Cloud Computing

Cloudlytics

Definition

- What is cloud ?
  - Defining it through industry standard terms

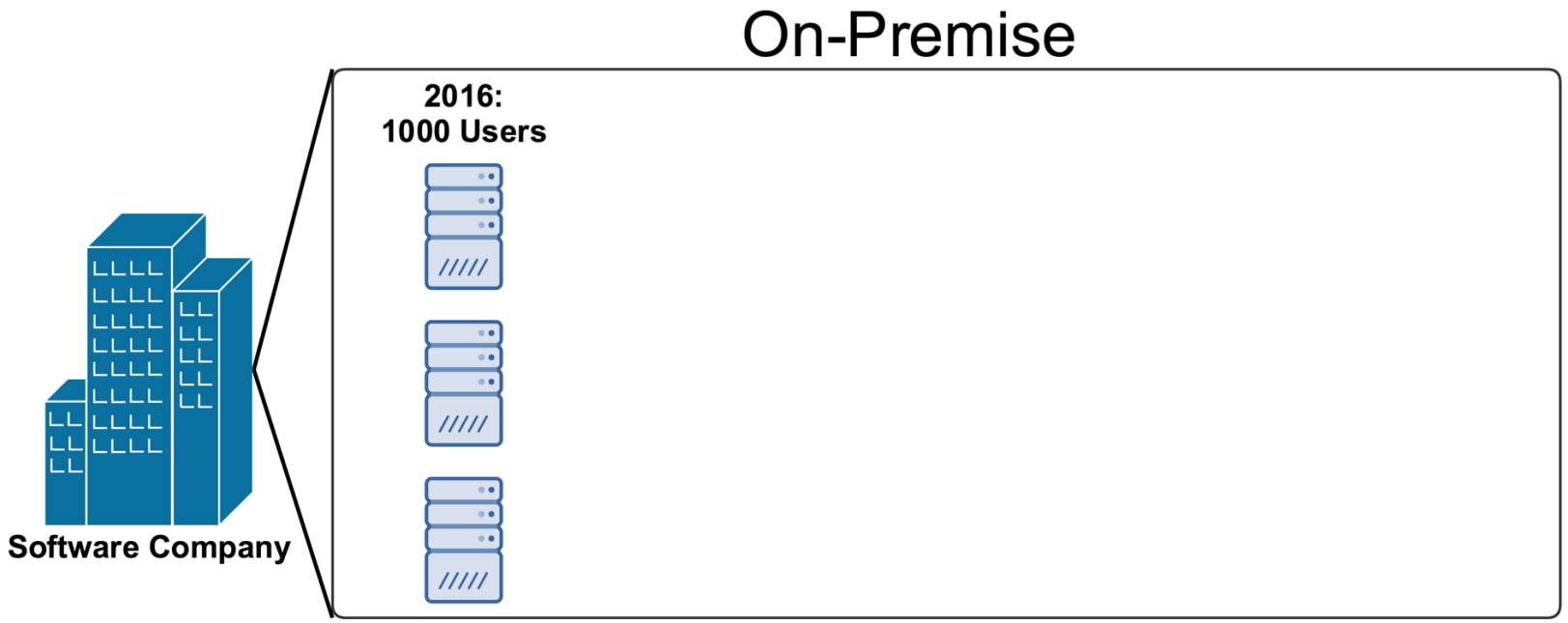
Essential Characteristics

- On demand self service
  - Comparison of traditional environment and cloud environment
- Broad network access

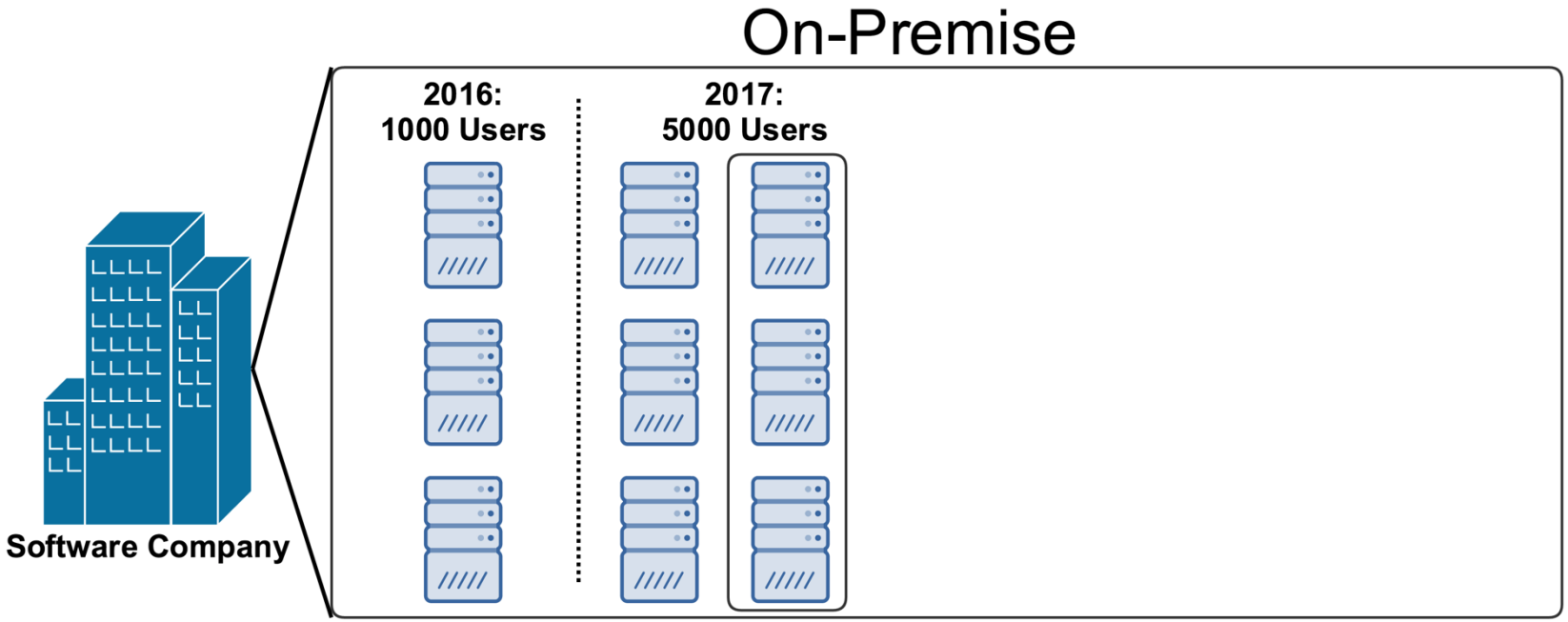
Essential Characteristics

- Resource pooling
- Rapid Elasticity

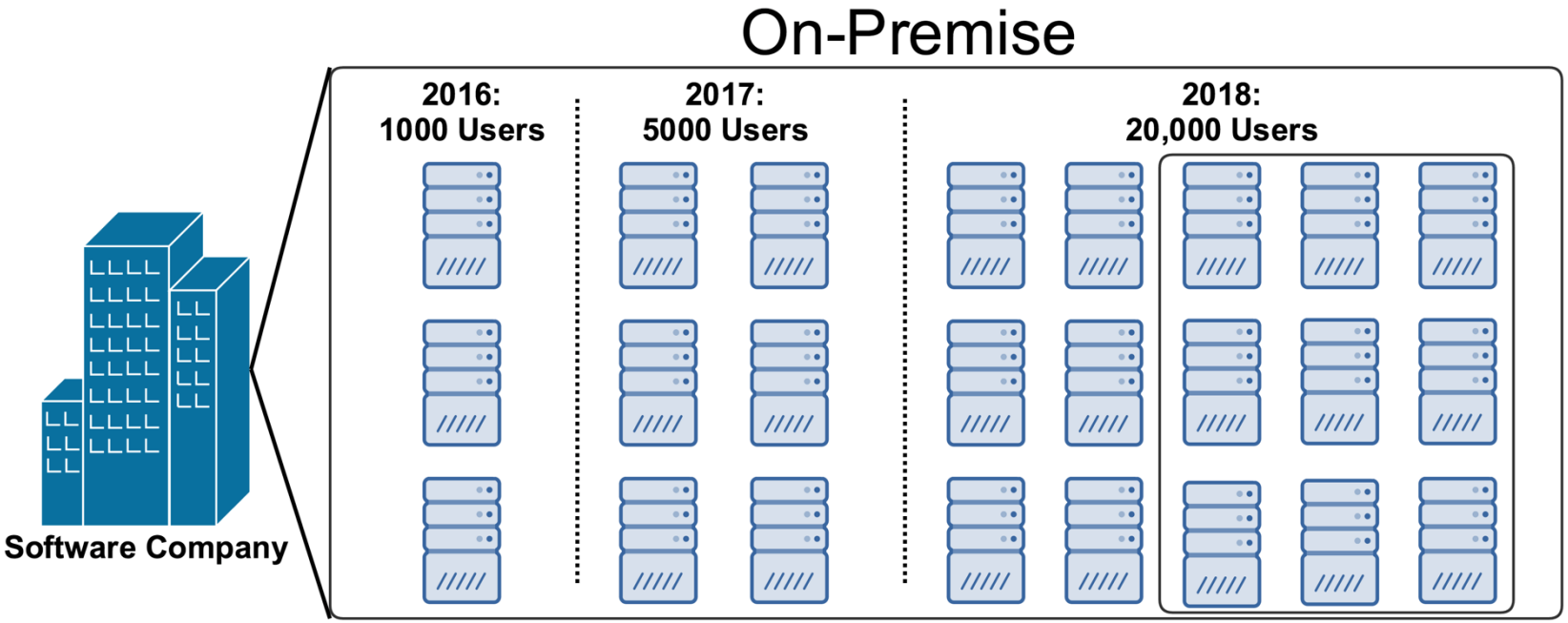
Common **Enterprise**
Uses of Cloud Services



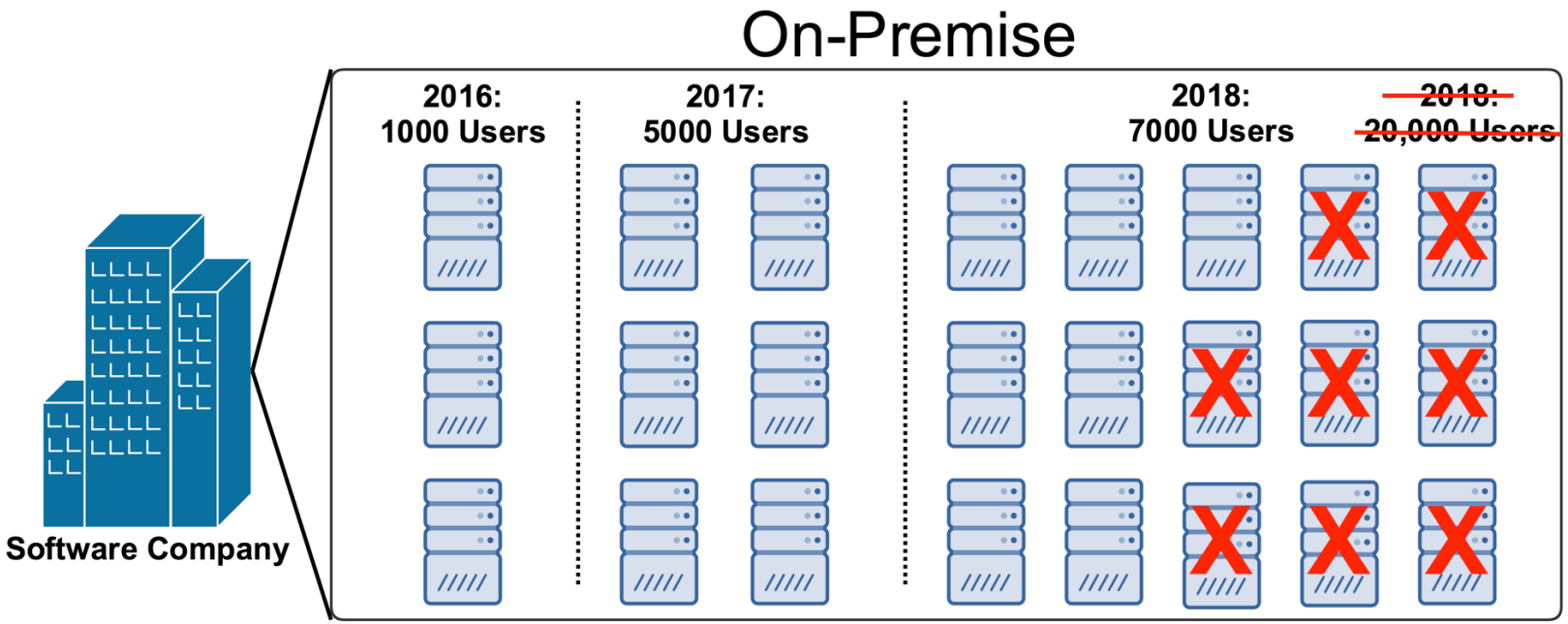
Common **Enterprise** Uses of Cloud Services



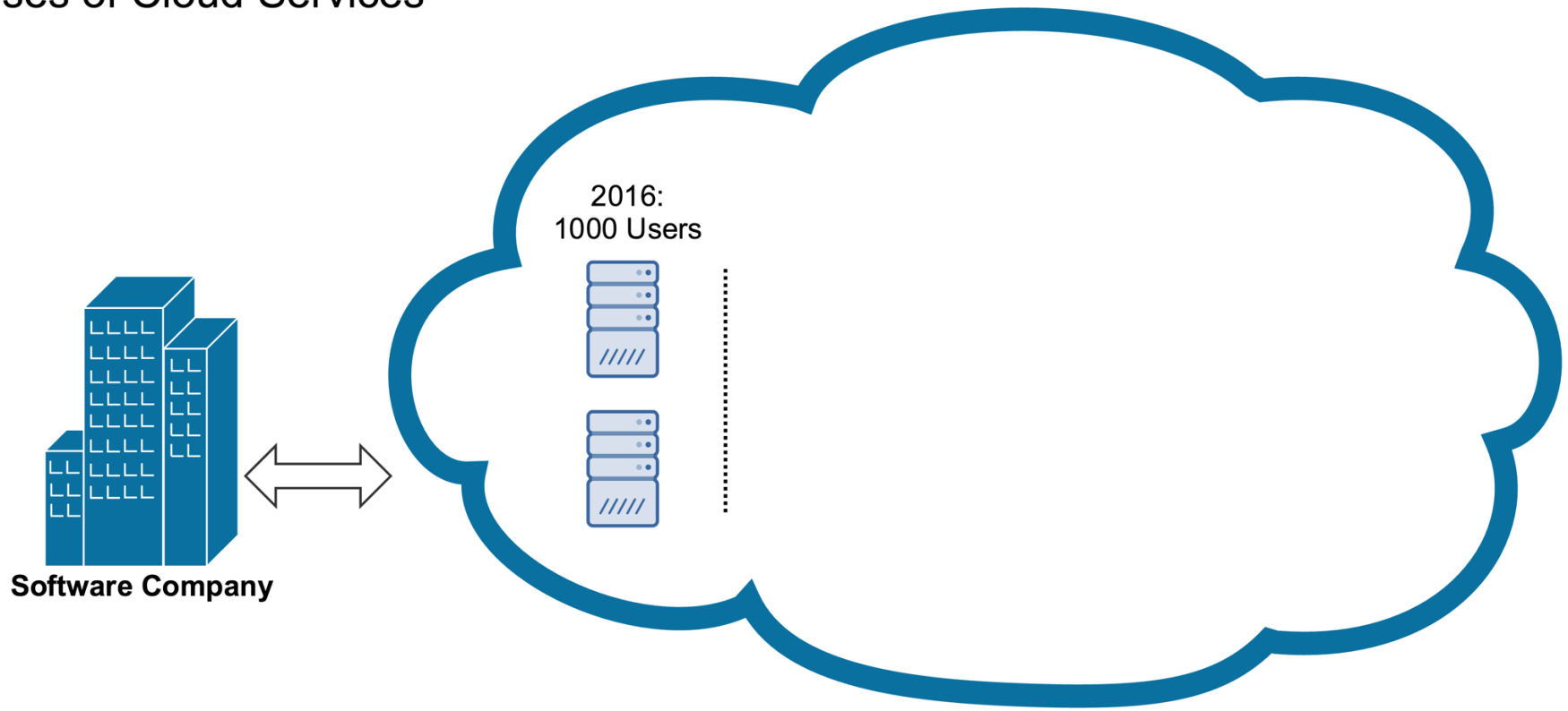
Common **Enterprise** Uses of Cloud Services



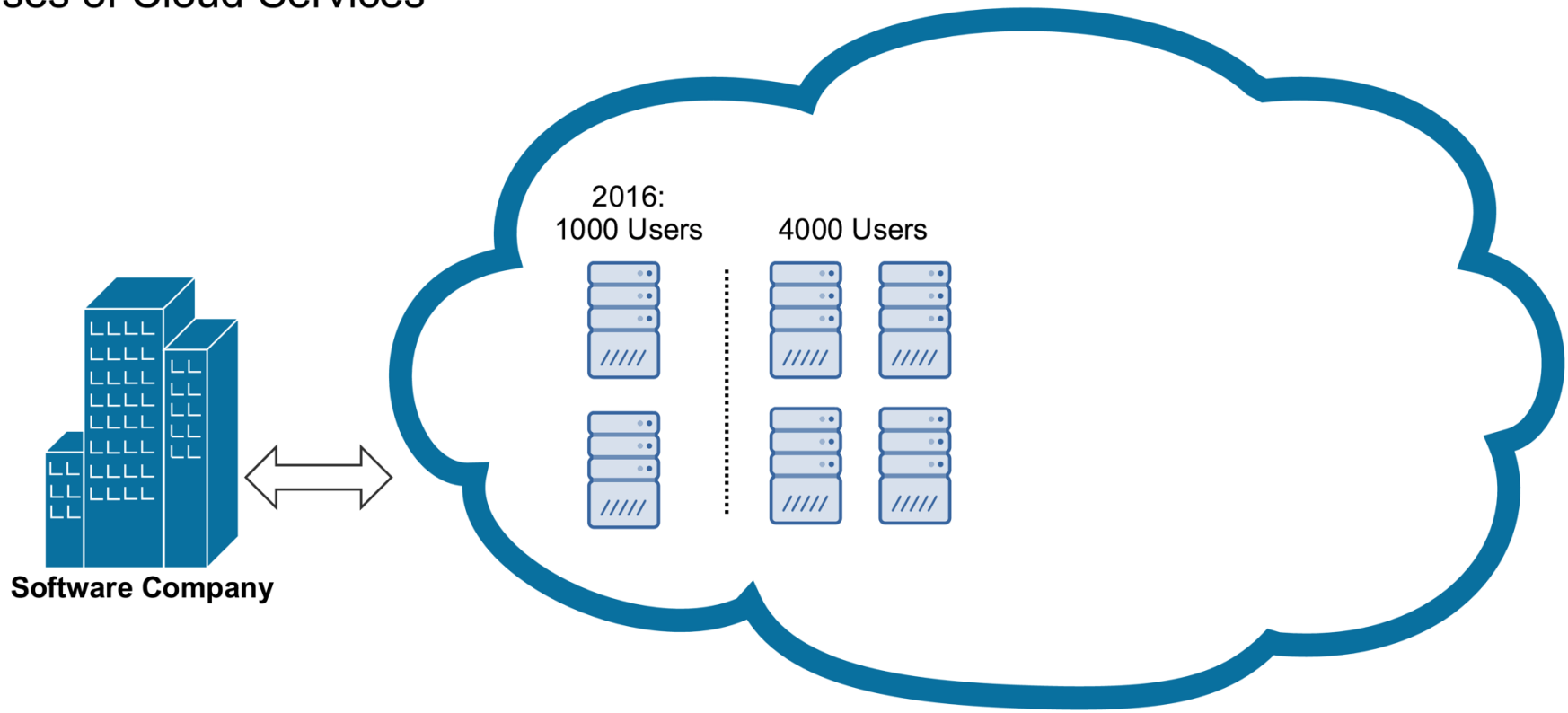
Common **Enterprise** Uses of Cloud Services



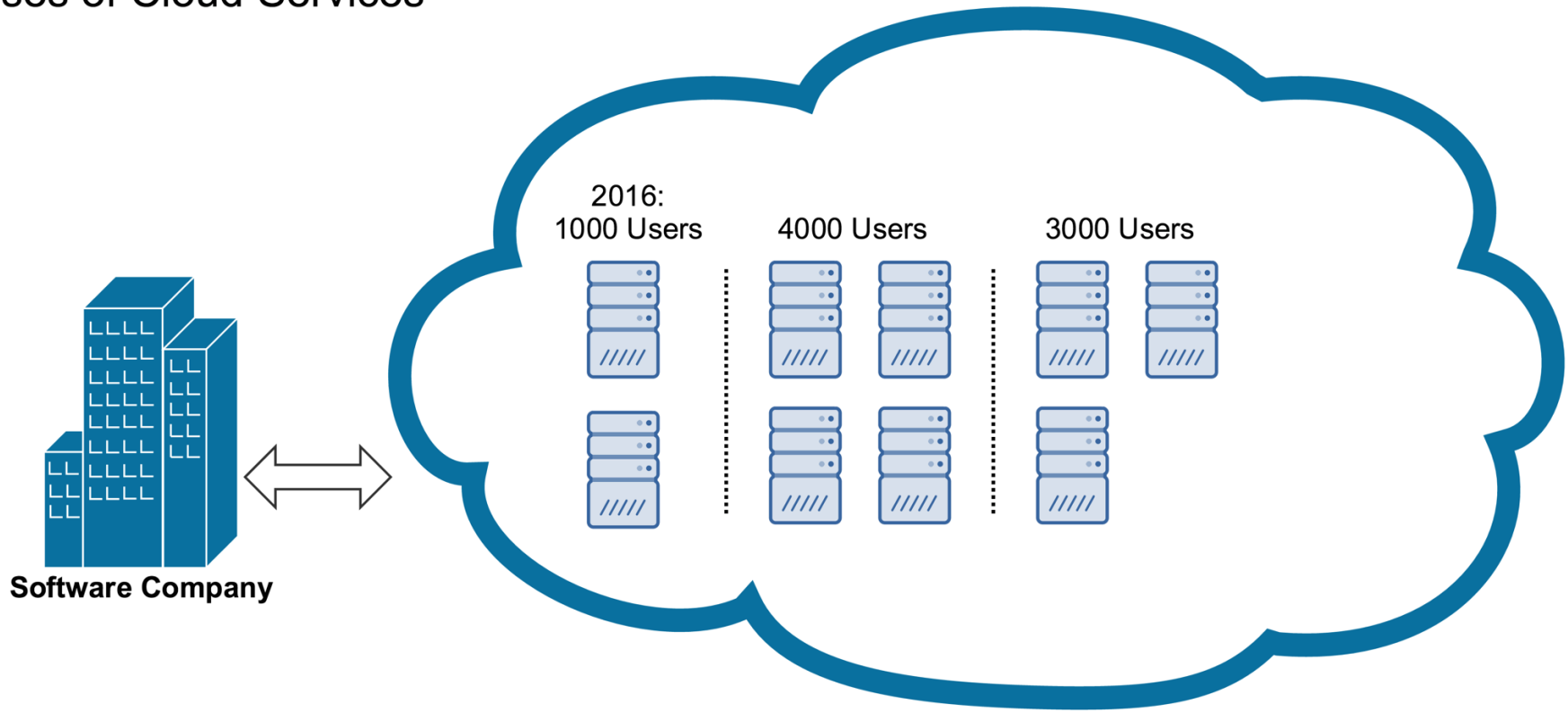
Common **Enterprise** Uses of Cloud Services



Common **Enterprise** Uses of Cloud Services



Common **Enterprise** Uses of Cloud Services



Cloud Elasticity

- Peak hours
- Cloud size can be increased and decreased on-demand
- We say here clouds are auto scalable

Interview point : Earlier we used to take care of peak hours, now we go for elastic services

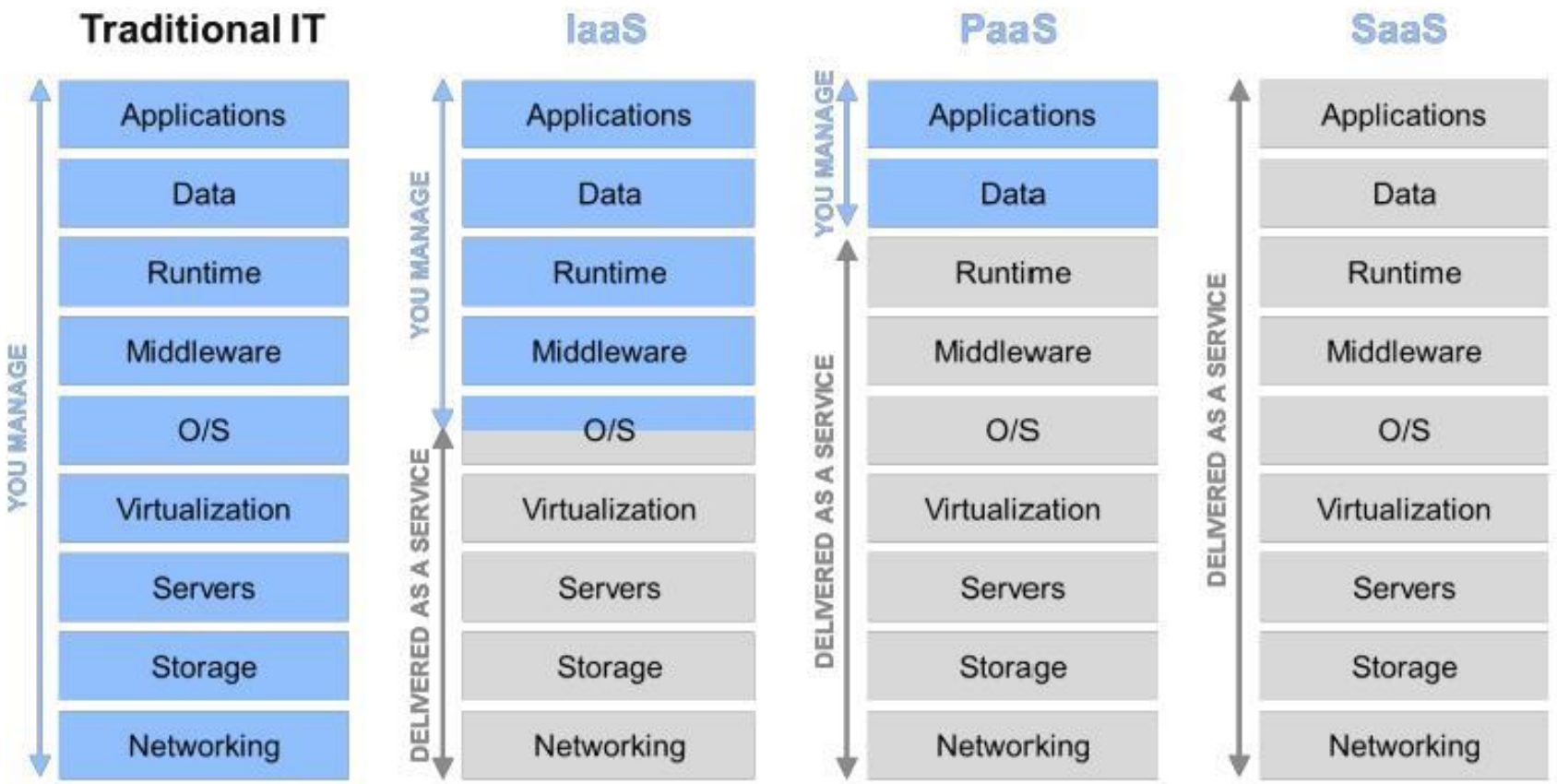
Why Cloud ?

- Business media says 73% of the companies are going for cloud
- Understanding the TCO
- 80 – 20 rule
- Direct cost-comparisons between the Cloud and on-premise are difficult, calculations of in-house costs fail to take into account:
    - The direct costs that accompany running a server: power, floor space, storage, and IT operations to manage those resources.
    - The indirect costs of running a server: network and storage infrastructure and IT operations to manage the general infrastructure.
    - The overhead costs of owning a server: procurement and accounting personnel, not to mention a critical resource in short supply: IT management and its attention.

Four main areas in which Cloud Computing allows businesses to break from the past:

- Virtualization – The ability to increase computing efficiency
- Democratization of Computing – Bringing enterprise scale infrastructure to small and medium businesses
- Scalability and fast provisioning – Bringing web scale IT at a rapid pace
- Commoditization of infrastructure – Enabling IT to focus on the strategic aspects of its role

Service models or Cloud Stack

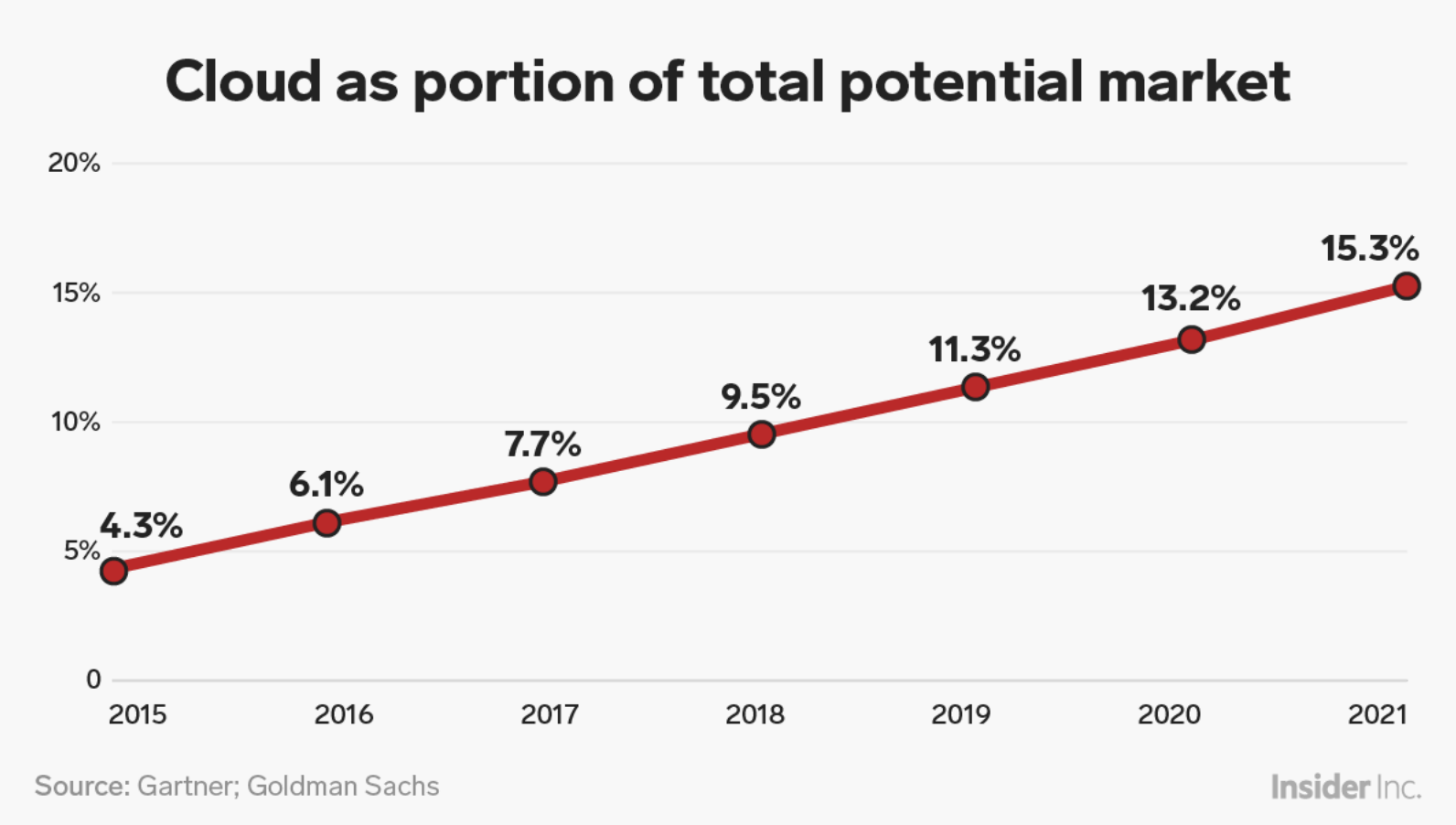


Source: Microsoft.

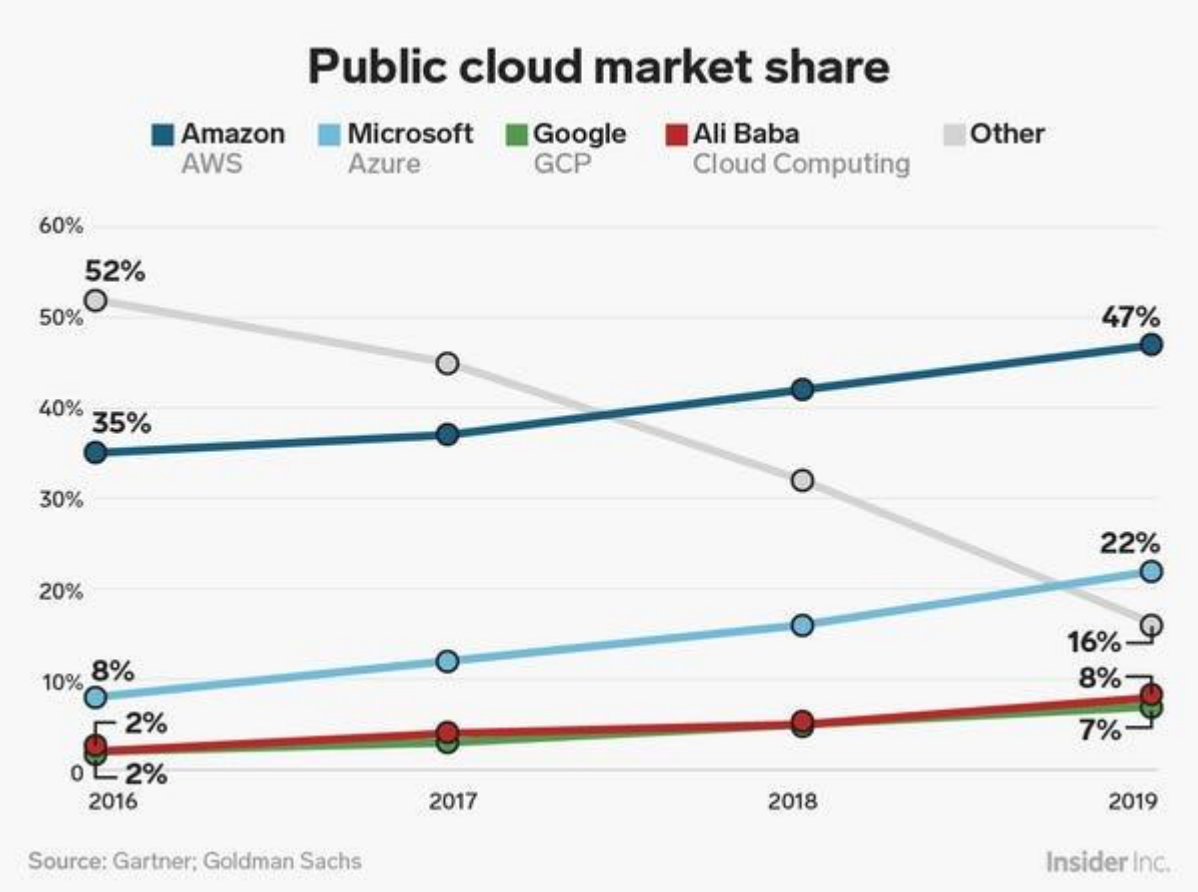
Deployment Models

- Private Cloud
- Community Cloud
- Public Cloud
- Hybrid Cloud
  - Virtual private cloud
  - Cloud Burst

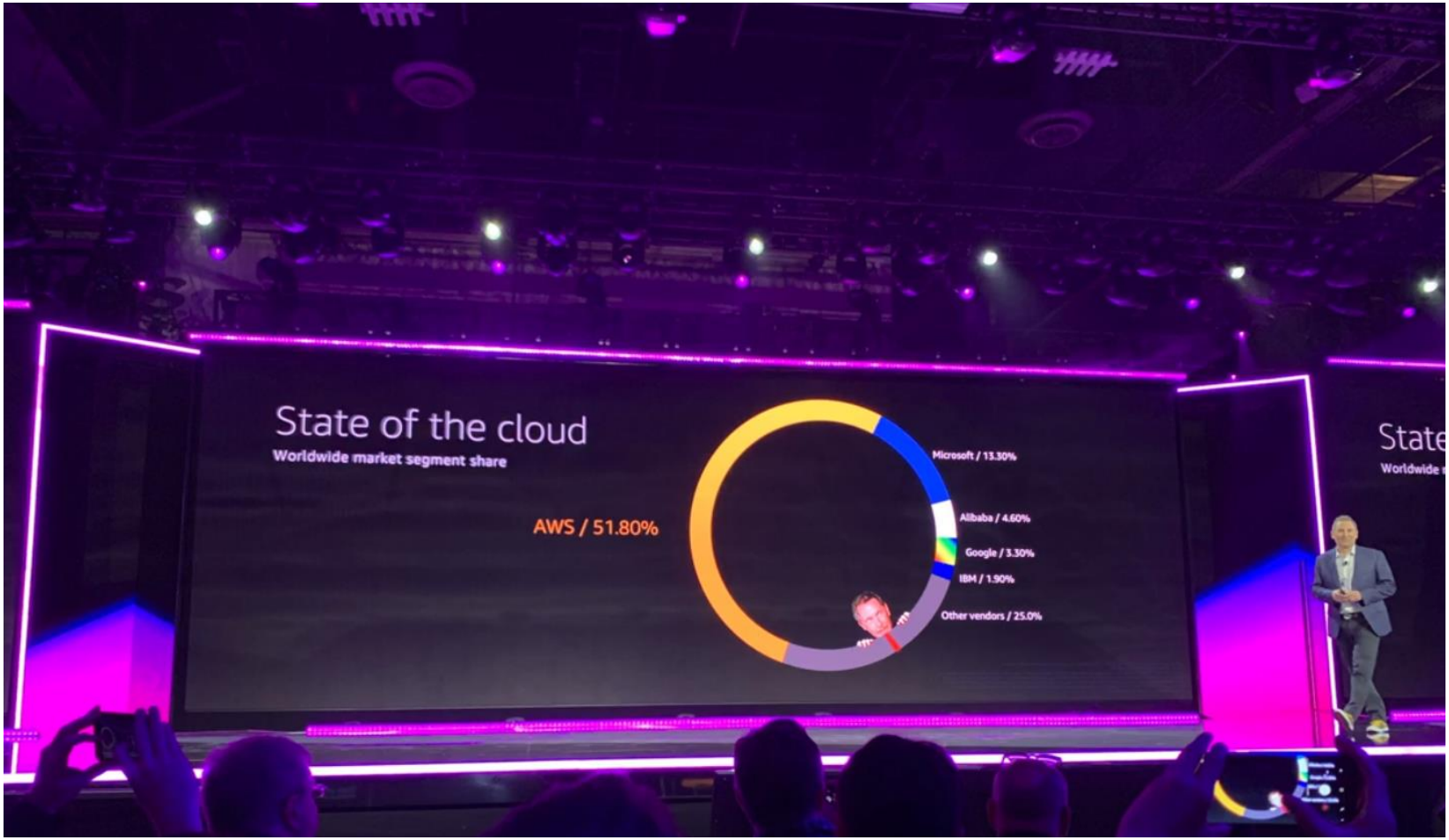
Why learn AWS ?



Why learn AWS ?



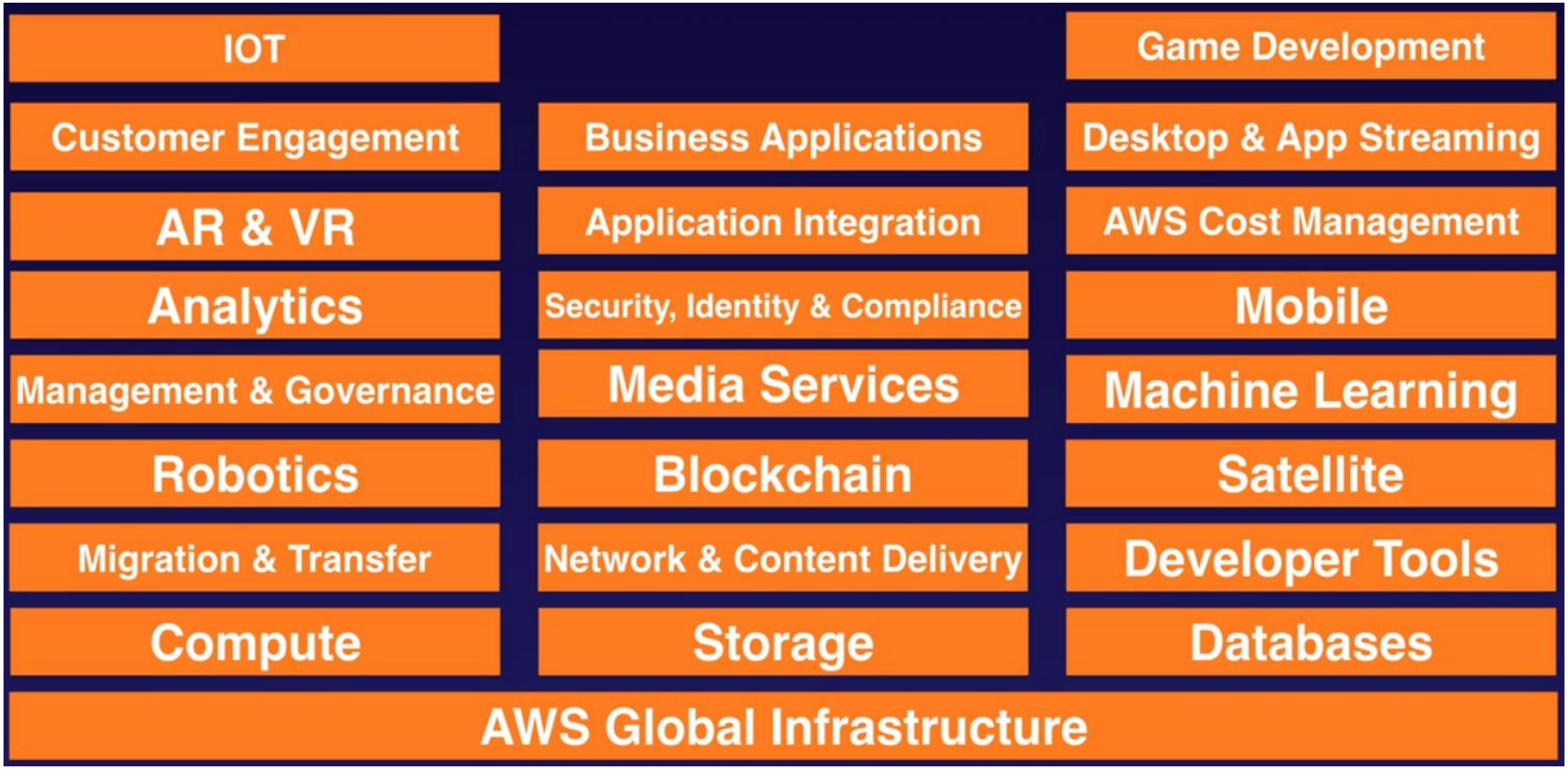
Why learn AWS ?



Why learn AWS ?

- Fastest growing cloud computing platform on the planet
- Largest public cloud computing platform on the planet
- More and more organizations are outsourcing their IT to AWS
- The AWS certifications are the most popular certifications right now
- The safest place to be in IT right now
- Gartner report

AWS Platform



S3 – Simple Storage Service

- S3 is object based storage i.e. it allows you to upload files
- Files can be from 1B to 5TB
- There is unlimited storage
- Files are stored in buckets
- Bucket names are unique globally
- Versioning
- Encryption

S3

- S3 is simple key value store
  - Key (name of the object)
  - Value (data)
- Amazon guarantees 99.99% availability for S3
- Amazon guarantees 11 9's of durability for S3
- S3 provides Lifecycle management
- Secure data using ACL and Bucket policies

S3 - Storage Tiers/Classes

- S3 (Standard) – 99.99% availability and 11 9's durability, stored redundantly across multiple devices in multiple facilities and is design to sustain loss of 2 facilities concurrently
- S3 IA (Infrequently accessed) – For data that is accessed less frequently , but requires rapid access when needed. Lower cost than S3, but requires a retrieval cost
- S3 One Zone – IA : For where you want a lower-cost option for infrequently accessed data, but do not require the multiple data center availability data resilience

S3 - Storage Tiers/Classes

- S3 Intelligent Tiering – Designed to optimise costs by automatically moving data to the most cost-effective access tier, without performance impact or operational overhead
- S3 Glacier – secure durable and low cost storage class for data archiving. Retrieval time configurable from minutes to hours
- S3 Glacier Deep Archive – AWS lowest cost storage class where a retrieval time of 12 hours is acceptable

	S3 Standard	S3 Intelligent-Tiering*	S3 Standard-IA	S3 One Zone-IA†	S3 Glacier	S3 Glacier Deep Archive**
Designed for durability	99.999999999% (11 9's)	99.999999999% (11 9's)	99.999999999% (11 9's)	99.999999999% (11 9's)	99.999999999% (11 9's)	99.999999999% (11 9's)
Designed for availability	99.99%	99.9%	99.9%	99.5%	N/A	N/A
Availability SLA	99.9%	99%	99%	99%	N/A	N/A
Availability Zones	≥3	≥3	≥3	1	≥3	≥3
Minimum capacity charge per object	N/A	N/A	128KB	128KB	40KB	40KB
Minimum storage duration charge	N/A	30 days	30 days	30 days	90 days	180 days
Retrieval fee	N/A	N/A	per GB retrieved	per GB retrieved	per GB retrieved	per GB retrieved
First byte latency	milliseconds	millseconds	milliseconds	milliseconds	select minutes or hours	select hours

Getting data on S3

- Upload to website
- Data Import/export
- Data pipeline

EC2 – Elastic Compute Cloud

- Backbone of AWS
- Provides resizable compute capability in cloud
- You pay for the capacity you use

Amazon EC2

Family	Speciality	Use case
F1	Field Programmable Gate Array	Genomics research, financial analytics, real-time video processing, big data etc
I3	High Speed Storage	NoSQL DBs, Data Warehousing etc
G3	Graphics Intensive	Video Encoding/ 3D Application Streaming
H1	High Disk Throughput	MapReduce-based workloads, distributed file systems such as HDFS and MapR-FS
T3	Lowest Cost, General Purpose	Web Servers/Small DBs
D2	Dense Storage	Fileservers/Data Warehousing/Hadoop
R5	Memory Optimized	Memory Intensive Apps/DBs
M5	General Purpose	Application Servers
C5	Compute Optimized	CPU Intensive Apps/DBs
P3	Graphics/General Purpose GPU	Machine Learning, Bit Coin Mining etc
X1	Memory Optimized	SAP HANA/Apache Spark etc
Z1D	High compute capacity and a high memory footprint.	Ideal for electronic design automation (EDA) and certain relational database workloads with high per-core licensing costs.
A1	Arm-based workloads	Scale-out workloads such as web servers
U-6tb1	Bare Metal	Bare metal capabilities that eliminate virtualization overhead

IAM – Identity and Access Management

- IAM allows you to manage users and their level of access to the AWS console
- Provides centralized control for your AWS account
- Granular permissions
- Identity federation
- Multifactor authentication
- Allows to set up password rotation policies

Volumes and Snapshots

- Volumes exist on EBS (Virtual HDD)
- Snapshots exists on S3
- You can take a snapshot of a volume, this will store that on S3
- Snapshots are point in time copies of volumes

EFS – Elastic File System

- A file storage service for EC2
- Central storage repository
- Supports NFSv4 protocol
- Pay as you use (30 cents per GB)
- Can support thousands of concurrent NFS connections

EC2 labs

- Web server on EC2
- Bootstrapping
- EC2 metadata (curl http://169.254.169.254/latest/meta-data/)
- Load Balancer
- AWS CLI
- Using roles to access S3
- Auto scaling

CloudWatch

- Monitoring
- Dashboards
- Alarms
- Events
- Std monitoring - 5 min
- Detailed monitoring – 1 min

VPC

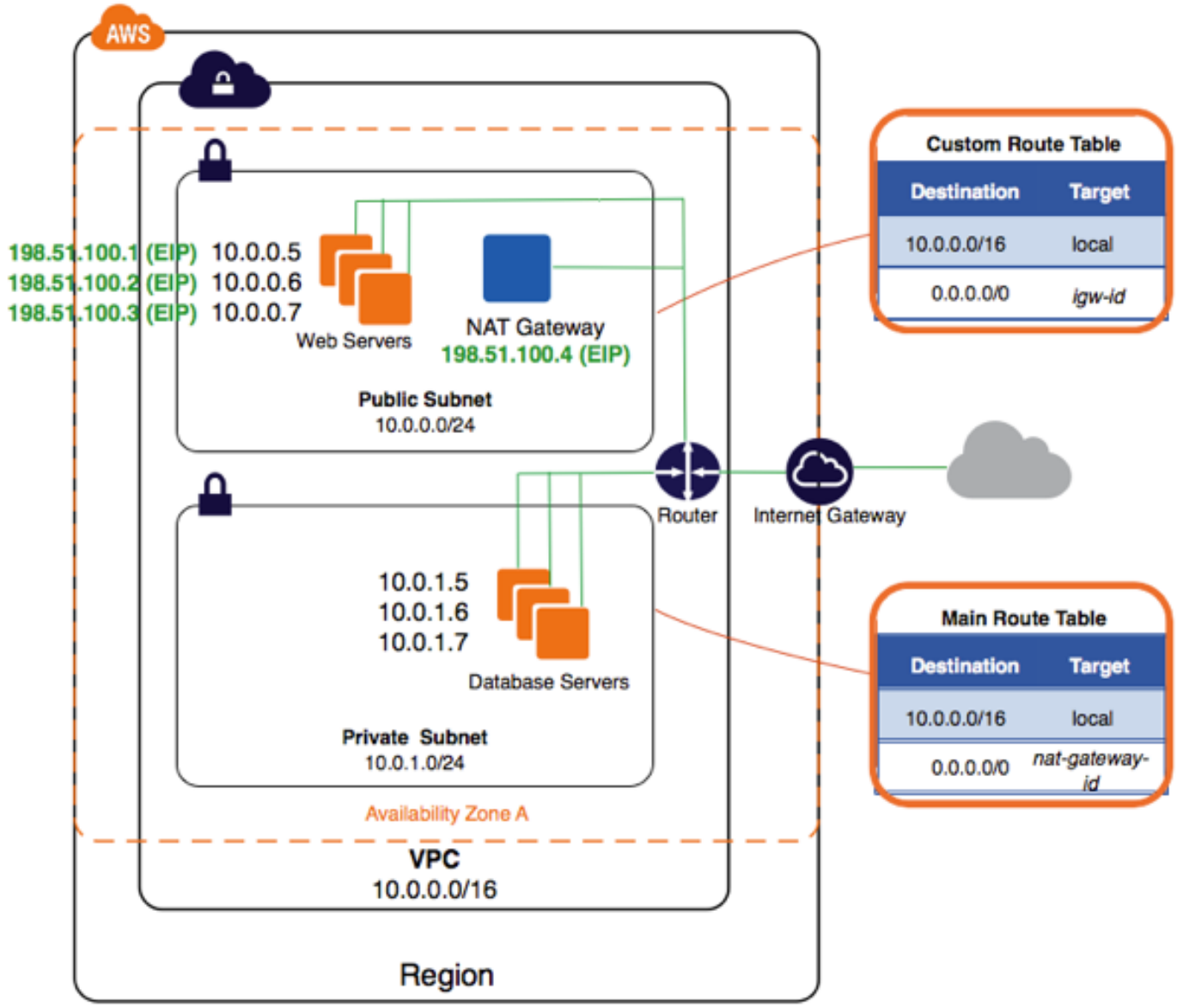
- Amazon Virtual Private cloud (VPC) lets you provision a logically isolated section of AWS cloud where you can launch AWS resources in a virtual network that you define

- You have complete control over virtual networking environment, including selection of your IP range, creation of subnet, configuration of route tables and network gateways

- Think of VPC as a logical datacentre in AWS

What can you do with a VPC

- Launch instances in a subnet of your choice
- Assign custom IP addresses
- Configure route tables between each subnet
- Create Internet Gateway and attach it to our VPC
- Much better security control over your AWS resources



Thank you