CLOUD COMPUTING: Evolution to Hyperscale Computing

Avery P. & Antti Vahvaselka

Brief history on Cloud Computing

- Beginning of cloud history starts in 1960
  - John MacCharty introduces cloud concept.
  - Two projects headline: ARPANET and MAC
- Shift to virtualization in 1970s
  - “Mother of all Demos”
  - VPN, idea of Internet
  - IBM popularizes VMs
- Gaining popularity in late 1990s
  - Defined as “spaces between users and provider”
  - SalesForces utilizes cloud like capabilities
  - Use of VMs comes in as well

Brief History of Cloud Computing (cont)

- 2000s a lot is happening:
  - Amazon introduces web-based retail services (2002)
  - AWS is released (2006)
  - Google launches “Google Docs” (2006)
  - Google started providing cloud computing services (2009)
    - At this point all big players have joined Microsoft, IBM, HP etc..
- 2010s and beyond
  - Private clouds were being released as of 2010 by big players.
  - Concept of Hybrid Clouds was introduced (2011)
  - Icloud was launched, Microsoft began to advertise cloud capabilities. (2011)
  - CloudBolt founded in 2012 created a hybrid cloud environment
  - By 2014 Cloud Computing has really developed its foundation

Cloud Basics!

- ELI5: Explain it like I am 5
  - Use of Virtual Servers over physical servers.
  - Utilizing other infrastructure rather than your own.
- Pros
  - Scalability
  - Reducing infrastructure cost
  - Defend against Disaster

Cons

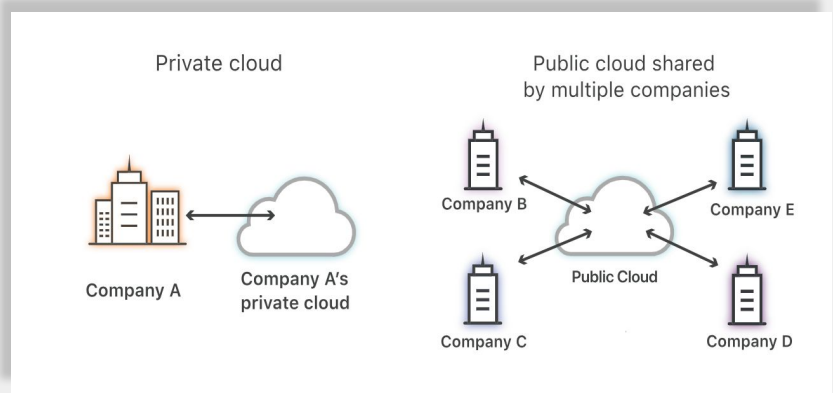
-Release of control

-Security

-Internet reliance

Cloud Types

- **Public Cloud**:
  - Offered to multiple customers by a cloud provider
  - Includes the following services:
    - SaaS, IaaS, PaaS
- **Private Cloud**:
  - Exact opposite of public cloud
    - Not shared with any other customers
    - All resources are controlled.



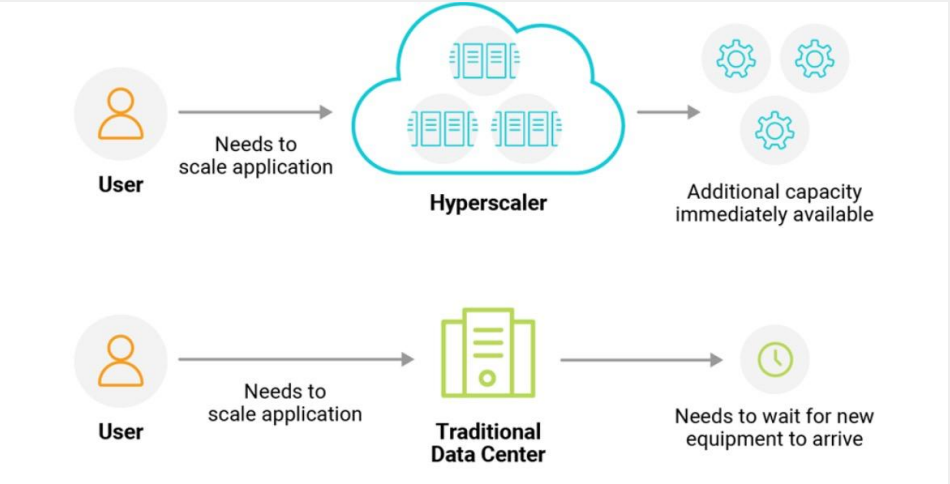
Hyperscale Computing

Now for what you all came for!

- ELI5: Ability of an infrastructure to to scale rapidly to handle a huge amount of workload.
- We can now refer big players as “Hyperscalers”
    - All big players are investing in this structure.
- Essentially as business demands grow so does the need.
    - Conveniency: no investment into infrastructure

Hyperscale Computing

- Provide all in one Services
  - IaaS, SaaS, PaaS
- Designed to handle large amounts of data with some key characteristics:
  - Distributed Architecture
  - High Energy Efficiency
  - Large Scalability
  - Running Globally
- Datacenter Infrastructure is next level!
  - EDCs vs HDCs



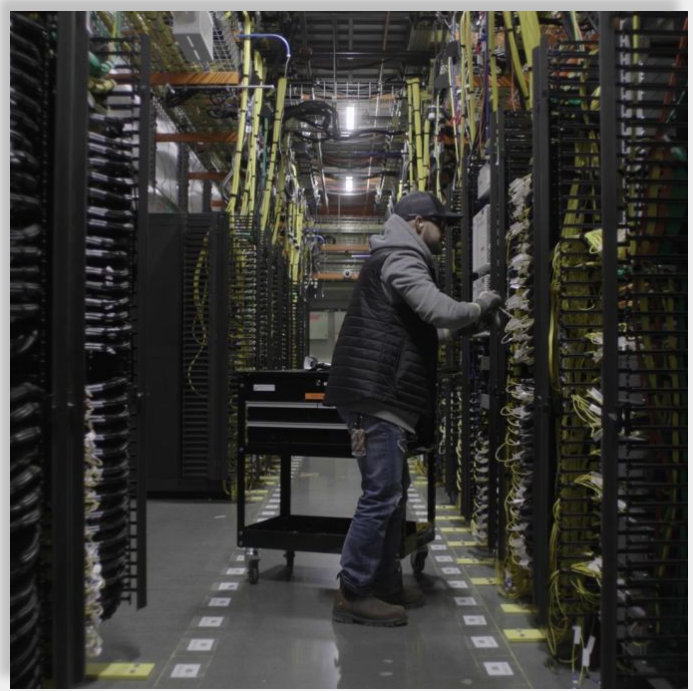
What is a Hyperscale Data Center?

- **Super basic definition:** They house the computers used to power the servers companies will use to run their web applications.
- **What makes for a Hyperscale Data Center?**
- Each data center will have to negotiate for a contract of up to 200 megawatts of power per year in order to keep the machinery running.
- Nearly half of the electricity usage will go towards powering the cooling systems.
- **Cooling can be achieved by many means...** (Project Natick)

Fun figures: Average cost of AWS main data center: 2.37B$

Electricity costs up to 2B$ / year

Costs vary greatly depending on location!

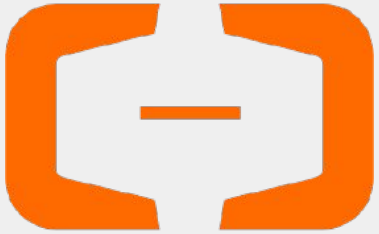


Physical Locations

Let’s take a look at the AWS ecosystem of Data Centers!

- **Regions** are geographical locations distant from each other. (US-east, US-west)
- **Availability Zones** are clusters of one or more data centers. (Oregon)
- **Local Zones** extend the region closer to end users to reduce latency.
- **Edge Caches and Edge Location** are a part of the “Cloudfront network”, content even closer to the end user.

Biggest Players



- AWS
- Google Cloud Platform (GCP)
- Microsoft Azure
- Alibaba Cloud (Aliyun)

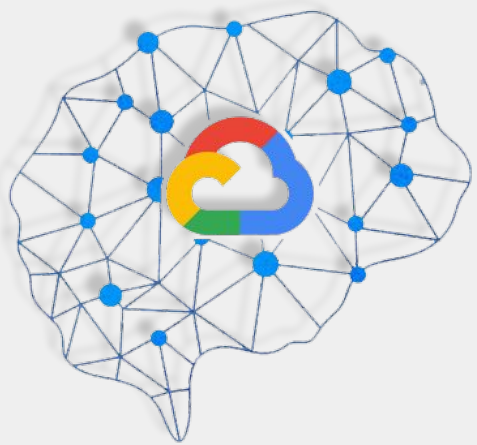
Google Cloud

Amazon Web Services

- Spearheaded by Andy Jassy (2021) and his team.
- AWS infrastructure is deeply integrated with the Amazon.com, Amazon Prime and Alexa products and services. -> Redundancy
- Has the most comprehensive portfolio of services out of all cloud providers.
- Around 100+ “Availability zones” and around 600+ “edge locations”.

Competitive Advantage: Entered the market very early (2006) providing users and companies with (Paas) and (IaaS) as opposed to just software

Google Cloud Platform



- Started offering services in 2011, 5 years behind AWS.
- Like AWS, Google integrates with all Google products and services including: Google search, Maps and Analytics
- Emphasizes data security and protection
- Has a reputation for innovation in machine learning and data analytics
- Has caused controversy due to ethical concerns regarding racial bias in face recognition algorithms

Competitive Advantage: Expertise in Kubernetes (Orchestrator of Applications)

Alibaba Cloud

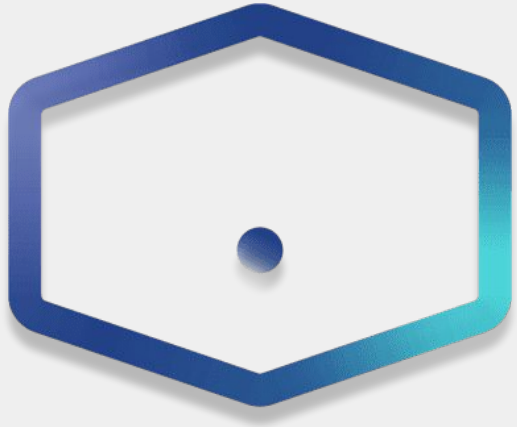
- Subsidiary of the Alibaba Group, established in 2009 to support the Alibaba Groups ecosystem
- June of 2023 new CEO appointed, Eddie Yongming Wu
- By far the most popular service provider among Chinese businesses and enterprises
- The entire Alibaba Group along with its subsidiaries gets **major funding** from the Chinese government, thus AC faces **very little competition**
- Real Life Application next! → Smart Cities

ET City Brain project by Alibaba Cloud

- Provides the city with IoT devices (sensors and actuators) and the **tools** to process the data streams coming in.
- Intelligent system that provides big data computing power and networks to store and process massive amounts of data.
- Traffic congestion and resource management

WHY?

- Without strategy in implementing technology, traffic control cameras and such are pointless, ETCB has implemented algorithms that detect problems in complex situations.



How does ET City Brain help?

Here are some of the ways Alibaba Cloud provides services to city governments with ET City Brain:

- **Smart light solutions**, providing intelligent street lights that adapt to weather conditions and reduce light pollution.
- **Traffic management solutions**, big data and AI-based used to help optimizing traffic signals and street cameras.
- **Structural integrity monitoring**, sensors designed to detect vibrations and monitor the conditions of buildings and/or bridges.

How does ET City Brain help?

Here are some of the ways Alibaba Cloud provides services to city governments with ET City Brain:

- **Accident and disaster management:** analyzing video footage of accidents and traffic congestion in real time will speed up clean up, as well as automatically alert authorities.
- **Smart healthcare:** streamlining healthcare operations by reducing manual errors. Alibaba Cloud will provide algorithms to analyze medical institution processes.
- **Mass transit management:** Alibaba Cloud will draw data from navigation applications, WiFi probes and phone providers in order to direct traffic during large social gatherings.

Conclusion

- Demand is there!
  - Infrastructure costs.
  - Smaller markets are seeing increase in Hyperscale activity.
- Market is being dominated.
  - New opportunities
  - Reshaping industry
- Growth of New Technologies = Growth of Hyperscalers

References:

- https://www.alibabacloud.com/blog/how-et-city-brain-is-transforming-the-way-we-live-one-city-at-a-time_593745
- https://www.venturousgroup.com/resources/city-brain-a-new-model-of-urban-governance-catalyzed-by-big-data/
- https://aws.amazon.com/what-is-aws/
- https://www.alibabacloud.com/about
- https://www.pluralsight.com/resources/blog/cloud/what-is-google-cloud-platform-gcp
- https://www.uopeople.edu/blog/pros-and-cons-of-cloud-computing/
- https://www.cloudflare.com/learning/cloud/what-is-a-public-cloud/#:~:text=The%20private%20cloud%20user%20has,hidden%20from%20other%20cloud%20customers.
- https://natick.research.microsoft.com/
- https://dataspan.com/blog/data-center-cooling-costs/
- https://www.youtube.com/watch?v=SVLPjGiZivc
- https://www.dynatrace.com/knowledge-base/hyperscale-computing/
- https://www.dataversity.net/brief-history-cloud-computing/
- https://datacentremagazine.com/articles/the-future-of-hyperscale-data-centres-exerting-influence
- https://aws.amazon.com/about-aws/global-infrastructure/regions_az/?p=ngi&loc=2