COS 109 – Cloud Computing Concepts

Name : Chaw Thiri Win

Contents

1. Introduction

This report explores cloud computing, its service models (IaaS, PaaS, SaaS), and deployment models. It identifies key characteristics, identifies leading global providers, and discusses Service-Level Agreements (SLAs). The report also discusses cloud security, data classification, and multi-tenancy implications. The aim is to provide a comprehensive understanding of cloud computing concepts and their practical relevance in today's IT environment.

2. TASK 1: Cloud Computing

2.1. Three primary service models in cloud computing

Infrastructure as a Service (IaaS) is a flexible cloud computing model that allows customers to rent virtualized computer resources online, offering scalability, affordability, and on-demand availability. It is popular among large-scale businesses for data backup, development environments, website hosting, and high-performance computing jobs.

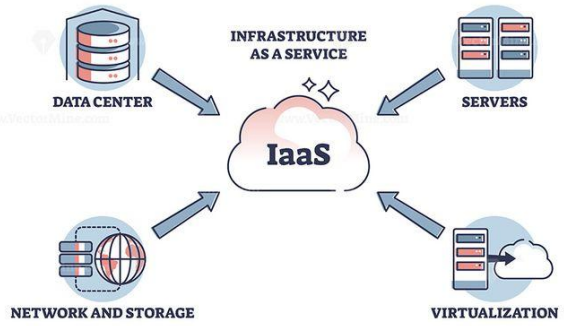

Figure 2a.1 : Infrastructure as a Service (IaaS)

Platform as a Service (PaaS) is a cloud computing architecture that offers a comprehensive development and deployment environment for applications, including web servers, database management systems, programming languages, and middleware. It increases operational efficiency and speeds up the development lifecycle, but has limitations like restricted infrastructure control and dependency on provider availability.

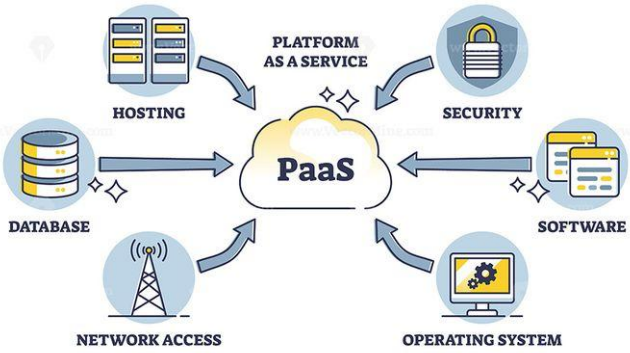

Figure 2a.2 : Platform as a Service (PaaS)

Software as a Service (SaaS) is the most popular cloud computing layer, offering managed software applications like Microsoft Office 365, Gmail, Salesforce, Dropbox, and Oracle ERP Cloud. Despite its affordability, usability, security features, and low cost, SaaS faces challenges like customization, internet reliance, and data control issues.

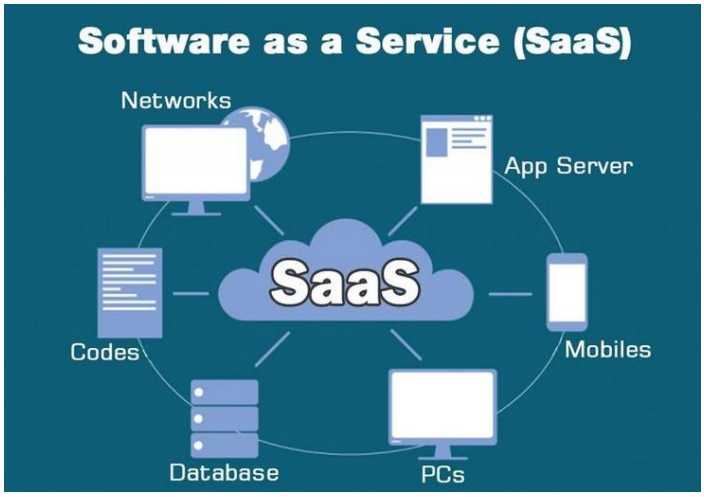

Figure 2a.3 : Software as a Service (SaaS)

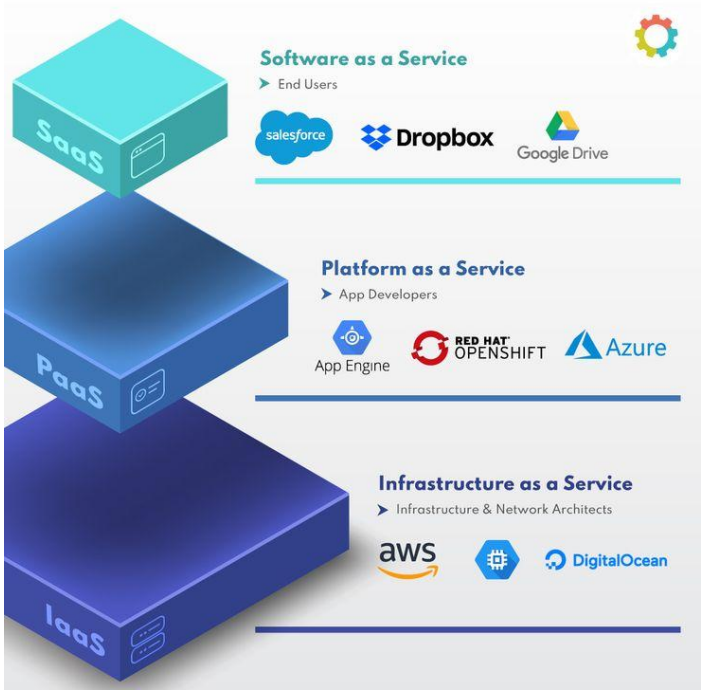

Figure 2a.4 : Real-World Example of Iaas, Paas, Saas

Feature	IaaS (Infrastructure as a Service)	PaaS (Platform as a Service)	SaaS (Software as a Service)
Description	Provides virtualized computing resources (servers, storage, networking) over the internet. Users manage OS, applications, and data.	Offers a complete platform for developing, running, and managing applications without handling underlying infrastructure.	Delivers ready-to-use software applications over the internet, managed entirely by the provider.
User Controls	Applications, data, runtime, middleware, OS	Applications and data	Only application settings and user data
Provider Manages	Virtualization, servers, storage, networking	Runtime, middleware, OS, virtualization, servers, storage, networking	Everything: application, data, runtime, OS, servers, storage, networking
Typical Use Cases	Data backup, development environments, website hosting, high-performance computing	Application development, API creation, analytics, IoT services	Email, CRM, collaboration tools, file sharing
Advantages	High scalability, flexibility, cost savings, on-demand resources	Faster development, reduced complexity, automatic scalability	Easy access, minimal setup, cost-effective, automatic updates
Limitations	Requires technical expertise, regulatory challenges	Limited infrastructure control, potential vendor lock-in	Limited customization, data privacy concerns, dependency on provider
Real-World Example	Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform	AWS Elastic Beanstalk, Google App Engine, Heroku	Salesforce, Microsoft Office 365, Dropbox

2.2. Four main cloud deployment models

Cloud computing is a method of providing clients with IT infrastructure. To qualify as a cloud service, five requirements must be met: on-demand self-service, broad network access, dynamic resource pooling, rapid elasticity, and measured services. Clients can make decisions about service start and stop without directly communicating with providers. Providers must also track usage and bill appropriately, ensuring better service quality.

2.2.1 Public Cloud

A public cloud is a cloud deployment strategy in which infrastructure and services are made available to anybody who pays for them via the internet by third-party providers. It has no startup fees and operates on a pay-per-use basis, which makes it affordable and widely available. Public clouds are usually multinational, providing services all over the world with the opportunity to choose where data is stored, however doing so may impact redundancy and availability. Because the provider handles all operations, consumers are not required to manage infrastructure or maintenance under this approach.

Advantages:

Economical and expandable: Pay as you go, minimal upfront costs, and dynamic scalability.

Simple to use: No hardware upkeep is required, and it can be accessed from any location with an internet connection.

High availability: Consistent uptime with upgrades and infrastructure handled by the provider.

Disadvantages:

Security issues: Less control over data protection in a shared setting.

Limited customization: There is less freedom to choose the infrastructure and configuration.

Internet-dependent: Not suitable for highly classified material; requires good connectivity.

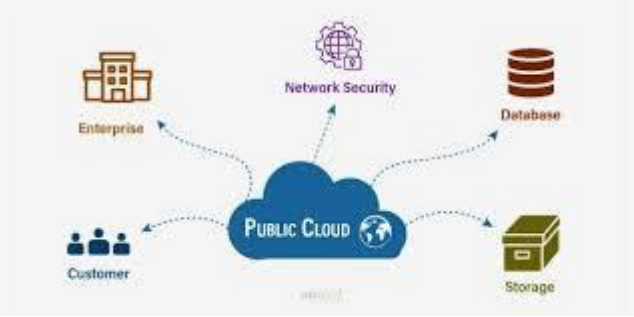

Figure 2b.1 : Public Cloud

2.2.2 Private Cloud

A private cloud is a deployment model that is exclusive to one company and is usually housed on-premises or on a private network. It offers more security, control, and customisation and is frequently managed by the company's own IT division. Private clouds are ideal for processing extremely sensitive or classified data since they provide many of the advantages of public clouds on a smaller scale and in a secure setting. Although it needs the same degree of isolation as a public cloud, a community cloud is a variation that is used by companies with comparable requirements.

Advantages:

Improved security and control: Better data governance and protection are guaranteed by dedicated infrastructure.

Environment that can be altered to satisfy particular legal or corporate needs.

Support for legacy systems: Integration with previous apps and systems is made easier.

Disadvantages:

Costlier: Needs a large outlay of funds for management, upkeep, and hardware.

Limited scalability: In contrast to public clouds, scaling resources is more difficult and takes longer.

Complex administration necessitates constant supervision and qualified staff.

Figure 2b.2 : Private Cloud

2.2.3 Hybrid Cloud

A hybrid cloud deployment strategy allows for the smooth transfer of data and applications between environments by combining public and/or private cloud services with on-premises infrastructure. By fusing the scalability and cost-effectiveness of public clouds with the security and control of private clouds, it offers businesses the best of both worlds. By incorporating cloud-native capabilities into pre-existing systems, hybrid clouds frequently enhance security, scalability, performance, and accessibility.

Sometimes, hybrid cloud is mistaken for multicloud, which uses services from several different cloud providers. Not all multicloud configurations are hybrid, but if many public cloud providers are used, a hybrid cloud can be multicloud. All things considered, when managed effectively, hybrid clouds provide less danger of vendor lock-in, increased control, and the best possible workload distribution.

Advantages:

Flexibility and control: combines the advantages of private and public clouds to create customized solutions.

Cost effectiveness: Infrastructure expenses are decreased by scalable public cloud resources.

Improved security: Private clouds can house sensitive data, while public clouds can be used for less important operations.

Disadvantages:

Complex management: Operational complexity rises while coordinating across several settings.

Latency problems: Performance problems and delays may arise during data transit across clouds.

Infrastructure planning is necessary; it calls for cautious adoption of RBAC, availability, and bandwidth.

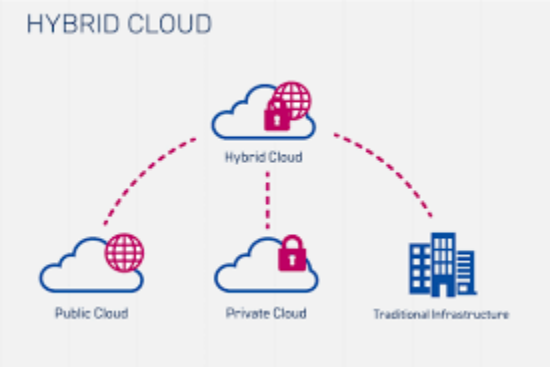

Figure 2b.3 : Hybrid Cloud

2.2.4 Community Cloud

A community cloud is a type of cloud deployment strategy created to assist a particular set of enterprises who have similar objectives or issues. It combines services from several clouds into a dispersed system that meets the demands of a business sector, industry, or community as a whole. The participating organizations share the infrastructure, which is usually administered by a third party or by the member organizations working together.

Advantages:

Cost-effective: Participating firms' individual expenses are decreased by shared infrastructure.

Enhanced cooperation: Promotes data exchange and cooperation between groups with related objectives.

Improved security: Designed to meet community needs, it provides superior protection over public clouds.

Disadvantages:

Limited scalability: The capacity to grow as needed may be constrained by shared resources.

Less customization: Since custom modifications must be in line with the community as a whole, they are more difficult to implement.

Cooperation is necessary for performance and upgrades; thus, everyone must work together.

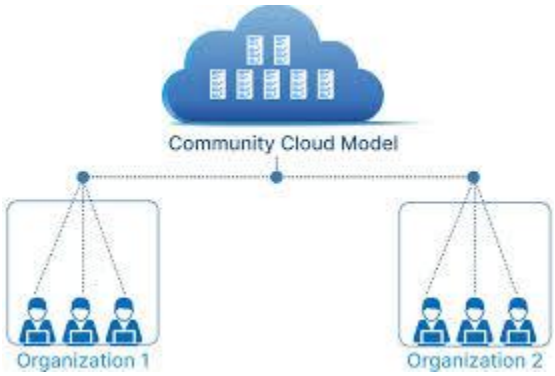

Figure 2b.4 : Community Cloud

2.3. Five essential characteristics of cloud computing

1. On-Demand Self-Service

Cloud computing eliminates the need for human communication with service providers by allowing users to autonomously deploy computing resources like storage or virtual machines. This self-service approach decreases service deployment delays and boosts flexibility. Usually, a web-based dashboard or portal is used by users to control these operations.

2. Broad Network Access

Cloud services can be accessed online using a variety of common devices, including desktops, laptops, tablets, and smartphones. Users can now access services from anywhere at any time, increasing their productivity and mobility. Performance and quality of access are affected by variables including bandwidth and latency.

3. Resource Pooling

Cloud providers employ a multi-tenant model in which several users share and pool real and virtual resources. These resources are used efficiently because they are dynamically allocated and redistributed in response to demand. Although the precise location of the resource is abstracted, users can select more general location parameters such as region or data center.

4. Rapid Elasticity

Cloud services can quickly and frequently scale up or down to accommodate changing demands. This guarantees that users are never over-provisioned and always have the capacity they require. Additionally, it removes the requirement for initial hardware purchases, which lowers the cost of scalability.

5. Measured Service

Providers are able to charge according to actual use since resource usage is tracked, managed, and reported. This metering enable pay-as-you-go pricing and transparency. By learning more about their consumption habits, users will be able to control and maximize their expenses.

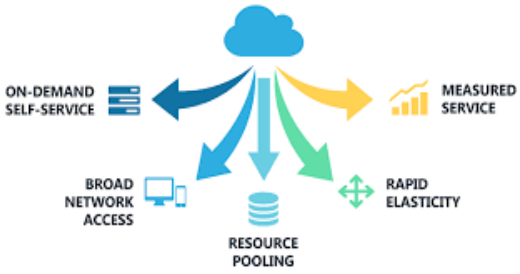

Figure 2c.1 : Essential Characteristic

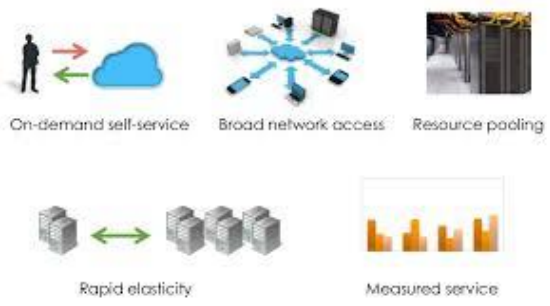

Figure 2c.2 : Details of Essential Characteristics

6. TASK 2: Cloud Service Provider

3.1. Three global cloud service providers

Cloud computing has become an essential component of modern IT infrastructure. And global cloud service providers is playing a critical role in enabling organizations to scale, innovate, and operate efficiently. According to today's data Amazon Web Services (AWS), Microsoft Azure, and Google Cloud Platform (GCP) remain the top three providers in the global market. Each of these companies offers a wide range of cloud services, as they continue to expand their infrastructure across the world to meet growing demand. This section compares their market share, strengths, and regional availability.

1. Amazon Web Services (AWS)
   i. Market Share: ~29% (slightly decreased from 31% in 2024).
   ii. Geographic Availability: Operates in 99 availability zones across 31 global regions, with further expansion planned.
   iii. Strengths: Industry leader with the most mature and comprehensive set of cloud services, including compute (EC2), storage (S3), and advanced ML tools. Known for innovation, global infrastructure, and scalability.
2. Microsoft Azure
   i. Market Share: ~22% (slightly down from 25% in early 2024).
   ii. Geographic Availability: More than 60 regions are covered, which is the most of any supplier.
   iii. Strengths: Azure Stack and Arc provide powerful hybrid cloud capabilities. outstanding compatibility with Microsoft products (Windows Server, Office 365). Due to its security and compliance features, it is highly preferred in the government and business sectors.
3. Google Cloud Platform (GCP)
   i. Market Share: ~12% (consistent growth from 11%);
   ii. Geographic Availability: Active in 38+ cloud regions and 100+ availability zones, with expanding presence in Asia, Europe, and South America;
   iii. Strengths: Supports open-source and multi-cloud tools like Kubernetes and Anthos; specializes in AI/ML and data analytics (BigQuery, Vertex AI); and is well-liked by developers and data-driven businesses.

   Reliable cloud services are provided internationally by AWS, Azure, and GCP, catering to both small and large businesses. They are growing across the Asia-Pacific area, especially in Japan, Singapore, and India. As providers build additional data centers, the Middle East, Africa, and Latin America are becoming important growth regions, especially in industries like government, energy, and education.

Figure 3a.1 : Top Three Global Service Providers

3.2. The concept of service-level agreements (SLA)

A formal, legally enforceable contract created between a cloud service provider (CSP) and a customer is known as a Service-Level Agreement (SLA). It outlines the expected levels of service quality, including metrics for responsibility, performance, and availability. In order to control expectations and guarantee that cloud providers provide dependable and consistent services, SLAs are essential.

Key Components of an SLA:

1. Service Availability
   - Indicates the percentage of uptime that is guaranteed, usually between 99.9% and 99.999%.
   - A 99.9% uptime rate, for instance, corresponds to a maximum of 8.76 hours of downtime annually.
2. Measures of Performance
   - Latency
   - Response time
   - Throughput
   - System accessibility

   For example, Google Cloud guarantees a latency of less than 300 milliseconds worldwide.
3. Compliance and Security
   - explains the security guidelines that are used for infrastructure and consumer data.
   - Guarantees compliance with pertinent legal and regulatory mandates, including GDPR, ISO/IEC standards, and industry-specific rules.
4. Penalties and Remediation

   Describes what happens if service levels are not fulfilled, which could include:
   - Service credits
   - Refunds in part
   - The escalation protocols
5. Ownership and Management of Data

   explains the management of consumer data, including:
   - Data migration or return procedures
   - Guidelines for erasing or destroying data upon service termination

A service-level agreement's (SLA) main goals are to guarantee the cloud provider's accountability, establish explicit performance standards for the client, and offer legal protection in the event that the service levels are not fulfilled.

Figure 3b.1 : Service-Level Agreement (SLA)

3.3. The key factors to consider when selecting a cloud service provider

To guarantee dependable, safe, and economical cloud adoption, enterprises must consider a number of crucial factors when selecting a cloud provider. Organizations may reduce operational risks and increase return on investment in the cloud environment by using an intelligent selection process.

a. **Cost and Pricing Models**

Effective spending management is involved transparent pricing models like pay-as-you-go, reserved instances, and subscription-based plans. Pay-as-you-go allows the user for actual usage payments, while reserved instances offer longer terms discounts for user . Subscription models offer fixed pricing for bundled services, with tools like pricing calculators helping estimate total cost of ownership.

b. **Security and Compliance**

Cloud providers must ensure of the industry standards and data protection requirements such as encryption, access controls, and certifications like ISO 27001, GDPR, and HIPAA, particularly in regulated sectors like healthcare and finance, to ensure the protection of sensitive information.

c. **Availability and Scalability**

Features such as auto-scaling, various availability zones, and a wide regional presence all contribute to the high availability. Because of this, business reliability is maintained and workloads may be dynamically adjusted to demand spikes without causing service interruptions.

d. **Integration and Compatibility**

Organizations must integrate with third-party platforms about their current IT infrastructure in order to maximize cloud utilization while preserving stable, legal, and adaptable operations. Being able to do this, SDKs, APIs, and CI/CD tools—which speed up development processes and take care of things such as cost models, security, availability, and compatibility—must be seamlessly supported.

A strategic choice, choosing a cloud service provider affects performance, compliance, and long-term business agility. Prioritizing important elements like price, security, scalability, and support helps businesses guarantee a cloud solution that will satisfy present and future needs.

4. TASK 3: Cloud Security

4.1. Three Common Cloud Security Threats and Solutions

1. **Data Breaches**

   Software flaws, improperly configured permissions, or inadequate authentication can all lead to unauthorized access to private data stored in the cloud.

   Solutions:

- Use encryption for both in-transit and at-rest data.
- Multi-factor authentication and strict access controls should be used.
- Regularly evaluate security systems and train employees on phishing threats.

2. **Insecure APIs**

   Cloud services may be vulnerable to illegal access, data leaks, or service interruptions due to inadequately secured APIs.

   Solutions :

- Using API authentication and secure coding techniques.
- Test APIs for vulnerabilities on a regular basis.
- Use rate-limiting and keep an eye on API consumption.

3. **Denial-of-Service (DoS/DDoS) Attacks**

   These types of attacks cause enormous traffic to overwhelm cloud services, rendering them unavailable.

   Solutions:

- Make use of DDoS defense tools, such as Cloudflare and AWS Shield.
- Turn on anomaly detection and auto-scaling.
- Implement load balancing and traffic filtering.

4.2. Importance of Data Classification in Cloud Risk Management

Data classification is the process of grouping information according to its value and level of sensitivity (public, confidential, restricted, etc.). It helps businesses to:

- Implement appropriate security measures (such as encryption and access control).
- Assure adherence to laws such as HIPAA and GDPR.
- Make safeguarding important data a top priority and enhance incident response.
- Allocating resources and expenditure on security should be optimized.

4.3. How Multi-Tenancy Can Impact Cloud Security

Multiple users can share the same cloud infrastructure while maintaining logical isolation thanks to multi-tenancy. Risks associated with this model include:

- Data Leakage: Inadequate separation could make data from one tenant available to another.
- Co-Tenant Attacks: Malevolent tenants may take advantage of communal spaces.
- Resource Interference: If one tenant uses a lot of resources, it may affect other tenants' service.

Mitigation Measures:

The following are mitigation strategies:

- Implement stringent physical and logical isolation (e.g., VLANs, hypervisor security).
- Perform routine threat monitoring and security audits.
- Put access controls and regulations tailored to each tenant in place.

Conclusion

Cloud computing offers scalable, cost-efficient, and flexible service models like Infrastructure as a Service (laaS), Platform as a Service (PaaS), and Sofware as a Service (SaaS) to meet operational needs. Its benefits include resource pooling, measurable service, on-demand self-service, wide network connectivity, and quick adaptability. Prominent international suppliers including Google Cloud, Microsoft Azure, and AWS offer extensive services. Priority one should be given to cloud security since it promotes innovation, company agility, and safe digital transformation.

Reference list

Ali, E. (2024). *A Service-Level Agreement (SLA) is a pivotal document in the realm of cloud computing, acting as a formalized contract between a cloud service provider (CSP) and a cloud service customer (CSC). This agreement delineates the expected quality and performance of cloud services, framed within a taxonomy*. [online] Linkedin.com. Available at: https://www.linkedin.com/pulse/understanding-service-level-agreements-slas-cloud-computing-ali-ah7bf/ [Accessed 4 May 2025].

AWS (2023). *AWS Service Legal Agreements*. [online] Amazon Web Services, Inc. Available at: https://aws.amazon.com/legal/service-level-agreements/?aws-sla-cards.sort-by=item.additionalFields.serviceNameLower&aws-sla-cards.sort-order=asc&awsf.tech-category-filter= [Accessed 4 May 2025].

Bhatia, V. (2022). *Essential Cloud Computing Characteristics | Synopsys Blog*. [online] www.synopsys.com. Available at: https://www.synopsys.com/blogs/chip-design/essential-cloud-computing-characteristics.html [Accessed 4 May 2025].

Brown, E.A. (2025). *Final Version of NIST Cloud Computing Definition Published*. [online] NIST. Available at: https://www.nist.gov/news-events/news/2011/10/final-version-nist-cloud-computing-definition-published [Accessed 4 May 2025].

Darktrace.com. (2024). *Darktrace*. [online] Available at: https://www.darktrace.com/cyber-ai-glossary/the-most-common-cloud-security-threats [Accessed 5 May 2025].

GeeksforGeeks. (2023). *Difference Between Public Cloud and Private Cloud*. [online] Available at: https://www.geeksforgeeks.org/difference-between-public-cloud-and-private-cloud/ [Accessed 4 May 2025].

Google Cloud. (n.d.). *Google Cloud Platform Service Level Agreements*. [online] Available at: https://cloud.google.com/terms/sla [Accessed 4 May 2025].

HulHub (2024). *Importance of Data Classification in Cloud Security*. [online] HulHub. Available at: https://www.hulhub.com/data-classification-in-cyber-security [Accessed 5 May 2025].

ISO (2016). *ISO/IEC 19086-1:2016*. [online] ISO. Available at: https://www.iso.org/standard/67545.html [Accessed 4 May 2025].

Microsoft.com. (2025). *Licensing Documents*. [online] Available at: https://www.microsoft.com/licensing/docs/view/Service-Level-Agreements-SLA-for-Online-Services?lang=1 [Accessed 4 May 2025].

Ncsc.gov.uk. (2025). *Service and deployment models*. [online] Available at: https://www.ncsc.gov.uk/collection/cloud/understanding-cloud-services/service-and-deployment-models#hybridcloud [Accessed 4 May 2025].

Tetiana Fydorenchyk (2016). *AWS, Azure, Google Cloud and Jelastic: Choose Your Cloud Hosting by Location*. [online] Virtuozzo Blog | News & insights for cloud service providers. Available at: https://www.virtuozzo.com/company/blog/aws-azure-google-cloud-and-jelastic-choose-your-cloud-hosting-by-location/ [Accessed 4 May 2025].

UMATechnology (2025). *15 TOP Cloud Service Providers (2025) - UMA Technology*. [online] UMA Technology. Available at: https://umatechnology.org/15-top-cloud-service-providers-2025/ [Accessed 4 May 2025].