



# Virtualization Concepts

Basic Data Engineering

Taught by Pichaya Tandayya

18 June 2025

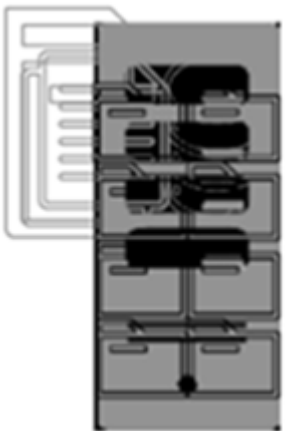
# OS Facts

- Generally, applications need to run inside an OS.
- Only one OS can run on a physical server at a time.

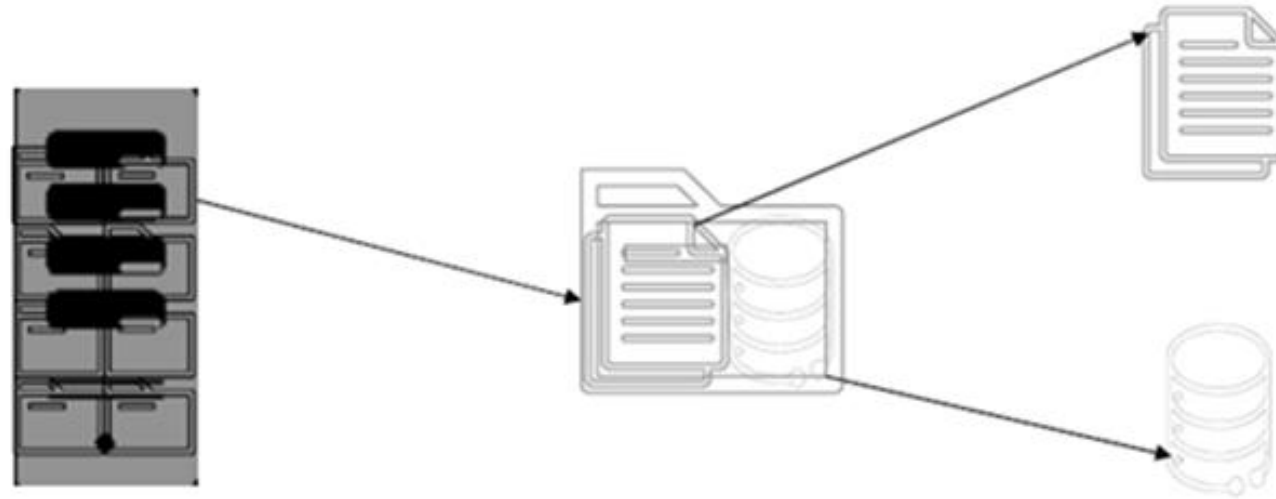
# What's Virtualization?



- A technology that **simulates hardware functionalities** and **creates multiple VMs** on a physical server.
- Virtualization **allows Virtual Machines (VMs)** that reside on the same physical server **to run independent OSs**.
- This way, **multiple OSs** can concurrently run on the same physical server.

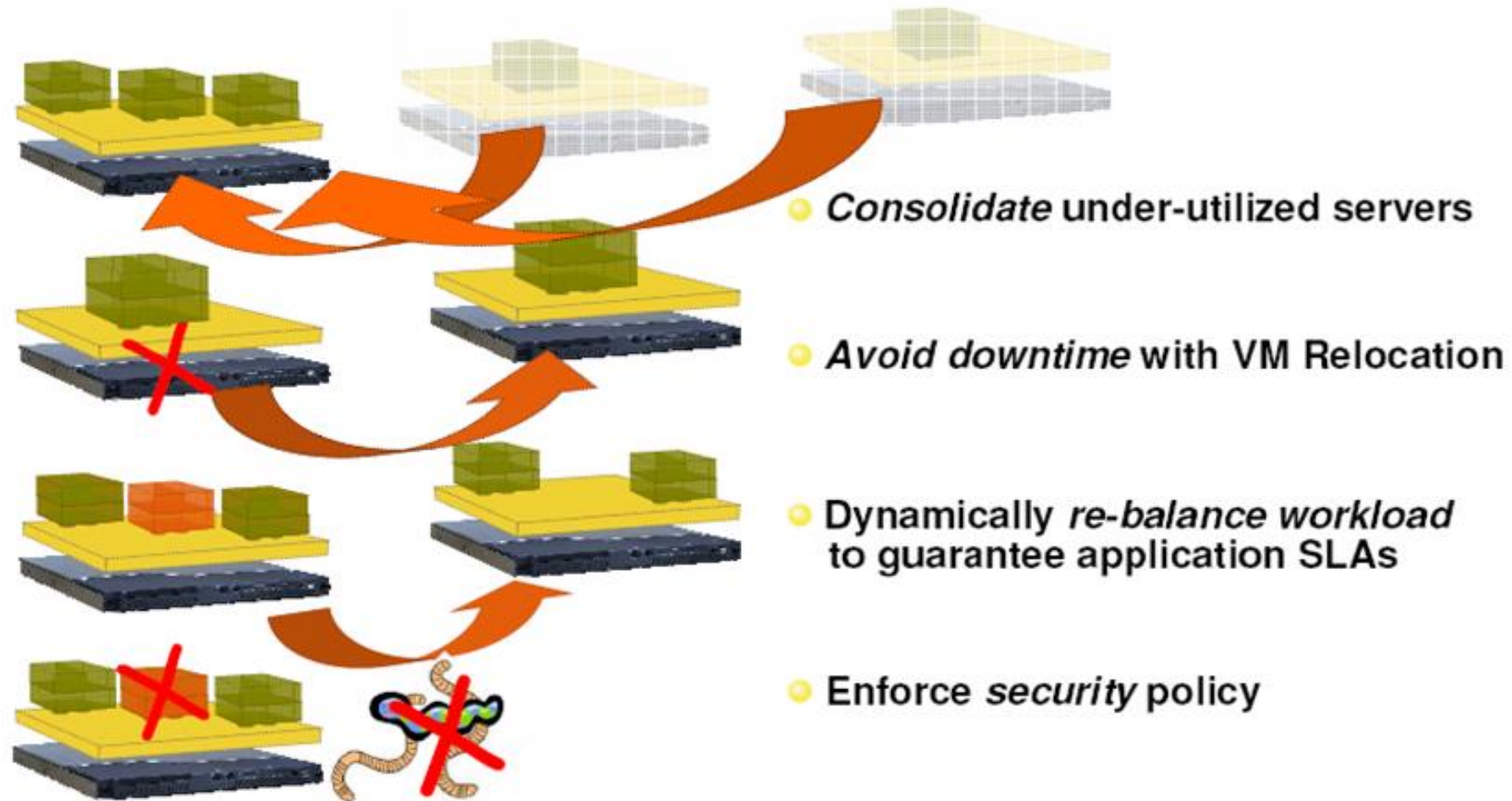


# What's Virtualization?



The essence of virtualization is to separate software from hardware by converting "physical" devices into "logical" folders or files.

# Virtualization Benefits



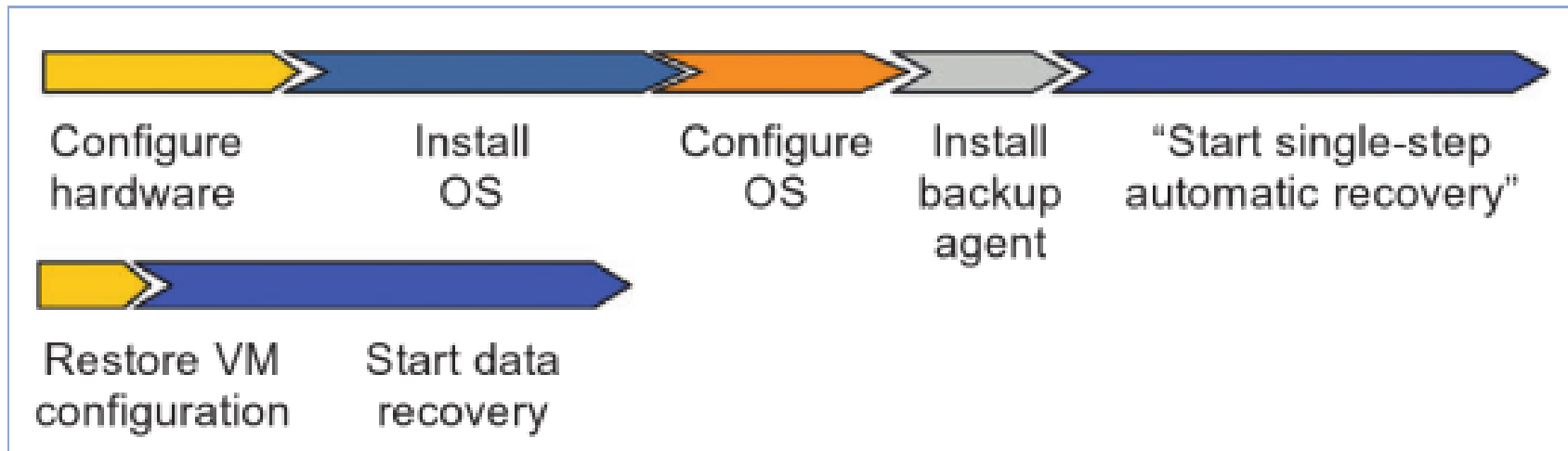
# Virtualization Benefits

- Separating the OS from the hardware
  - Users no longer forced to upgrade OS to run on latest hardware
- Device support is part of the platform
  - Write one device driver rather than N
  - Better for system reliability/availability
  - Faster to get new hardware deployed
- Enables “Virtual Appliances”
  - Applications encapsulated with their OS
  - Easy configuration and management

# Recovery Overhead

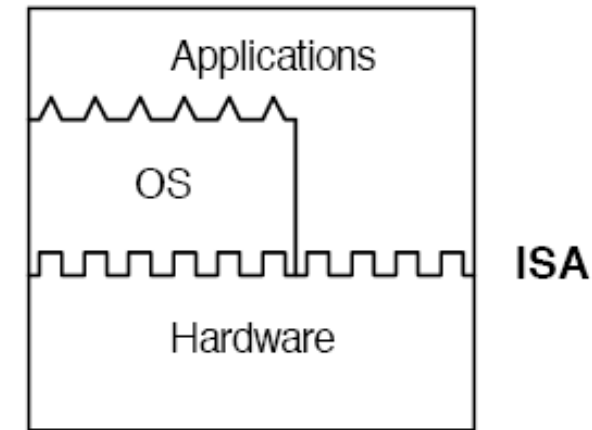
## ค่าโสหุ้ยในการกู้คืน

เปรียบเทียบการกู้คืนระบบใน Physical Machine กับ การใช้ live migration ของ Virtual Machine



# โมเดลฮาร์ดแวร์ดั้งเดิม

- Application ต่าง ๆ เข้าถึง ฮาร์ดแวร์ (memory, i/o) ผ่าน system calls ไปยังระบบปฏิบัติการโดย ชุดคำสั่งพิเศษ
- ข้อดี
  - สามารถแยกการออกแบบและพัฒนาระบบปฏิบัติการ และฮาร์ดแวร์ออกจากกันได้ โดยมีข้อตกลงตาม ISA
  - สามารถ upgrade ฮาร์ดแวร์และซอฟต์แวร์ได้ โดยไม่ต้องแจ้ง หรือไม่กระทบ Application programs
- ข้อเสีย
  - Application ที่คอมไพล์บน ISA รูปแบบหนึ่งจะ ไม่สามารถรันบน ISA อีกรูปแบบหนึ่งได้
  - Applications ที่คอมไพล์ สำหรับ Mac ใช้การเรียนรู้ system call แตกต่างจาก windows.
  - ISA ต้องสนับสนุน ซอฟต์แวร์รุ่นเก่า
    - เป็นข้อจำกัดที่ส่งผลต่อการพัฒนาความสามารถให้เพิ่มขึ้นไปจากเดิม
  - ซอฟต์แวร์พัฒนาแยกจากฮาร์ดแวร์ ซึ่งอาจไม่ optimized สำหรับฮาร์ดแวร์นั้น ๆ



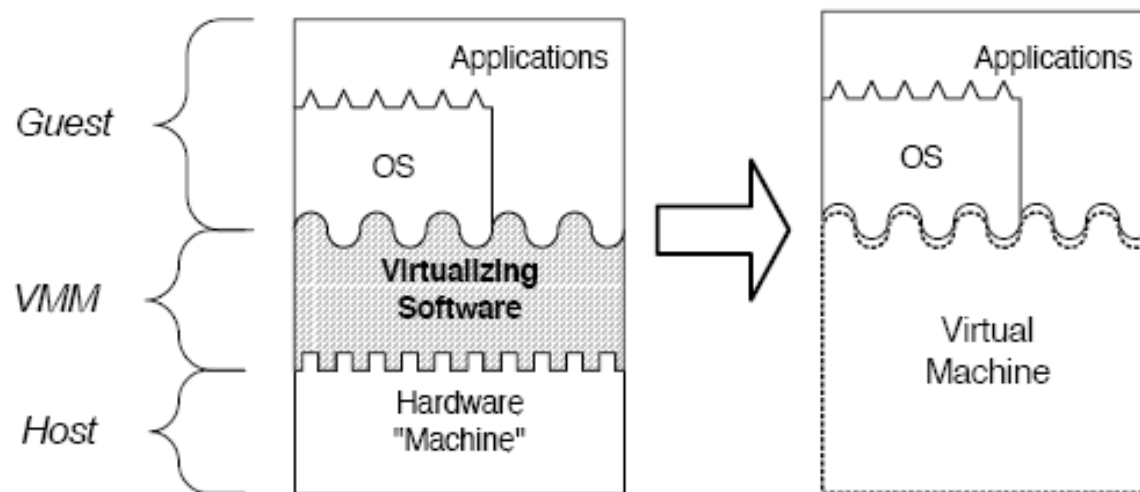
*ISA: Industry Standard Architecture*

Source: K. Hwang, G. Fox and J. Dongarra, Distributed and Cloud Computing

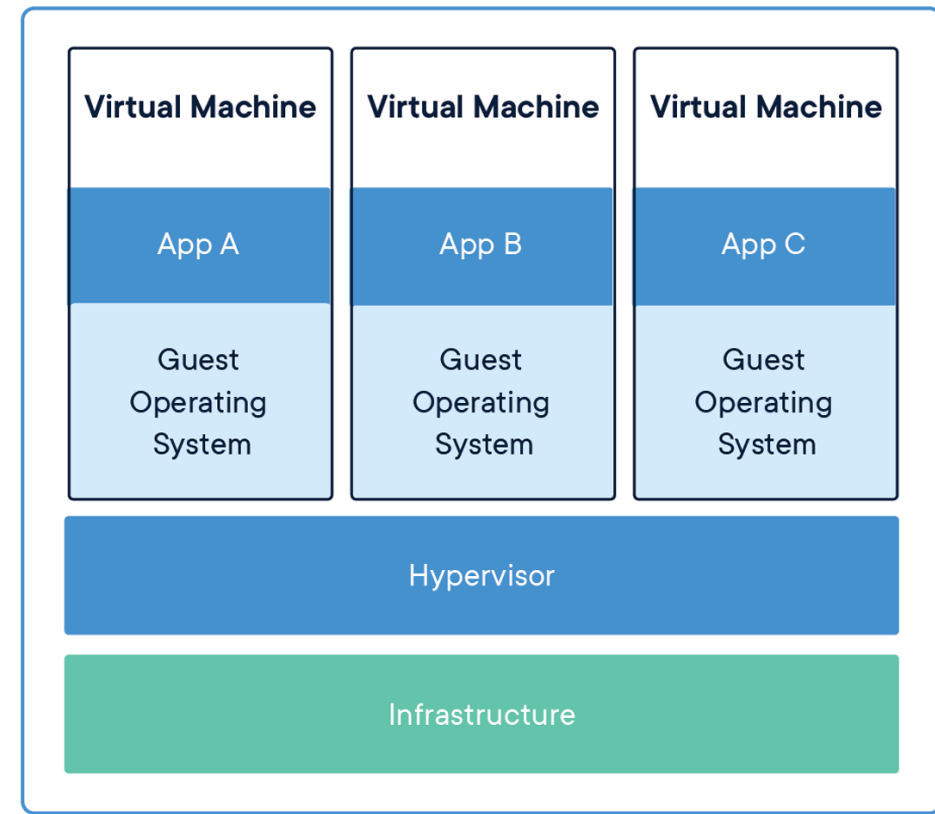
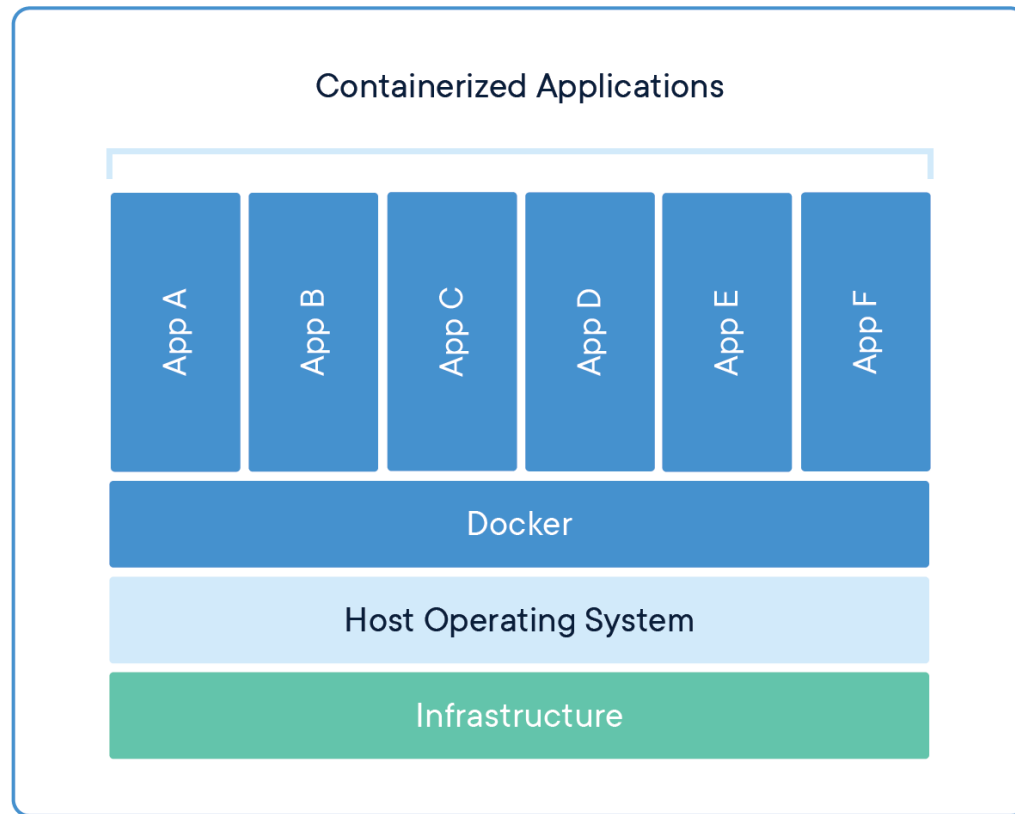


# พื้นฐาน Virtual Machine Basics

- ซอฟต์แวร์เสมือนอยู่ระหว่างเครื่องและซอฟต์แวร์งานประยุกต์แบบดั้งเดิม และทำให้ 'มองเห็น' และใช้งานได้ตาม ISA ที่ซอฟต์แวร์นั้นถูกคอมไพล์มา (executable file)
- กระบวนการ Virtualization เกี่ยวข้องกับ:
  - Mapping of virtual resources (registers และ memory) เข้ากับ ทรัพยากรฮาร์ดแวร์ของเครื่องจริง
  - ใช้ชุดคำสั่งของเครื่องจริง ทำกิจกรรมที่ระบุมาในชุดคำสั่งของเครื่องเสมือน

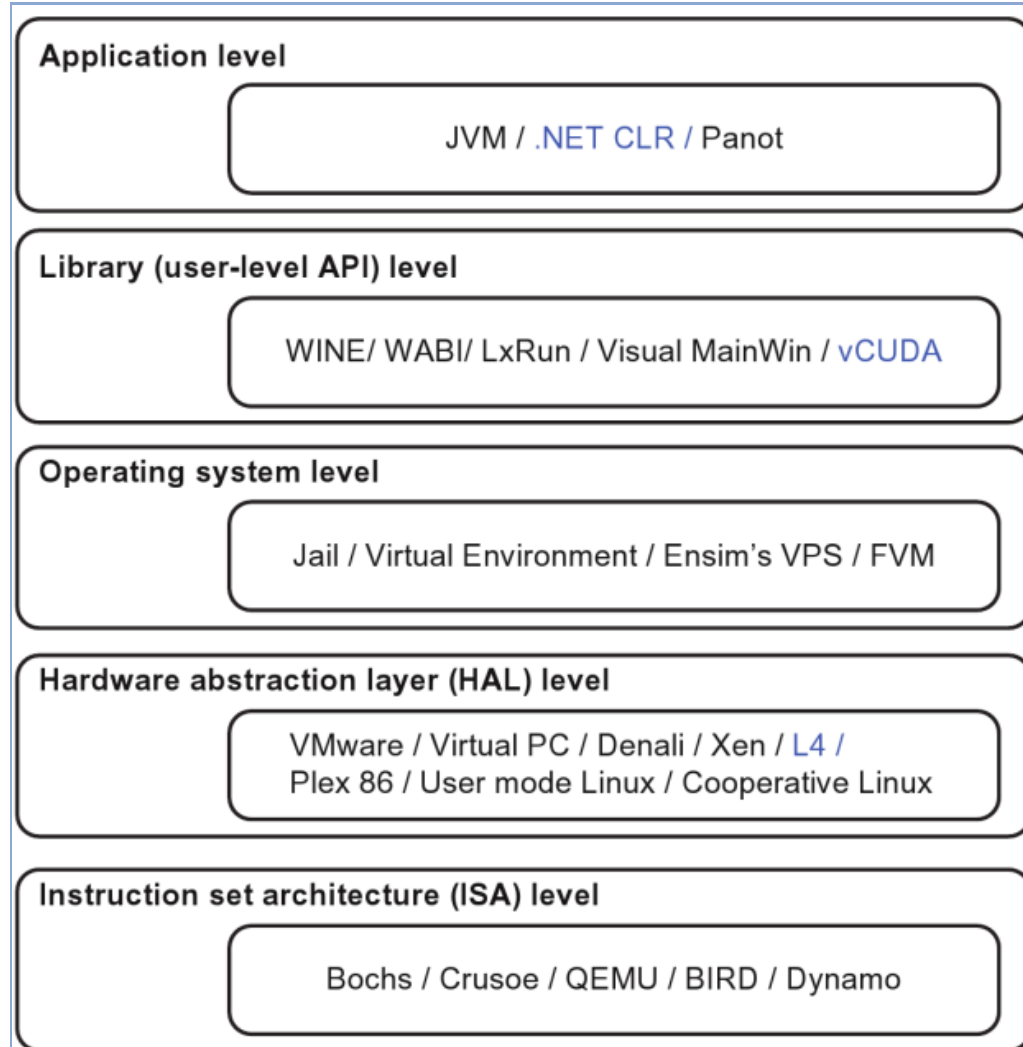


# Containers vs. Virtual Machines



Source: <https://www.docker.com/resources/what-container>

# 5 Abstraction Levels of Virtualization



Containers/Dockers

# Virtualization at Operating System (OS) Level

- เป็น abstraction layer ระหว่าง ระบบปฏิบัติการแบบดั้งเดิม และ user applications ซึ่งมักเรียกว่า containers
- virtualization นี้สร้าง isolated containers บน single physical server
- ทุก user application แשר kernel เดียวกัน ผ่าน guest OS ซึ่งอาจแตกต่างกันได้ในแต่ละพื้นที่ของผู้ใช้ (container) เช่น อาจมี Linux ได้หลาย distributions ภายใต้ kernel เดียวกัน
- containers ทำตัวเหมือนเซิร์ฟเวอร์/เครื่องจริง

# Virtualization at Operating System (OS) Level

- ข้อดี

- ใช้เวลาในการ startup/shutdown น้อย, ใช้ทรัพยากรของระบบน้อย มี scalability สูง
- ทุก OS-level VM บน physical machine เดียวกัน แชร์ OS kernel เดียวกัน

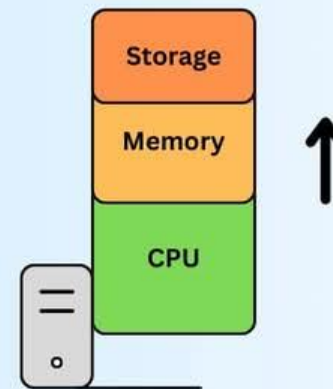
- ข้อบกพร่องและข้อจำกัด

- ทุก VM ใน operating system level ต้องมี guest OS ประเภทเดียวกัน
- Poor application flexibility and isolation

# Scalability

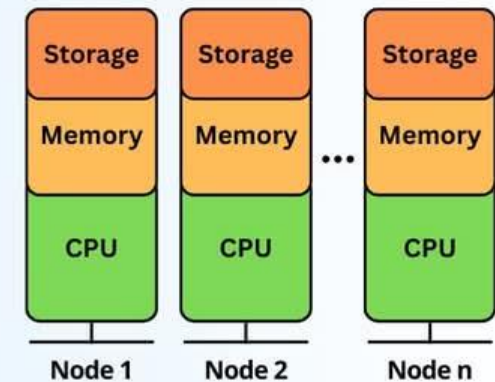
- Vertical scaling - scale up server
  - CPUs, Memory, etc.
- Horizontal scaling - clustered servers

การสเกลแนวตั้ง  
**Vertical scaling**



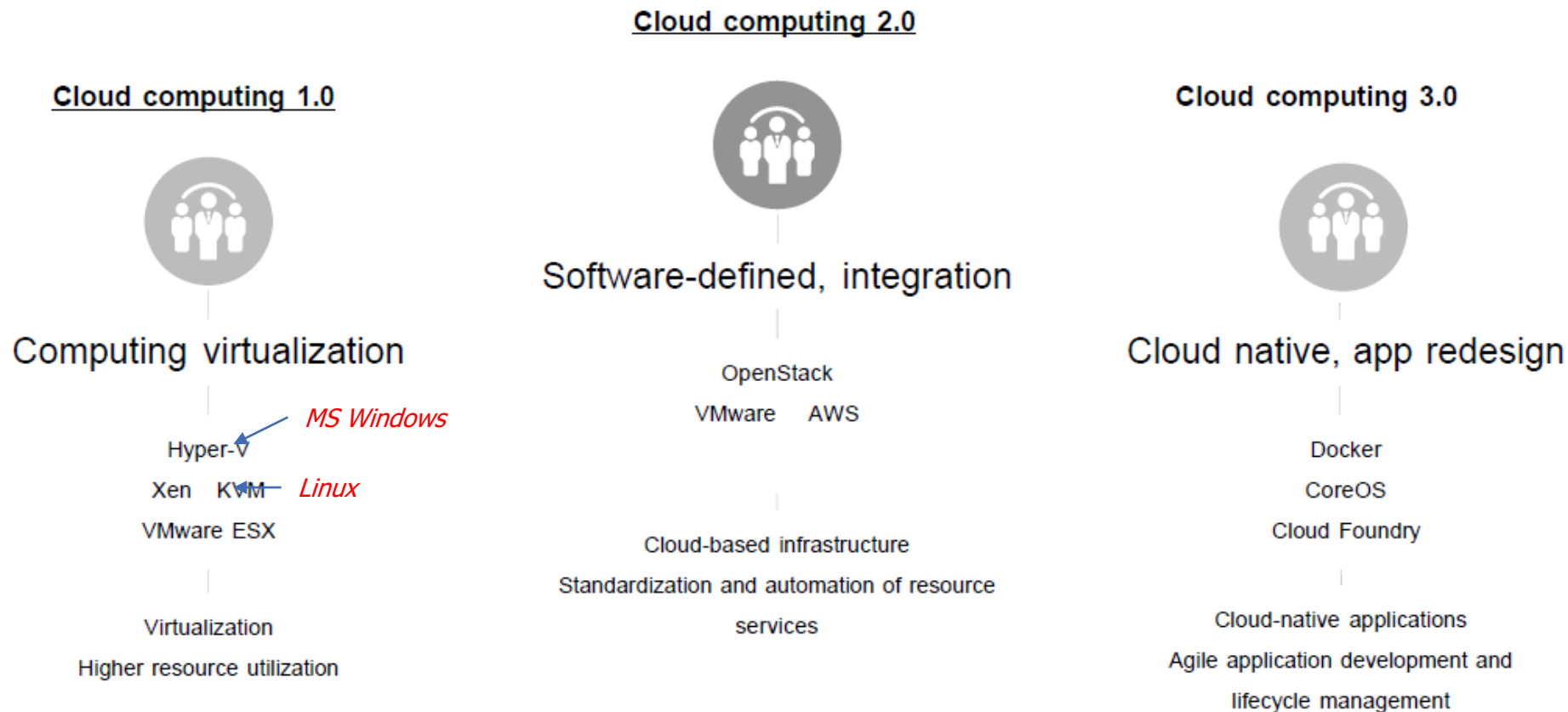
Single machine

การสเกลแนวนอน  
**Horizontal scaling**



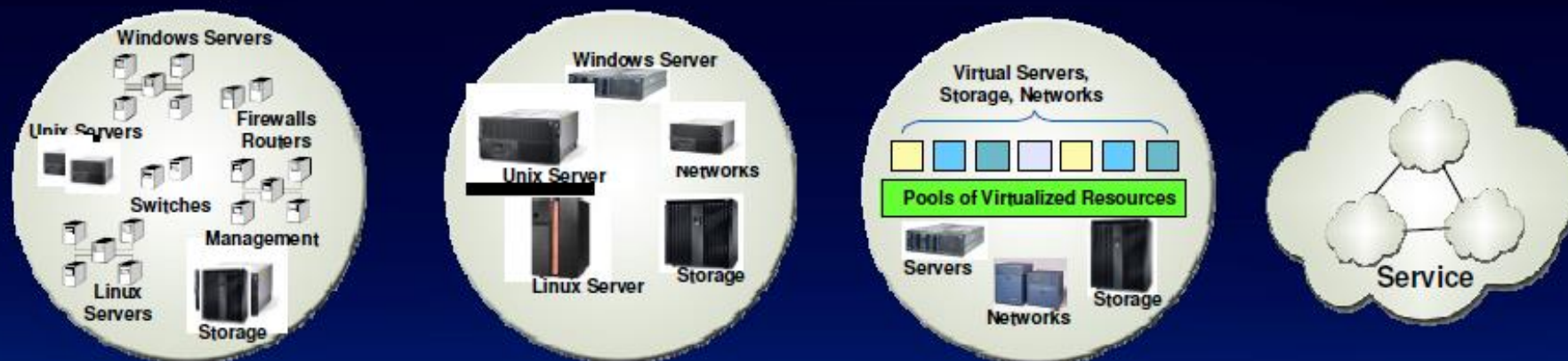
Multiple nodes

# Cloud Computing Evolution





# Technology Roadmap to Cloud



## Complex Infrastructure Sprawl

- IT assets & datacenters kept growing
- Desperate system tools
- Inconsistent processes
- Soaring IT & energy costs

## Physical Consolidation

- Consolidate IT assets & datacenters
- Standardize and centralize management
- Streamline processes with ITIL best practices
- Energy saving - Phase out inefficient HW

## Virtualization

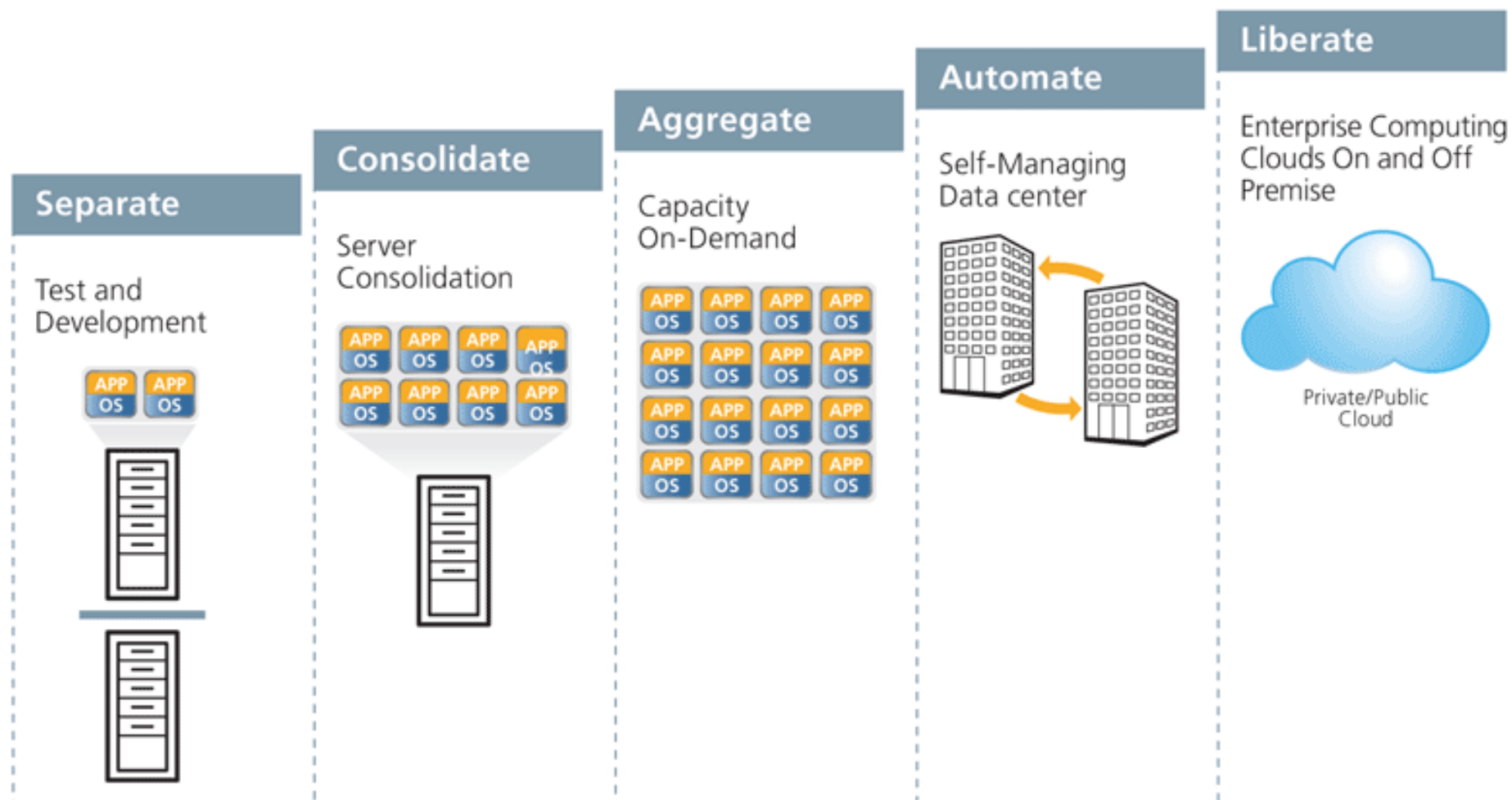
- Virtualize infrastructure - increased system utilization
- Unify virtual & physical mgmt
- Promote resource sharing across organization
- Energy saving – maximize effective use

## Cloud

- Service oriented architecture infrastructure
- Rapid provisioning of IT resources, massive scaling
- Dynamic service mgmt
- Energy saving via auto workload distribution



# Enterprise virtualization to cloud maturity model



Source: <http://westconsolutions.com/content/vendors/f5-networks/virtualization-and-cloud-computing-solutions>

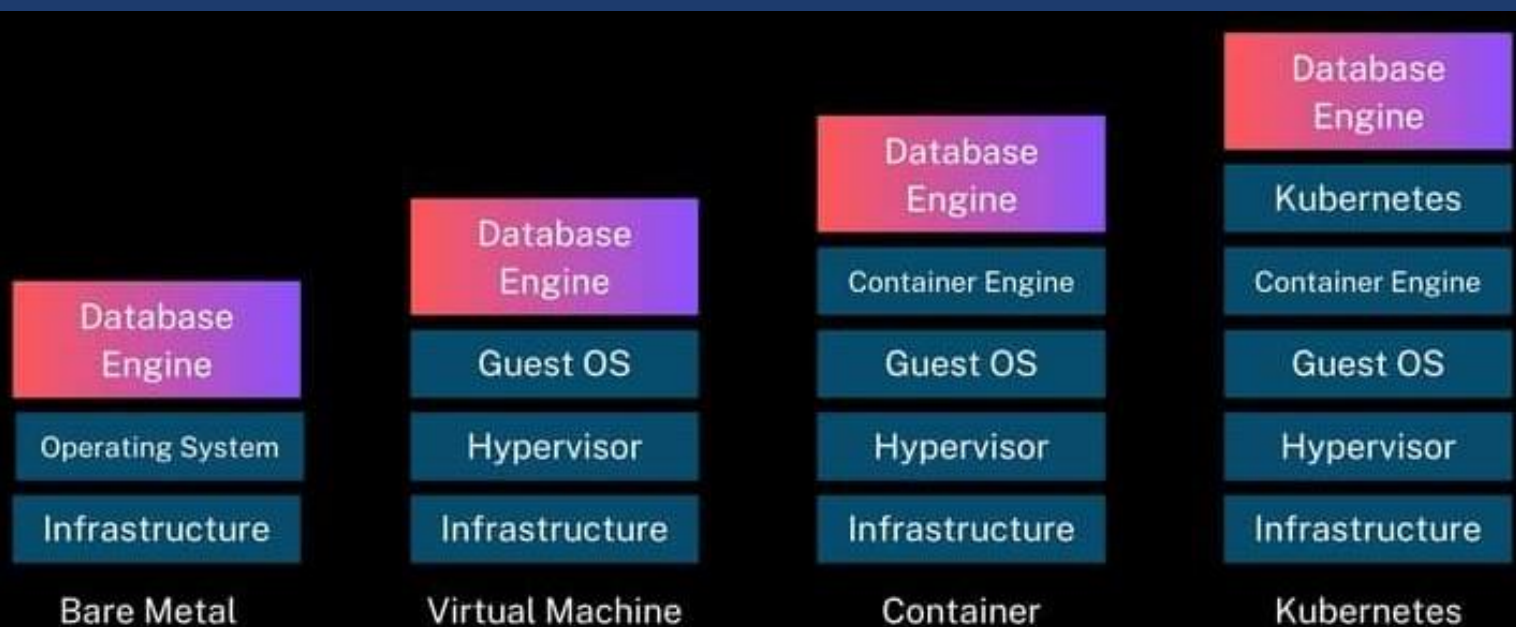
# Virtualization for Datacenter Automation to serve millions of clients, simultaneously

- Server Consolidation in Virtualized Datacenter
- Virtual Storage Provisioning and Deprovisioning
- Cloud Operating Systems for Virtual Datacenters
- Trust Management in virtualized Datacenters

# 5 คุณลักษณะของ Cloud Computing

- On-Demand Self-Service → • เพิ่มลดทรัพยากรได้ด้วยตนเอง
- Broad network access → • เข้าใช้งานได้จากทุกที่ ทุกอุปกรณ์ ทุกเวลา
- Resource pooling → • รวมทรัพยากรไว้ที่เดียวกัน แชร์กันได้ตามแต่ใครใช้มากใช้น้อย
- Rapid elasticity → • ยืดขยายได้ตามต้องการอย่างรวดเร็ว
- Measured service → • วัดการใช้งานได้ละเอียด เพื่อการจ่ายตามที่ใช้

# Running Database in a Container



Types of database running

Example Database Application on Kubernetes



yugabyteDB



# Cloud Store: Storage Options

#GCPSketchnote

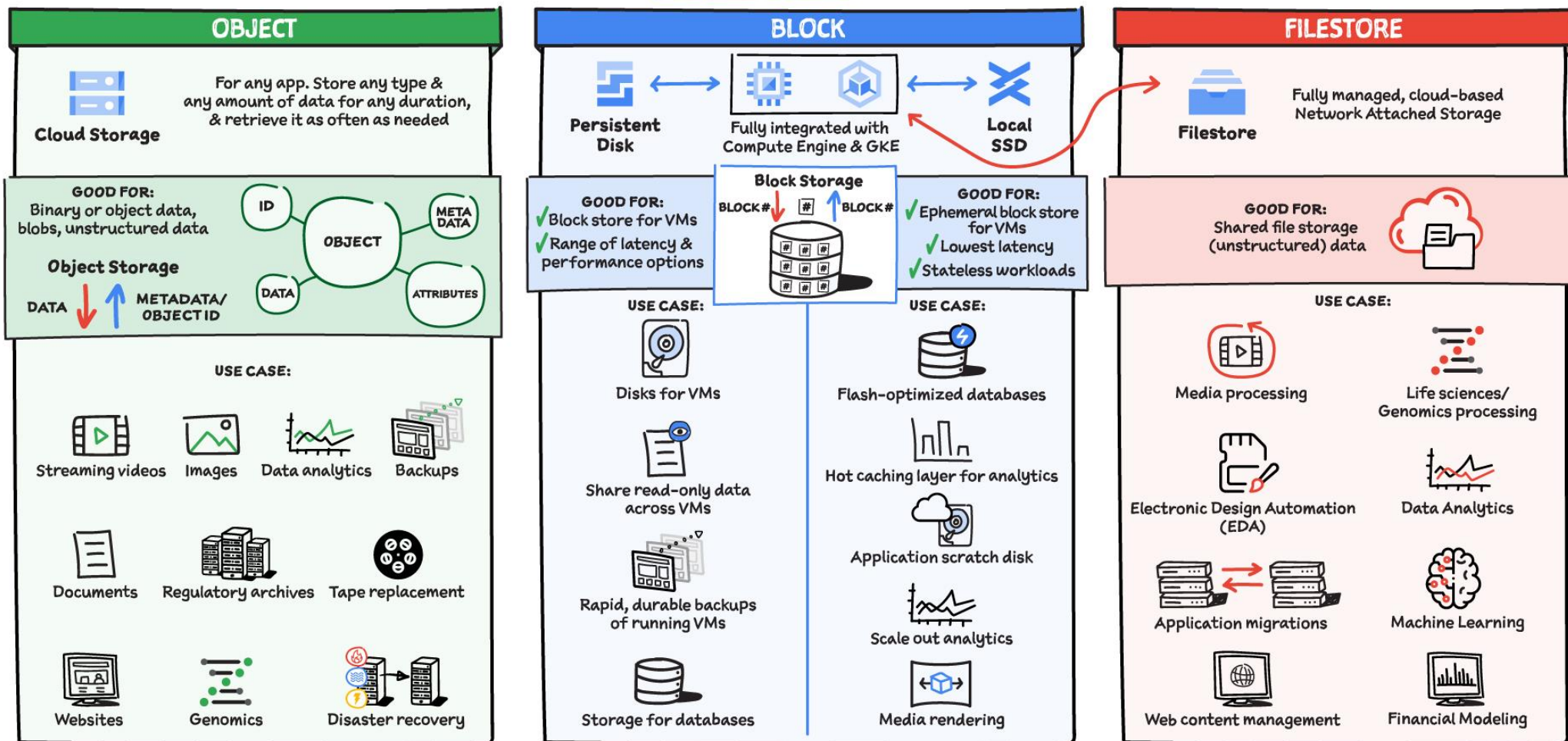
@PVERGADIA

THECLOUDGIRL.DEV

04.23.2021

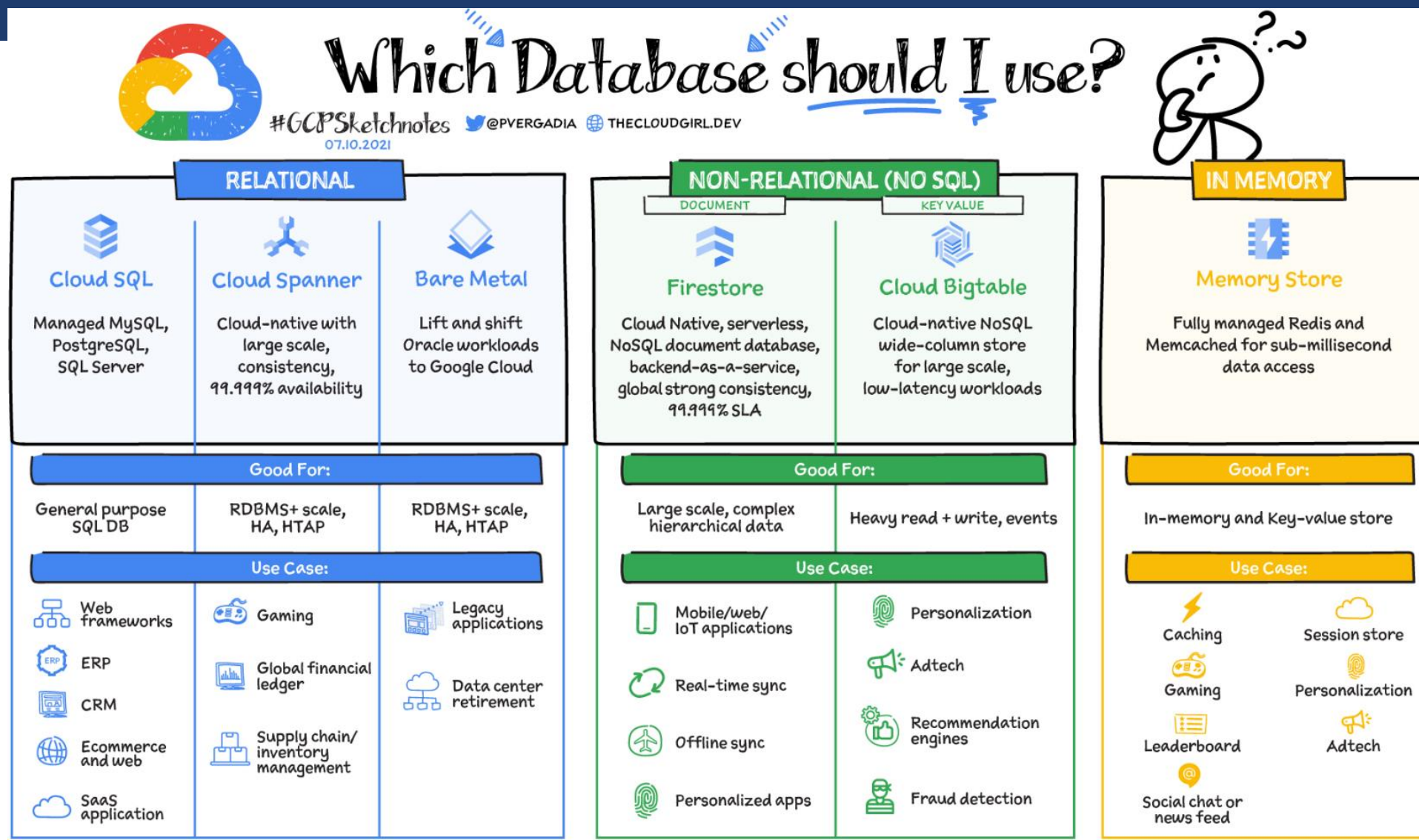


## Which Storage Should I Use?



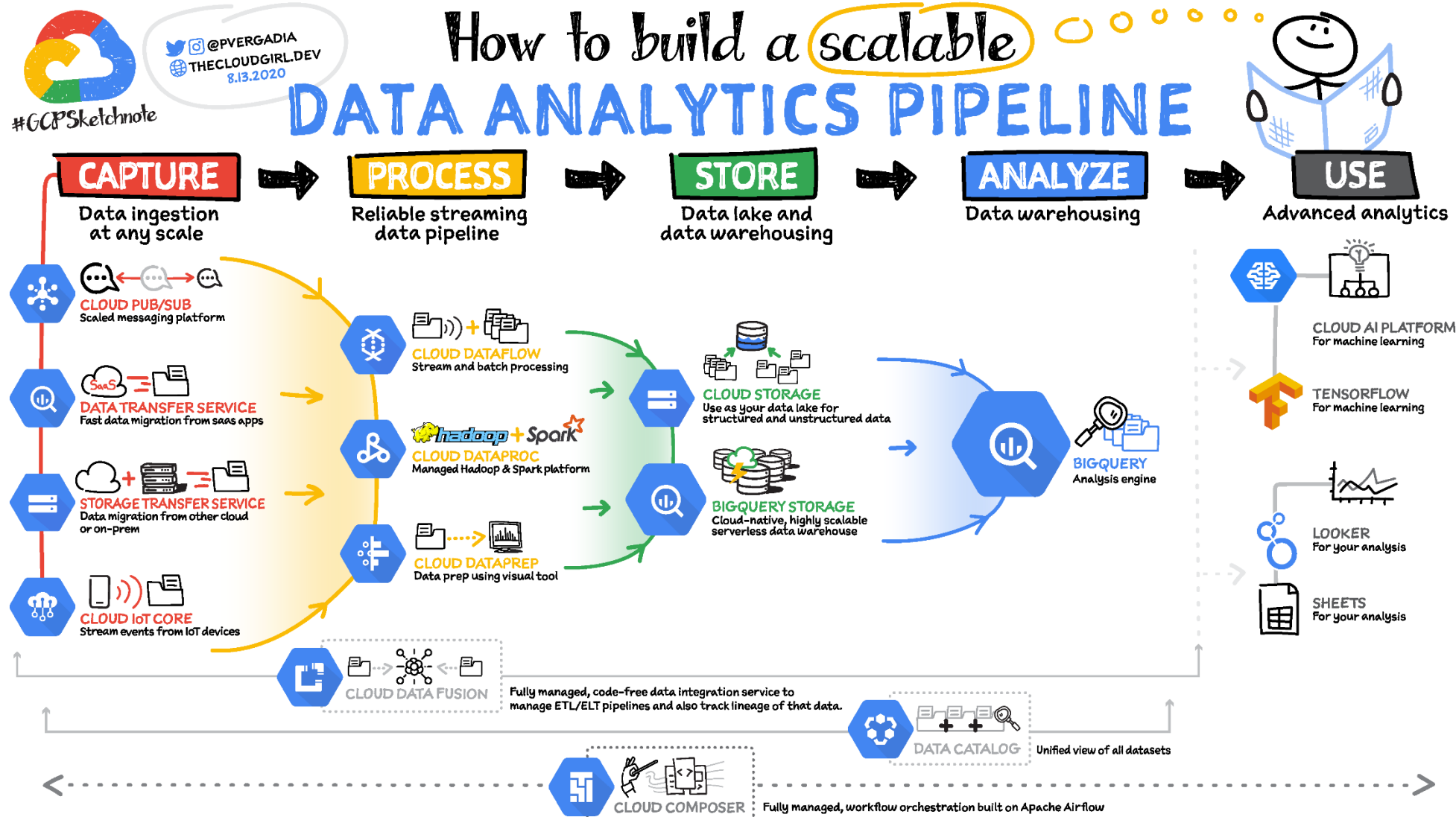


# Which Database should I use?



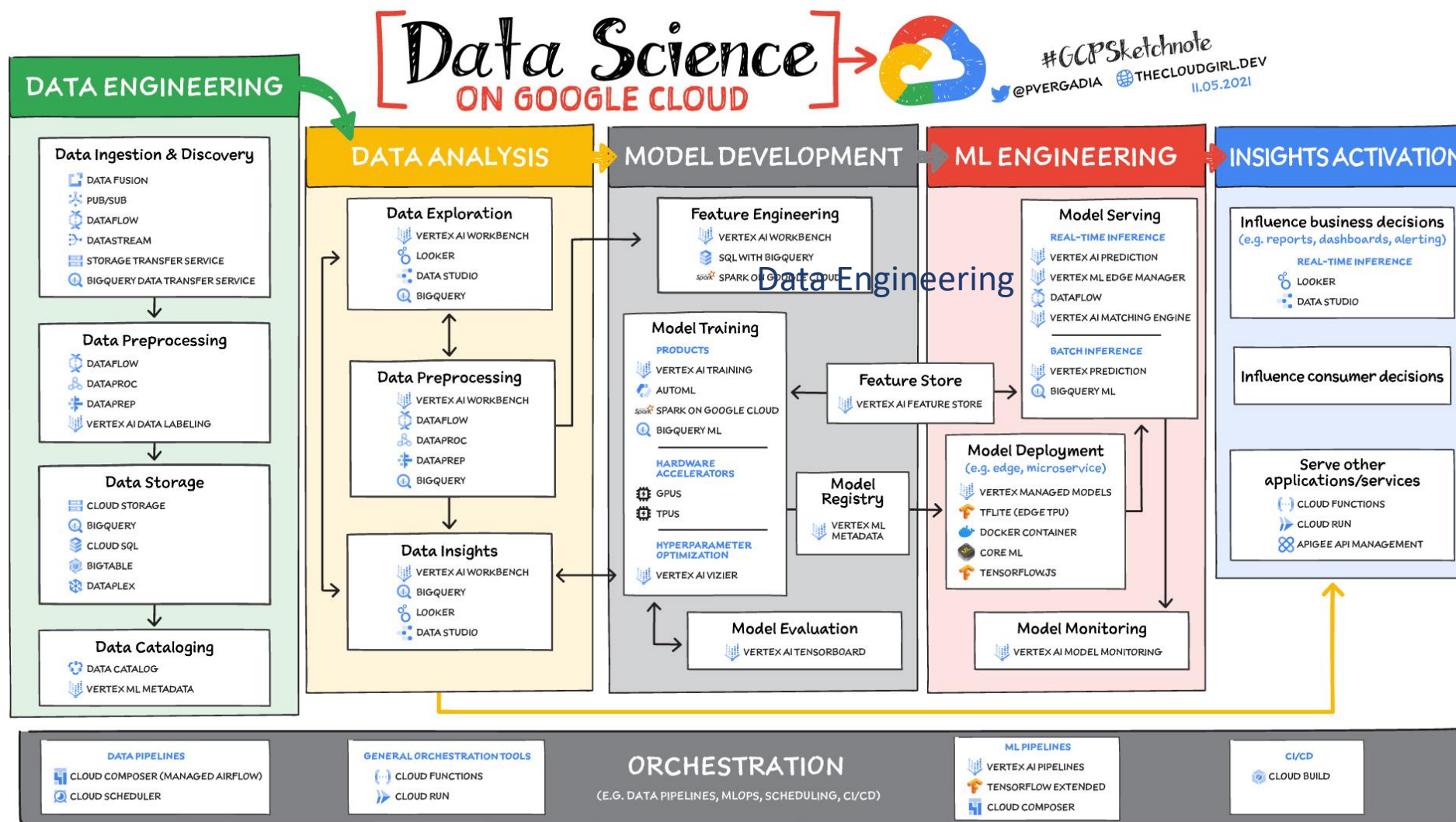
- Relational
- Non-relational
- In Memory

# Build a Scalable Data Analytics Pipeline on the Cloud



- Capture
- Process
- Store
- Analyze
- Use

# Data Science on the Cloud



- Data Engineering
- Data Analysis
- Model Development
- ML Engineering
- Insights Activation



