Jet Propulsion Laboratory
California Institute of Technology

Cloud 101

Mike Gangl, Caltech/JPL, michael.e.gangl@jpl.nasa.gov

Outline

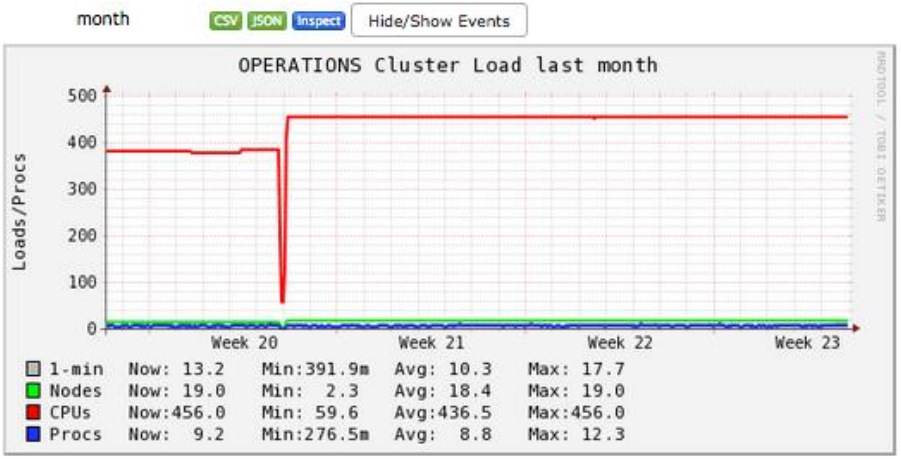
- What is cloud computing?
- Cloud service models
- Deployment models
- Storage
- Use cases
- Economics
- Caveats and issues

What is cloud computing?

- 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. –NIST

What is cloud computing?

- Key Features:
  - On-demand self service
  - Broad network access
  - Resource pooling
  - Rapid elasticity
  - Measured service
- It is not:
  - Just like mainframes
  - Just server virtualization
  - Just a really, really big cluster
  - A panacea

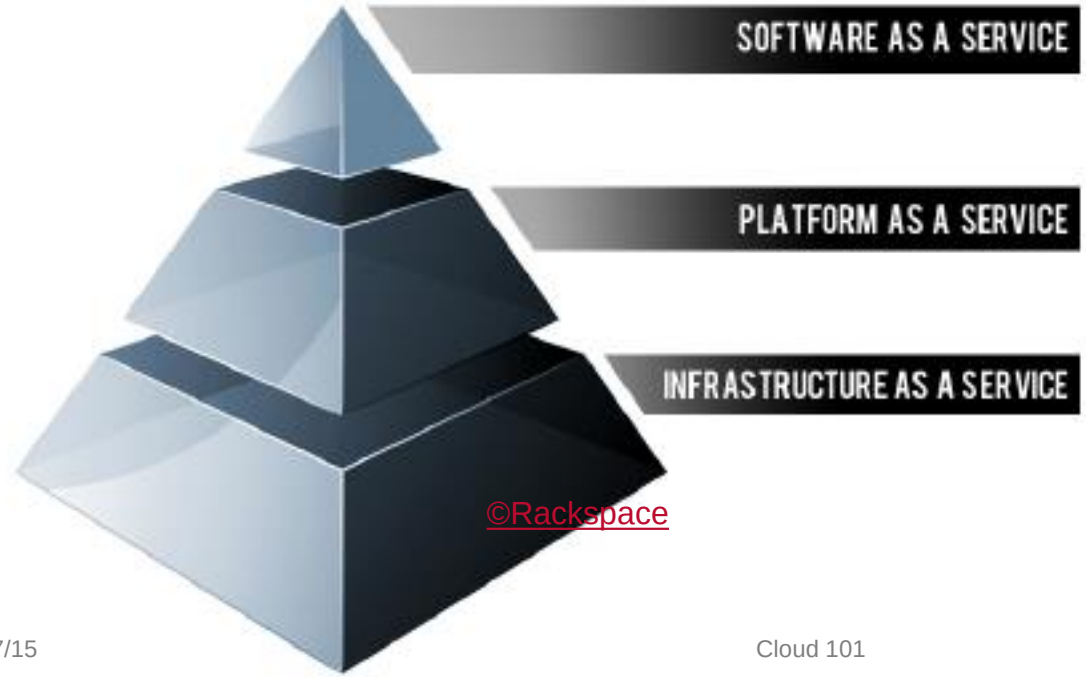


What is cloud computing?

Service Models

Cloud Services Models

- Clouds offer three models of service:
  - Software-as-a-Service (SaaS)
  - Platform-as-a-Service (PaaS)
  - Infrastructure-as-a-Service (IaaS)



Cloud Service Models

- Software-as-a-Service
  - Application deployed over the internet
  - Most visible to end users
  - Great for large user bases or burst usage
    - Gmail, flickr, Dropbox
    - Burst usage examples:
      - TurboTax: Busiest from January to April
      - CBS: NCAA March Madness Streaming
      - JPL: MSL Landing

Cloud Service Models

- Platform-as-a-Service
  - Suite of services to create web applications quickly
  - Infrastructure and software is provided
    - Databases, user information/authentication, messaging, load balancing, scalability
  - Targets software development
  - Examples
    - Google App Engine
    - AWS Elastic Beanstalk

Cloud Service Models

- Infrastructure-as-a-Service
  - Users allocate resources, often virtualized, such as servers, network, IP address and disk space
  - Variable cost based on size/capacity of resources
  - Allocation can be on-demand, reserved, or ‘spot’ instance pricing.
  - Replaces the need of buying hardware. Easily “upgradeable”.
  - Examples
    - Amazon Web Services
    - Rackspace

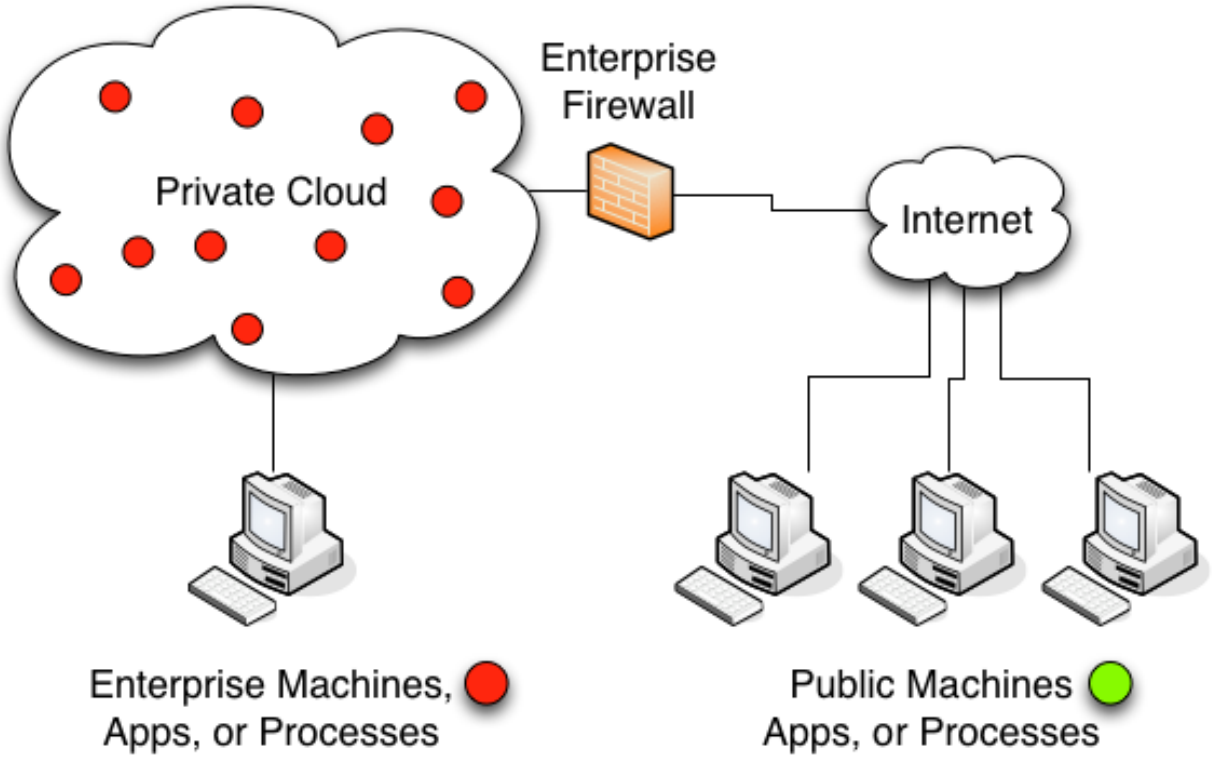
Deployment Models

Deployment models

- Deployment depends on:
  - Sensitivity of data/software
  - Use case being addressed
  - Costs
- Cloud deployments come in 4 models
  - Private
  - Public
  - Hybrid
  - Community

Cloud Deployment Models

- Private Cloud

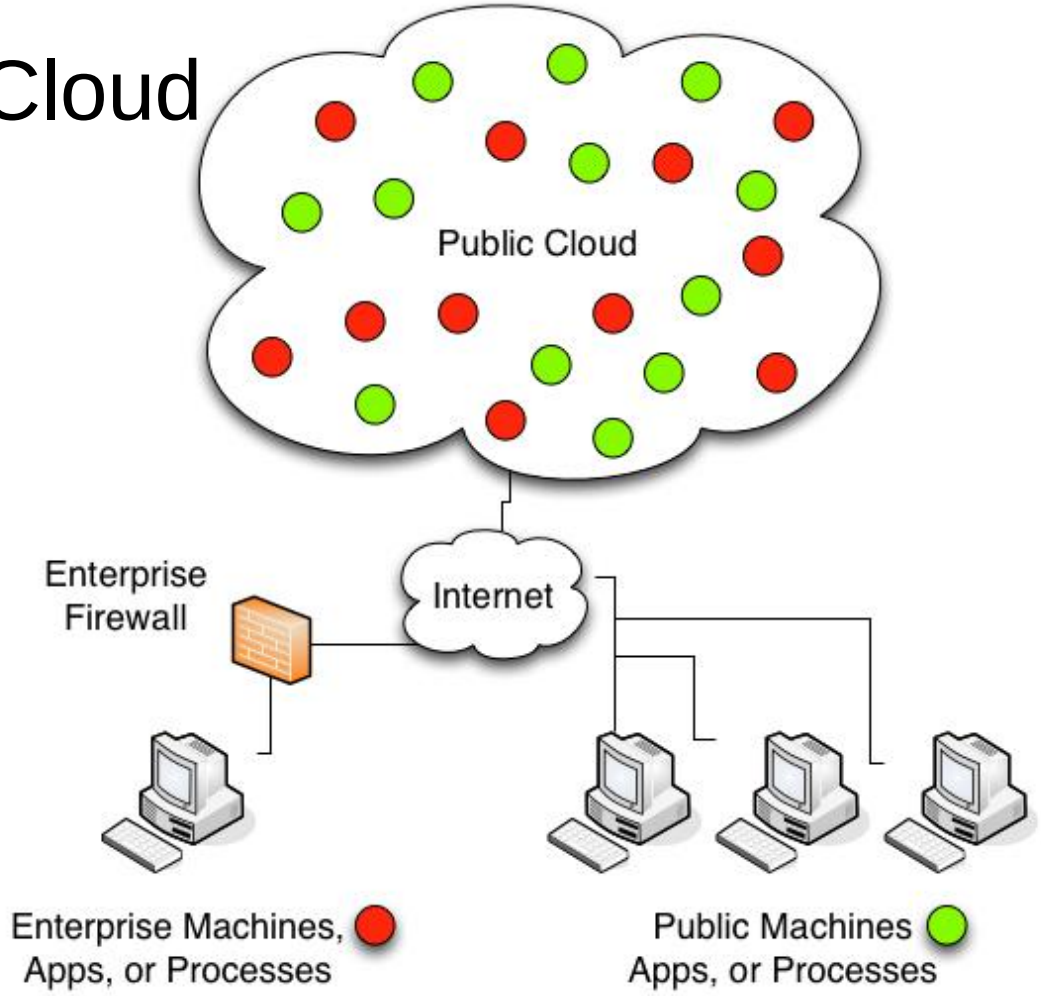


Cloud Deployment Models

- Private Cloud
  - Cloud infrastructure usually managed locally (intranet vs internet)
  - Resources are not shared outside of the cloud owner
  - Useful when security is the key issue:
    - Access within local firewalls or through secure VPN
    - SOX, HIPPA, trade secrets
  - Or based on other needs
    - Public cloud SLAs aren’t guaranteeing enough
  - Beware:
    - Cloud technology companies are here today, gone tomorrow
    - Upgrading cloud stacks (Openstack, Eucalyptus) is not trivial
    - Users don’t like downtime

Cloud Deployment Models

- Public Cloud

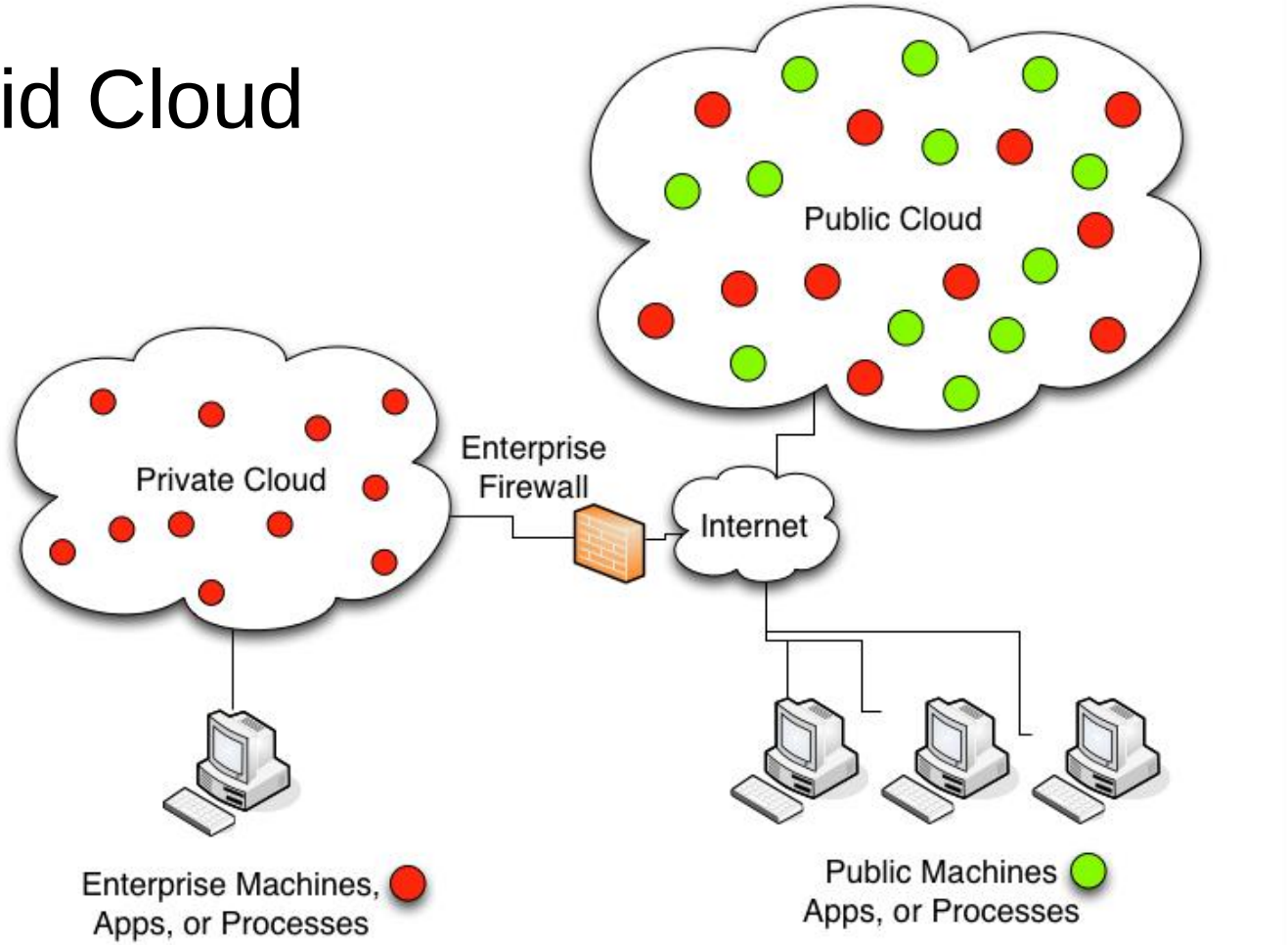


Cloud Deployment Models

- Public Cloud
  - Cloud infrastructure provided by third party; Quintessential cloud (internet vs intranet)
  - Provider hosts multiple clients
  - Useful for many applications:
    - Provide SAAS applications
    - Handle load spikes
    - Test out server/hardware architectures without large capital investments
  - Examples:
    - Amazon
    - Rackspace
    - Google

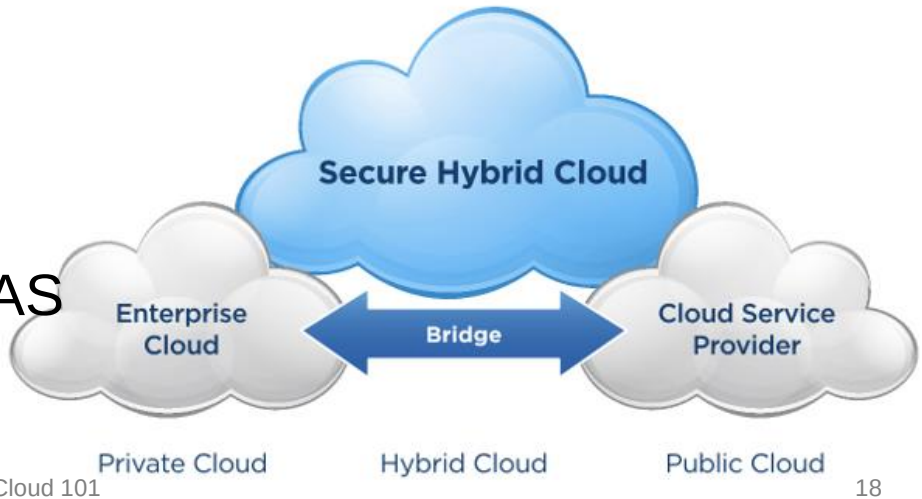
Cloud Deployment Models

- Hybrid Cloud



Cloud Deployment Models

- Hybrid Cloud
  - "Best of both worlds" model
  - On premise cloud for normal load and functionality to reduce nominal cost
  - "cloud bursting" to public cloud to handle large demands spikes
  - Needs manual or automatic orchestration
  - Examples
    - Microsoft Azure
    - (Sales)Force.com PAAS



Cloud Deployment Models

- Community Cloud
  - "private cloud"++
  - Jointly owned private cloud by groups with similar compliance and policy considerations
  - Examples
    - Local first response agencies might share a cloud that handles route, hospital and population information instead of creating their own
    - ESDIS cloud for all DAACs to use

Storage

Storage

- Storage in the cloud has different properties
  - Instance Storage
  - Volume Storage
  - Object Storage
- Public clouds generally make you pay for storage and download costs, not input cost
- Storage has differing speeds of access

Storage

- Instance Storage
  - Disk space that comes with procured instances
  - Exists only for the life of the instance!!!
  - Free with instance pricing
- Volume (Block) Storage
  - Most like a traditional hard drive
  - Amazon EBS, Rackspace CBS, Open stack ‘Block Storage’, Eucalyptus volumes)
  - Can be attached to different instances, but only one at a time
  - Data persists on the device even when not in use

Storage

- Object Storage
  - Network accessible persistent storage (usually http/ReST)
  - Amazon S3, Rackspace Cloud Files, Eucalyptus Buckets, OpenStack Object store
  - Scalability and redundancy built in
  - Private and/or public access
  - Economy of scale cost savings

Storage

- Object Storage Continued
  - Some providers have multiple tiers (AWS and Google most notably)
  - Normal Object Storage ($.03/GB example only)
  - Reduced Redundancy ($.024/GB)
    - Make sure you know what is meant by redundancy
  - Archive Storage ($.01/GB)
    - Long access times (up to 4 hours!)

Storage

	Instance Storage	Volume/Block Storage	Object Storage
Speed	High	Medium/High	Medium/Low
Redundancy	No	No	Yes
Chance of data loss	High*	low	low
Cost	Free**	High	Low
Scalable	No	No	Yes

* All data in instance store is lost when the instance crashes, dies, or is terminated
** usually this comes free with each instance, but amount of space varies

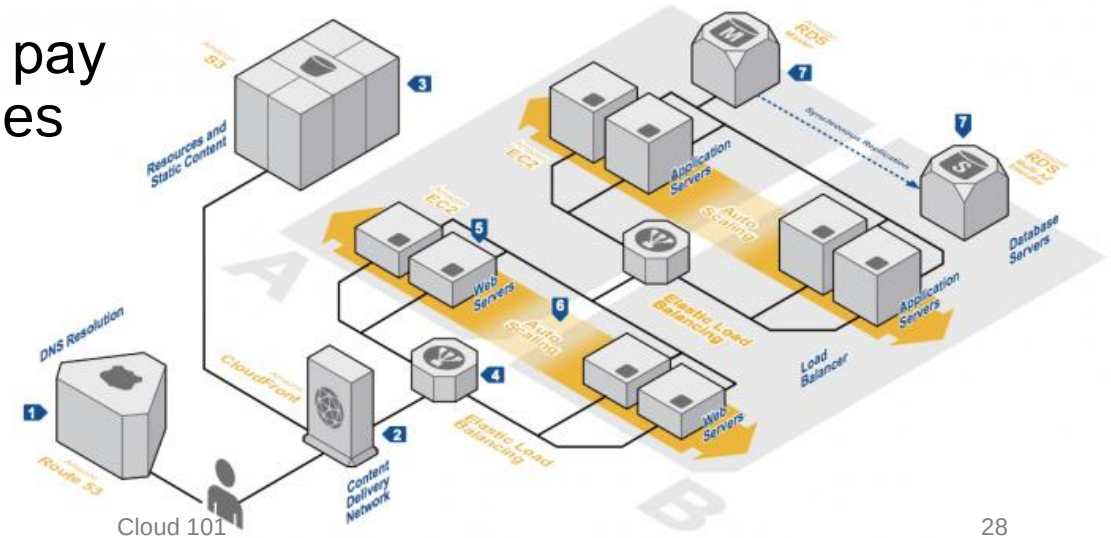
Use Cases

Use cases

- What are common use cases for the cloud?
  - Mitigating load spikes
  - Scalability
  - Batch processing
  - Data Locality

Use cases

- Load Spikes & Scalability
  - Websites and Content distribution
    - Load balancers in front of expandable bank of ‘servers’
    - Many “cloud” technologies (e.g. NoSQL databases) come easy-to-scale as data volume increases
    - TurboTax
  - Storage Capacity
    - “infinite” storage, pay as usage increases
    - Allow “unlimited” download bandwidth



Use cases

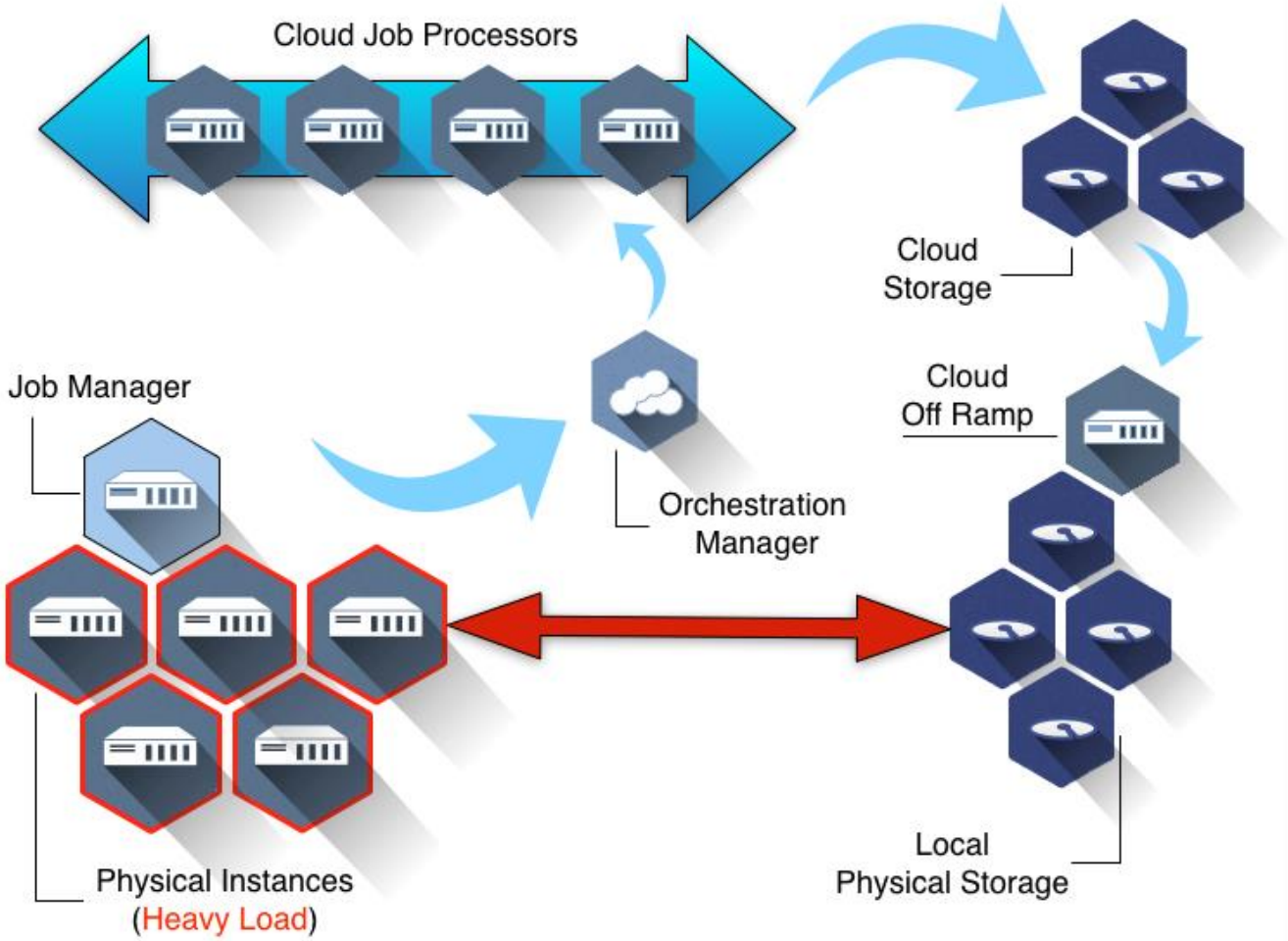
- Batch Processing
  - Provision hardware for one-off processing
  - Save results back on institutional hardware
  - Examples
    - Backfilling Metadata or running analysis on millions of files might take months. Trade $ for time and have it done in days.
    - CPU Intensive tasks; e.g. data mining
    - Large Map/Reduce frameworks jobs
    - eHarmony & your perfect match

Use cases

- Data Locality
  - Data volume is expanding, users can no longer have the entire dataset on their own machine or even a single server
  - Bring the processing to the data!
    - Store all data on cloud storage
    - Have users pay for instances (local to the data) to do their work
    - Provide tools in the environment to assist users (search, visualize, datamine)

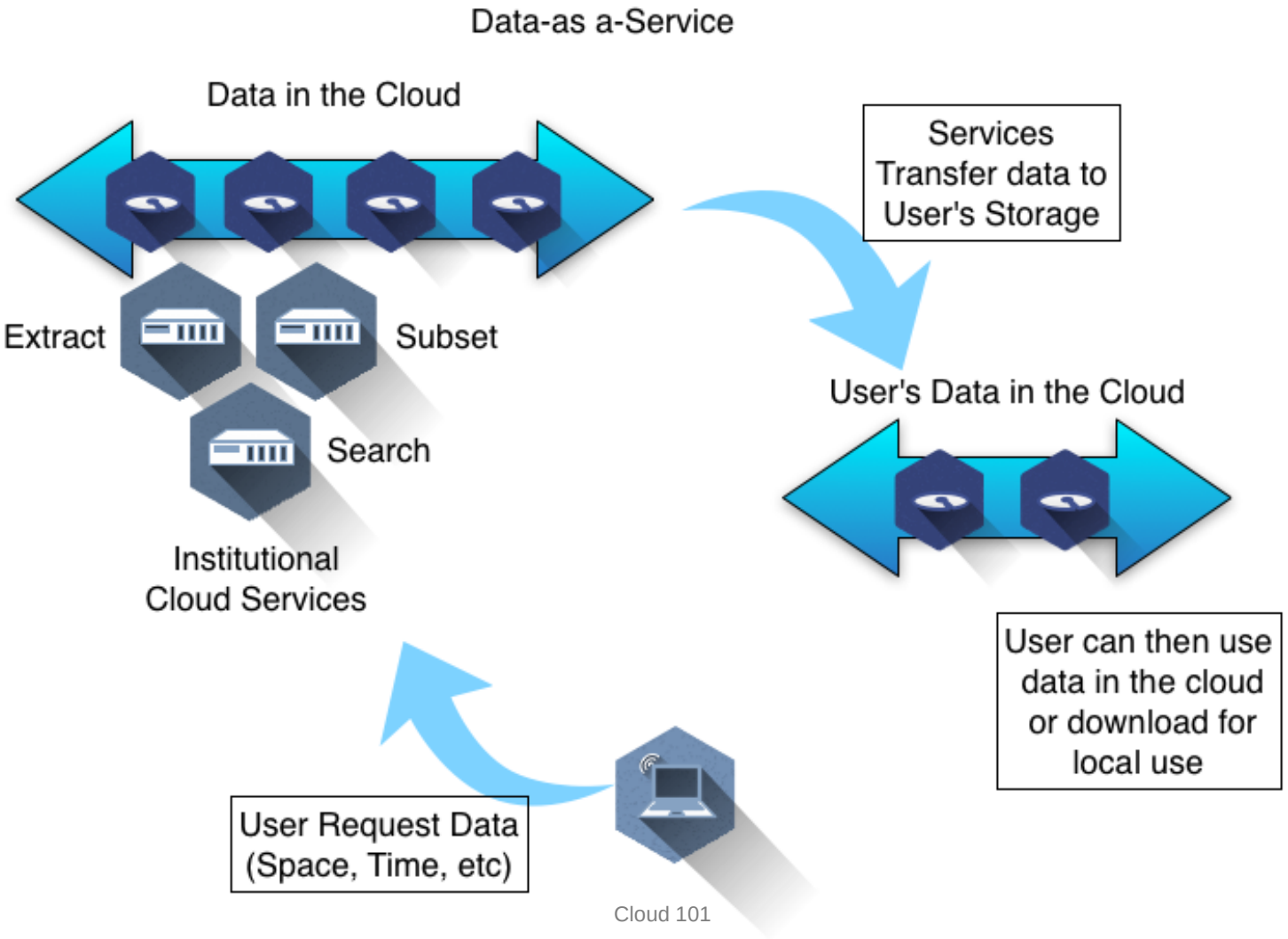
Use Cases

- Dealing with heavy loads (Reprocessing)



Use Cases

- Data-as-a-Service



Cloud Economics

Economics

- Where “We can do that?” meets “It’s going to cost how much?”
- Take advantage of economies of scale
- Cloud economics inherently tricky:
  - Can’t look only at capital costs
    - Admins
    - Facility charges (power, cooling)?
    - Redundancy
    - Failure scenarios
- We will focus mainly on public clouds

Economics

- Instances and servers
  - On demand pricing (Pay as you go)
    - Pay more for better hardware
    - $1.20/hour/instance
  - Reserve pricing
    - Flat reservation fee + lower hourly
    - Pay for a collection of resources up front, use and deploy however you'd like
  - Spot pricing
    - Bid on pricing instances
    - If "rate" falls below your bid, your instances start. They stop when the rates exceed bid
    - Good for non-timely processing
- Storage/Object Stores
  - Pay per GB sometimes in, always out
  - Bulk storage "discounts"
  - Different Storage Tiers

Economics

- Other costing
  - Some providers aren't as flexible as Amazon or Rackspace
    - Monthly pricing for a set allocation of resources
    - Use it or lose it
  - Bare metal clouds
    - More expensive, but more predictable performance
  - Who needs SAs?
    - I do.

Economics

- S3 costs

	Standard Storage	Reduced Redundancy Storage	Glacier Storage
First 1 TB / month	$0.0300 per GB	$0.0240 per GB	$0.0100 per GB
Next 49 TB / month	$0.0295 per GB	$0.0236 per GB	$0.0100 per GB
Next 450 TB / month	$0.0290 per GB	$0.0232 per GB	$0.0100 per GB
Next 500 TB / month	$0.0285 per GB	$0.0228 per GB	$0.0100 per GB
Next 4000 TB / month	$0.0280 per GB	$0.0224 per GB	$0.0100 per GB
Over 5000 TB / month	$0.0275 per GB	$0.0220 per GB	$0.0100 per GB

Issues and Caveats

Caveats

- All those cool services being offered…
  - Quickly allow you to build scalable apps
  - Lock you in to their custom tools and APIs
- Unchecked, costs can become a huge issue
  - Your datasets become extremely popular and users are downloading terabytes each day… you're paying for that!
  - Not using the XL1.xlarge instance to it's fullest? You're paying for those unused cycles!

Caveats

- Developer paradigm shifts
  - Must refactor programs and applications to scale
  - Plan for failures
  - NIH*. Does. Not. Work.
- The Promised Land
  - Cloud orchestration and dynamic scalability aren’t always easy
  - One touch solutions will vary from provider to provider

*Not Invented Here

Questions?

Jet Propulsion Laboratory
California Institute of Technology

jpl.nasa.gov

Backup Slides

- Useful links
  - Understanding the Cloud Computing stak (IaaS, PaaS, SaaS) by Rackspace
  - http://aws.amazon.com/architecture/
  - http://www.ibm.com/cloud-computing/in/en/what-is-cloud-computing.html
  - http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf [NIST Definition]
  - http://aws.amazon.com/s3/pricing/
  - http://aws.amazon.com/ec2/pricing/
  - http://www.rackspace.com/en-us/cloud/public-pricing