Topics you will find in this deck

- What is inside a Computer and Server ?
- How are Enterprise Servers Different?
- What are Application, Web and Data Servers?
- What is a Hybrid environment? (On-premise and Cloud)?
- What are Development, Test and Production Environments ?
- What is Hardware, Software and OS virtualization?
- What is an Use Case?
- What is an Universe(semantic layer)?
- What is Normalization?
- What is Denormalization?
- What are Primary and Foreign Keys?
- What is ETL?
- What is IoT Data?
- Data Wrangling
- Data Massaging
- Data Curator
- Data Ethics and Security
- Data Governance and Management
- Data Masking and Anonymization

You are free to

- Share — copy and redistribute the material in any medium or format
- Adapt — remix, transform, and build upon the material for any purpose, even commercially.

Under the following terms*

Attribution — You must give appropriate credit, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the Author endorses you or your use

Please refer for the full terms: *https://creativecommons.org/licenses/by/4.0/*

*The author does not provide any guarantees, warranties for the content and accuracy. The content may or may not be updated. By using this material you indemnify the author against any liabilities including copyrights and/or damages of any sort.

Application server

What it is

A middleware software deployed on premise or on cloud

It like a store where people come and get what they want

Can run multiple programming languages

Responsible for High Availability, Traffic management, Relability

Application

Orchestrates the user interaction with the business logic, Transformations, database and other IT resources

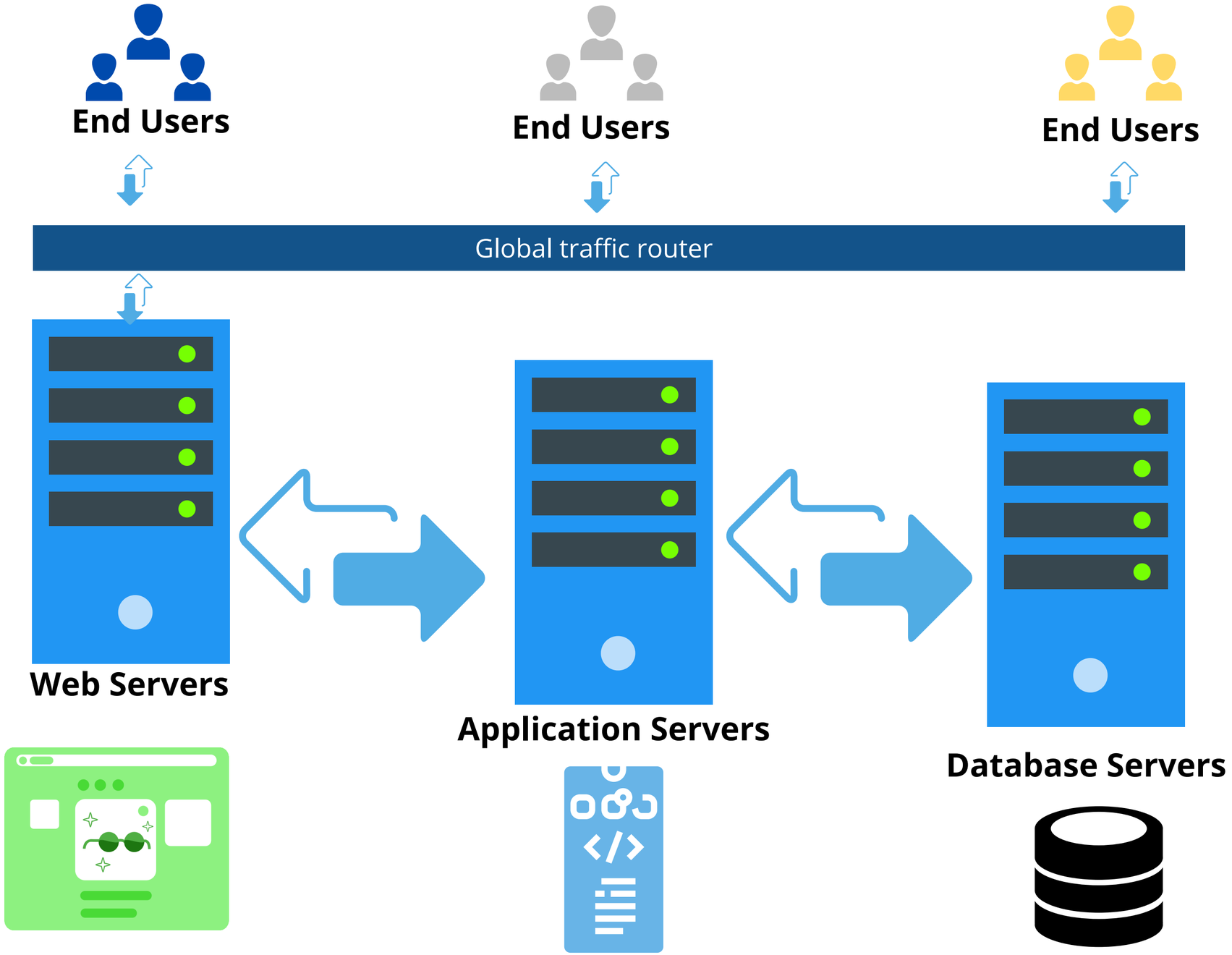
Most major application run on a Application server. Many come embedded in the OS. If not you have to install to run the application

Examples

Most enterprise applications use it to render some part of it

Vendors

Tom Cat
IBM Websphere
Glass fish
Oracle Web logic


End Users
End Users
End Users
Global traffic router
Web Servers
Application Servers
Database Servers

On premise / Cloud / Hybrid

What it is

A method of storing data, applications and other IT assets in a distributed environment

On premise - Your own Data center

Cloud - Private, Public, Multi

Application

- Lower over all costs
- No installation, upgrade and maintenance
- Ad hoc scaling capability
- Stronger Security
- Connectivity

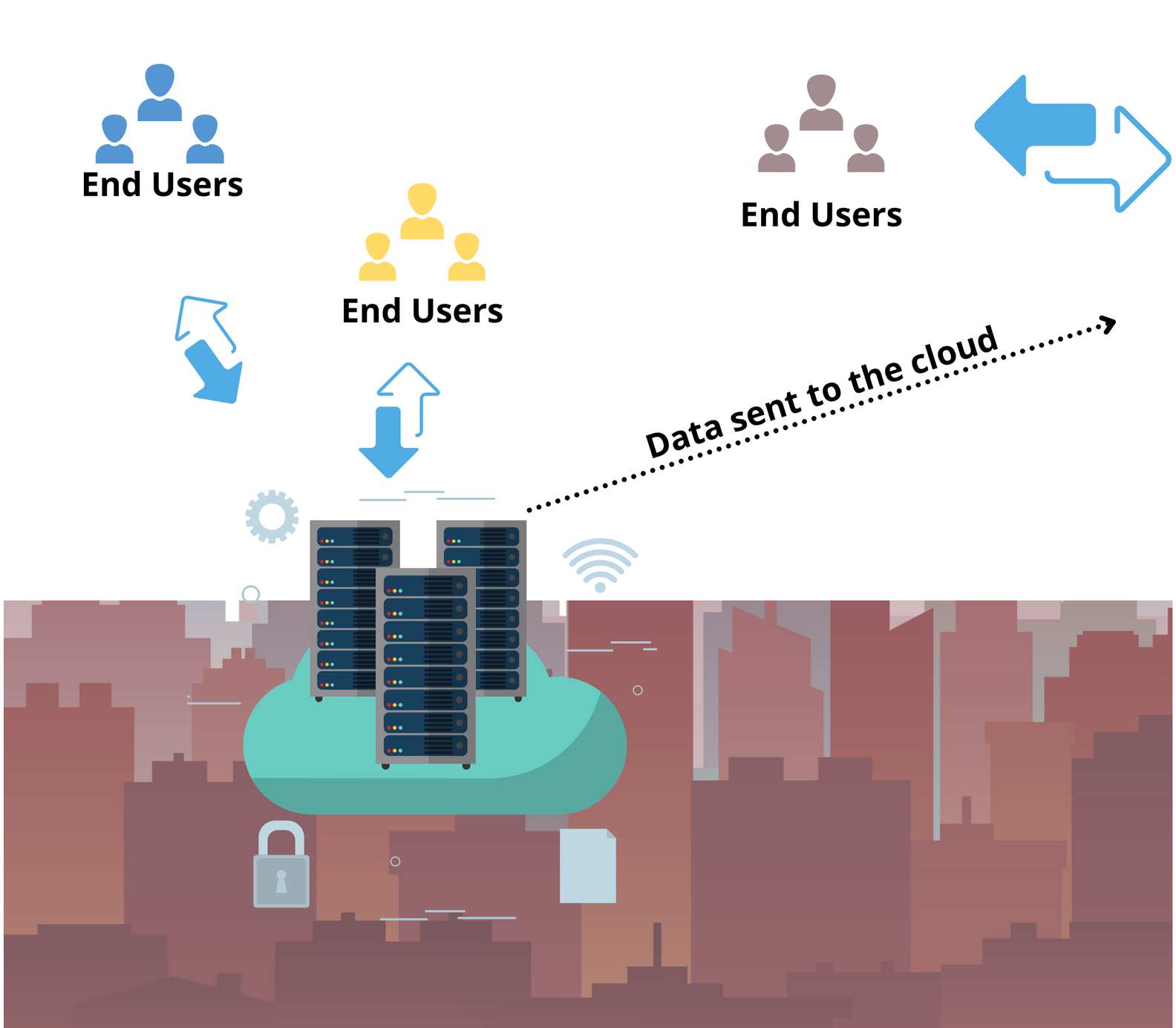
Data and application near the consumer

Examples

Most companies today have some proportion of their application and data in hybrid mode

Vendors

AWS, GCP, SAP, IBM, Microsoft, Alibaba



- Data moved from on premise to cloud on a regular basis
- Application and data originates in the cloud and stays there
- Data is shared across multiple clouds
- Data is sent to on premise

Prod/Dev/Test Environment

What it is

A different IT infrastructure for different purposes

Production - This is the live environment in which applications are hosted and end users access these applications

Test - This environment simulates a small subset of the actual production environment.

Development - For development

Application

- Only deploy ready applications
- Reduce downtime
- Improve user satisfaction
- Security
- Reliability
- Performance

Decoupled development, testing and production

Examples

Most applications are developed in the development environment , tested in the test environment and then moved to production

Vendors

The hardware software is similar in all environments.

Developers/Testers

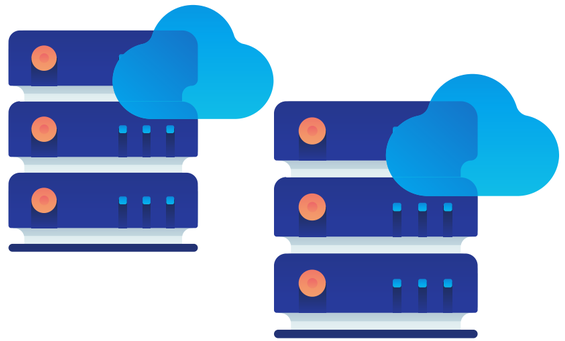
Power Users

End Users

End Users

End Users

Dev



Test

Prod

Virtualization

What it is

Software abstraction of hardware components

Simulated hardware functions on top of actual hardware

Application

- Run multiple O/s on the same hardware and a single copy of O/S
- Easy of deployment, maintenance, security
- Reduce hardware dependence for critical functions

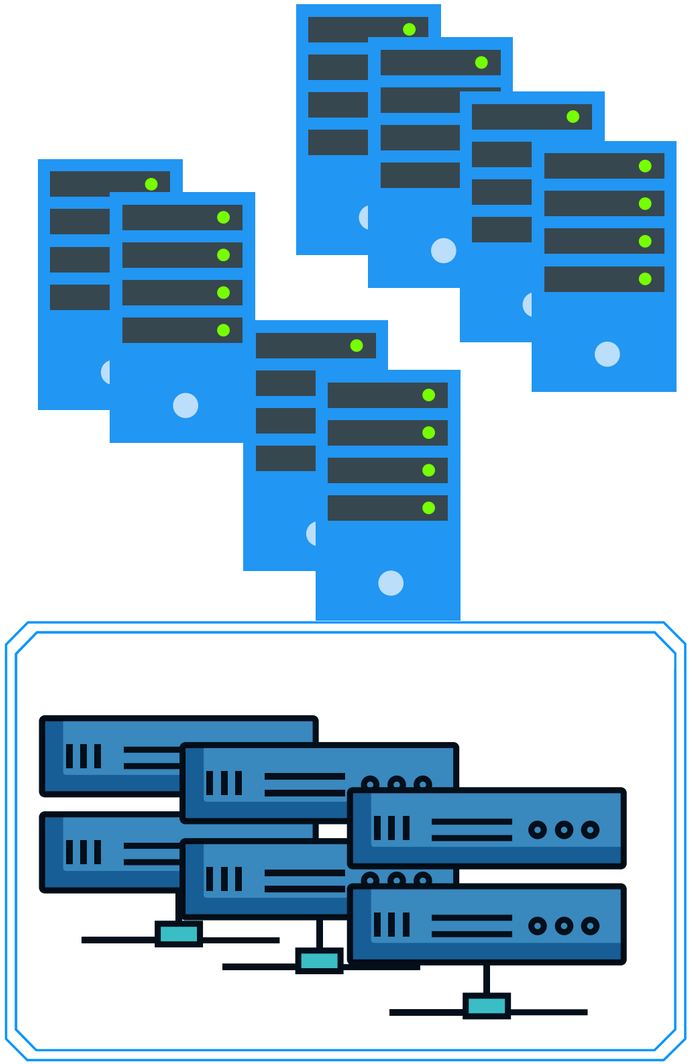
Examples

- O/S Virtualization
- Desktop Virtualization
- Network Virtualization

Vendors

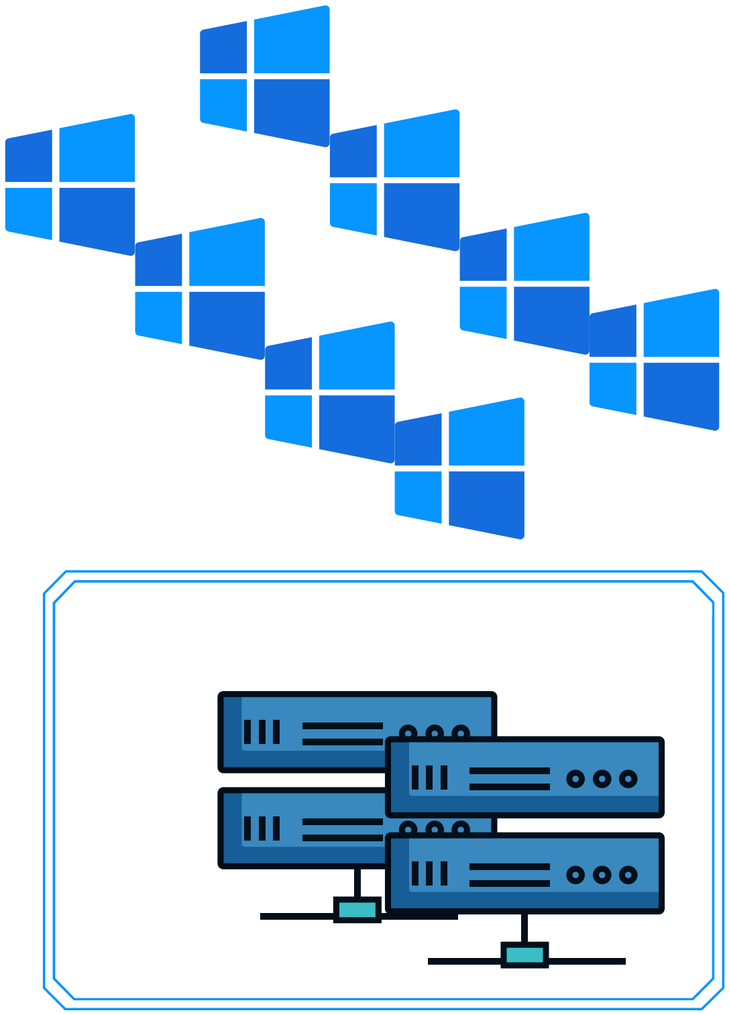
VMWare, Oracle Virtual box, Microsoft Hyper-V and many more

Different O/S Environments



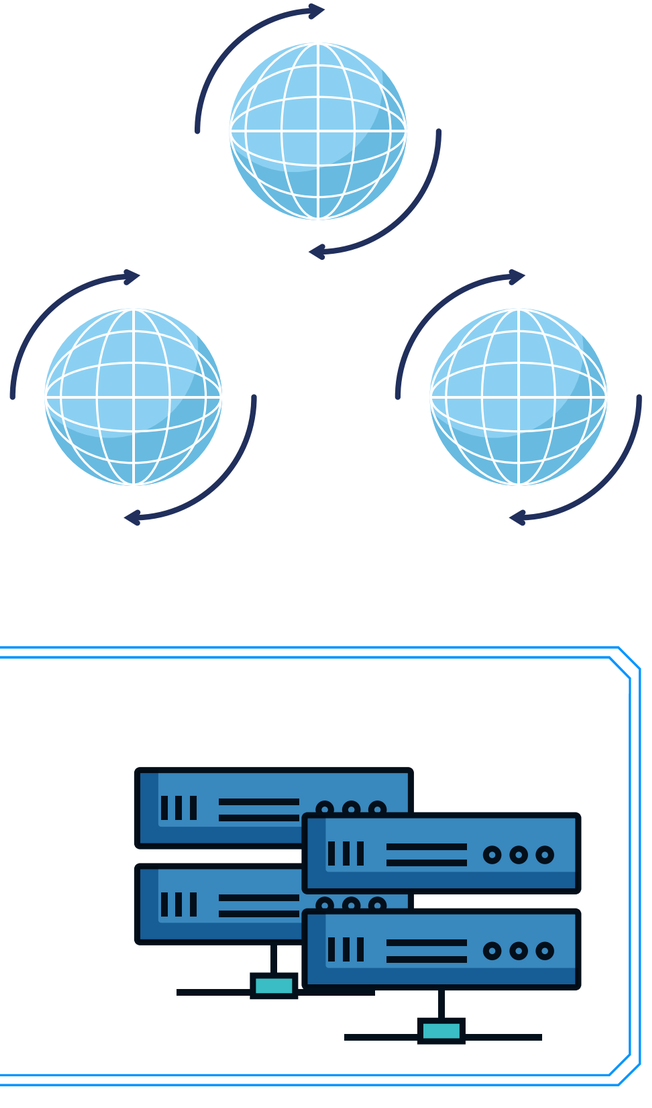
Hardware Virtualization

Different Desktop Environments



O/S Virtualization

Different Ip's



Network Virtualization

Use cases

What it is

A method to capture a scenario and possible requirements

It is a document, can be written on a napkin of sophisticated software's

Application

- To identify and isolate major requirements, actors
- To capture user interactions, data, user flow, alternatives, acceptance criteria
- Test scenarios
- Error handling

Uses the UML diagram unified modelling language

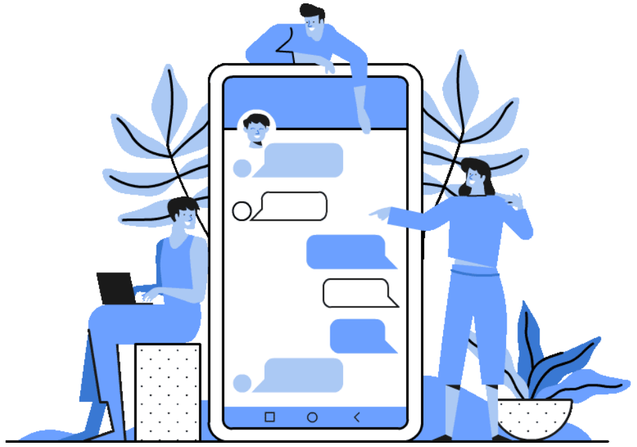
Examples

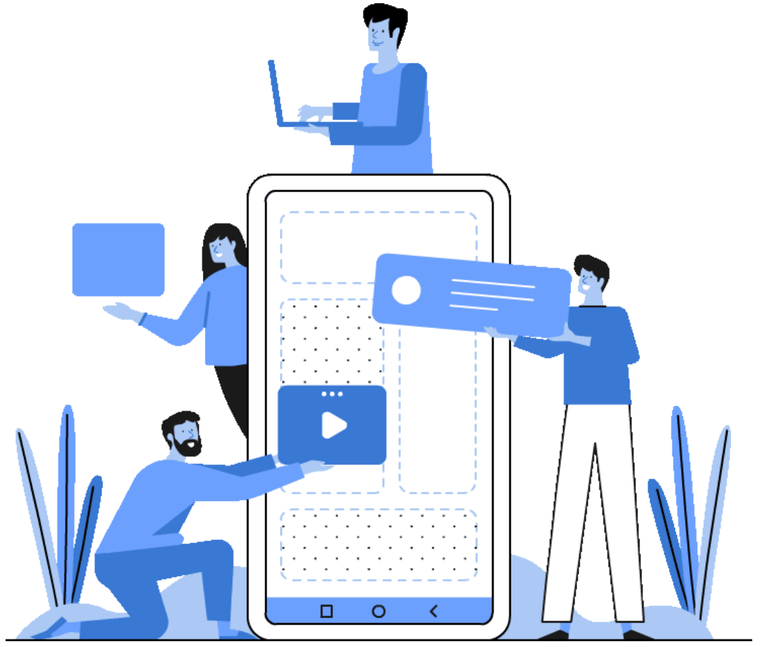
IT requirements

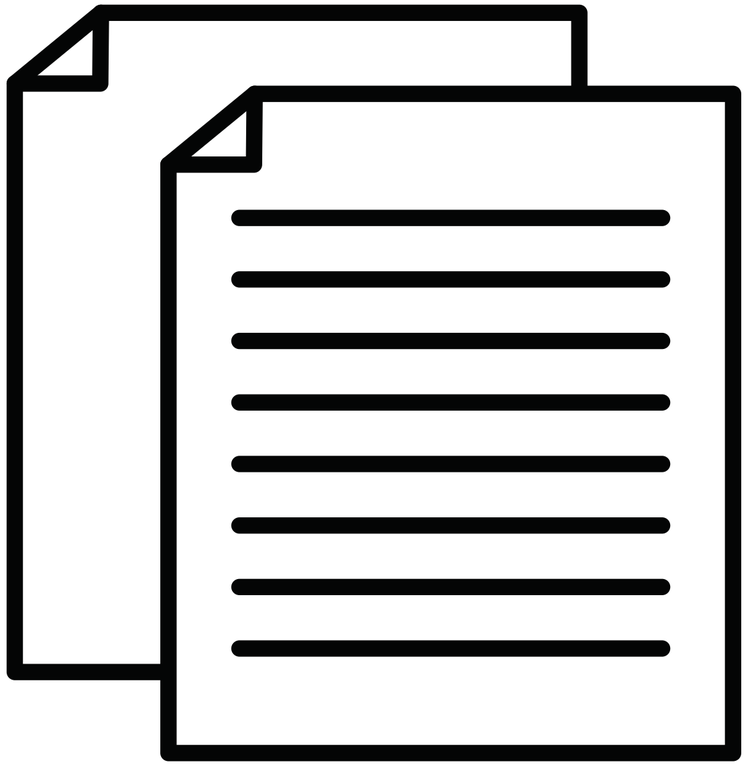
- User registration
- Payment process
- Top 5 Sales

Non IT requiremnts

- New employee onboarding
- DL for all different domains
- New customer acquisition
- Crypto currency options







https://www.usability.gov/how-to-and-tools/methods/use-cases.html

Cloud architecture

What it is

A method a store, manipulate, transform, manage and retrieve data

Deployment of Applications on the cloud

Cloud - Private, Public, Multi

Application

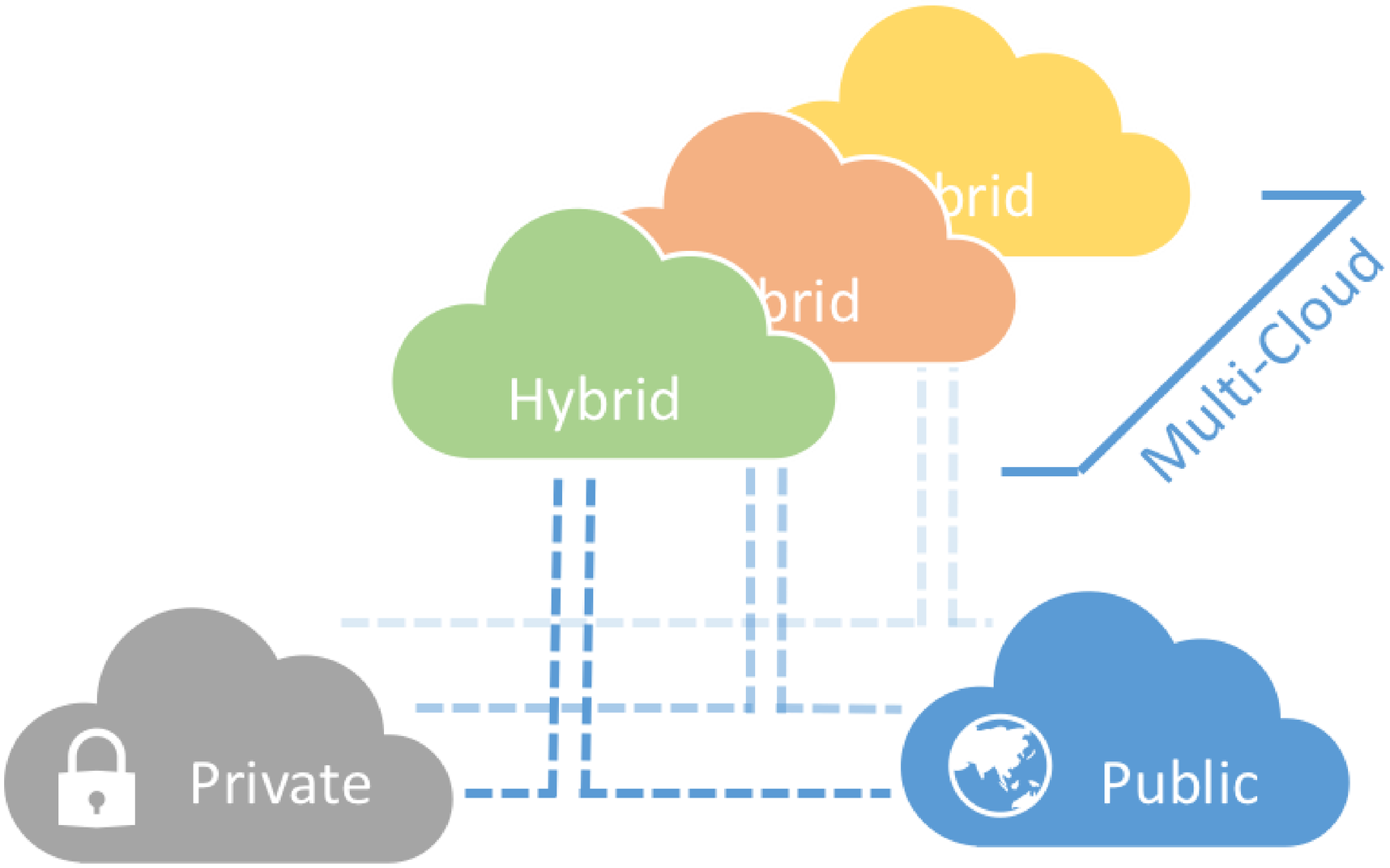
- Lower over all costs
- No installation, upgrade and maintenance
- Ad hoc scaling capability
- Stronger Security
- Connectivity

Examples

Most companies today have some proportion of their application and data in hybrid mode

Vendors

AWS, GCP, SAP, IBM, Microsoft, Alibaba


brid
brid
Hybrid
Multi-Cloud
Private
Public

Universe

What it is

A semantic layer where data is collated and curated for the reporting tools including multi dimensional models.

A term specific to some tools like SAP

Application

- Single layer of access for the reporting tools
- Heavy logic can be pushed to this layer
- Security can built from this layer
- Abstraction of technical details from the business

Examples

Tools like SAP have a universe concept

Vendors

SAP and others

Normalization

What it is

A method of database design to store information.

A set of governing rules to manage data

Application

- Reduce Redundancy
- Better manage insert, updates and deletes
- Improve data relevance
- Improve maintainability

Examples

Most OLTP and transaction systems are normalized

Vendors

It can be implemented in any RDBMS

Denormalization

What it is

A method to build meaningful information for reporting, monitoring and decision making

Required complex set of tools and, techniques

Application

- Single access to usable data
- Maintain historic qualified data
- Enable the business in building descriptive, diagnostic and predictive reports
- Reduce burden on conventional source systems for reporting purposes

Examples

Most companies today have Built data lakes.

Vendors

It can be implemented in any RDBMS

Primary key

What it is

A column in the table with unique values for each record.

Application

- Used to identify each record uniquely
- Maintain data integrity

Examples

Every table has at least one primary key

Vendors

can be implemented in any database

Foreign Key

What it is

A column in the table with unique values for each record.

If table A and table B need to be related then one table will store the primary key of the other table which can be used later to establish joins

Application

- It is used to make joins with other related tables
- Used to identify each record uniquely
- Maintain data integrity

Examples

All database designs involves a foreign key concept

Vendors

can be implemented in any database

ETL/ELT

What it is

A method and corresponding tools to extract, transform and load data

Required complex set of tools and, techniques

Application

- Useful in extracting data from multiple sources of different varieties and at various frequencies
- Can implement most transformation on data
- Can be used to load data to various targets at varies frequencies

Examples

Most companies today use these tools

Vendors

SAP, IBM, Snowflake, Oracle, Talend, Abinitio, informatica, AWS, GCP and many others

IoT data

What it is

Data generated by physical devices enabled for intranet and internet communication

Generally high velocity low volume data

Application

- Monitoring and controlling connected physical devices remotely
- Can be implemented on any physical entity

Examples

- Digital Twins
- City services management
- Self driving cars

Vendors

There are many Iot chip manufacturers from the popular arduino to now even intel. Most cloud providers now can handle IoT data

2

Datatically

What it is

A thinking methodology based on data

A new way of looking and making sense of things

Application

- Discovering new insights
- Understanding your own business, products, customers and more better
- Ensure business continuity
- Profitability
- Innovation
- Risk Management
- Proactive planning
- Add value to discussions
- Make premium Salaries

Examples

Thinking in terms of data is a basic skill all professional need

Data ware house

What it is

An Application on top of RDBMS to collate data from different parts of the organization for reporting, analytics and predictions

Can be on premise or on the cloud

Application/Purpose

- Single access to usable data
- Maintain historic qualified data
- Enable the business in building descriptive, diagnostic and predictive reports
- Reduce burden on conventional source systems for reporting purposes

Examples

Most companies today have a data ware house. It is also called a business warehouse.

Vendors

SAP, IBM, Microsoft, AWS, GCP, and many others

Data Mart

What it is

A subset of the data warehouse. Designed for specific lines of business or domains.

Can be on premise or on the cloud

Application

- Domain specific access to usable data
- Maintain historic qualified data
- Enable the business in building descriptive, diagnostic and predictive reports
- Reduce burden on conventional source systems for reporting purposes
- Reduce burden on the Data warehouse

Examples

Most companies today have multiple data marts to service the reporting requirements of different domains

Vendors

SAP, IBM, Microsoft, AWS, GCP, and many others

Data lake

What it is

A repository to store all formats of data close to its original format

Deep historical cold data

Application

- Build models on historic data
- Build All layers of reporting artifacts on historical data with current data
- Cheap storage for cold data

Examples

Most companies today have Built data lakes.

Vendors

AWS, GCP, SAP, IBM, Microsoft, Alibaba and many others

Hot, warm and cold data

What it is

A method to categorize data based on its usage

The client decides which is hot, warm and cold based on their unique requirements and circumstances

Application

- Gives a criteria to keep the most used data in the most high performance systems
- An important concept in in-memory databases
- Achieve cost reduction and optimum performance

Examples

data used every hour or day = Hot
data used every day or week = Warm
Most historical data = cold

Vendors

can be implemented in any database or cloud storage

Data Wrangling and Massaging

What it is

This describes man

Required complex set of tools and, techniques

Application

- Single access to usable data
- Maintain historic qualified data
- Enable the business in building descriptive, diagnostic and predictive reports
- Reduce burden on conventional source systems for reporting purposes

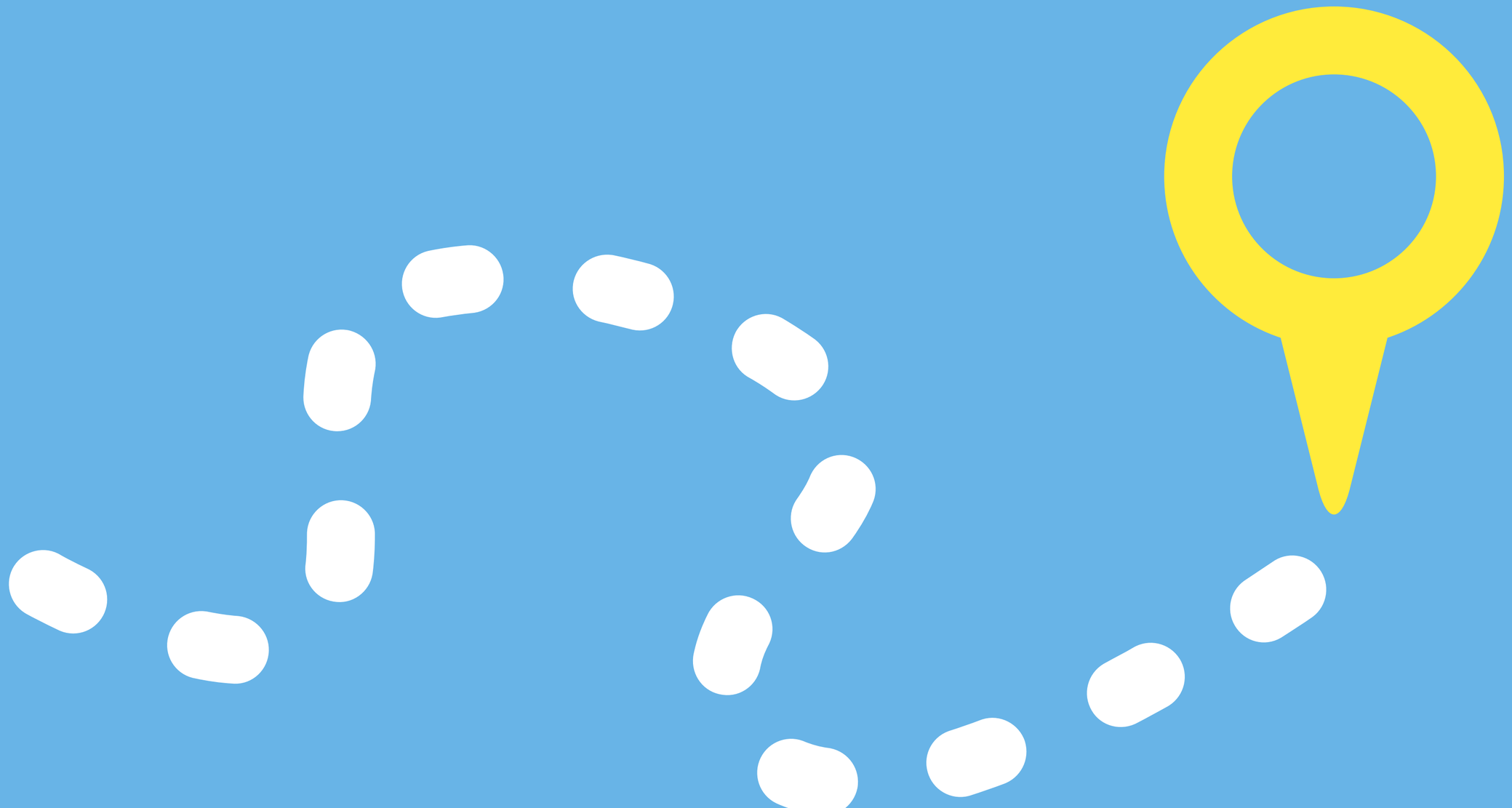
Examples

Most companies today have Built data lakes.

Vendors

AWS, GCP, SAP, IBM, Microsoft, Alibaba and many others

DATA JOURNEY IN AN ENTERPRISE



DATA JOURNEY

Documentation/Meta data Management

Generation → Extraction → Transform → Clean → Split/Filter → Combine

Data Wrangling

Data Massaging

Data Curation (Curator specific)

DATA JOURNEY

Documentation/Meta data Management

Sanitize

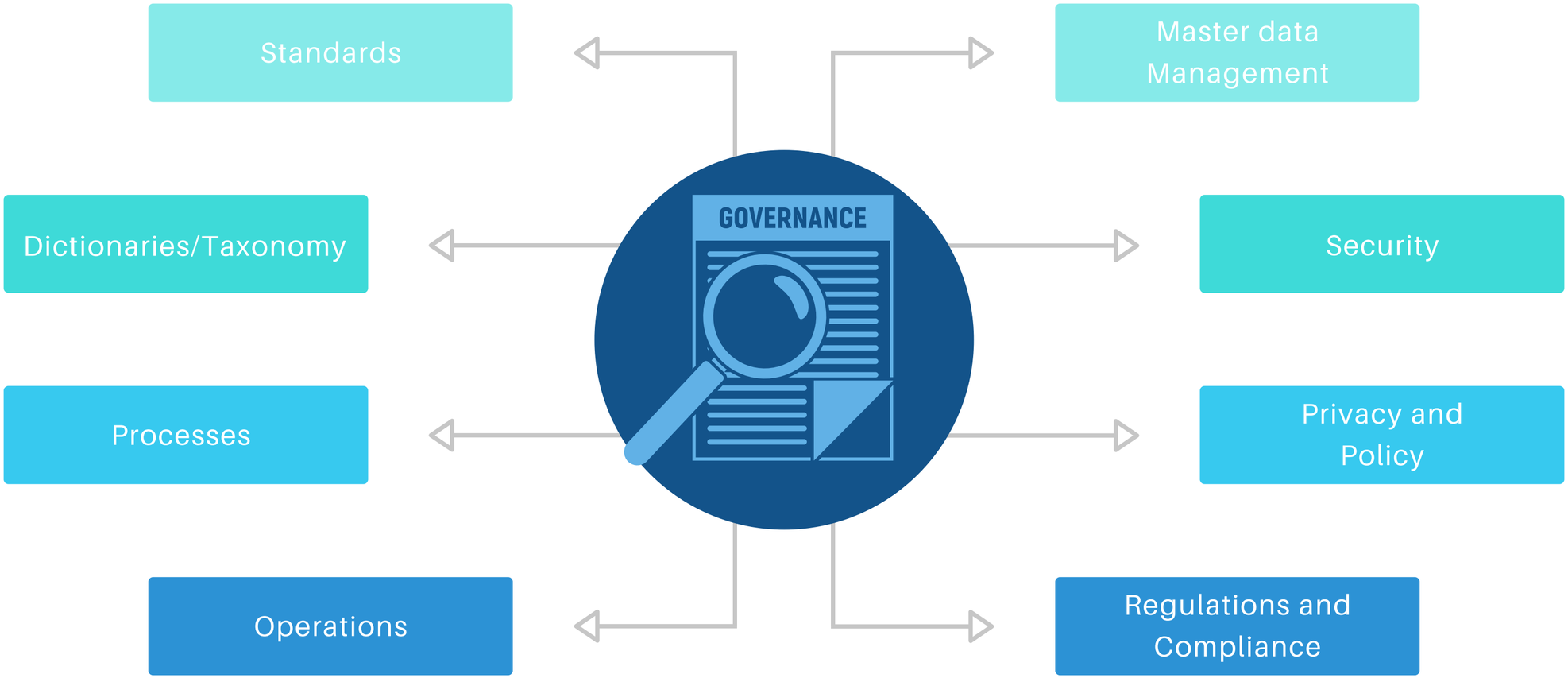
Data Massaging

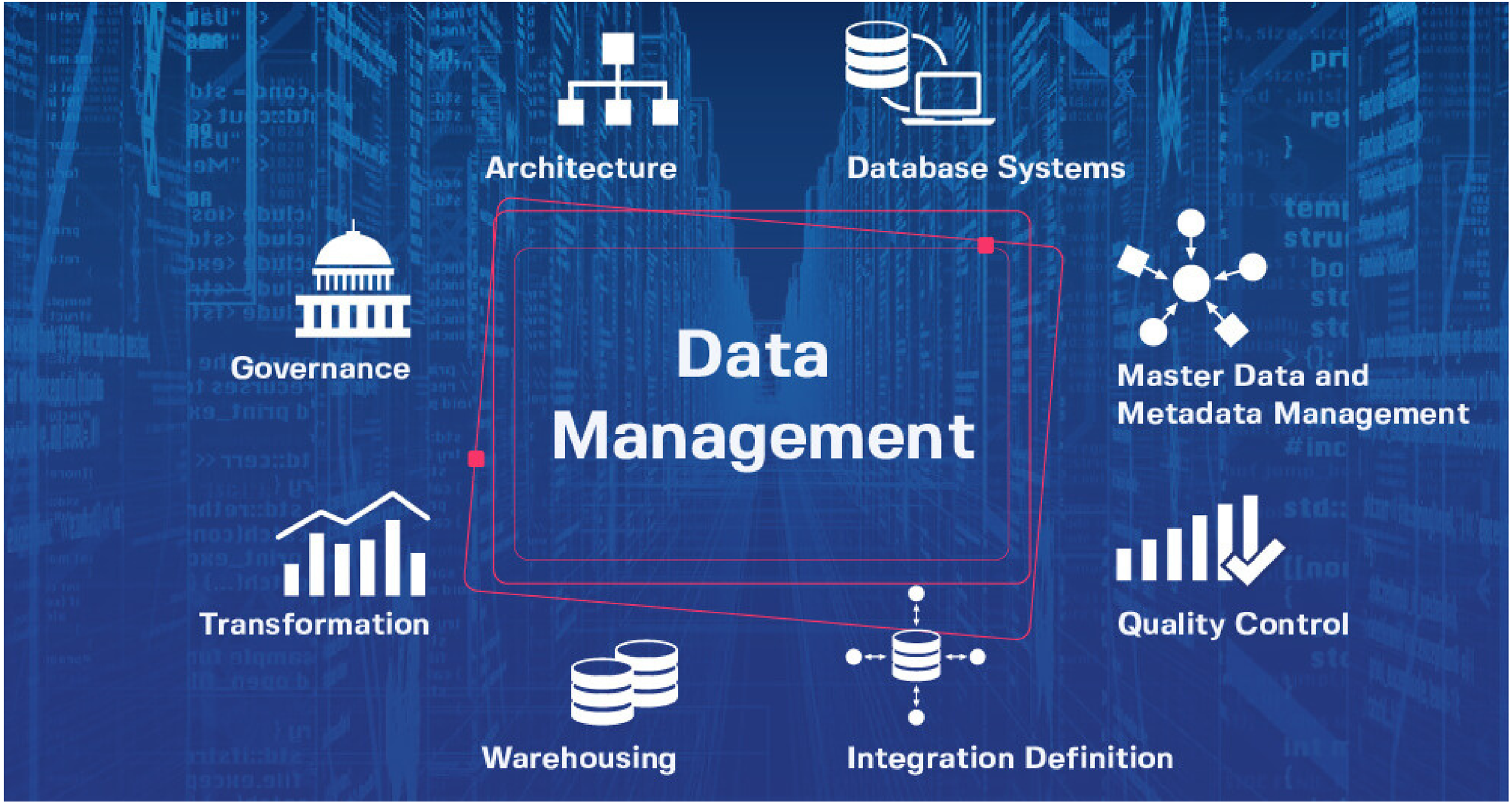
Data Ethics

Data Access/Security

Data Curation (Curator)

Data Governance





Source: https://www.divectors.com/blog/data-management-strategy/

Data Masking

What it is

Obfuscation of sensitive data

Multiple methods simple and complex can be used

Application

- Securing row level information
- Saves complex individualized role based implementations

Examples

Actual value of the cell can be replaced by XXXXXXX or some standard value or anything else

Data Anonymization

What it is

Obfuscation of sensitive data

Multiple methods simple and complex can be used

Application

- Securing row level information
- Saves complex individualized role based implementations

Examples

Actual value of the cell can be replaced by some algorithm generated data