The following definition is the result of direct and indirect feedback from more than fifty technology experts, practitioners, consultants, analysts and business leaders. The methodology used to develop this definition was to put forth a premise about a new architectural framework called supercloud and iterate the definition with community input over a nine month period. [...]
The following definition is the result of direct and indirect feedback from more than fifty technology experts, practitioners, consultants, analysts and business leaders. The methodology used to develop this definition was to put forth a premise about a new architectural framework called supercloud and iterate the definition with community input over a nine month period. The primary input for this definition was in-depth interviews with experts. These interviews were transcribed and reviewed to identify common patterns and disagreements to expand the depth and defensibility of the definition. Secondary research was also conducted by reviewing enhancements, criticism and commentary on the definition primarily in the form of independent blogs, videos and social media posts.

The three main objectives of this definition are to: 1) Provide the basis for discussion about a new architectural framework for cloud computing; 2) Iterate on the first and second versions of the definition; and 3) Include enough specificity to determine what is and is not included in the category.

1.0 Introduction

Supercloud is a term introduced to describe an evolving architecture in computing. This definition identifies specific essential characteristics and is intended to provide a foundation for understanding the future of cloud computing. The service and deployment models comprise a set of guidelines intended to include platforms that conform to the essential supercloud properties, with a minimum number of constraints in business and delivery models.

Supercloud is a natural evolution of today’s multi-cloud and hybrid computing models. Most enterprises deploy applications and data on multiple cloud platforms for a variety of beneficial reasons including feature sets, risk avoidance, negotiation leverage and other factors. However, the deployment of multiple cloud platforms creates friction, data silos and requires a diversity of skills to manage tools that are unique to each cloud. This increases costs and makes governance and security more challenging.

The business value of supercloud is it has the potential to deliver significant simplification and reduced cost to deploy or move applications between clouds. Supercloud provides secure and facile access to data spread over multiple clouds and a common experience for developers and operators. The business value of supercloud will increase proportional to the number of specialized clouds and edge compute locations enterprises deploy.

1.1 Audience.

The intended audience for this definition includes business technology practitioners, architects, developers, and other consumers of cloud computing services.

2.0 Definition

Supercloud is an emerging computing architecture that comprises a set of services abstracted from the underlying primitives of hyperscale clouds (e.g., compute, storage, networking, security, and other native resources) to create a global system spanning more than one cloud.

Supercloud has three essential properties, three deployment models, and three service models.

2.1 Essential Properties

**Runs a set of services across more than one cloud.** A supercloud consumer can access service elements that run on more than one cloud without requiring specific expertise and knowledge of the underlying cloud infrastructure.

**Purpose-built SuperPaaS Layer.** A *capability that* abstracts the underlying primitives of the native PaaS layer within each cloud and creates a common experience across clouds for developers, operators, users and/or ecosystem partners. In addition, the superPaaS layer acts as a cloud interpreter tailored explicitly for the supercloud’s objectives.

**Metadata Intelligence.** A metadata capability optimized for specific supercloud services runs workloads
efficiently across federated cloud platforms. For example, the supercloud has an awareness of cost, latency, bandwidth, governance, security, data sovereignty, or other attributes in each supported cloud platform and explicitly serves the intended purpose of the supercloud.

2.2 Deployment Models

**Single Cloud Instantiation.** A control plane runs its service on one cloud but supports data plane interactions with more than one other cloud. An example is a Kubernetes cluster management service that runs on one cloud but can deploy and manage clusters on other clouds.

**Multi-cloud, Multi-region Instantiation.** A full stack of services is instantiated on individual clouds and regions. A unified interface supports interactions across more than one cloud. An example is a set of data protection services (e.g., backup, restore, archive, data analytics, data management) installed in multiple clouds and cloud regions and controlled through a unified platform interface. Cohesity is an example of this deployment model.

**Global Instantiation.** A single global instantiation of services spans multiple cloud provider regions. An example is a data platform that enables governed and secure data sharing across clouds and regions. Snowflake and Oracle Database Service for Microsoft Azure are examples.

2.3 Service Models

**Infrastructure as a Service (IaaS).** The ability to provision a service including compute, storage, networking, security or other computing resources across multiple clouds and on which workloads can be provisioned without knowledge of the underlying cloud infrastructure. An example is a data storage service with a single interface that spans multiple clouds and cloud regions. NetApp's Cloud Volumes service is such an example.

**Platform as a Service (PaaS).** DevOps professionals are able to create and deploy applications using programming languages, libraries, services, and tools supported by the cloud provider. The developer and operational experience is identical across clouds with no need to manage the underlying compute, storage, network and security controls of the cloud provider. VMware Cloud Foundation is an example.

**Software as a Service (SaaS).** Users access applications from a Web browser or mobile application that invokes services running in more than one cloud. The user has no knowledge or control over the underlying cloud infrastructure. An example is an ERP system that runs on a private cloud infrastructure and invokes analytics and machine learning services from a public cloud provider in a seamless interaction with no user context switching. SAP HANA Cloud is such an example.

3.0 Frequently Asked Questions (FAQ)

3.1 Why does the industry need another term?

There is broad agreement that clouds in the 2020s are different from clouds of the 2010s. That something new is happening within the AWS (and other) cloud ecosystems, beyond traditional IaaS and PaaS and isn’t just SaaS running in the cloud.

Supercloud is an attempt to describe a new architecture that integrates infrastructure, unique platform attributes and software to solve specific problems that public cloud vendors aren’t directly addressing.

Supercloud is an evocative term that catalyzes debate, conversation and thought.

3.2 What Problems do superclouds solve?

It’s generally accepted that customers are using multiple clouds. For the most part, these clouds operate in disconnected silos where operating in each cloud requires different skills. The development environment is different as is the operating environment. With different APIs and primitives and management tools that are optimized for each respective hyperscale cloud. Their functions and value props don’t extend to their
As a result, there’s friction when moving between different clouds. It’s hard to share data, move work, secure and govern data, enforce organizational policies, and edicts across clouds.

Supercloud is an architecture designed to create a single environment that enables management of workloads and data across clouds in an effort to remove complexity, accelerate application development, streamline operations and share data safely irrespective of location.

3.3 Why aren’t hyperscale clouds considered superclouds?

Hyperscale clouds are walled gardens and generally cloud providers want to keep data in their clouds. The more homogeneous hyperscalers can make their clouds, the better the experience for their customers in terms of performance, security and cost.

However, not all data will end up in public clouds and solutions such as Amazon Outposts, Azure Arc and Google Anthos are early indications that hyperscalers will extend their clouds. It’s likely they will pursue supercloud opportunities in earnest but they have plenty of work to do within their own clouds. And today at least they appear to be primarily focused on providing the tools that will enable others to build superclouds on top of their platforms.

3.4 Isn’t supercloud just another term for multi-cloud?

Some might call this trend multi-cloud 2.0. But many believe that multi-cloud has been a natural outcome of shadow cloud, multiple development teams choosing different clouds, mergers and acquisitions. Historically multi-cloud has created a unique experience within each cloud with little or no connection between the cloud silos. And a distinct developer and operational experience within each cloud.

Supercloud sets out to build incremental value across clouds and on top of hyperscale CAPEX that goes beyond cloud compatibility within each cloud. Supercloud is a term that connotes something deliberately different from multi-cloud by default.

3.5 Isn’t plain Old PaaS already supercloud?

Supercloud and its corresponding superPaaS layer gives the freedom to store, process, manage, secure and connect islands of data along a continuum with a common developer and operational experience across clouds. This is different from traditional PaaS offerings from cloud providers.

Importantly, the sets of SuperPaaS services are designed to support a specific supercloud’s objectives – e.g. data sharing or data protection or storage and retrieval or cost optimization or ultra low latency, etc. In other words, the services offered are specific to that supercloud and will vary by each supercloud offering. Supercloud and its inherent superPaaS and corresponding cloud interpreter will be optimized to solve specific problems not previously addressed.

3.6 What is the difference between supercloud and SaaS?

SaaS is a subset of supercloud. Most SaaS platforms either run in their own cloud or run in a specific public cloud. But the cloud services are normally independant and the SaaS applications and data are siloed.

However, SaaS vendors are evolving and adopting supercloud to offer distributed solutions across cloud platforms and stretching out to the near and far edge.

3.7 What are some examples of superclouds?

**Analytics.** Snowflake is an example with its data cloud in our view. It’s a supercloud optimized for data sharing, governance, query performance, security, ecosystem enablement and ultimately monetization. Snowflake is now bringing in new data types and conforms with the definition above.

**Converged Databases.** Running transaction and analytic workloads on a single database. An example is
Couchbase with Capella, which enables stretching the cloud to the edge with Arm-based platforms and optimizing for low latency across clouds and out to the edge via a synchronization capability.

**Document Database Workloads.** MongoDB is a developer-friendly platform that, with Atlas, is moving to a supercloud model running document databases very efficiently. Accommodating analytic workloads and creating a common developer experience across clouds.

**Data Science Workloads.** Databricks is bringing a common experience for data scientists and data engineers driving machine intelligence into applications and fixing the broken data lake with the emergence of the lakehouse and Delta Sharing.

**General Purpose Workloads.** VMware is attempting to create a common operating environment across clouds and on-prem and out to the edge. VMware’s focus is on managing and moving workloads, balancing workloads and being able to recover rapidly across clouds.

**Network Routing.** Aviatrix is building what appears to be a supercloud to optimize network performance and automating network security across clouds.

### 4.0 Sources

The following individuals have contributed directly or indirectly to this definition. Much of the feedback that was received and considered came in the form of criticisms. The inclusion of the individuals in this list of contributors is not meant to imply each has endorsed the term and concepts around supercloud. Some have rejected it. Nonetheless, their contributions are noted because they provided useful input, catalyzed thought and furthered the research.

Notably, the supercloud initiative is meant to be open and inclusive. The work here is not complete and more input and constructive criticism is welcome.

David Vellante is co-CEO of SiliconANGLE Media, as well as co-founder and Chief Analyst at The Wikibon Project, the world’s leading open source technology research community. Dave is a long-time tech industry analyst, entrepreneur, writer and speaker. As co-host of theCUBE - “The ESPN of Tech,” Vellante has interviewed over 5,000 experts since 2010. He is also a co-founder of CrowdChat, an angel funded startup based in Palo Alto using big data techniques to extract business value from social data. Prior to these exploits, Dave founded a CIO consultancy and spent a decade growing and managing IDC’s largest business unit. He lives in Massachusetts with his wife and four children where he is active in town activities including serving as the president of his town’s local “Kiddie Sports” association. Dave holds a B.S. in Applied Mathematics from Union College.

David Vellante
@dvellante
david.vellante@wikibon.org