-cloud PRACE. (n.d.). Edge computing: An overview of framework and applications. PRACE. https://prace-ri.eu/wp-content/uploads/Edge-Computing-An-Overview-of-Framework-and-Applications.pdf

[16] Ramya Mohana Krishnan . (2021, August 5). What is community cloud? Spice works. https://www.spiceworks.com/tech/cloud/articles/what-is-community-cloud/

[17] Rittinghouse, J.W., & Ransome, J. F. (2017). Cloud computing: Implementation, management, and security. CRC Press. https://doi.org/10.xxxx

[18] Statista Research Department.(2022) .Number of IoT connected devices worldwide 2020–2025. Statista. https://www.statista.com

[19] Statista. (2023).Global IaaS market revenue from 2018 to 2023.Statista.https://www.statista.com

[20] Vaquero, L.M., Rodero-Merino, L.,Caceres, J.,& Lindner,M.(2008).A break in the clouds: Towards a cloud definition. ACM SIGCOMM Computer Communication Review, 39(1), 50–55. https://doi.org/10.xxxx