Glossary

Term	Definition
Abstraction layers	Abstraction layers are mechanisms that hide the complexity of underlying systems, providing a simplified interface for user and developer interaction. They allow programmers to implement functionality without dealing with the intricate details of lower-level operations.
Accelerated compute	Accelerated compute refers to powerful accelerator chips, often used in conjunction with high bandwidth networking.
Accelerator chips	Accelerator chips refers to high performance computer chips – such as graphics processing Units (GPUs),) and application -specific integrated circuits (ASICs) – used to accelerate a range of computation tasks.
Adaptors	Small, focussed pieces of software that facilitate communication between two or more components that cannot directly interoperate with each other.
AHB	The Azure Hybrid Benefit allows customers with existing on-premises Windows Server or SQL Server core licences with Software Assurance subscriptions to migrate these licences onto Azure at a discount.
Amazon Elastic Container Service (ECS)	Amazon Elastic Container Service is a fully managed container orchestration service that helps you to more efficiently deploy, manage, and scale containerized applications.
Amazon Elastic Kubernetes Service (EKS)	Amazon EKS is a managed Kubernetes service run in the AWS cloud and on-premises data centres.
Ancillary services and tools	This includes both cloud services and related software tools (such as directory software) that support the management of

	the cloud, including IAM, observability, billing, provisioning and orchestration.
Application Programming Interface (API)	A software interface that allows two or more pieces of software to communicate with each other.
Application-Specific Integrated Circuits	Accelerator chips which are hardwired for a specific application. ASICs developed for FM training and inference are typically hyper-specialised GPUs, with modifications that increase the efficiency of specific AI workloads.
Bring Your Own Licence (BYOL)	A licensing model that allows customers to transfer their existing software licences to a new environment.
Cloud credits	Cloud credits are a form of discounting by cloud providers. Cloud providers offer customers credits (rather than cash) as an incentive to switch clouds or use their cloud more. Customers can redeem cloud credits against spend on cloud services.
Cloud services	Services that provide access to processing, storage, networking, and other raw computing resources (often referred to as Infrastructure as a Service, IaaS) as well as services that can be used to develop, test, run and manage applications in the cloud (often referred to as Platform as a Service, PaaS.)
Cloud-agnostic services and tools	These aim to provide a consistent set of capabilities and interfaces across different clouds.
Cloud-native workloads	Cloud native workloads refer to workloads created on the cloud, not migrated from on-premises.
Committed Spend Agreement/ discount (CSA/CSD)	This is a negotiated agreement between a Cloud Provider and a customer in which the customer commits to spend a minimum amount across the Cloud Provider's services over a period of years, and in return, receives an individually negotiated percentage discount.

Containerised	A containerised application is a software application that has been packaged in a container.
Container	A software tool that enables the packaging and isolation of applications with their entire runtime environment – all the files necessary to run.
Content delivery network (CDN)	A geographically distributed network of servers aiming at fast delivery of internet content, including HTML pages, scripts, stylesheets, images, audio files and videos. Serving content from a CDN, located closer to the end-user, may improve load times, reduce bandwidth costs and increase availability.
Core services	Core services are the main IaaS and PaaS services that contribute to the key objectives of the customers workload.
CPU	Central Processing Unit.
CSP	Cloud solution provider
DaaS	VDI on the cloud can also be referred to as Desktop-as-a-Service.
Data portability	Ability to easily move data from one cloud to another or from a customer's IT environment to a cloud and have that data usable in the target cloud with minimal disruption
Data transfer fees	These comprise internal transfer fees and egress fees
Database as a service (DBaaS)	A managed cloud service that provides customers with access to a database.
Digital native business	Business that exists primarily or entirely online and uses cloud technologies throughout its operation.
Duplicated multi-cloud (cloud duplication)	This multi-cloud architecture occurs whenever customers aim to mirror their cloud architecture on two or more public clouds,

	so that all or some of their applications and data can run equivalently on all of them.
EC2	Amazon Elastic Compute Cloud.
Egress	The act of data being transferred from the cloud provider's cloud network to an external location.
Egress fees	Fees charged by cloud providers to customers for data transferred from the provider's cloud network to an external location.
Enterprise agreements (EA)	Enterprise agreements, which are individually negotiated, are generally reserved for larger customers with higher spending.
Enterprise workloads	Workloads which support 'line of business' applications including enterprise planning, business intelligence, resource management (ERP) or Customer Relationship Management (CRM).
Foundation Model (FM)	Foundation Models (including, among others, language models or LLMs) are a type of AI technology that is trained on vast amounts of data that can be adapted to a wide range of tasks and operations.
Graphical Processing Unit (GPU)	An accelerator chip which has many small, specialised processing cores that run in parallel to perform computations simultaneously.
Hybrid cloud	A cloud deployment model in which public and private clouds are combined.
Identity and Access Management (IAM)	Security frameworks that authenticate and authorise individuals to access specific resources within an organisation's network. They use policies, protocols, and technologies to manage digital identities, granting appropriate access levels to systems, applications, and data.
Independent software vendors (ISVs)	Suppliers of PaaS and/or SaaS, that do not own any of the underlying raw computing resources.

Infrastructure as a code (IaC)	IaC is the use of high-level descriptive coding language to automate and standardise the provisioning and deployment of IT infrastructure such as networks, virtual machines, load balancers, and connection topologies required by any application.
Infrastructure as a service (IaaS)	Cloud Services that provide access to raw computing resources (compute, storage, and network) for processing Workloads and storing data. These computing resources are in the form of servers and networking equipment owned and managed by the IaaS provider (and typically held on racks in a remote data centre). To allow and manage that access, IaaS also includes some necessary software, including networking and Virtualisation. The customer has the highest level of control over the cloud stack, including over the operating system, applications, and data. IaaS should be distinguished from Bare Metal services.
Ingress	The act of data entering the provider's cloud network from an external location.
Integrated development environment (IDE)	IDEs are a type of software containing a range of tools that software engineers use to write code, build applications, web pages or services.
Integrated multi-cloud	A cloud architecture where customers can mix and match cloud services from different public cloud providers and there is a degree of integration between these services (eg, data is stored on one public cloud but analysed on a different one).
Internal (data) transfers	The transfer of data by a cloud customer within a cloud provider's infrastructure.
Internal transfer fees	The cost incurred by a customer when transferring data within and out of a cloud provider's infrastructure.
Interoperability	The ability of computer systems or software to communicate with one another.

Kubernetes	Open-source platform designed for automating the deployment, scaling and management of containerised applications.
Latency	The time it takes for data to travel between any two points on a network.
Managed service providers (MSPs)	Companies that provide IT services, such as app development and app management, e.g. Capgemini, Deloitte. MSPs may be considered ISVs that only offer services on request as opposed to launching public offerings.
Marketplace	A website operated by a cloud provider where customers can purchase services (supplied by the marketplace owner and third parties) that are compatible with their clouds.
Multiple cloud/multi-cloud	The use by a customer of two or more public clouds.
Open-source software	Software released under a license in which the copyright holder grants the users the right to freely use, change, and distribute the software and its source code.
Origin cloud	The cloud from which a customer is switching when switching between clouds.
OS	Operating system
Platform as a Service (PaaS)	Cloud Services that provide access to a virtual environment for customers to develop, test, deploy and run applications. These include application development computing platforms and pre-built application components and tools which customers can then use to build and manage full applications. The customer has less control over the cloud stack compared to IaaS – they still manage applications and data but not the PaaS computing platform (including its operating system) and the pre-built application components and tools.
Private cloud/private cloud services	A cloud deployment model in which computing resources (like the hardware) are used exclusively by one customer.

Relational Database Management System (RDBMS)	A RDBMS is a type of Database Management System which manages and stores data in separate tables and defines relationships between those tables.
Server	A computer being used to provide a service to other computers, as opposed to human users. Servers are usually, but not always, built from specialised hardware (eg redundant power supplies) and used with a server-specific operating system (eg Linux).
Server OS	Server Operating System. Software is designed to run a server's hardware and provide a platform for the use of application software.
Serverless	Serverless computing enables developers to build applications faster by eliminating the need for them to manage infrastructure. With serverless applications, the cloud provider automatically provisions, scales, and manages the infrastructure required to run the code.
Siloed multi-cloud	This multi-cloud architecture occurs where the customer runs different customer applications, stores different customer data sets and/or uses different cloud services hosted on two or more public clouds with no or minimal integration between these clouds (i.e., different applications are 'siloed' on different public clouds).
Software as a Service (SaaS)	Complete applications hosted in the cloud. These cloud applications can be offered by the Cloud Provider that owns the underlying raw compute resources or by an Independent Software Vendor. The service provider(s) manages all hardware and software.
Stack	A set of hardware and software components that work together to create a computing platform for running applications.
Standard compute infrastructure	Compute infrastructure that relies on general-purpose CPU processors.

Switching	The act of a customer changing provider of cloud infrastructure services or moving Workloads from one provider to another.
System for Cross-domain Identity management (SCIM)	A standard protocol that enables the management of user accounts, groups, and some level of 'access' across different digital (and cloud) services.
TFlops	Teraflops are used to measure computational performance. One TFLOPs is equal to one trillion floating point operations per second.
Third-party cloud service	A cloud service provided by a party other than a cloud provider, such as an ISV or other supplier of Cloud Services.
Traditional / on-premises IT	Dedicated physical computing resources that are not part of the cloud. These are typically owned by, and located on the premises of, the customer.
Vcore hours	Virtual core hours are hours of usage normalised for the number of core processing units being used to run a particular instance or operating system environment.
Virtual desktop infrastructure (VDI)	A VDI allows a customer to run applications on remote virtual machines, which are accessed over the internet by end users 'thin client' software.
Virtual machines (VMs)	A software-defined computer that is created by running a guest operating system on top of the host operating system of the physical server. Each virtual machine runs its own operating system and behaves like an independent server, even though it is running on just a portion of the actual underlying server hardware. The software that creates, runs and manages virtual machines is called a hypervisor.
Virtualisation	The process of using software to create an abstraction layer over servers that allows the hardware elements of a single server to be divided into multiple virtual servers, commonly called virtual machines. Each virtual machine runs its own operating system and behaves like an independent server, even though it is running on just a portion of the actual

	underlying server hardware. The software that creates, runs and manages virtual machines is called a hypervisor.
Visual Studio	Microsoft Visual Studio is a type of Integrated Development Environment.
Workload	A specific application, service, capability or a specific amount of work that can be run on an IT resource.