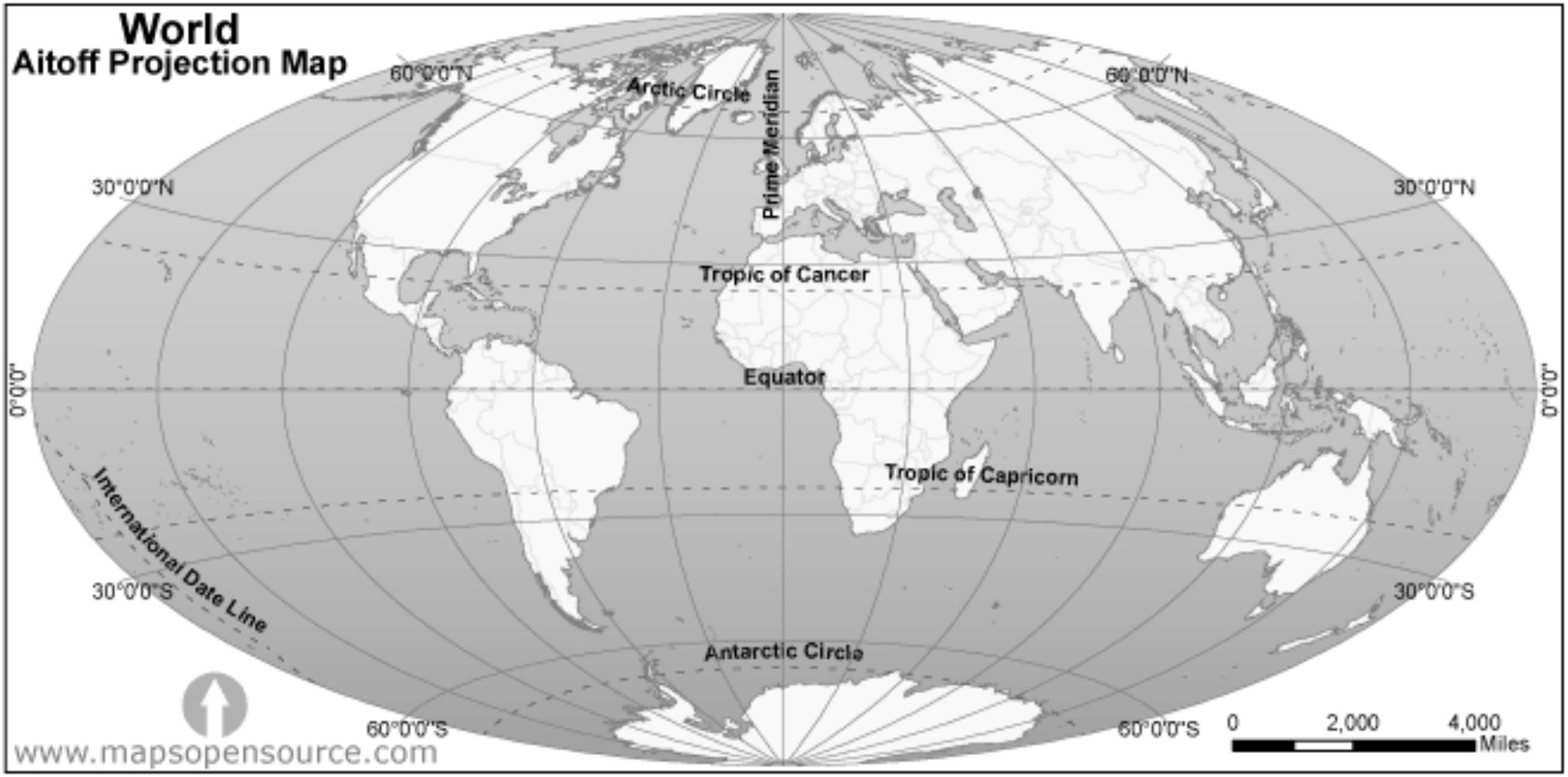
ECE 1779
Introduction to Cloud Computing

Eyal de Lara

delara@cs.toronto.edu

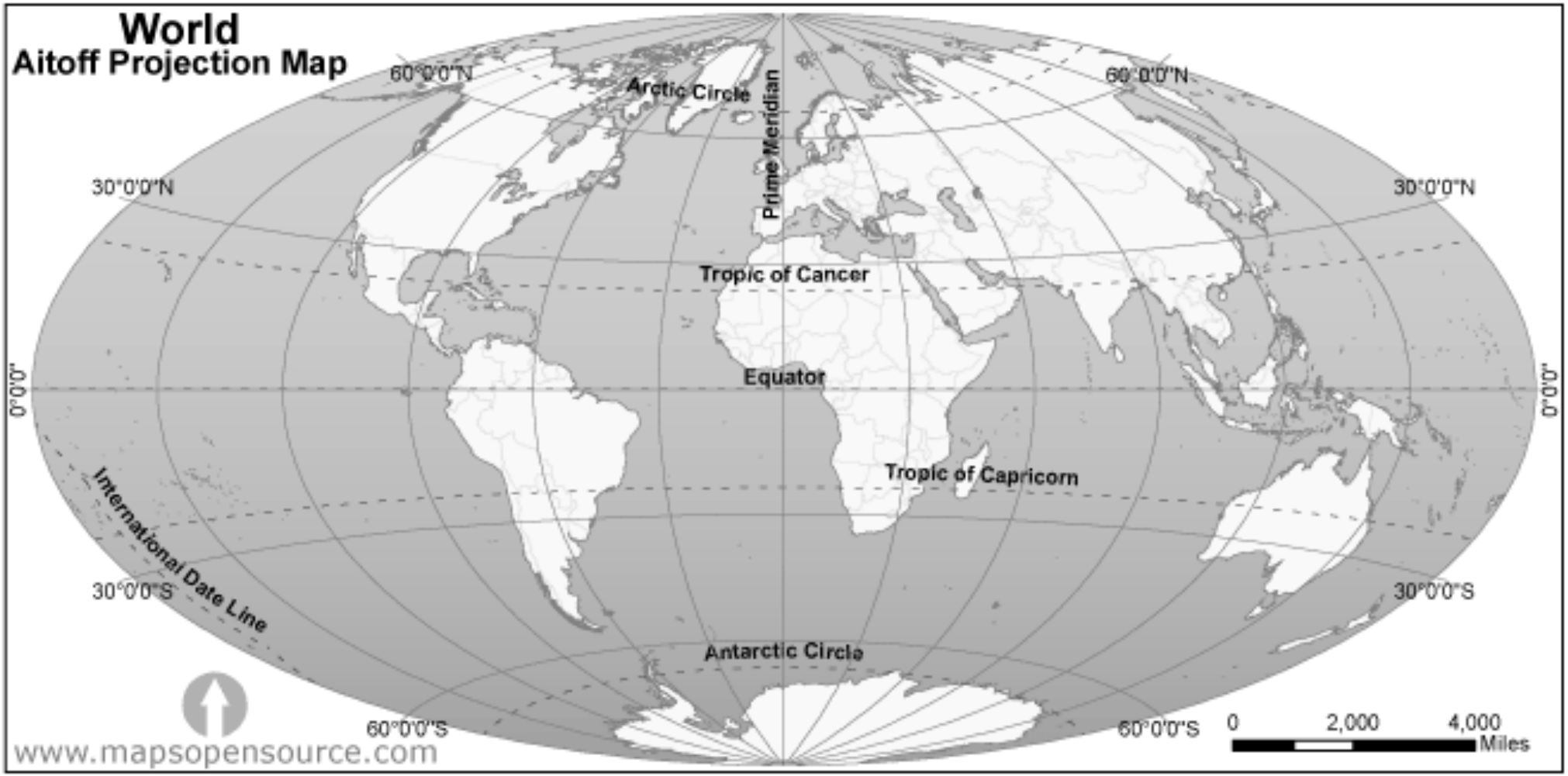
www.cs.toronto.edu/~delara/courses/ece1779

Where in the World is Carmen Sandiego?



Where in the World is Carmen Sandiego?

are you?



Course Overview

Introduction
Python
Web Development
Databases
Amazon Web Services
Cluster Computing
Function as a Service
No SQL Databases
Virtualization
Batch Processing
Stream Processing

Marking

	Weight	Due
Class Participation	10%	
A1 (Web Development)	20%	Feb 22
A2 (Amazon EC2)	35%	Mar 22
A3 (Amazon Lambda)	35%	Apr 21-23

Class Mechanics

- Prepare
  - Tasks listed on web site, under “Calendar and Lecture Notes”
  - Papers, book chapters, videos
- Lecture
  - Zoom based presentation
  - Hands-on group exercises

Week 2: Python Overview

- Prepare

  Option 1:

  An Introduction to Computer Science Using Python 3

  By Paul Gries, Jennifer Campberll, Jason Montojo

  Read chapters 1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 14

  Option 2:

  PCRS

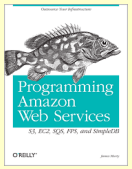
  View videos in *prepare* sections for weeks 2

- Lecture Friday 10am-12pm

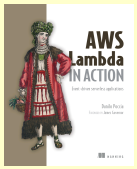
Recommended Reading

The Cloud at Your Service
Jothy Rosenberg and Art Mateos

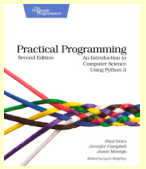
Cloud Computing Bible
Barrie Sosinsky



Programming Amazon Web Services
James Murty



AWS Lambda in Action
Danilo Poccia



Practical Programming (2nd edition)
An Introduction to Computer Science Using Python 3
by Paul Gries, Jennifer Campbell, Jason Montojo

Wolf, goat, and a cabbage river crossing puzzle

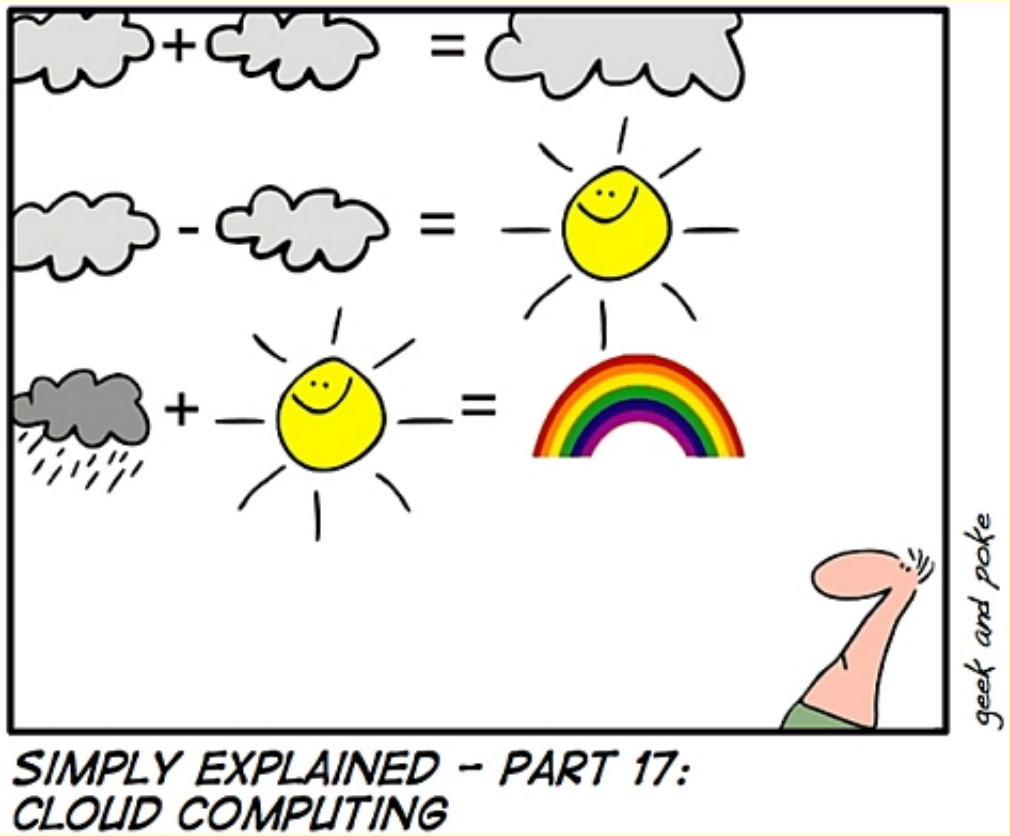
- Once upon a time a farmer went to a market and purchased a wolf, a goat, and a cabbage. On his way home, the farmer came to the bank of a river and rented a boat. But crossing the river by boat, the farmer could carry only himself and a single one of his purchases: the wolf, the goat, or the cabbage.
- If left unattended together, the wolf would eat the goat, or the goat would eat the cabbage.
- The farmer's challenge was to carry himself and his purchases to the far bank of the river, leaving each purchase intact.
- Question: What is the minimum number of times that the farmer needs to row the boat between river banks to accomplish the task?

Source: https://en.wikipedia.org/wiki/Wolf,_goat_and_cabbage_problem

Instructions

- Working on the problem:
  - Introduce yourself to the other students in your breakout room
  - On zoom, start a shared whiteboard by having one person in your breakout room click on "Share Content" and then "Whiteboard"
  - The rest of the group can also draw on the whiteboard using the screen annotation tool
- Once the problem is solved each member should submit their answer on PCRS:
  - https://pcrs.teach.cs.toronto.edu/ECE1779-2021-01
  - Navigate to Week 1: Perform, River Crossing Puzzle
  - Answer the multiple-choice question
  - You can attempt answering multiple times without penalty
  - Exit the breakout room when done.

What is Cloud Computing?



Definition

Cloud computing is the delivery of computing as a service rather than a product, whereby shared resources, software, and information are provided to computers and other devices as a metered service over a network (typically the Internet).

Wikipedia

Why call it “Cloud” Computing?

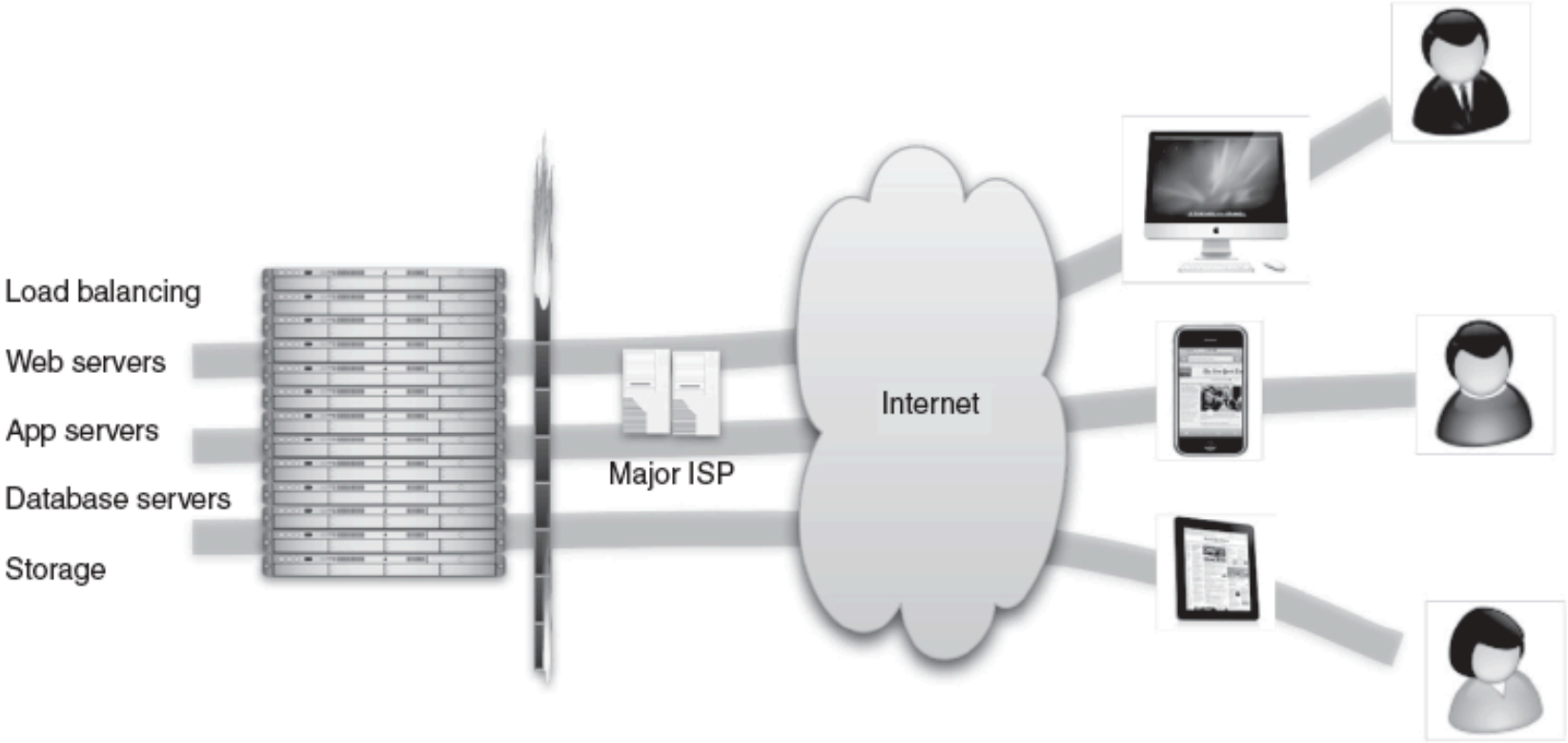
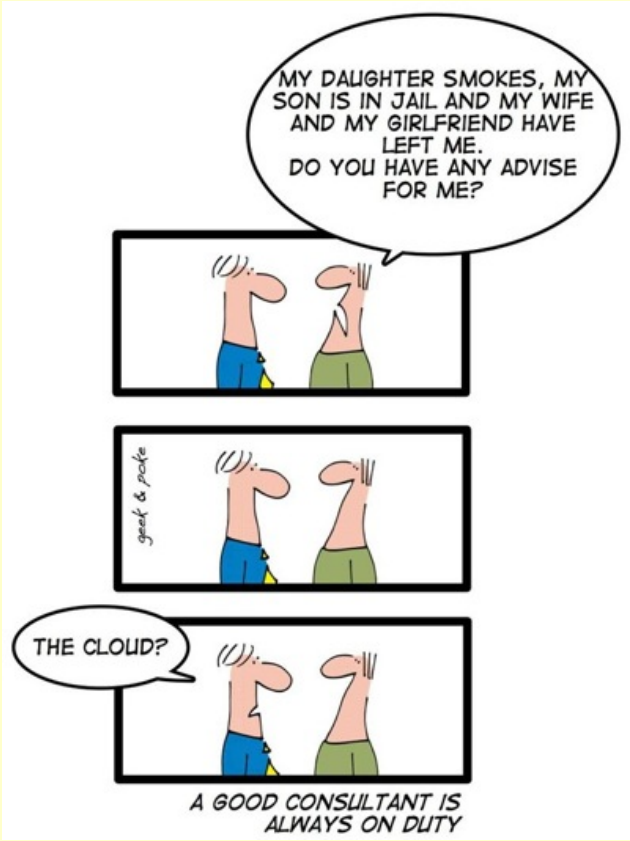
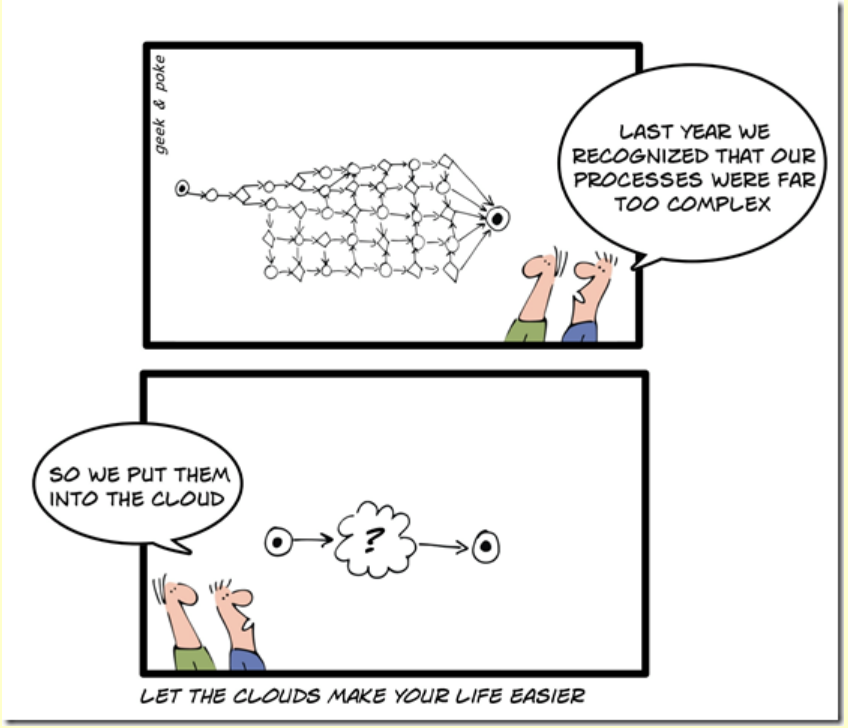


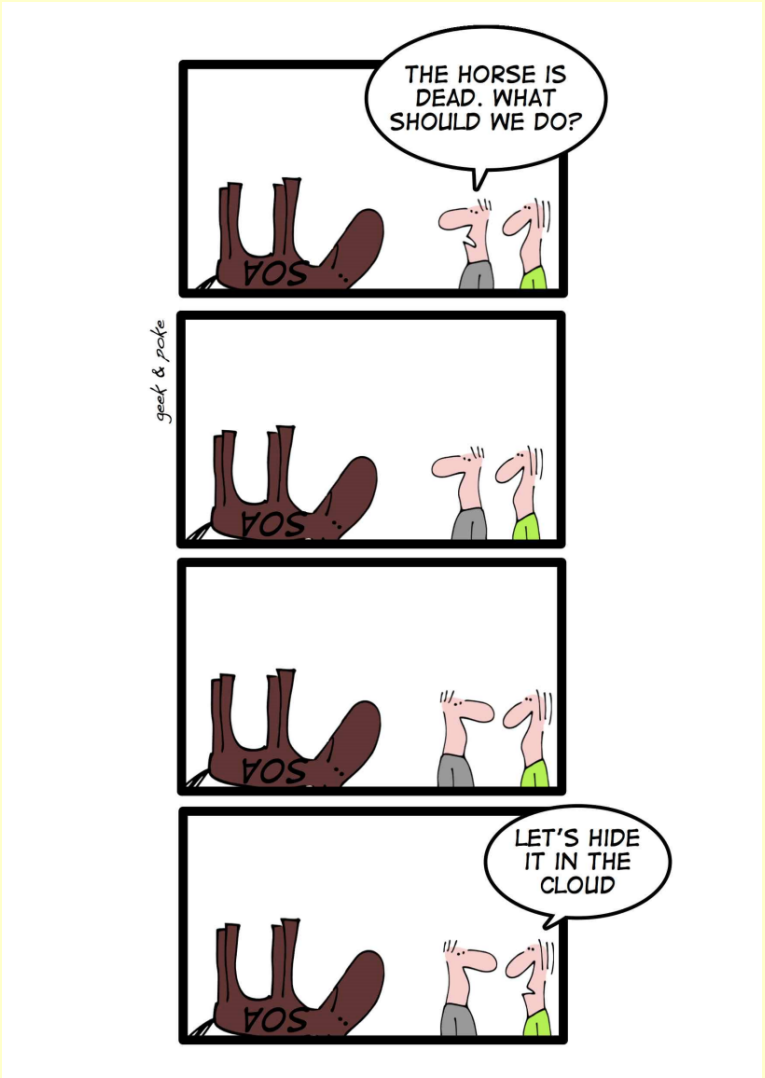
Figure 1.4 A picture of a cloud is a ubiquitous representation of the internet and is used almost universally in discussions or drawings of computer architecture.

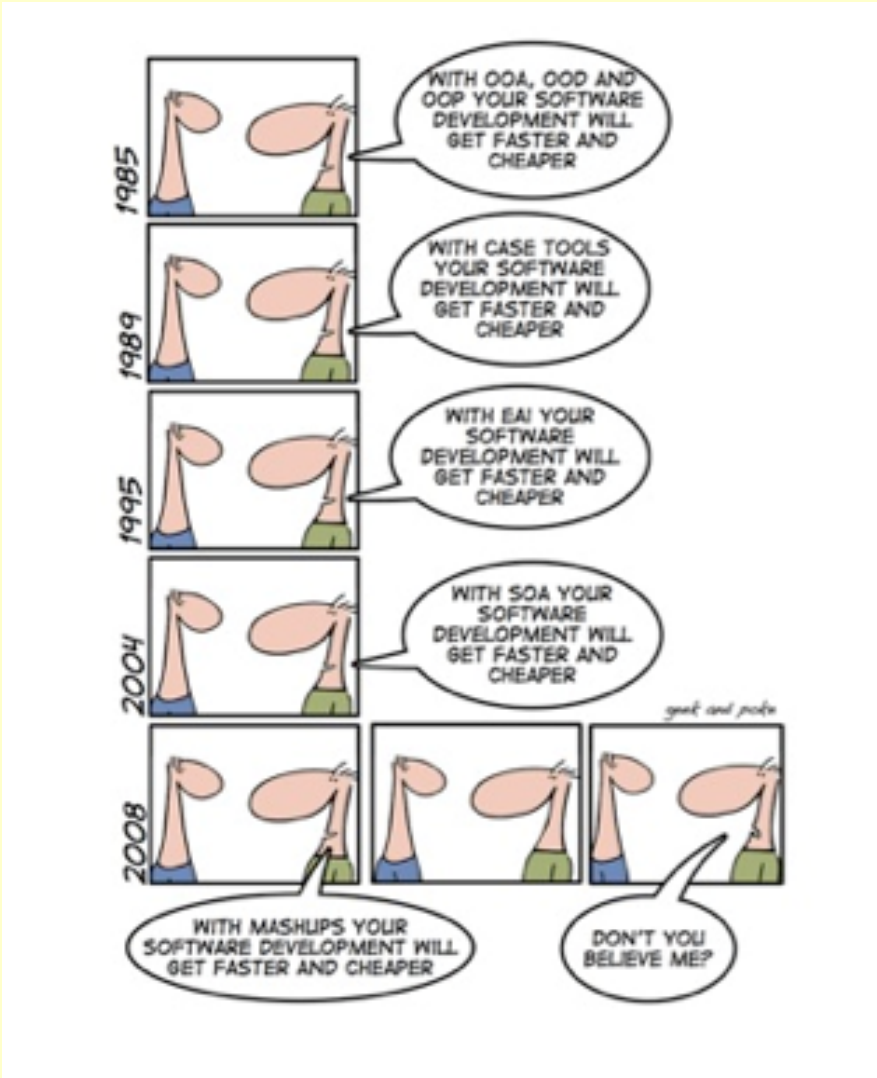
The Cloud will Solve ALL Problems





Haven’t I heard this story before?

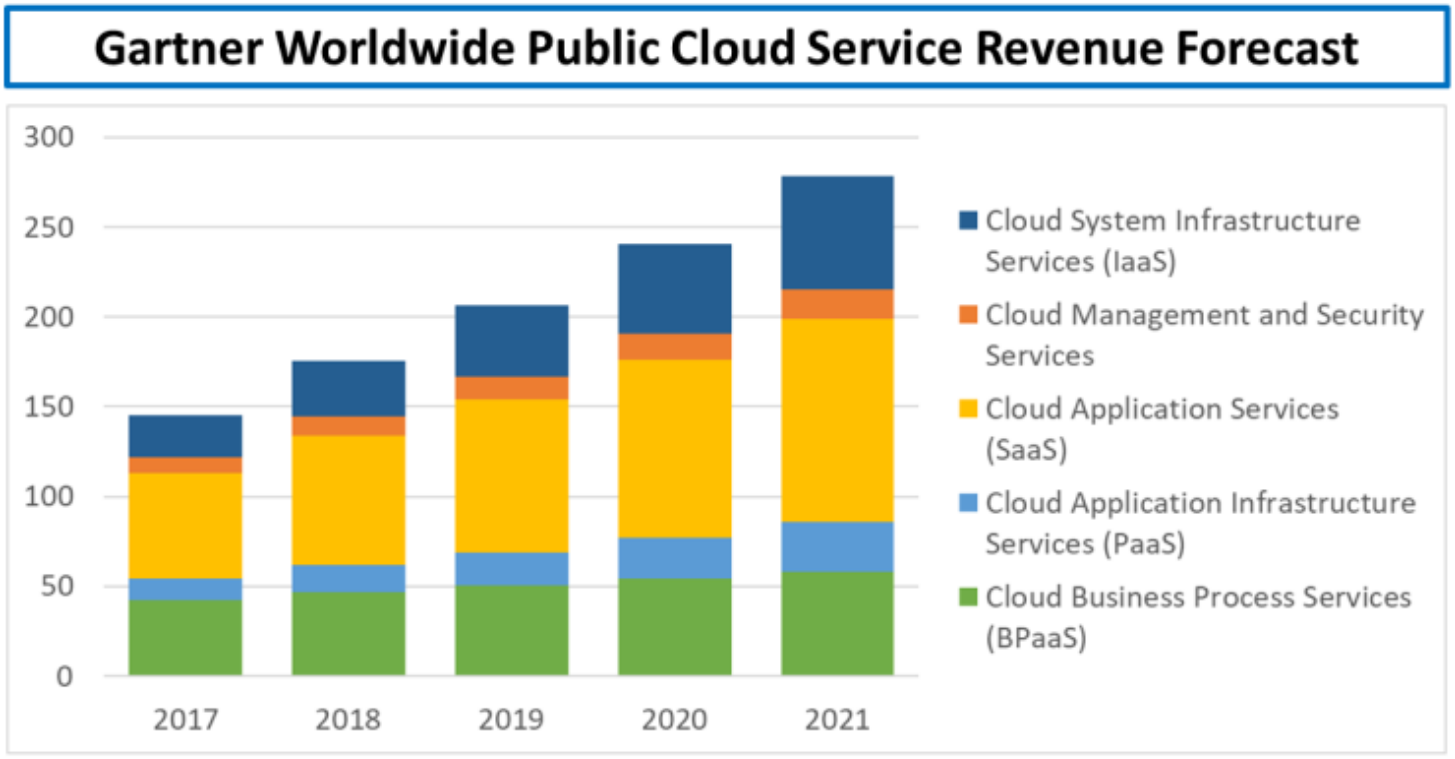




What is New?

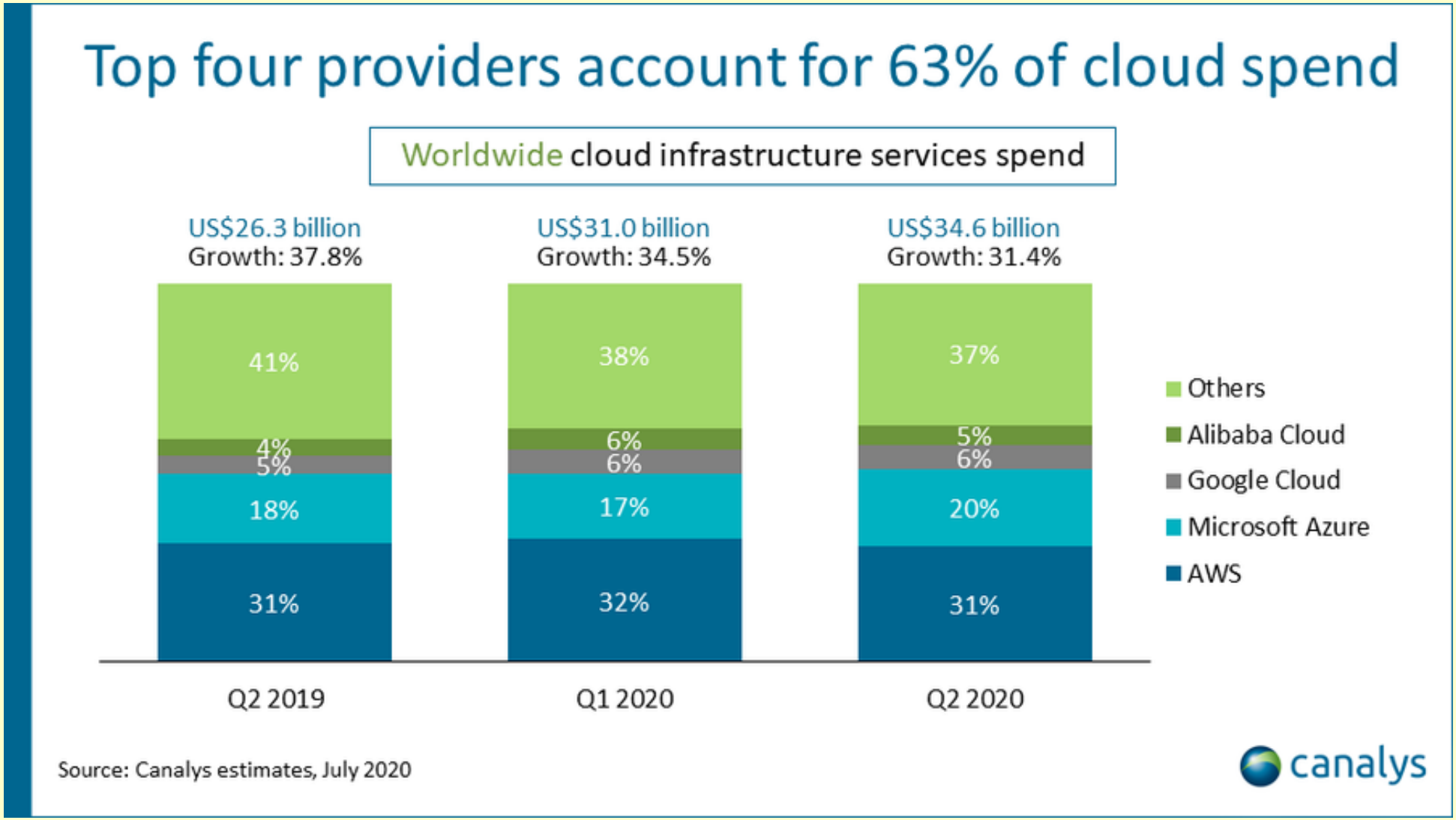
- No up-front commitment by Cloud users
  - Anyone with a credit card can rent resources
- Metered billing
  - Pay-as-you-go model
  - Can pay for use on a short-term basis (processors by hour, and storage by day)
- Elasticity
  - Infinite computing resources available on demand
  - Pooled resources
  - No need to plan far ahead on provisioning
- Virtualization
  - User gets controls application
  - Provider can still safely multiplex customers on same HW
- Automation
  - APIs to allocate and release resources

Why care about Cloud Computing?



Source: Gartner

Main Players



Advantages

- Shift from CAPEX to OPEX
  - Lowers barrier for starting a new business/project
- Can be cheaper even in the long run
  - Economies of scale
  - Geography

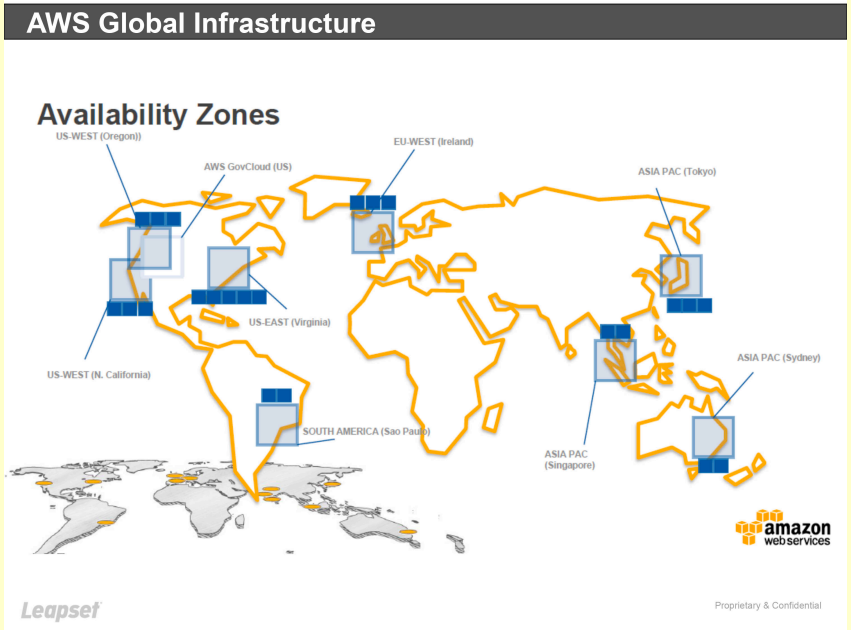
Technology	Cost in Medium-sized DC	Cost in Very Large DC	Ratio
Network	$95 per Mbit/sec/month	$13 per Mbit/sec/month	7.1
Storage	$2.20 per GByte / month	$0.40 per GByte / month	5.7
Administration	≈140 Servers / Administrator	>1000 Servers / Administrator	7.1

Advantages

- Instant agility
  - No need to procure, provision and configure new HW
  - No need to plan far ahead on provisioning
- Security
  - Cloud provider can afford better physical and logical security
  - Bank vs. company vault

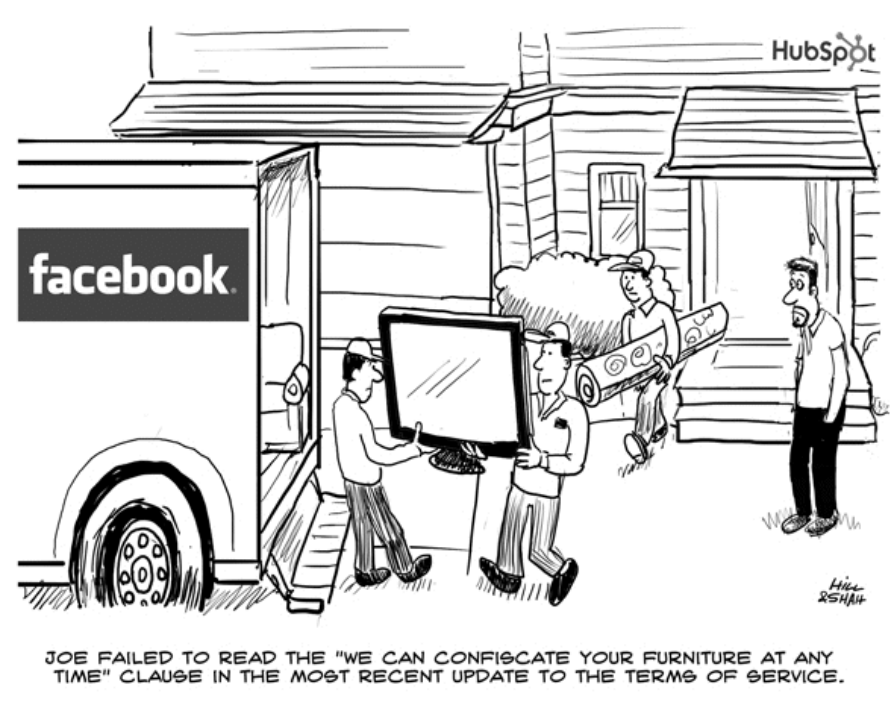
Advantages

- Reliability
  - Redundancy, geographic replication



Challenges

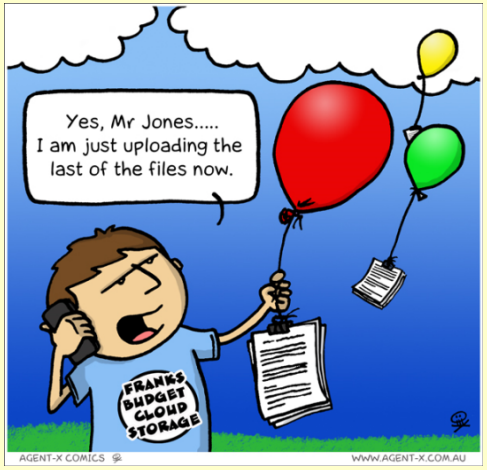
- Performance predictability (QoS)
  - Fair sharing of I/O resources among VMs
- Vendor lock-in



Challenges

- Security and privacy

- Data transfer bottleneck
  - Uploading 10 TB over 20 Mbps link takes 45 days!

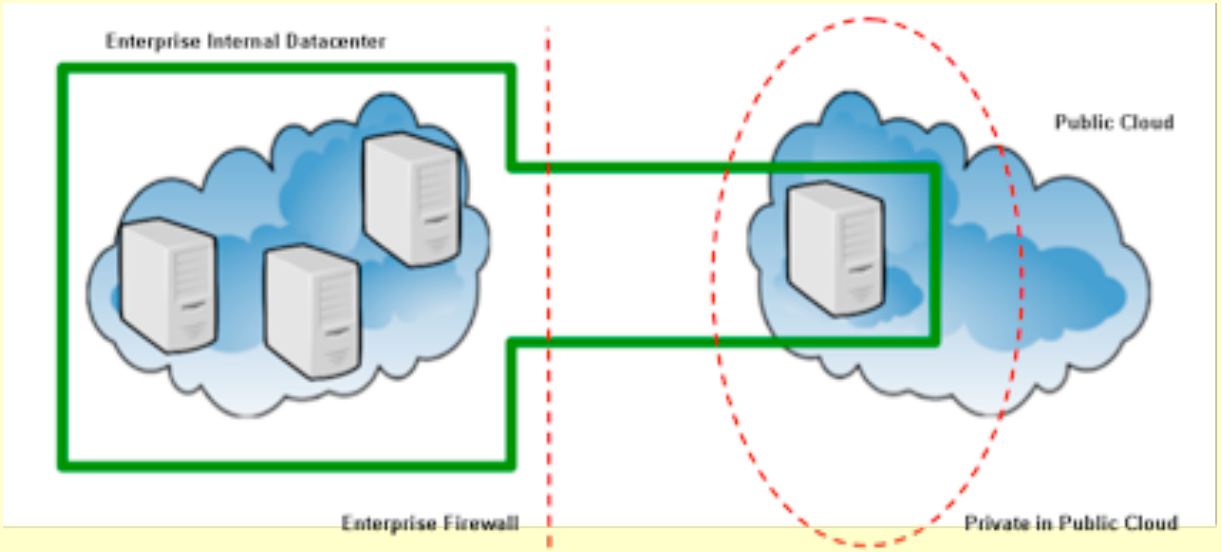


Cloud-friendly Applications

- Existing apps that do not require specialized hardware resources
- Internet scale apps
  - Geographical proximity to user base.
  - Potentially large user base
- Apps with elastic demands
  - For most services peak demand exceeds average by factor of 2 to 10apps
- Burst computing
  - Testing
  - Bach processing

Cloud Types

- Public
- Private
- Hybrid



Cloud Computing Taxonomy

- Infrastructure/hardware as a Service (IaaS)
- Platform as a Service (PaaS)
- Software/Framework as a Service (SaaS/FaaS)

Infrastructure as a service (IaaS)

- Also know as Utility Computing
- VM images of different OS flavors
- Can run any application natively on chosen OS
- User administers VMs
  - User in charge of elasticity and failover
- Almost like buying your how HW
- Billing:
  - CPU hour
  - Gigabyte of storage per month
  - Network traffic in and out of datacenter
- Examples:
  - Amazon Elastic Compute Cloud (EC2), Rackspace

Example: Amazon EC2 Instance Types

- **Small Instance**
  2 GB memory
  1 EC2 Compute Unit
  I/O Performance: Moderate

- **XXLarge Instance**
  32 GB memory
  8 EC2 Compute Units
  I/O Performance: High

- **GPU Compute**
  732 GiB of memory
  64 EC2 Compute Units
  NVIDIA K80 GPUs
  I/O Performance: Very High (10 Gigabit Ethernet)

- One EC2 Compute Unit provides the equivalent CPU capacity of a 1.0-1.2 GHz 2007 Opteron or 2007 Xeon processor.

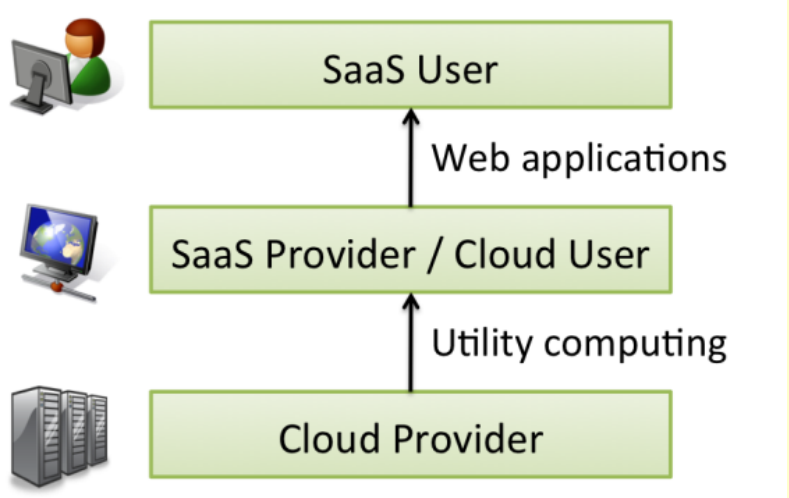
Platform as a Service (PaaS)

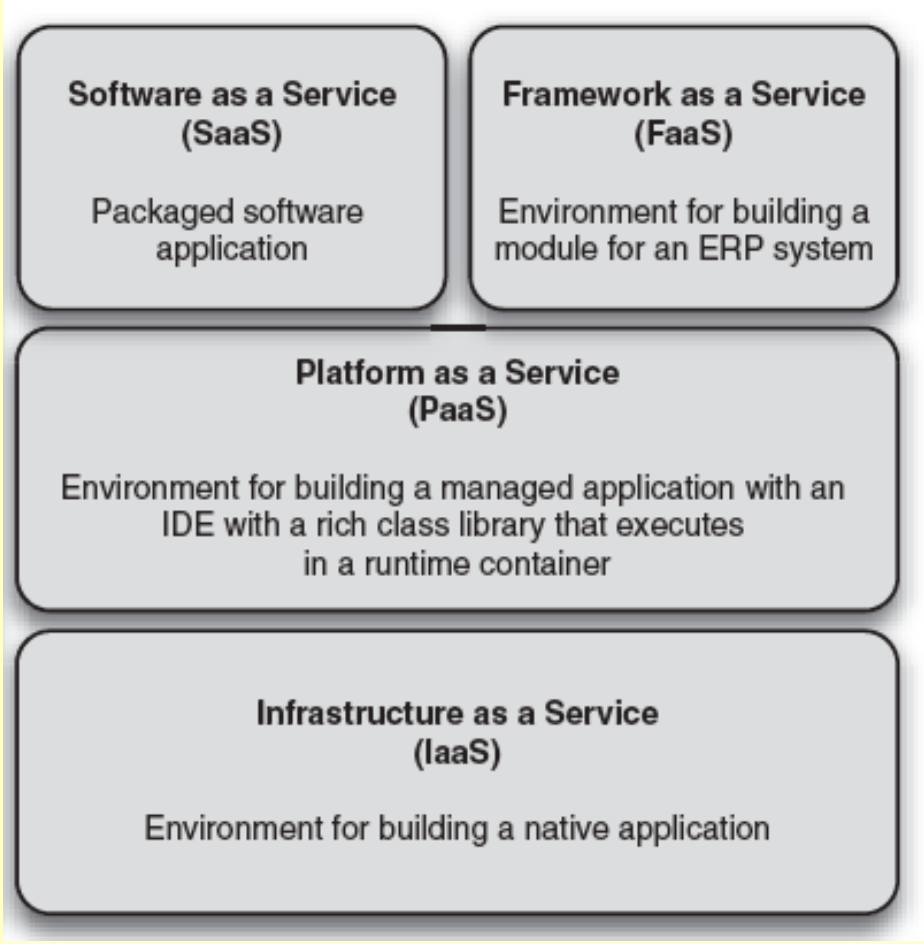
- Application need to conform to provider's API
- VM managed automatically by provider
  - Automatic scaling and failover
- Billing:
  - Per-request
  - Gigabyte of storage per month
  - Network traffic in and out of datacenter
- Examples:
  - Google App Engine
  - Microsoft Azure
  - AWS Lambda

Software/Framework as a Service (SaaS/FaaS)

- SaaS: services and applications available on an on-demand basis.
- FaaS: a configurable SaaS
- Billing:
  - Application specific
  - Per hour
  - Per transaction
- Examples:
  - Salesforce.com
  - Google Docs

How do they stack up?



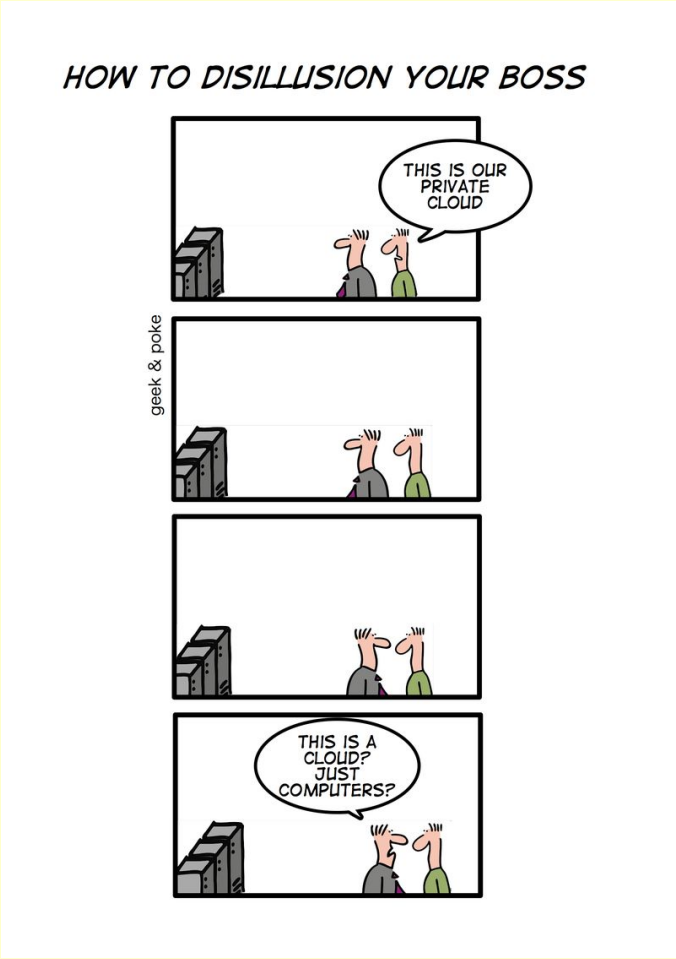


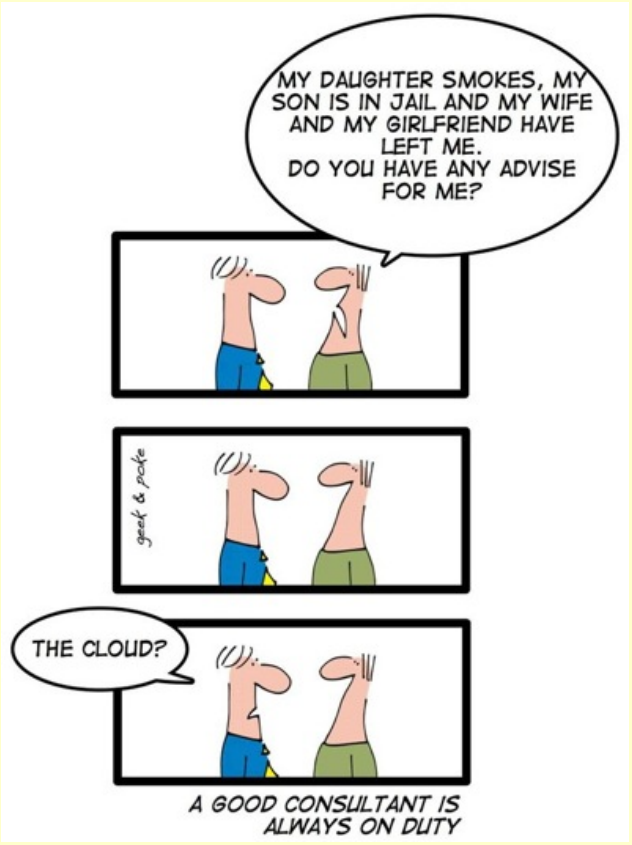
Choosing between IaaS and PaaS

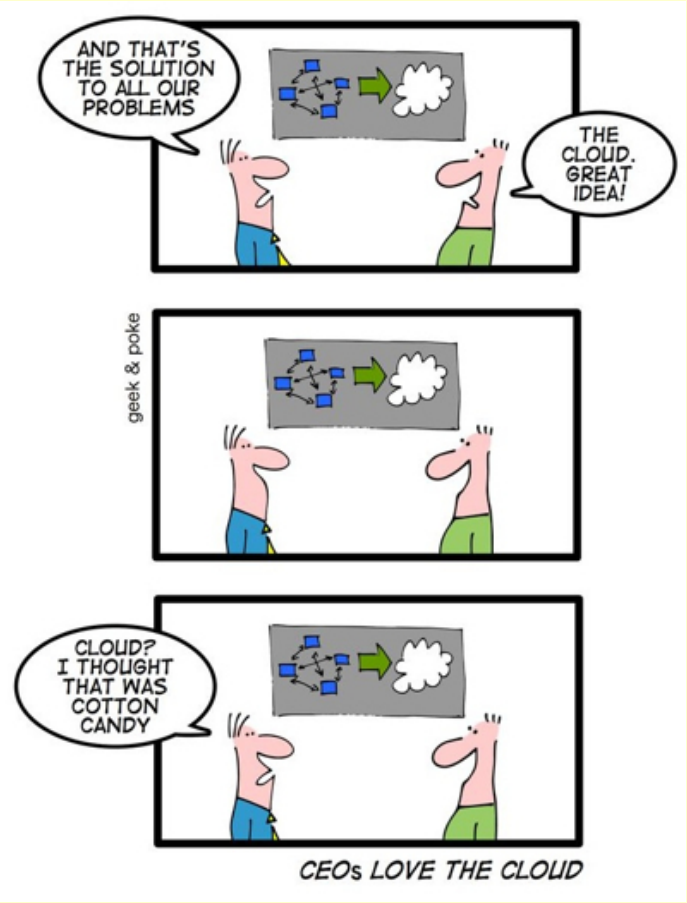
- IaaS
  - Use third party sw or other programming languages
  - Have existing code
  - Care about lock in
    - Transfer web app to your own server or a different provider
  - Want complete control
  - Are OK with the extra work needed to manage the system

- PaaS
  - Provider's API provider all needed functionality
  - Have no issues with lock-in
  - Don't want to or cannot afford to manage the servers

Questions?


HOW TO DISILLUSION YOUR BOSS
THIS IS OUR PRIVATE CLOUD
THIS IS A CLOUD? JUST COMPUTERS?
geek & poke


MY DAUGHTER SMOKES, MY SON IS IN JAIL AND MY WIFE AND MY GIRLFRIEND HAVE LEFT ME. DO YOU HAVE ANY ADVISE FOR ME?
geek & poke
THE CLOUD?
A GOOD CONSULTANT IS ALWAYS ON DUTY


AND THAT'S THE SOLUTION TO ALL OUR PROBLEMS
THE CLOUD. GREAT IDEA!
geek & poke
CLOUD? I THOUGHT THAT WAS COTTON CANDY
CEOs LOVE THE CLOUD

Week 2: Python Overview

- Prepare

  Option 1:

  An Introduction to Computer Science Using Python 3

  By Paul Gries, Jennifer Campberll, Jason Montojo

  Read chapters 1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 14

  Option 2:

  PCRS

  View videos in *prepare* sections for week 2