

Cloud Computing Technologies

DAT515 - Fall 2025

Introduction

Prof. Hein Meling





CLOUD

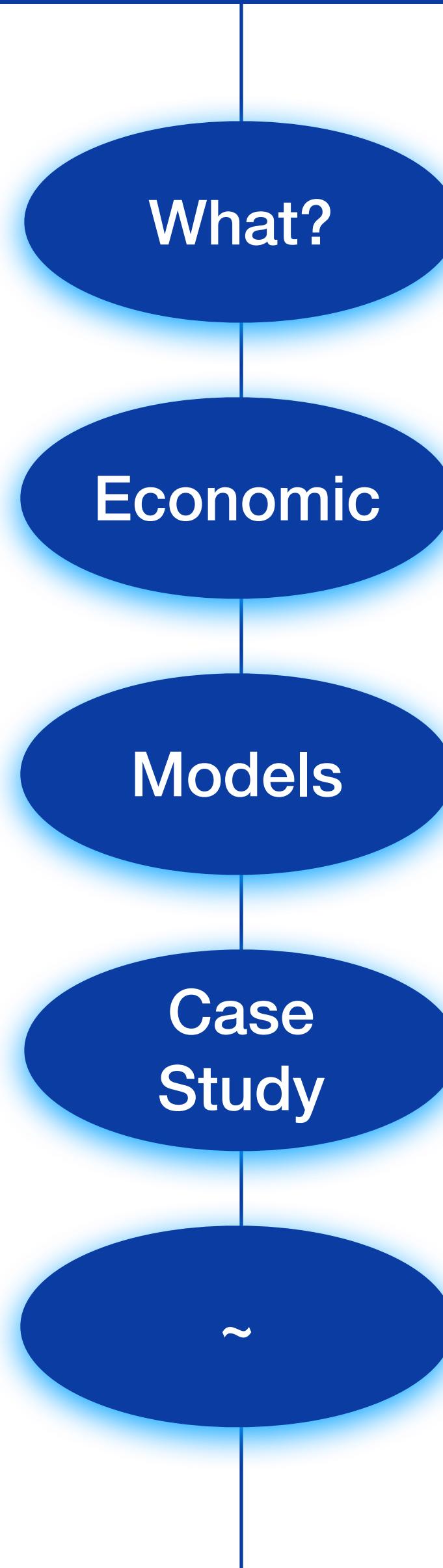
SERVER

SER

SER

SORVERS

CLOUD
BOSS



What is Cloud Computing?

Cloud Computing Economics

Cloud Service and Deployment Models

Case Study: Dropbox

Conclusions

Why is it called Cloud Computing?



Electric Power Grid ~ Grid / Utility Computing



Google Data Center ~ Cloud Computing

Why is it Called Cloud Computing?

The term "cloud computing"

- Emerged from using a **cloud symbol** in network diagrams to represent the internet or an abstracted network
- Became popular in the mid-2000s



What is Cloud Computing?

“The cloud is just
someone else’s computer

<https://www.drewlyton.com/story/the-future-is-not-self-hosted/>

Cloud computing is a delivery model for **on-demand computing resources** over a network. It typically involves **pay-as-you-go pricing** and **self-service access** to a pool of virtualized computing resources.

The cloud computing market is dominated by three major providers: **AWS**, **Microsoft Azure**, and **Google Cloud Platform**.

Cloud computing offers several advantages:

- Scalability**: Resources can be easily scaled up or down based on demand.
- Cost Efficiency**: Reduced capital costs for hardware and infrastructure.
- Reliability**: Built-in redundancy and failover mechanisms.
- Flexibility**: Quick deployment and configuration of resources.
- Global Reach**: Accessible from anywhere with an internet connection.

Cloud computing is built on several key technologies:

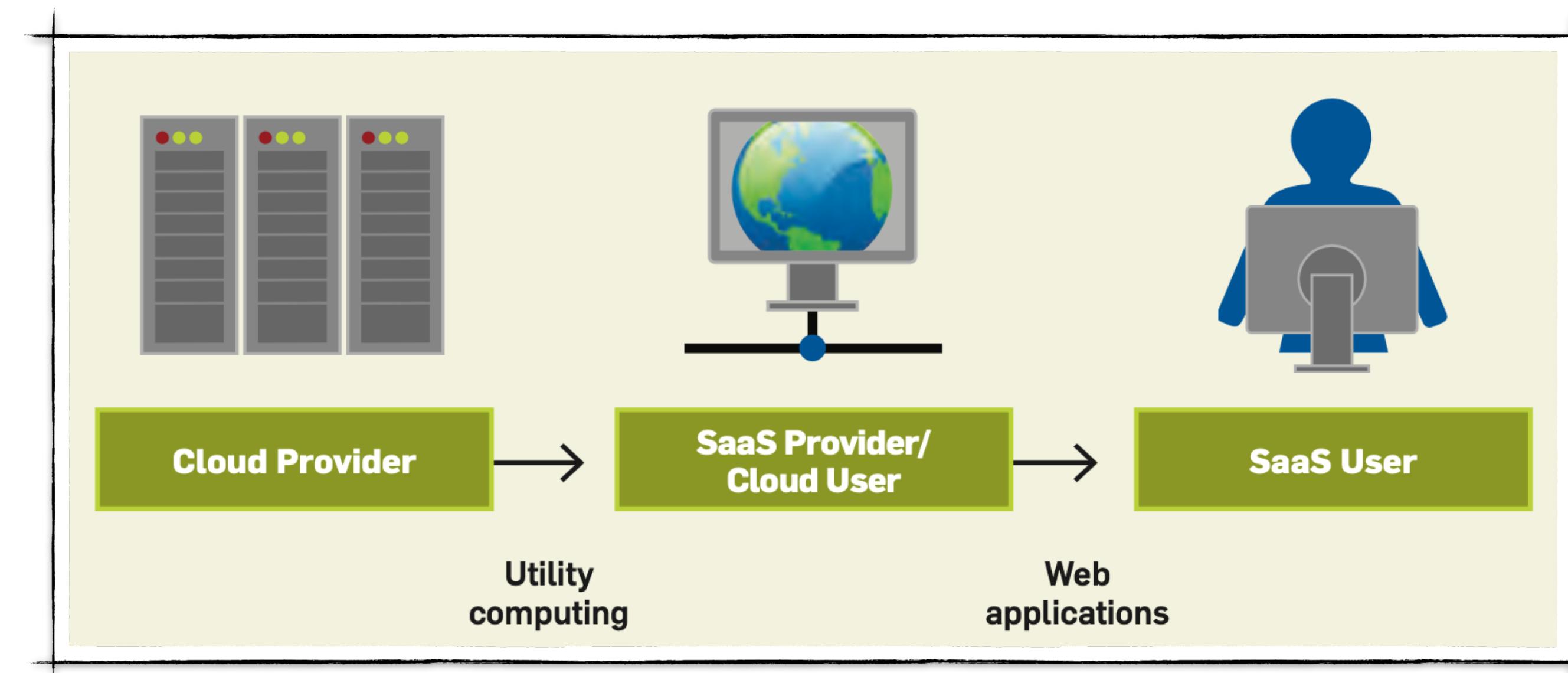
- Virtualization**: Creates multiple virtual machines from a single physical server.
- Cloud Storage**: Provides scalable storage solutions.
- Cloud Computing Services**: Offered as a service, such as **SaaS** (Software as a Service), **PaaS** (Platform as a Service), and **IaaS** (Infrastructure as a Service).
- Cloud Computing Models**: **Public Cloud** (shared infrastructure), **Private Cloud** (dedicated infrastructure), and **Hybrid Cloud** (a combination of public and private clouds).
- Cloud Computing Deployment Models**: **Deployment** (Serverless), **Containerization** (Containers), and **Virtualization**.
- Cloud Computing Technologies**: **Kubernetes** (orchestration), **Docker** (containers), and **OpenShift** (container orchestration).
- Cloud Computing Tools**: **CloudWatch Metrics** (monitoring), **CloudWatch Logs** (logging), and **CloudWatch Metrics Insights** (analytical insights).

Cloud computing has transformed various industries, including **Finance**, **Healthcare**, **Manufacturing**, and **Retail**, by providing them with cost-effective, reliable, and scalable computing resources.

What is Cloud Computing?

Definition: Cloud Computing

- Applications **delivered as services** over the Internet,
- Plus the **hardware** and **systems software** in the **data centers** that provide these services.



Source: A View of Cloud Computing

NIST's Definition

“Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.

Core Advantages of Cloud Computing

- Lower Costs: No large capital outlays for hardware
- Pay-per-use model
 - Align costs with actual usage, reducing waste
- Elastic resource allocation
- Rapid scalability

- Public cloud
 - A service made available to the general public on a pay-as-you-go basis
 - Utility computing
- Private cloud
 - Internal data centers of an organization not available to the public

Public vs. Private Clouds

Advantage	Public Cloud	Conventional Data Center
Appearance of infinite computing resources on demand	Yes	No
Elimination of an up-front commitment by Cloud users	Yes	No
Ability to pay for use of computing resources on a short-term basis as needed	Yes	No
Economies of scale due to very large data centers	Yes	Usually not
Higher utilization by multiplexing of workloads from different organizations	Yes	Depends on company size
Simplify operation and increase utilization via resource virtualization	Yes	No

Joke time

Because it found someone
with him ~~dead~~ up there!?

Cloud Computing Economics

Cost Example for Batch Task

- Large computing tasks can get results as quickly as their programs can scale.
- Consider
 - Using 1000 servers for **one hour**

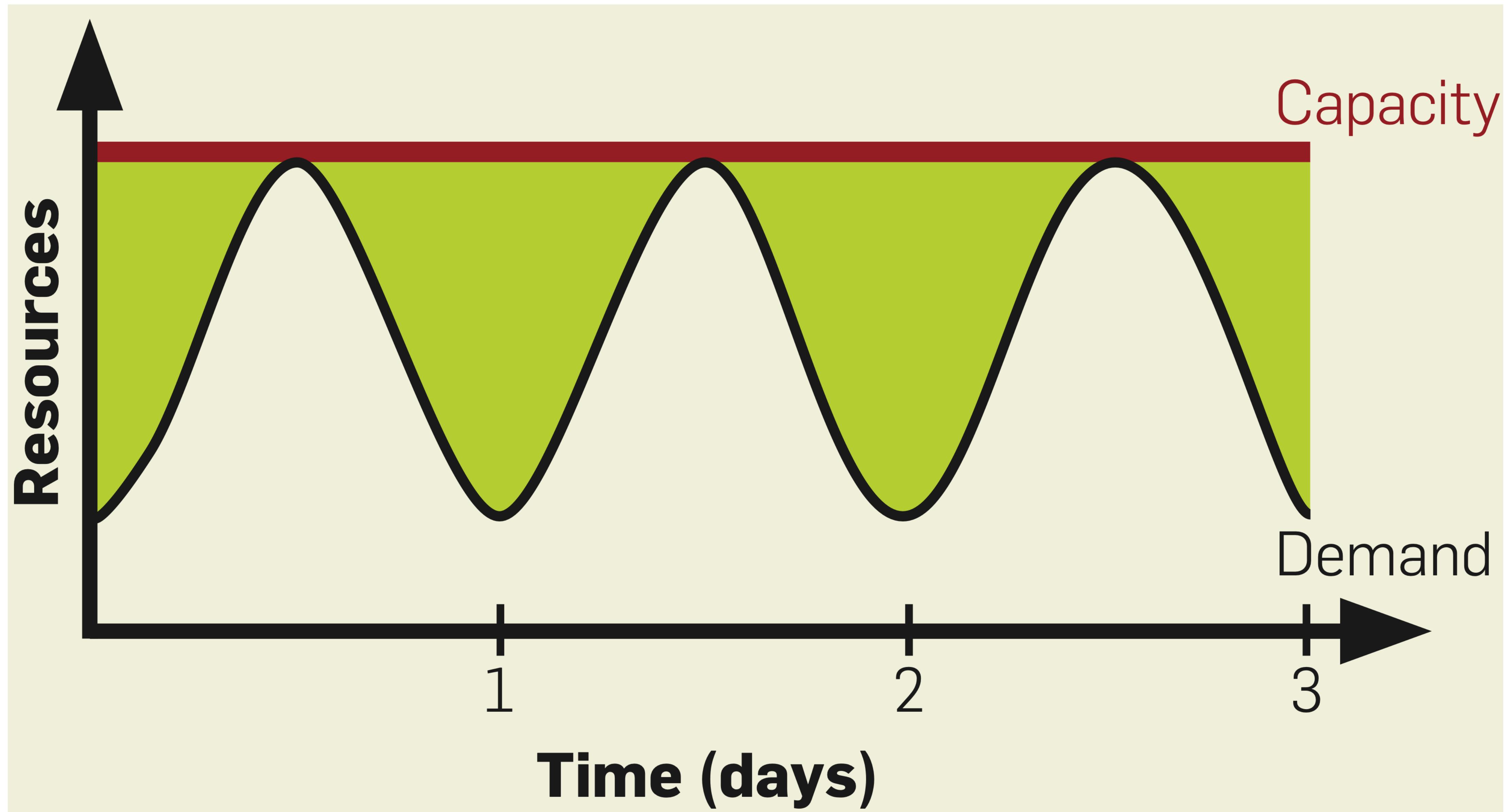
VS

- Using **one server** for 1000 hours
- *Cost the same for the user*

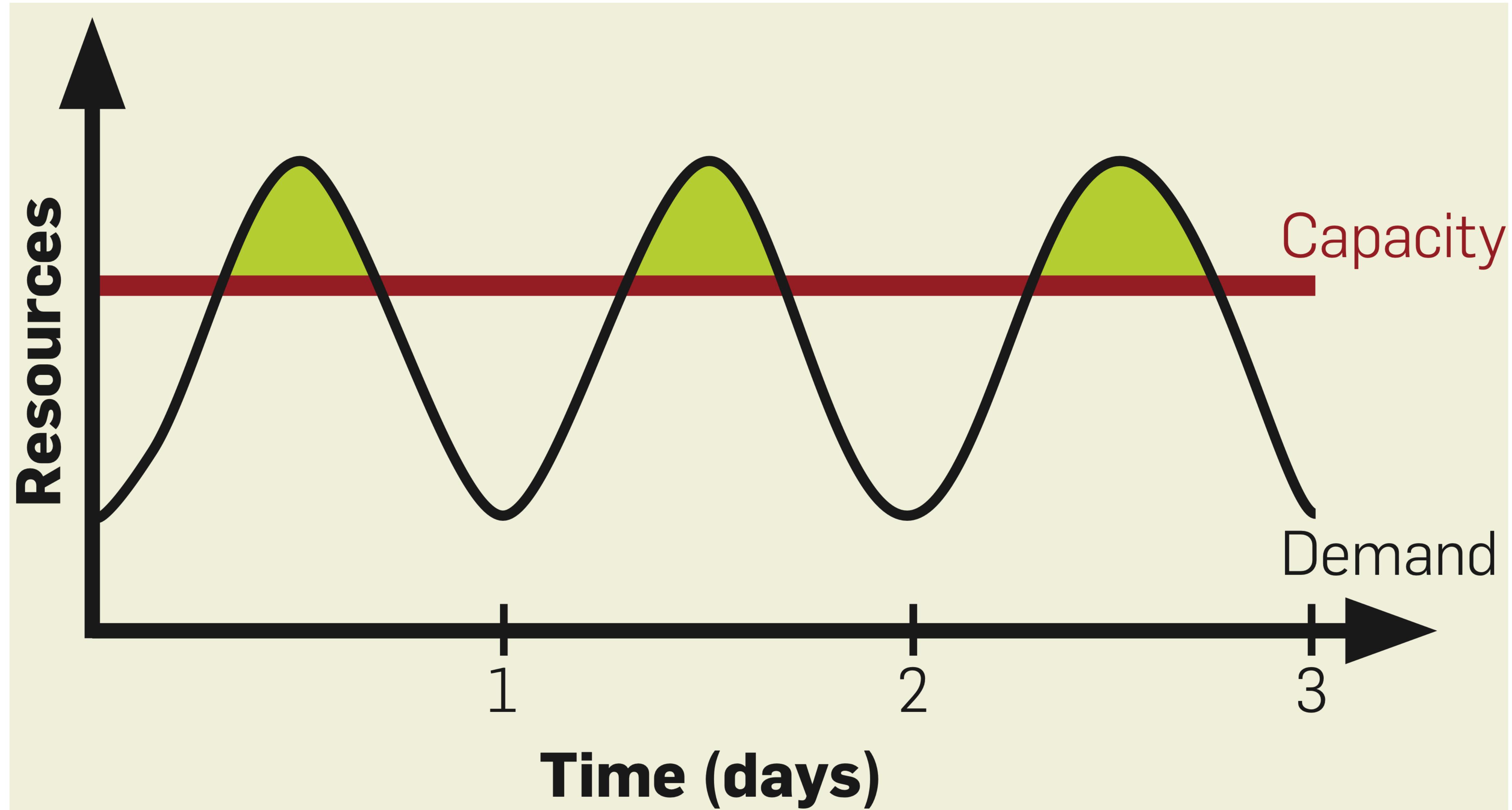
Cost Example for Predictable Demand

- Peak @ noon: 500 servers
- Trough @ midnight: 100 servers
- Average utilization over a day: 300 servers
- Actual cost per day: $300 \times 24 = 7,200$ server-hours
- Must provision for peak (500 servers) and pay: $500 \times 24 = 12,000$ server-hours
- *Utility computing is cheaper* if 3-year pay-as-you-go costs < 1.7x purchase cost

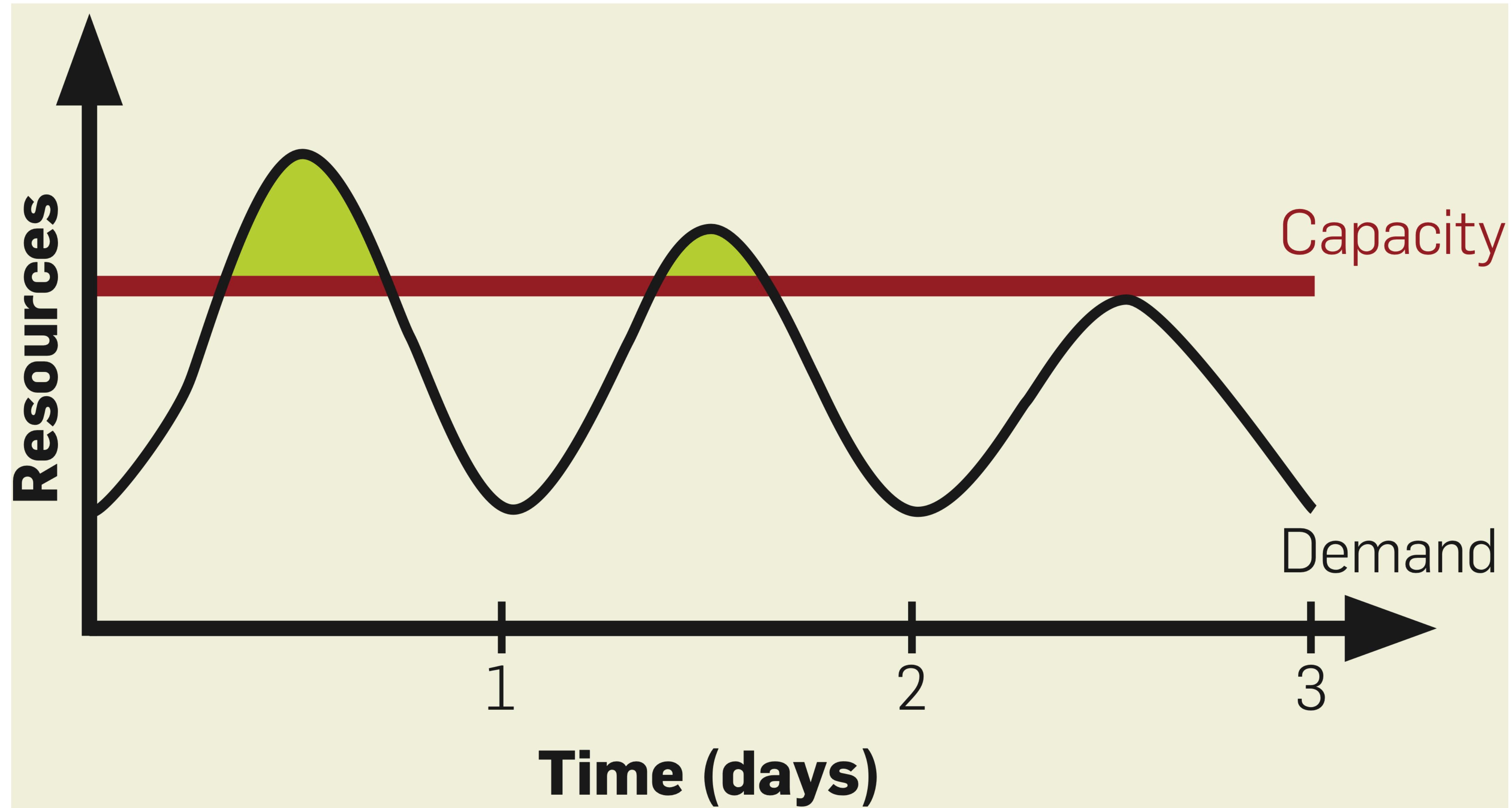
Provisioning for peak load



Underprovisioning Case 1



Underprovisioning Case 2



Reduced Costs

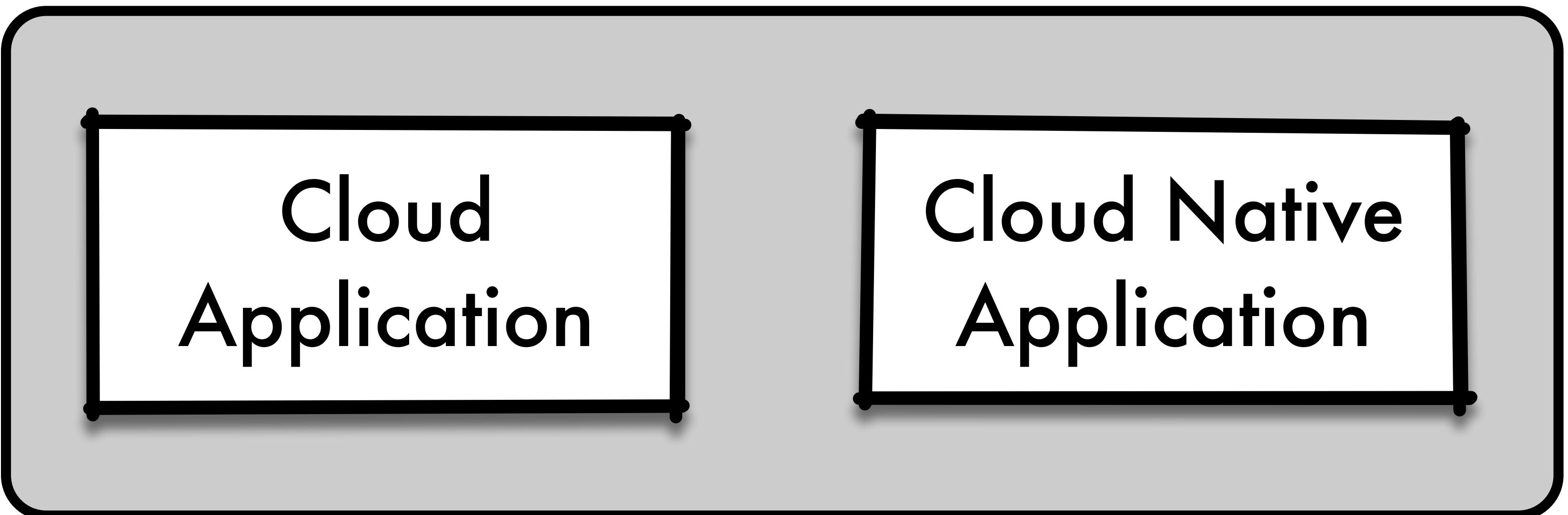
- Electricity: Install own power generation capacity (Data Center Solar)
- Bandwidth: Build own networks (Google's Global Network)
- Hardware: Build custom hardware (Google's TPU, custom NICs)
- Software: Custom data center software – Hire the best engineers
- Cloud providers have greater leverage to negotiate prices on all of the above.

Cloud Service Models

Not hosted on
a cloud platform



Hosted on a cloud platform



Not built for
a cloud platform

Built for
a cloud platform

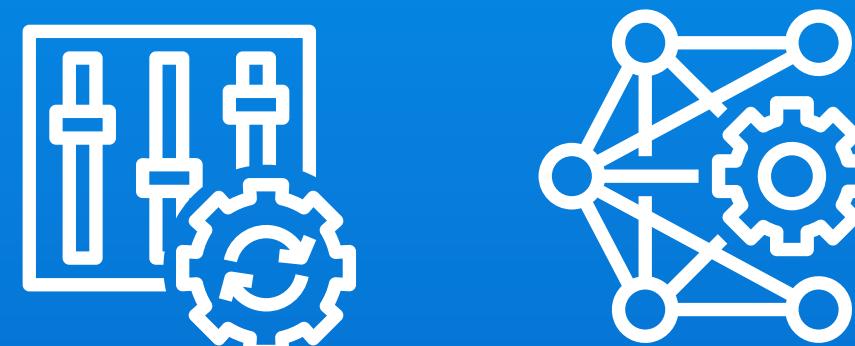
Software as a Service (SaaS)

Dropbox, Slack, GMail, Overleaf
(delivered over internet, browser-based UI)



Platform as a Service (PaaS)

Google App Engine, Windows Azure, AWS Lamda
(platform for software development)



Infrastructure as a Service (IaaS)

Amazon Web Services, Google Compute Engine
(rent out storage and compute capacity)



- Provides virtualized computing resources over the internet, including servers, storage, and networking
- Enabling users to run and manage operating systems, applications, and data **without managing physical hardware**
- Gives users more control

- Offers a development environment and tools to build, deploy, and manage applications, abstracting the underlying infrastructure
- Developers **focus on writing code** without worrying about hardware and software setup
- Less control to the user than IaaS
 - Issue: PaaS is less standardized across vendors

- Delivers software applications over the internet (often on a subscription basis), accessible via web browsers
- Eliminate the need for users to install, manage, or maintain **software on their local devices**
- SaaS users have no control over
 - Developer/Company of SaaS app decide where/how to host

Service Model Comparison

On-site	IaaS	PaaS	SaaS
Applications	Applications	Applications	Applications
Data	Data	Data	Data
Runtime	Runtime	Runtime	Runtime
Middleware	Middleware	Middleware	Middleware
Operating System	Operating System	Operating System	Operating System
Virtualization	Virtualization	Virtualization	Virtualization
Servers	Servers	Servers	Servers
Storage	Storage	Storage	Storage
Networking	Networking	Networking	Networking

An Incomplete List of Cloud Computing Providers

SaaS

- Salesforce
- Cisco WebEx
- Dropbox
- ZenDesk
- MailChimp
- Slack
- HubSpot
- DocuSign
- Google Apps
- Microsoft Office 365

PaaS

- Google App Engine
- OpenShift
- Heroku
- Force.com
- Windows Azure
- AWS Elastic Beanstalk
- Apache Stratos
- Magento Commerce
- AWS Lamda
- SAP Cloud

IaaS

- Amazon Web Services (AWS)
- Google Compute Engine (GCE)
- IBM Cloud
- Microsoft Azure
- Rackspace
- Linode
- Cisco Metacloud
- Digital Ocean
- Vultr
- Oracle Cloud

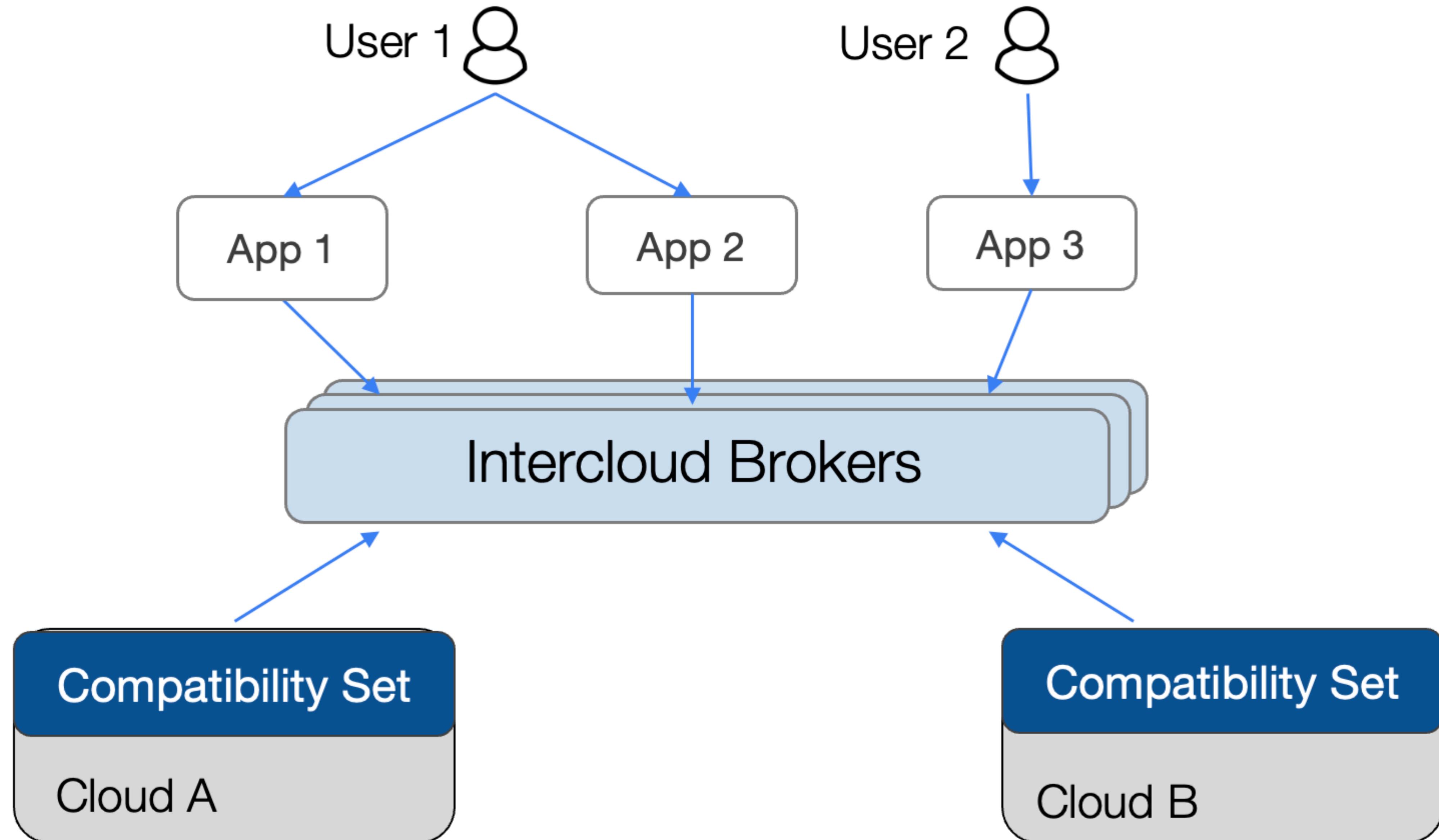
Source: <https://www.leanix.net/en/wiki/apm/iaas-vs-paas-vs-saas>

Cloud Deployment Models

Mixed Computing Environment

- Design systems/applications with same/similar technology
- Makes it easy to deploy to different environments
 - Public cloud providers
 - Private cloud
- Load surges can be offloaded to public cloud providers

- On-premise (owned by organization) or
- Off-premise (rented from infrastructure providers)
 - Vendor-owned data center
 - Dedicated resources to a single customer (isolated access)
- Example: Red Hat's OpenStack Platform



Case Study: Dropbox

- Dropbox: Popular file hosting and sharing service
- Challenge
 - Storing and synchronizing vast amounts of user data
 - Scale to accommodate growth

AWS Solution

- Initially, Dropbox relied heavily on
 - AWS Cloud Storage and Services
- Benefits:
 - **Scalability:** Quickly expand storage capacity with user growth (without investing in physical infrastructure)
 - **Data Redundancy:** Improved data availability and resilience
 - **Security:** Enhanced user data protection

AWS Solution

- Successfully handled storage needs of millions of users
- Streamlined operations and improved data security
- Focus on **building features** and enhancing **user experience**

Migration to In-House Infrastructure

- Project **Magic Pocket** completed in 2016
- Migrated most of their data storage from AWS S3 to their custom-built infrastructure
 - Driven by the desire to reduce costs
 - Gain more control over their hardware
 - Tailor the storage system specifically for their needs

Hybrid Approach

- After the migration, Dropbox continues to use AWS for certain services and workloads
- Primary data storage and most of their computing infrastructure are hosted on their own servers
- Allows Dropbox to leverage the benefits of both cloud and on-premises infrastructure

Questions?

DOI:10.1145/1721654.1721672

Clearing the clouds away from the true potential and obstacles posed by this computing capability.

BY MICHAEL ARMBRUST, ARMANDO FOX, REAN GRIFFITH, ANTHONY D. JOSEPH, RANDY KATZ, ANDY KONWINSKI, GUNHO LEE, DAVID PATTERSON, ARIEL RABKIN, ION STOICA, AND MATEI ZAHARIA

A View of Cloud Computing

CLOUD COMPUTING, THE long-held dream of computing as a utility, has the potential to transform a large part of the IT industry, making software even more attractive as a service and shaping the way IT hardware is designed and purchased. Developers with innovative ideas for new Internet services no longer require the large capital outlays in hardware to deploy their service or the human expense to operate it. They need not be concerned about overprovisioning for a service whose popularity does not meet their predictions, thus wasting costly resources, or underprovisioning for one that becomes wildly popular, thus missing potential customers and revenue. Moreover, companies with large batch-oriented tasks can get results as quickly as their programs can scale, since using 1,000 servers for one hour costs no more than using one server for 1,000

hours. This elasticity of resources, without paying a premium for large scale, is unprecedented in the history of IT.

As a result, cloud computing is a popular topic for blogging and white papers and has been featured in the title of workshops, conferences, and even magazines. Nevertheless, confusion remains about exactly what it is and when it's useful, causing Oracle's CEO Larry Ellison to vent his frustration: "The interesting thing about cloud computing is that we've redefined cloud computing to include everything that we already do.... I don't understand what we would do differently in the light of cloud computing other than change the wording of some of our ads."

Our goal in this article is to reduce that confusion by clarifying terms, providing simple figures to quantify comparisons between of cloud and conventional computing, and identifying the top technical and non-technical obstacles and opportunities of cloud computing. (Armbrust et al^a is a more detailed version of this article.)

Defining Cloud Computing

Cloud computing refers to both the applications delivered as services over the Internet and the hardware and systems software in the data centers that provide those services. The services themselves have long been referred to as Software as a Service (SaaS).^a Some vendors use terms such as IaaS (Infrastructure as a Service) and PaaS (Platform as a Service) to describe their products, but we eschew these because accepted definitions for them still vary widely. The line between "low-level" infrastructure and a higher-level "platform" is not crisp. We believe the two are more alike than different, and we consider them together. Similarly, the

^a For the purposes of this article, we use the term Software as a Service to mean applications delivered over the Internet. The broadest definition would encompass any on demand software, including those that run software locally but control use via remote software licensing.