



---

# **CLOUD COMPUTING APPLICATIONS**

Infrastructure as a Service

---

Reza Farivar

# Infrastructure as a Service

- The most fundamental of Cloud Computing Models
- Allows the user to “rent” computing resources
- The product is a “virtual” computer, that you can access remotely and do whatever you want
  - From the choice of the I/O specs for attached “hard drives” to the Operating System, middleware and applications
  - Network Connection
  - You are now responsible for managing everything running on the machine, including security of your server
- These resources are usually virtualized

# Virtualized Resources

- Different customers have different needs
  - Ephemeral needs
- The Cloud Provider cannot operate a pool of many different sized computers
- Solution: Cloud provider operates a fleet of similar, powerful, hardware
- Carve out chunks of resources through virtualization
  - CPU
  - Memory
  - Storage
  - Network
  - Accelerators

# Virtualized Resources

- Different customers have different needs
  - Ephemeral needs
- The Cloud Provider cannot operate a pool of many different sized computers
- Solution: Cloud provider operates a fleet of similar, powerful, hardware
- Carve out chunks of resources through virtualization: VM Instance
  - CPU
  - Memory
  - Storage
  - Network
  - Accelerators
- Metal as a Service (MaaS)

Dedicated Host SKUs (VM series and Host Type)	Available vCPUs	Available RAM	CPU
D4sv4_Type1	96	768 GiB	2.35 GHz AMD EPYC™ 7452
D4dsv4_Type1	80	504 GiB	Intel® Xeon® Platinum 8272CL (Cascade Lake)
D4sv4_Type1	80	504 GiB	Intel® Xeon® Platinum 8272CL (Cascade Lake)
D4sv3_Type1	64	256 GiB	2.3 GHz Intel® Xeon® E5-2673 v4 (Broadwell)
D4sv3_Type2	76	504 GiB	Intel® Xeon® Platinum 8171M (Skylake)
E4sv3_Type2	76	504 GiB	Intel® Xeon® Platinum 8171M (Skylake)
E4sv3_Type3	80	504 GiB	Intel® Xeon® Platinum 8272CL (Cascade Lake)
F4sv2_Type2	72	144 GiB	Intel® Xeon® Platinum 8168 (Skylake)
F4sv2_Type3	86	504 GiB	Intel® Xeon® Platinum 8272CL (Cascade Lake)
L4sv2_Type1	80	640 GiB	2.55 GHz AMD EPYC™ 7551
M4s_Type1	128	2,048 GiB	Intel® Xeon® Platinum 8280 (Cascade Lake)
M4sm_Type1	128	3,892 GiB	Intel® Xeon® Platinum 8280 (Cascade Lake)
M4smv2_Type1	416	11,400 GiB	Intel® Xeon® Platinum 8180M (Skylake)

# Virtualized Resources

- Different customers have different needs
  - Ephemeral needs
- The Cloud Provider cannot operate a pool of many different sized computers
- Solution: Cloud provider operates a fleet of similar, powerful, hardware
- Carve out chunks of resources through virtualization: VM Instance
  - CPU
  - Memory
  - Storage
  - Network
  - Accelerators
- Metal as a Service (MaaS)



# Advantages of IaaS vs. On-Prem

- No need to run a data center
  - No worries about space, power supplies, physical building security, network, failing components, ...
- OpEx vs. CapEx
- Use different instances when needed
  - Rapid innovation
  - Quick response to shifting business conditions

# IaaS Examples

- Microsoft Azure
- Amazon EC2 (Elastic Compute Cloud)
- Google Cloud Platform Compute Engine
- Oracle Cloud
- IBM Cloud
- Alibaba Cloud
- Rackspace
- Vultr
- ...

The screenshot shows the IBM Cloud 'Virtual server instance' configuration page. The interface is dark-themed with a blue header. At the top, there's a navigation bar with links for 'Catalog', 'Cost Estimator', and 'Docs'. A banner at the top offers a \$200 credit for new accounts. The main content area is titled 'Virtual server instance' and includes a 'Preview mode' warning. Below this, users can select the 'Type of virtual server' (Public, Dedicated, Transient, Reserved), set the 'Quantity' (1) and 'Billing' (Hourly), and choose a 'Location' (NA West, NA South, NA East, South America, Europe, Asia-Pacific). The 'Profile' section shows 'Balanced | B1.2x4' with 2 vCPUs, 4 GB RAM, and SAN storage at \$0.085 per hour. The 'Image' section offers various operating systems like CentOS, Debian, Red Hat, Microsoft, and Ubuntu. The 'Attached storage disks' section shows a 'Boot disk' of 25 GB (SAN) for \$0.000. The 'Network interface' section shows 'Uplink port speeds' and 'Public egress - bandwidth'. The 'Private security group' and 'Public security group' sections are at the bottom. A right-hand sidebar shows a 'Summary' of the configuration and a 'Total due per hour' of \$0.09, with buttons for 'Sign up to create', 'Save as quote', and 'Add to estimate'.

# Instance Pricing

- On-Demand
- Reserved
- Spot Pricing

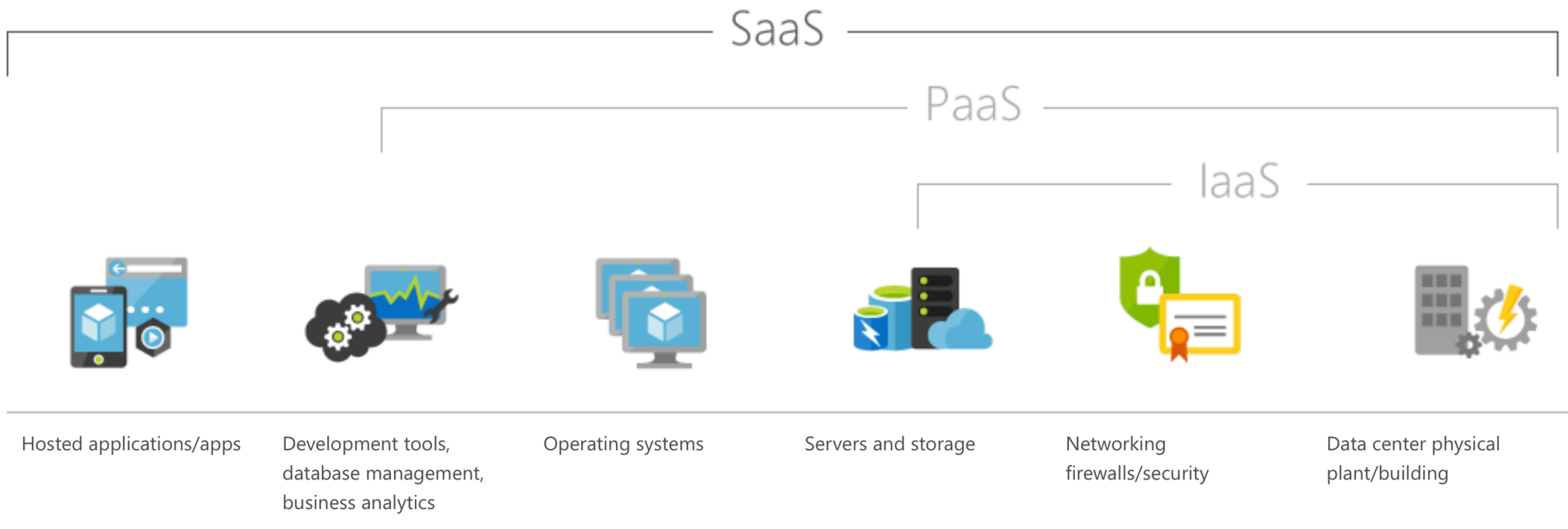
Cloud Computing Applications  
Copyright R. Farivar  
All Rights Reserved



# IaaS Sub-Category: Containers and Orchestration

- A subcategory of IaaS, or a place half-way between IaaS and PaaS (more towards the IaaS)
- You may think of a container as a light-weight Virtual Machine
  - Time to spin up a VM is tens of seconds to a few minutes
  - Time to start a container is fraction of a second to a few seconds
- Linux-Only

# SaaS in Perspective



*\* Image courtesy of Microsoft Azure*