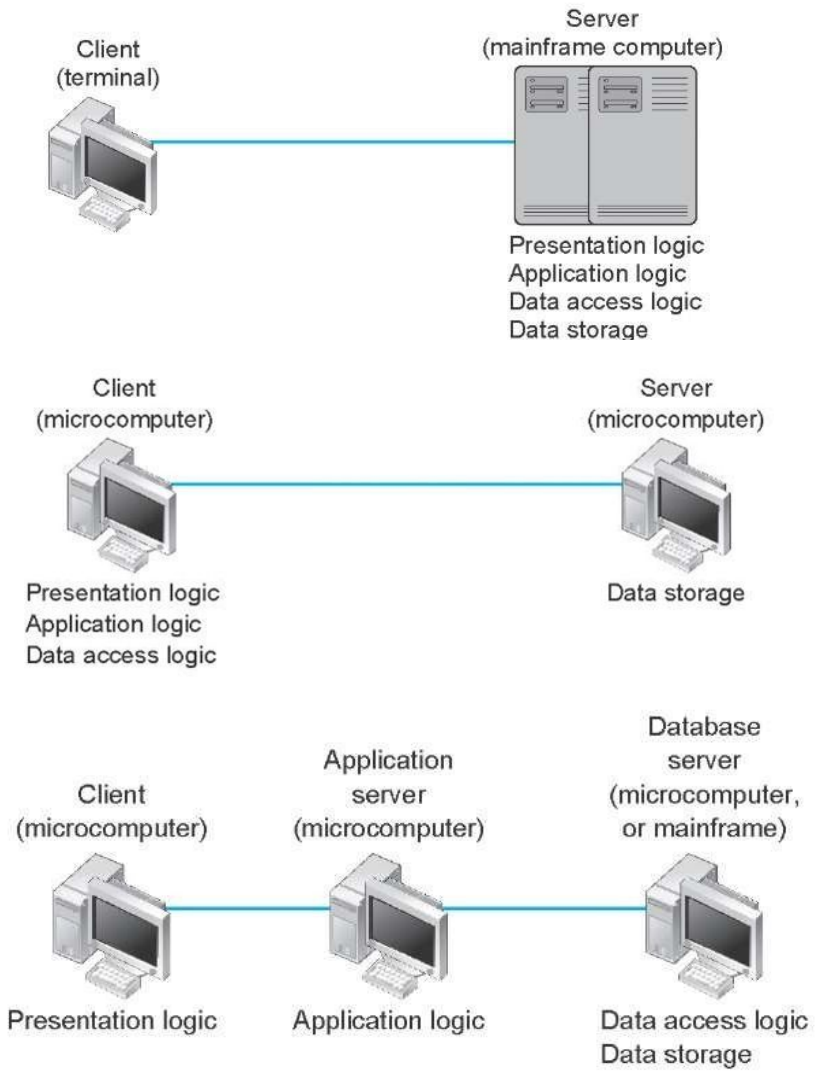
Chapter 1. Introduction

Bilkent University | CS443 | 2021, Spring | Dr. Orçun Dayıbaş

Motivation

- **History**
  - Mainframes
    - LAN-based
  - Client/server
    - 2-tier (fat client)
  - N-tier Applications
    - to be precise; 3 and more tiers
    - Thin client

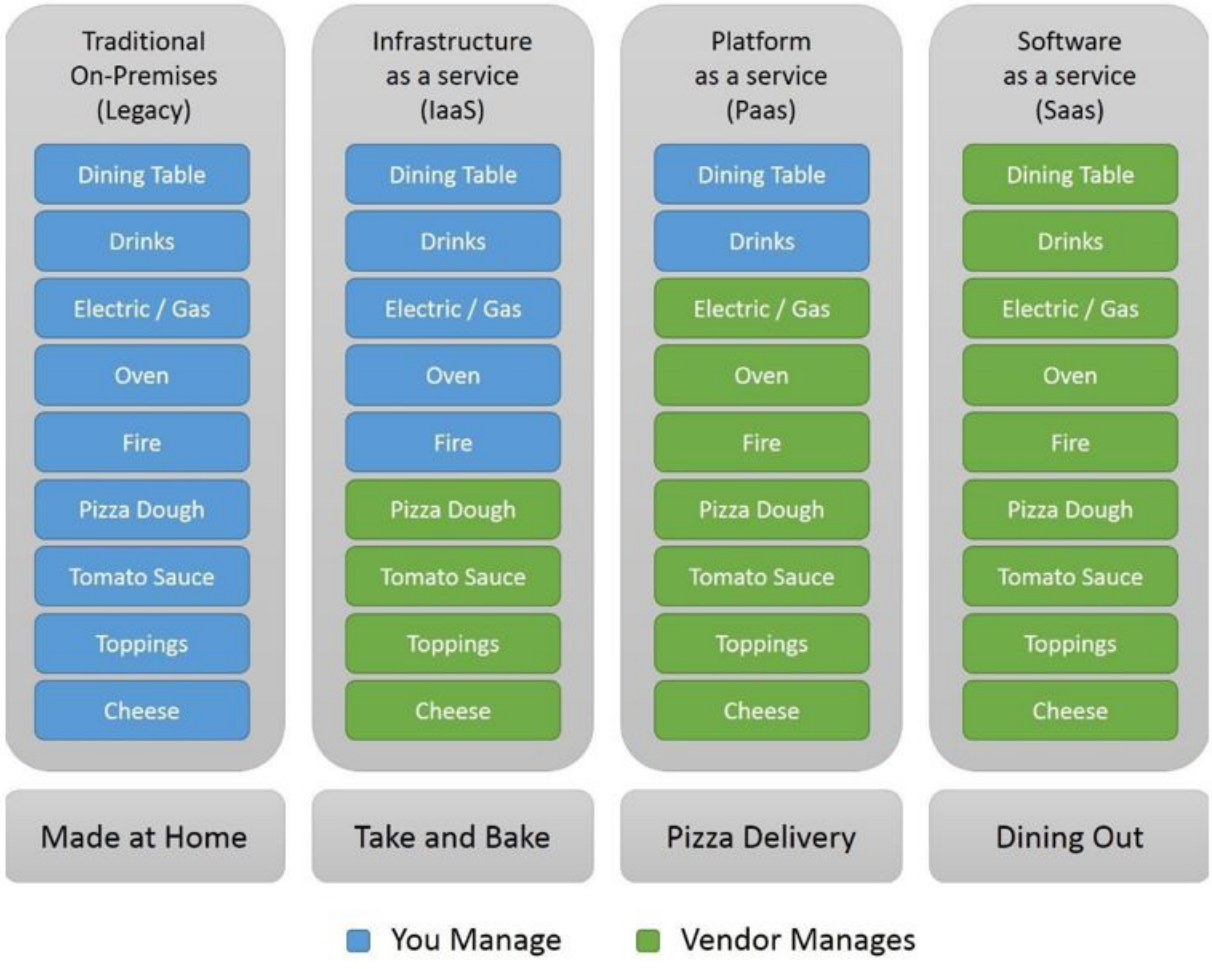


Motivation

- **Cloud computing**
  - NIST definition: “Cloud Computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction.
  - I don’t care about tiers, just give me a service for “X”
- **Service models (XaaS)**
  - IaaS / PaaS / SaaS

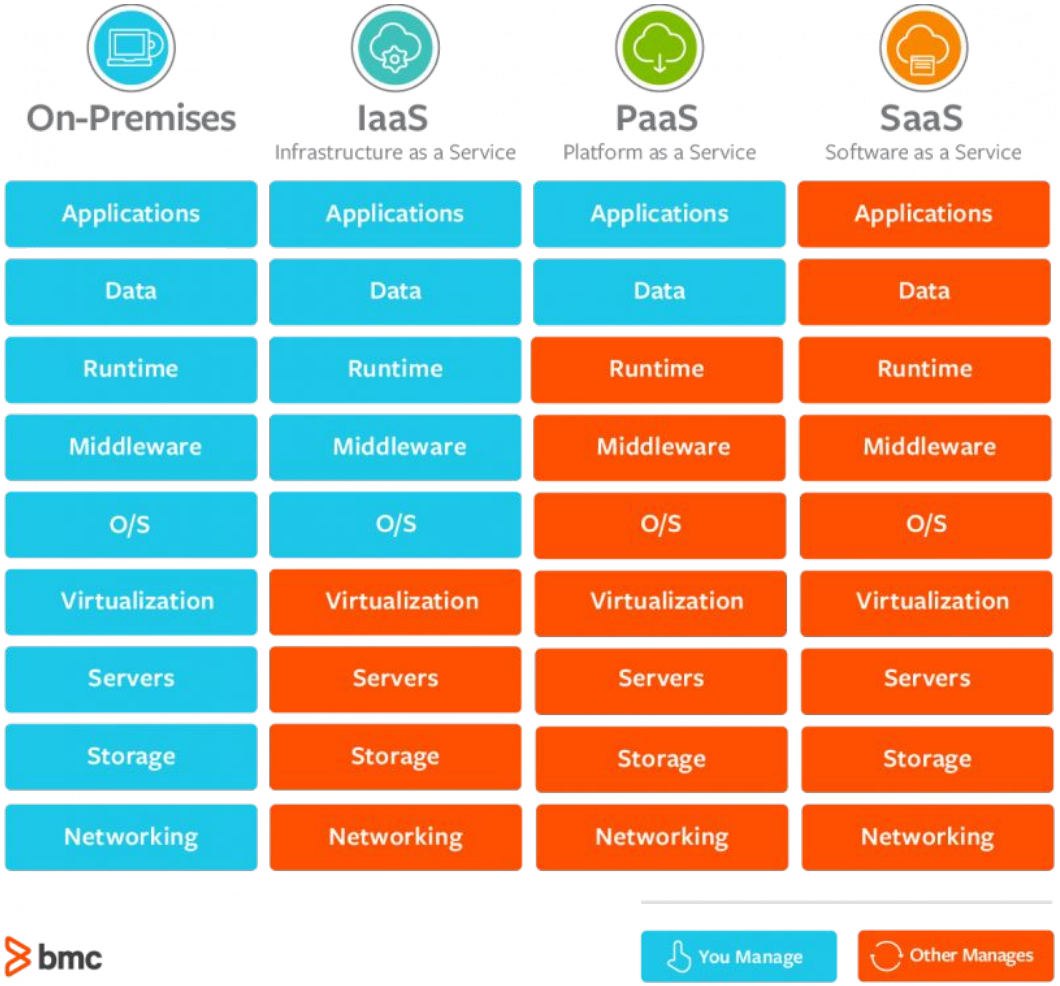
Motivation

- **Pizza as a Service**



Motivation

- **Service Models**



Platform Type	Common Examples
SaaS	Google Apps, Dropbox, Salesforce, Cisco WebEx, Concur, GoToMeeting
PaaS	AWS Elastic Beanstalk, Windows Azure, Heroku, Force.com, Google App Engine, Apache Stratos, OpenShift
IaaS	DigitalOcean, Linode, Rackspace, Amazon Web Services (AWS), Cisco Metapod, Microsoft Azure, Google Compute Engine (GCE)

Motivation

- **Management Models**
  - Public cloud
    - Services delivered across the public internet
    - The cloud vendor is resp. for managing (developing, maintaining, etc.) the pool of computing resource
    - **Pros**
      - No pre-investment, High scalability/flexibility, reduced complexity, flexible pricing based on diff. SLA offerings
    - **Cons**
      - The total cost of ownership (TCO) can rise exponentially for large-scale usage
      - Low visibility & control over infrastructure

Motivation

- **Management Models**
  - Private cloud
    - A solution dedicated for use by a single organization
    - Data center resources may be located on-prem or operated by 3th party vendor off-site (the computing resources are not shared with other customers)
    - **Pros**
      - With greater visibility and control into the infrastructure, organizations can operate compliance-sensitive IT workloads without compromising on the security and performance
      - High SLA performance and efficiency

Motivation

- **Management Models**
  - Private cloud
    - **Cons**
      - Expensive solution with a relatively high total cost of ownership as compared to public cloud alternatives for short-term use cases
      - The infrastructure may not offer high scalability to meet unpredictable demands if the cloud data center is limited to on-premise computing resources.

Motivation

- **Management Models**
  - Hybrid cloud
    - A mixture of public and private cloud solutions
    - Apps and data workloads can share the resources between public and private cloud deployment based on business and technical policies
    - **Pros**
      - Flexible policy-driven deployment to distribute workloads across public and private infrastructure
      - Scalability of public cloud environments is achieved without exposing sensitive IT workloads to the inherent security risks
      - High reliability as the services are distributed across multiple data centers across public and private data centers

Motivation

- **Management Models**
  - Hybrid cloud
    - **Cons**
      - Cost (it can be expensive due to additional investments)
      - Additional infrastructure complexity is introduced as organizations operate and manage an evolving mix of private and public cloud architecture

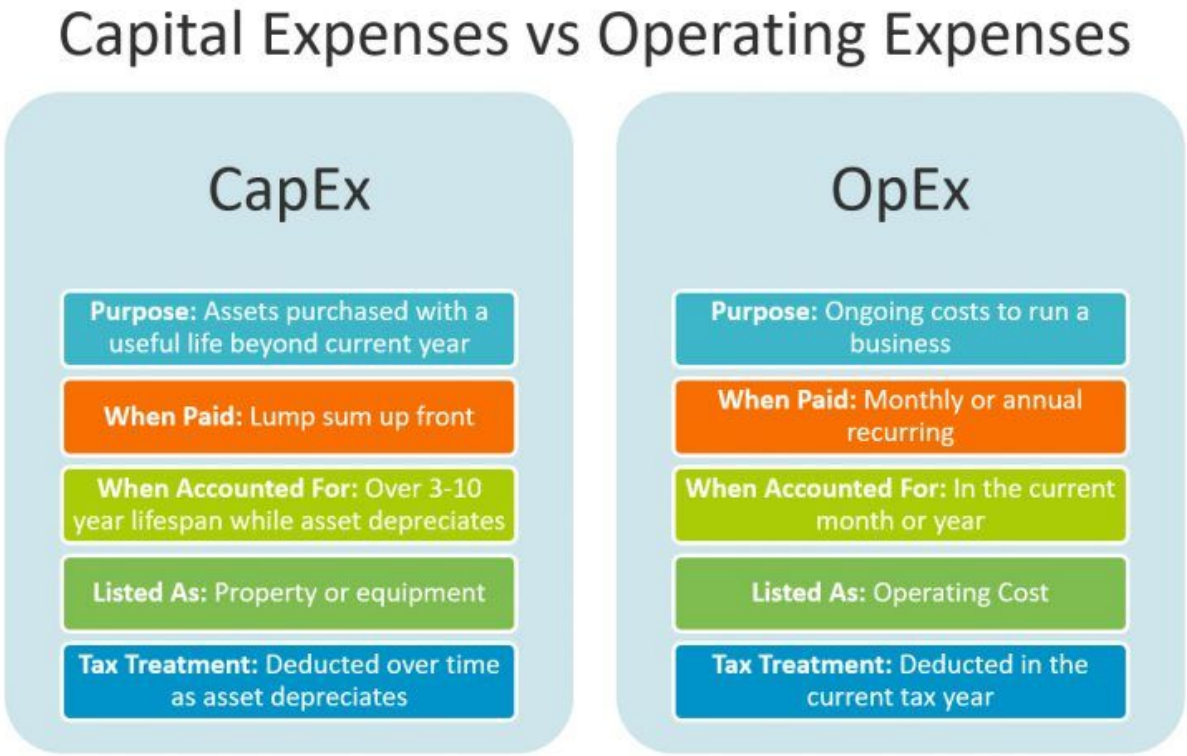
"Cloud Bursting" is a configuration that's set up between a private cloud and a public cloud to deal with peaks in IT demand. If an organization using a private cloud reaches 100 percent of its resource capacity, the overflow traffic is directed to a public cloud so there's no interruption of services.

Motivation

- **Criteria to select a cloud provider**
  - Workload attributes: Performance, security, data size
  - Business needs: time to deployment, compliance regulatory, geographical reach, SLA
  - Ecosystem availability: maturity of SaaS offerings, viability of alt. Services, avail. Of reseller/integrator

Essential Characteristics

- **Cloud Economics**



An IEEE study showed that most servers will show CPU utilization between 10 and 15% (averaged over 24 h, and 7 days a week), and the same is true of network bandwidth (storage is also underutilized)

What about your car?

bmc

Essential Characteristics

- **5 essentials**

Essential Characteristics

- **1. Rapid Elasticity**
  - The most essential one (it's at the core)
  - Ability to scale resources up/down as needed
  - To the consumers, the cloud appears to be **infinite**
    - They can purchase as much or as little computing power as they need
    - The cloud provider has to make it **transparent** to their customers

Essential Characteristics

- **2. Measured Service**
  - Aspects of the cloud service are controlled and **monitored** by the cloud provider
  - Crucial for;
    - Billing, Access control, resource optimization, capacity planning and other tasks

Essential Characteristics

- **3. On-demand Self-Service**
  - A consumer can use services as needed without any human interaction
  - NIST definition: “Cloud Computing is a model for enabling ubiquitous, **convenient**, **on-demand** network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction.

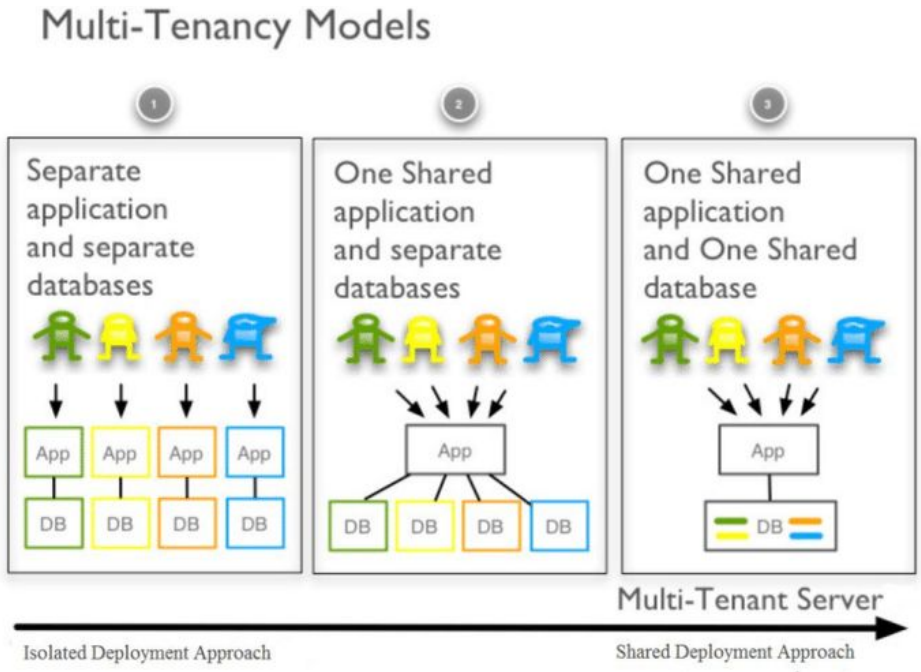
Essential Characteristics

- **4. Ubiquitous Network Access**
  - The cloud provider's capabilities are available over the network and can be accessed through standard mechanisms

"Ubiquitous computing", is the growing trend of embedding computational capability (generally in the form of microprocessors) into everyday objects to make them effectively communicate and perform useful tasks in a way that minimizes the end user's need to interact with computers as computers. Ubiquitous computing devices are network-connected and constantly available.

Essential Characteristics

- **5. Resource Pooling**
  - It allows a cloud provider to serve its consumers via a multi-tenant model



Considerations

- **Cloud computing benefits**
  - Shared infrastructure reduces cost
  - Pay as you go or only pay for what you use
  - On-demand elasticity (from large pool of res.)
  - Increased focus on the application layer
  - Let someone else worry about the HW
  - Avoid decaying investments like HW

Considerations

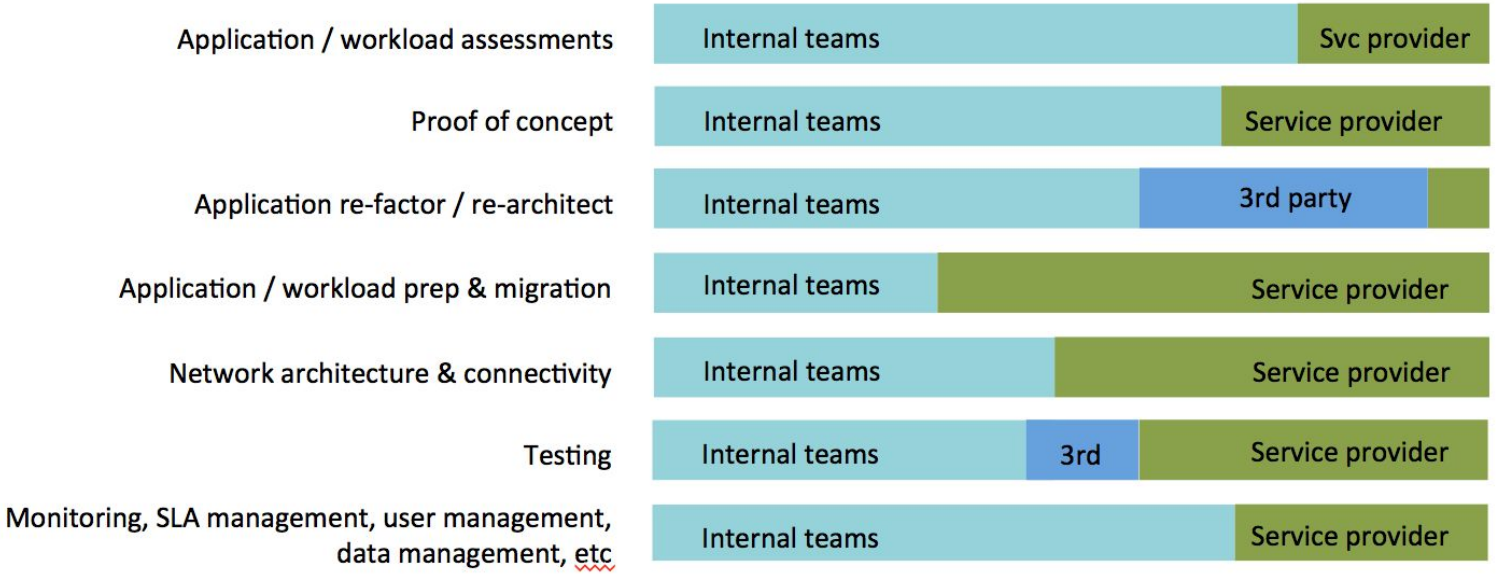
- **Cloud computing potential risks**
  - You lose direct knowledge and control on underlying HW
  - Noisy neighbors or other VMs sharing the same HW
  - Hard to diagnose performance issues, due to limited visibility and virtualization
  - Potential security risks of placing your mission critical data on remote servers
  - Vendor lock-in means getting stuck with a Cloud provider who has your data

How to select a provider?

- **Certification & standards**
  - e.g. security → ISO 27001 or ISO 27017

How to select a provider?

- **Technologies & roadmap**
  - Assess how much re-coding or customisation you may have to do
  - Ask about the provider's roadmap of service development (long term fit)



How to select a provider?

- **Data governance & security**
  - You should be aware of regulatory or data privacy rules governing personal data
  - The location your data resides in may be a key
  - Regulations
    - Europe: GDPR (The General Data Protection Regulation 2016/679) is for EU/EAA citizens
    - Turkey: KVKK (Kişisel Verilerin Korunması Kanunu) is for Turkish citizens
      - https://www.kvkk.gov.tr/
    - US: ???

How to select a provider?

- **Ecosystem & service dependencies**
  - Service providers may have multiple vendor relationships that are important to understand
  - SaaS CRM for instance – are there existing integrations with finance and marketing services?
  - For PaaS – is there a cloud marketplace from which to buy complimentary services that are preconfigured to integrate effectively on the same platform?

How to select a provider?

- **Contracts, Commercials & SLAs**
  - Agreements range from out of the box “terms and conditions” agreed online, through to individually negotiated contracts and SLAs
  - Each provider has a unique bundle of services / pricing models / price advantages

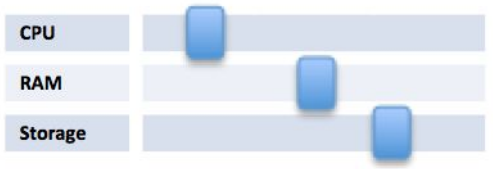
Consumption Period

	CSP - A	CSP - B	CSP – C
Minutes	✔	✔	✗
Hours	✗	✔	✗
Months	✗	✔	✔
Years	✔	✗	✔

Packaged

	Tier 1	Tier 2	Tier 3
Core	1	2	3
RAM	512MB	1GB	4GB
Storage	20GB	30GB	40GB
Network	3TB	6TB	9TB

Configurable



How to select a provider?

- **Vendor Lock-in, Company Profile**
  - Services that rely heavily on bespoke or unique proprietary components may impact your portability to other providers or in-house operations
  - Ensure you have a clear **exit strategy** in place at the start of your relationship.

Examples of vendor lock-in candidates

- CSP compatible application architecture
- Proprietary cloud management tools
- Customised geographic diversity
- Proprietary cloud APIs
- Customised cloud Web services (e.g. Database)
- Premium configurations
- Custom configurations
- Data controls and access
- Data formats (not standardised)
- Service density with one provider

SLA (Service Level Agreement)

- **A service-level agreement (SLA) is a commitment between a service provider and a client**
  - Particular aspects of the service (quality, availability, responsibilities) are agreed between the service provider and the service user
  - SLAs should contain 3 major components:
    - Service level objectives (SLO)
    - Remediation policies and penalties/incentives related to these objectives
    - Exclusions and caveats

SLA (Service Level Agreement)

- **SLAs are being used for the following purposes**
  - **Data persistence, system reliability, and business continuity** as individual consumers may be patient to wait for their search or e-mail results but businesses need predictability to meet their goals and deliver products in a timely manner
  - SLAs imply that Cloud Service Providers will need systems in place to ensure **redundancy** and **security** of their customers' information
  - SLAs also need to cover **performance** aspects of the Cloud Services being provided in terms of the required computing speed and network bandwidth.

SLA (Service Level Agreement)

- **Example**
  - More on this: https://cloud.google.com/blog/products/gcp/sre-fundamentals-slis-slas-and-slos

TERM	DEFINITION	EXAMPLE
Service level indicator (SLI)	The SLI is your core measurement of performance.	"Customers can log in and view their data ..."
Service level objective (SLO)	SLOs are the target values or goals for the performance of your system. SLOs represent an ongoing commitment.	"99.9% of the time ..."
Service level agreement (SLA)	The SLA defines what happens if you don't meet your SLI/SLO commitments.	"... or they can request a refund for losses incurred due to unavailability of the service."

https://newrelic.com/resource/devops-done-right

Q/A