Cloud Computing & Big Data

PARALLEL & SCALABLE MACHINE LEARNING & DEEP LEARNING

Prof. Dr. – Ing. Morris Riedel
Adjunct Associated Professor
School of Engineering and Natural Sciences, University of Iceland
Research Group Leader, Juelich Supercomputing Centre, Germany

LECTURE 8

Infrastructure-As-A-Service (IAAS)

November 1st, 2018
Room Stapi 108

UNIVERSITY OF ICELAND
SCHOOL OF ENGINEERING AND NATURAL SCIENCES
FACULTY OF INDUSTRIAL ENGINEERING, MECHANICAL ENGINEERING AND COMPUTER SCIENCE

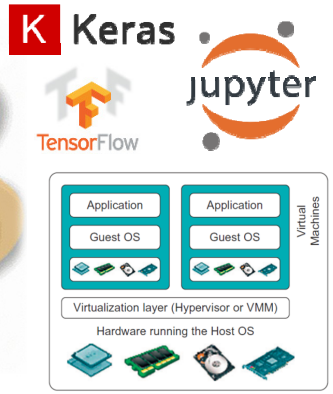
JÜLICH Forschungszentrum

HELMHOLTZ
RESEARCH FOR GRAND CHALLENGES

DEEP Projects

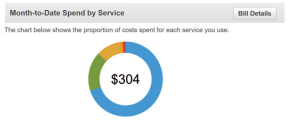
Review of Lecture 7 – Deep Learning Applications in Clouds

- Deep Learning in Clouds



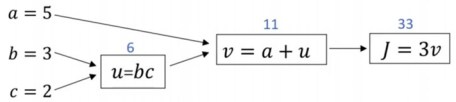
GPU Instances - Current Generation

p3.2xlarge	8	26	61 GiB	EBS Only	$3.06 per Hour
p3.8xlarge	32	94	244 GiB	EBS Only	$12.24 per Hour
p3.16xlarge	64	188	488 GiB	EBS Only	$24.48 per Hour
p2.xlarge	4	12	61 GiB	EBS Only	$0.90 per Hour
p2.8xlarge	32	94	488 GiB	EBS Only	$7.20 per Hour

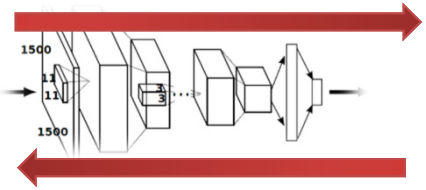


[5] AWS EC2 Pricing

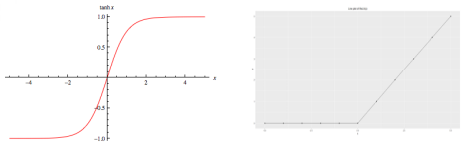
- Advanced Concepts



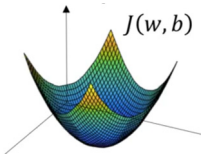
(using weights)



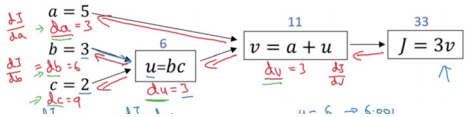
(updating the weights)



(activation functions)

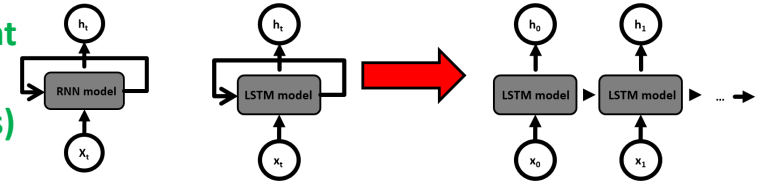


(optimization: e.g. stochastic gradient descent)



(Long Short-Term Memory for sequence data)

(Recurrent Neural Networks)



[2] Keras Python Deep Learning Library *[4] Jupyter Web page* *[6] Coursera Deep Learning Specialization*

[1] Distributed & Cloud Computing Book *[3] Tensorflow Deep Learning Framework* *[7] Mount Esja in Iceland*

Outline of the Course

1. Cloud Computing & Big Data
2. Machine Learning Models in Clouds
3. Apache Spark for Cloud Applications
4. Virtualization & Data Center Design
5. Map-Reduce Computing Paradigm
6. Deep Learning driven by Big Data
7. Deep Learning Applications in Clouds
8. Infrastructure-As-A-Service (IAAS)
9. Platform-As-A-Service (PAAS)
10. Software-As-A-Service (SAAS)
11. Data Analytics & Cloud Data Mining
12. Docker & Container Management
13. OpenStack Cloud Operating System
14. Online Social Networking & Graphs
15. Data Streaming Tools & Applications
16. Epilogue

+ additional practical lectures for our hands-on exercises in context

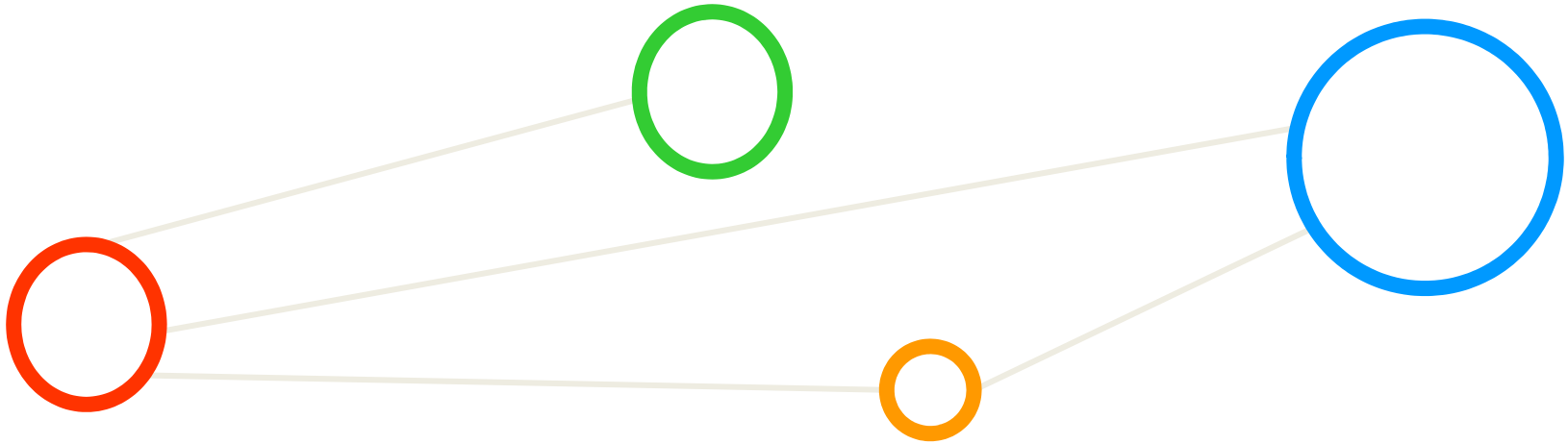
- Practical Topics
- Theoretical / Conceptual Topics

Outline

- Understanding IAAS Provisioning
  - Different Cloud Service Levels
  - Amazon Web Services (AWS) Examples
  - AWS Management Console & EC2 Instance
  - Configuration and Launching EC2 Instances
  - Payment Models & vCPU/Memory pricing
- Advanced IAAS Topics
  - Cloud Storage with Amazon S3 & EBS & EFS
  - AWS Marketplace & Users
  - Openstack and IAAS Deployment Options
  - Qstack Software & Advania Cloud Services
  - GoGrid/DataPipe & Enomaly/Virtustream

- **Promises from previous lecture(s):**
- ***Lecture 1:*** **Lecture 8 & 9 & 10 offer more insights into concrete cloud systems and their architectures today**
- ***Lecture 1 & Practical Lecture 5.1:*** **Lecture 8 provides more details about Amazon EC2 and its Infrastructure-as-a-Service models**
- ***Lecture 1:*** **Lecture 8 provides more details about Amazon S3 and its Infrastructure-as-a-Service models**
- ***Lecture 3:*** **Lecture 8 provides more detailed information about AWS products with Infrastructure-As-A-Service**
- ***Lecture 4:*** **Lecture 8 & 9 & 10 offer more insights into concrete cloud systems and their use of virtualization**
- ***Lecture 5 & Practical Lecture 5.1:*** **Lecture 8 will provide more examples using Infrastructure-As-A-Service (IAAS) & AWS Cloud Services**
- **Lecture 7: Lecture 8 will provide more pieces of information about IAAS features of AWS EC2 & others**

Understanding IAAS Provisioning



Internet Cloud Systems – Examples – Revisited (cf. Lecture 1)

- Selected Cloud Systems (aka 'Clouds') known today
  - Google AppEngine **(massive computing & storage & applications)**
  - Amazon Web Service **(massive computing & storage)**
  - Facebook **(online social networking & advertisement)**
  - SalesForce.com **(customer relationship management)**
  - Rackspace **(managed cloud provider & hosting)**
  - IBM Bluemix **(cloud platform)**
  - Enomaly **(elastic computing cloud)**

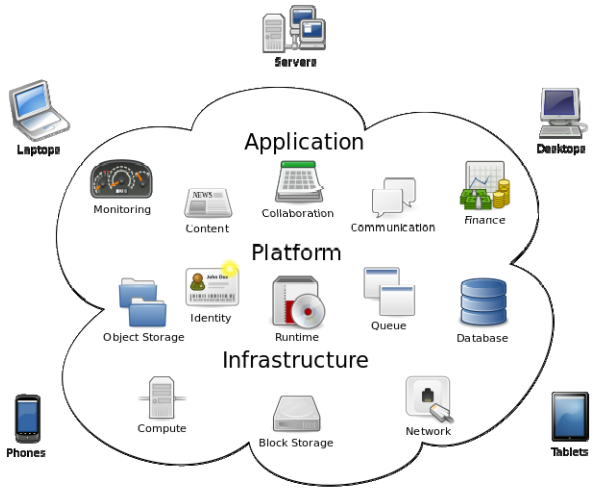

- Cloud systems play an increasingly important role in upgrading traditional Web services and Internet applications

[1] Distributed & Cloud Computing Book

- **Large-scale Internet applications have significantly enhanced the quality of life in society today**
- **Internet Cloud Systems often show overlaps in provisioning of various services (e.g. computing, storage, etc.), but also often have some unique selling proposition (USP) to pick one or another**

Three Levels of Cloud Service Models: *AAS

- Levels oriented towards different users
  - Full customization to direct usable applications
- Software as a Service (SAAS)
  - Provides specific 'ready-to-run applications'
  - Sometimes related to geographical location
- Platform as a Services (PAAS)
  - Virtual images ready to deploy your software
  - Includes a 'platform for creation of your services'



Cloud Computing

[8] Wikipedia 'Cloud computing'

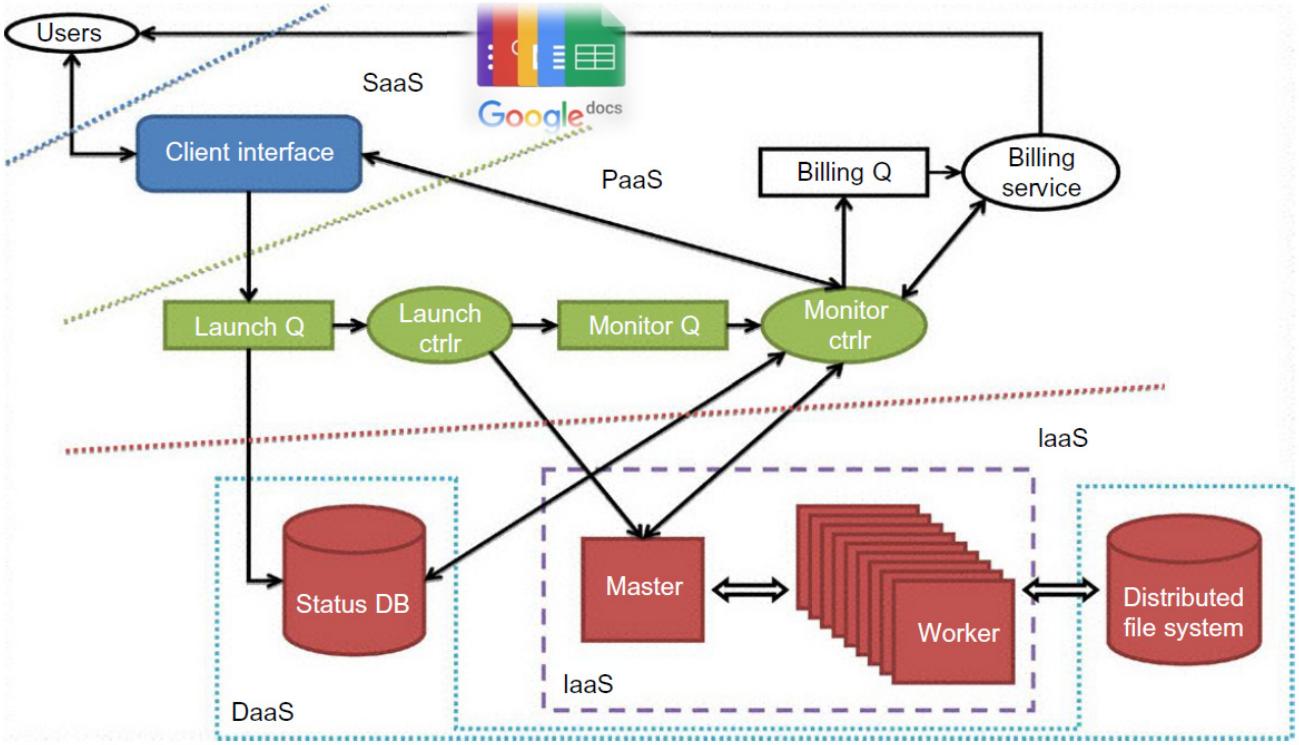
- Infrastructure as a Service (IAAS) — **focus in this lecture**
  - Provides 'bare metal infrastructure' & virtual IT resources (cf. Lecture 4)
  - Use and tune infrastructure as needed (compute, storage, networking, …)

- **Cloud computing infrastructures typically offer services on three different levels: Infrastructure as a Service (IAAS), Platform as a Service (PAAS), and Software as a Service (SAAS)**

Different Cloud Service Models – SAAS

- Software-As-A-Service (SAAS)
  - E.g. services for business processes, consumer applications related to geographical locations, …

- **The Conceptual ideas and key usage of the SAAS cloud service model include consume applications, avoid installations, special interfaces & ready-to-run applications**
- **SAAS is based on easy accessible software remotely accessed via Web browsers and centrally hosted in data centers**

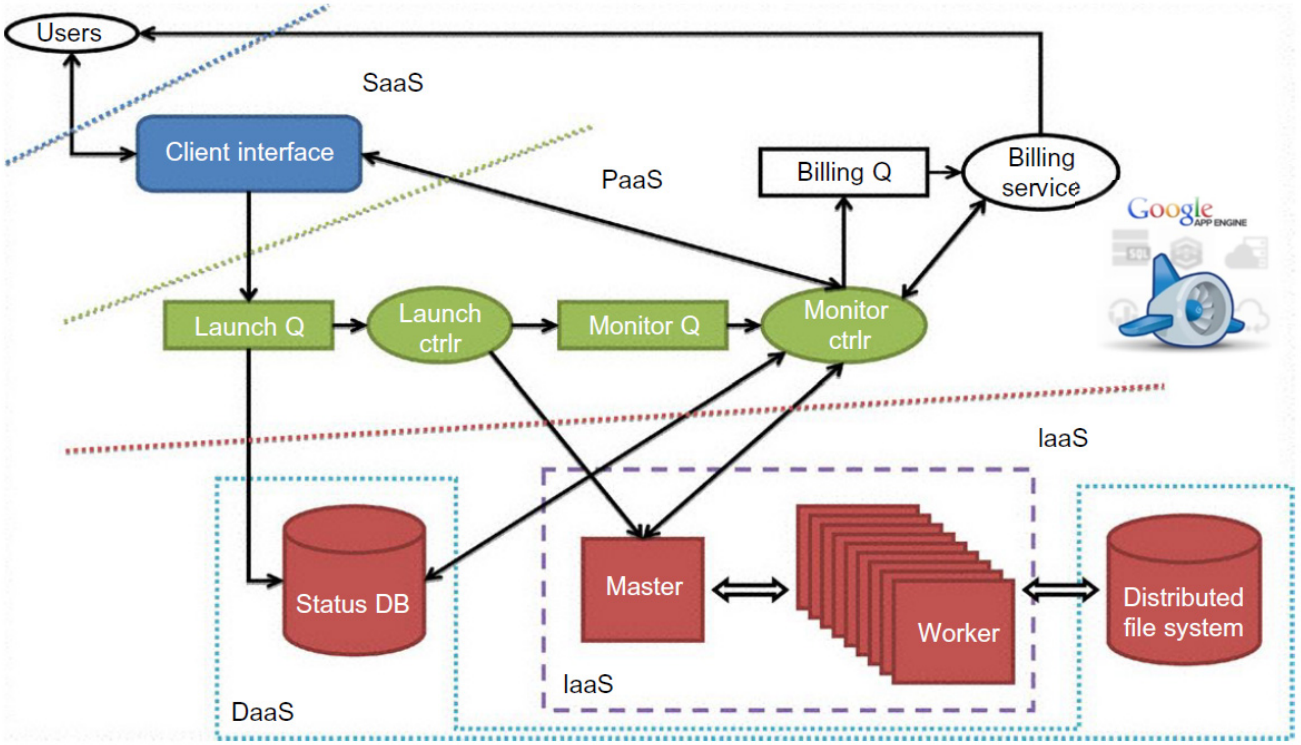


[1] Distributed & Cloud Computing Book

Different Cloud Service Models – PAAS

- Platform-As-A-Service (PAAS)
  - E.g. used to provision billing services, handle compute job queing, launching of images, and monitoring to support application developers

- **The Conceptual ideas and key usage of the PAAS cloud service model is building cloud applications with software development kits (SDKs) & application programming Interfaces (APIs) via basic services**
- **PAAS is based on known application frameworks similiar to ASP, J2EE, JSP and languages like Python, Java, Ruby, etc.**

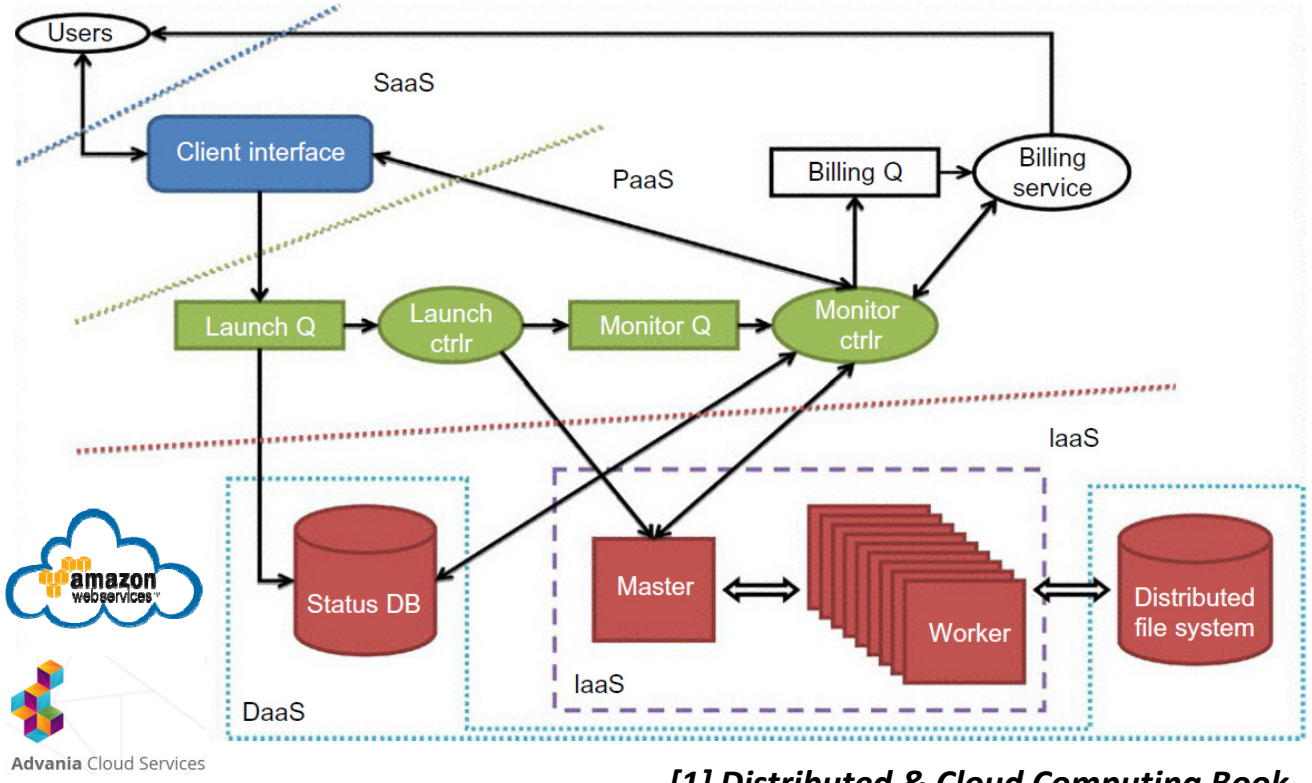


[1] Distributed & Cloud Computing Book

Different Cloud Service Models – IAAS

- Infrastructure-As-A-Service (IAAS)
  - E.g. used to provision databases (aka Database-As-A-Service), compute instances, distributed file systems, and storage to satisfy 'general' demands

- **The Conceptual ideas and key usage of the IAAS cloud service model include hosting, bare metal, and resource provisioning**
- **IAAS is based on virtual machines (cf. Lecture 4) that are used to flexible share computing and storage resources in a scalable and secure manner**



[1] Distributed & Cloud Computing Book

IAAS Provider Example – Amazon Web Services (AWS)

- On-Demand delivery of virtualized IT resources
  - Compute power, database storage, networking
  - Application services, databases, specific software
  - Offers ‘rapid access‘ to flexible and low cost IT resources
  - Access as many resources as you need (~instantly)
  - Using pay-as-you-go pricing (pay for what you use)
  - Provides tools for deployment and management of the infrastructure

[9] AWS Web page

- Selected Business Cases
  - Start-ups that need flexibility, have no capital investment, no experience
  - Run applications that share photos to millions of mobile end users
  - Support business critical operations (e.g. understanding customer needs)

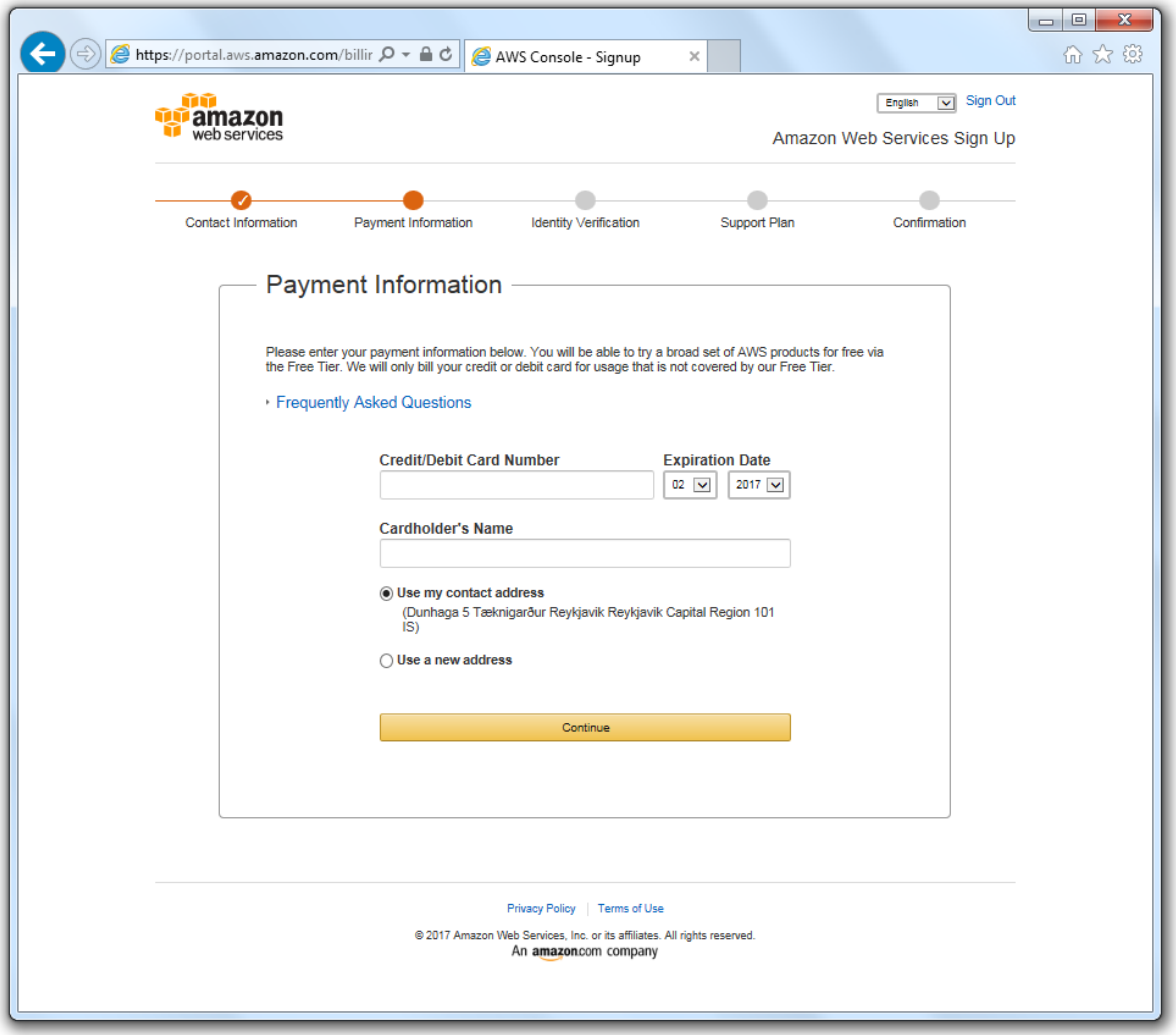
[10] Big Data Tips ‘Amazon Web Services’

- **Amazon Web Services (AWS) is a commercial provider using the IAAS cloud service model in order to provide specifically tuned IT resources with roughly 70 different cloud services to choose from**
- **AWS also provides a wide variety of other services using the PAAS and SAAS cloud service models**

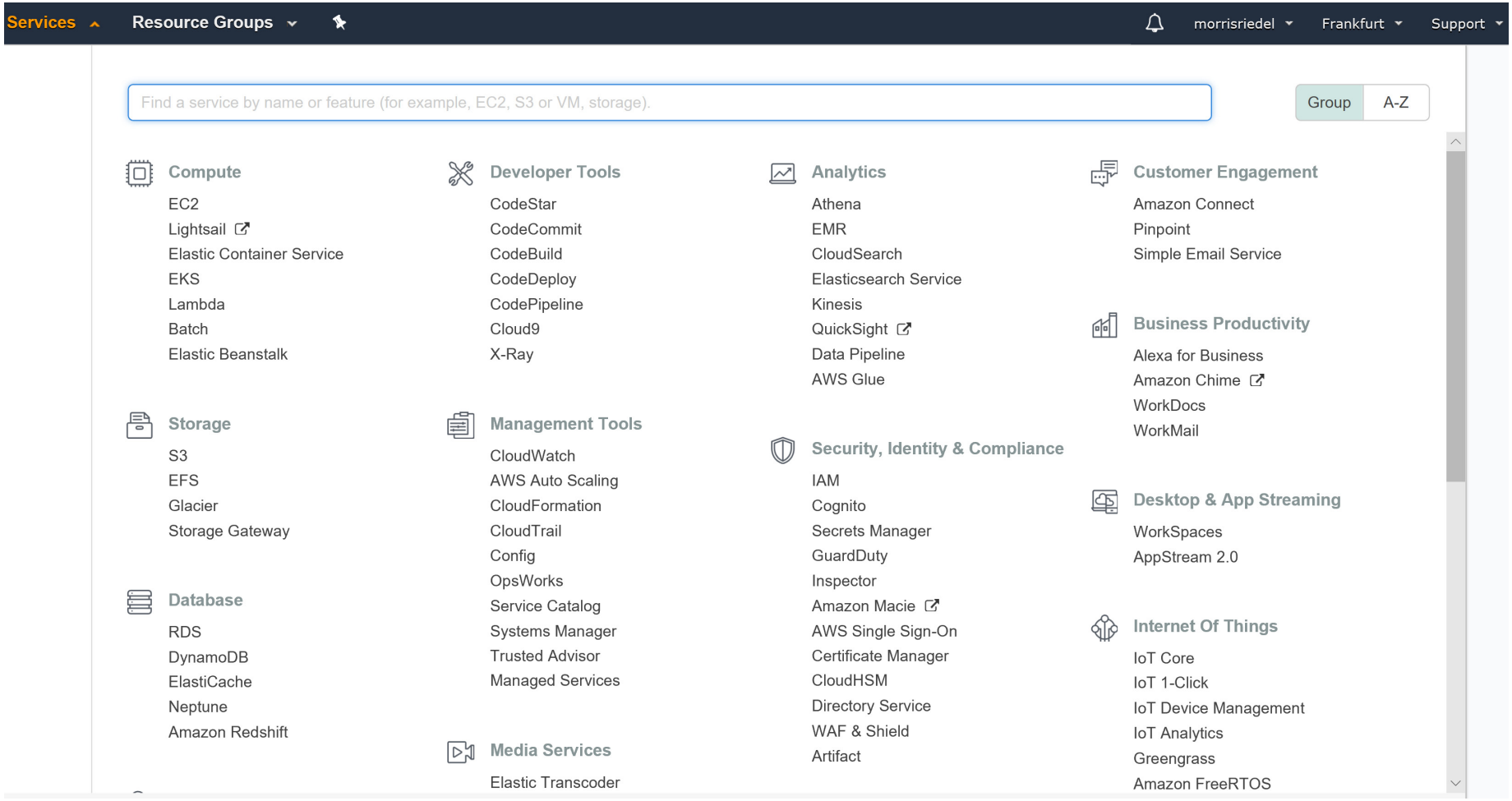
AWS Sign-Up Process requires Payment Information

- Account
  - Name, Email
  - Password
  - Address, Phone
  - Company or private person
- Payment(!)
  - Credit Card

  (informally known as 'swipe the credit card')
- Verification
  - Auto call & PIN
- Support plan
  - How you plan to use AWS



AWS Management Console – Services – Revisited



IAAS Provider Example – AWS Services Overview

Compute
- Amazon EC2
- Amazon EC2 Container Registry
- Amazon EC2 Container Service
- Amazon Lightsail
- Amazon VPC
- AWS Batch
- AWS Elastic Beanstalk
- AWS Lambda
- Auto Scaling
- Elastic Load Balancing

Storage
- Amazon Simple Storage Service (S3)
- Amazon Elastic Block Storage (EBS)
- Amazon Elastic File System (EFS)
- Amazon Glacier
- AWS Storage Gateway
- AWS Snowball
- AWS Snowball Edge
- AWS Snowmobile

Database
- Amazon Aurora
- Amazon RDS
- Amazon DynamoDB
- Amazon ElastiCache
- Amazon Redshift
- AWS Database Migration Service

Migration
- AWS Discovery Services
- AWS Database Migration Service
- AWS Server Migration Service
- AWS Snowball
- AWS Snowball Edge
- AWS Snowmobile

Networking & Content Delivery
- Amazon VPC
- Amazon CloudFront
- Amazon Route 53
- AWS Direct Connect
- Elastic Load Balancing

Developer Tools
- AWS CodeCommit
- AWS CodeBuild
- AWS CodeDeploy
- AWS CodePipeline
- AWS X-Ray
- AWS Command Line Interface

Management Tools
- Amazon CloudWatch
- Amazon EC2 Systems Manager
- AWS CloudFormation
- AWS CloudTrail
- AWS Config
- AWS OpsWorks
- AWS Service Catalog
- AWS Trusted Advisor
- AWS Personal Health Dashboard
- AWS Command Line Interface
- AWS Management Console
- AWS Managed Services

Security, Identity & Compliance
- AWS Identity and Access Management (IAM)
- Amazon Inspector
- AWS Certificate Manager
- AWS CloudHSM
- AWS Directory Service
- Amazon Cloud Directory
- AWS Key Management Service
- AWS Organizations
- AWS Shield
- AWS WAF
- AWS Artifact

Analytics
- Amazon Athena
- Amazon EMR
- Amazon CloudSearch
- Amazon Elasticsearch Service
- Amazon Kinesis
- Amazon Redshift
- Amazon QuickSight
- AWS Data Pipeline
- AWS Glue

Artificial Intelligence
- Amazon Lex
- Amazon Polly
- Amazon Rekognition
- Amazon Machine Learning

Mobile Services
- AWS Mobile Hub
- Amazon API Gateway
- Amazon Cognito
- Amazon Pinpoint
- AWS Device Farm
- AWS Mobile SDK

Application Services
- AWS Step Functions
- Amazon API Gateway
- Amazon Elastic Transcoder
- Amazon AppStream

Messaging
- Amazon SQS
- Amazon Pinpoint
- Amazon SES
- Amazon SNS

Business Productivity
- Amazon WorkDocs
- Amazon WorkMail

Desktop & App Streaming
- Amazon WorkSpaces
- Amazon AppStream 2.0

Software
- AWS Marketplace

Internet of Things
- AWS IoT Platform
- AWS Greengrass
- AWS IoT Button

Game Development
- Amazon Lumberyard

[9] AWS Web page

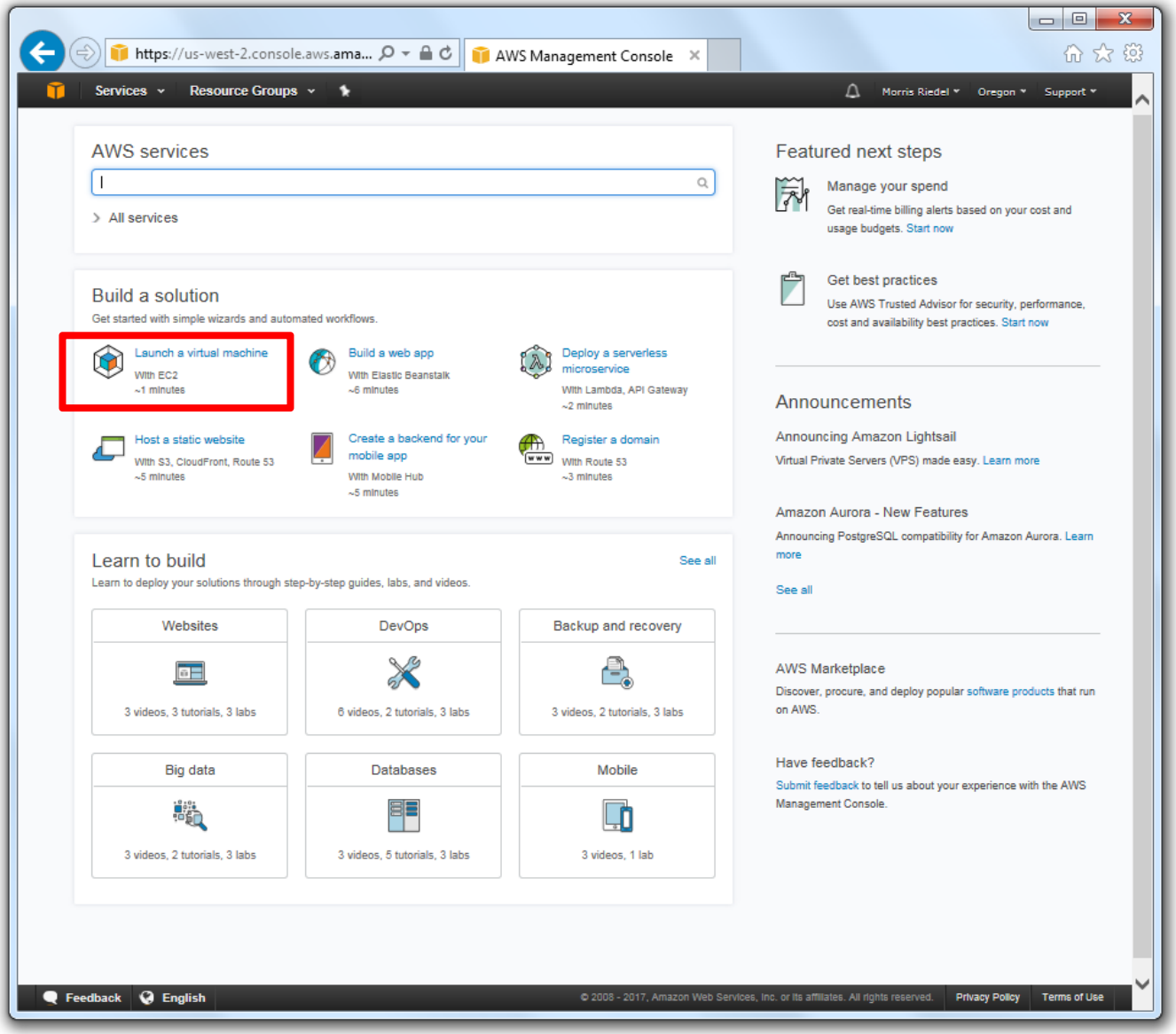
Understand AWS Cloud Service Portfolio – Compute

- Multiple compute products
  - Deploy, run, and scale applications as virtual servers, containers, or code
- Products & Usage
  - Amazon EC2: Virtual servers in the cloud
  - Amazon EC2 Container Registry: Store and retrieve Docker images
  - Amazon EC2 Container Service: Run & manage Docker containers
  - Amazon Lightsail: Launch and manage virtual private servers
  - Amazon VPC: Isolated cloud resources
  - AWS Batch: Run batch jobs at any scale
  - AWS Elastic Beanstalk: Run and manage Web Apps
  - AWS Lambda: Run your code in response to events
  - Auto Scaling: Automatic Elasticity

[9] AWS Web page

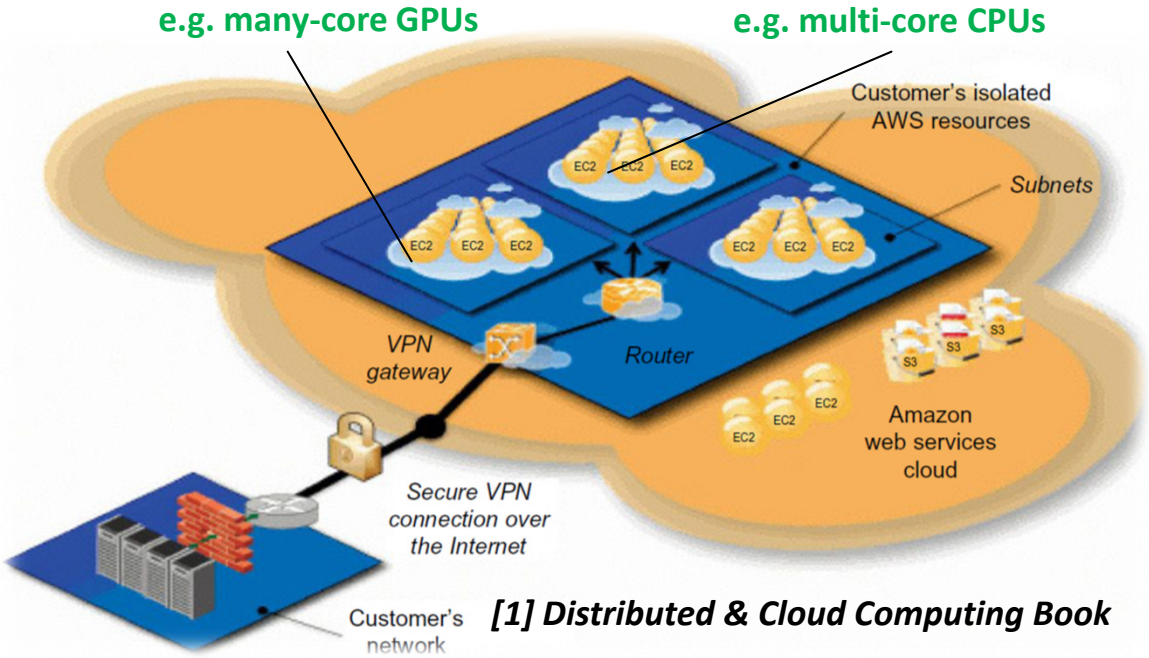
➢ **Lecture 12 provides more details about containers, images, and registries including tool Docker**

AWS Management Console



IAAS Compute Provisioning – AWS EC2 – Revisited

- Provides an elastic compute cloud (EC2)
  - Elastic load balancing services and so-called auto-scaling
  - E.g. great during peak times in business (e.g. x-mas shopping, etc.)
  - E.g. use exactly the right type and size of computing a start-up needs for newest bright ideas
  - E.g. New York Times use it to quickly retrieve pictorial information from millions of articles

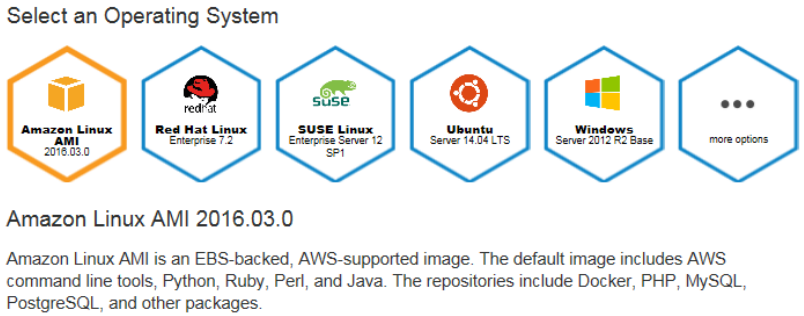


[1] Distributed & Cloud Computing Book

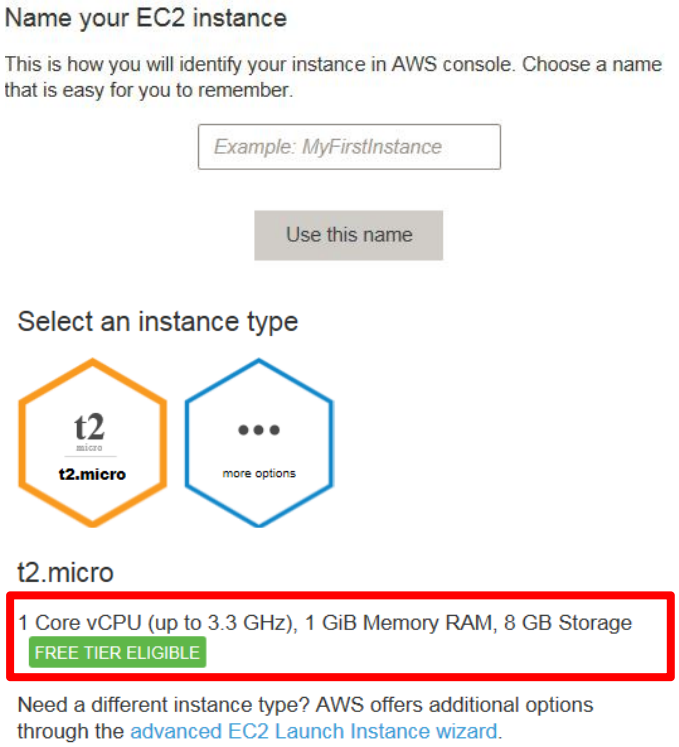
- **AWS offers the IAAS Amazon EC2 service that provides an elastic compute cloud (EC2) that ensures that a sufficient number of EC2 instances are provisioned to meet expected performance**

AWS Amazon EC2 Instances – Quick Launch Wizard

- Launch a EC2 virtual machine via console
  - E.g. Quick launch wizard lets one easily create an EC2 instance (AWS-recommended default configuration, advanced wizard available too)
- Quick launch wizard
  1. Provide EC2 instance name
  2. Select operating system

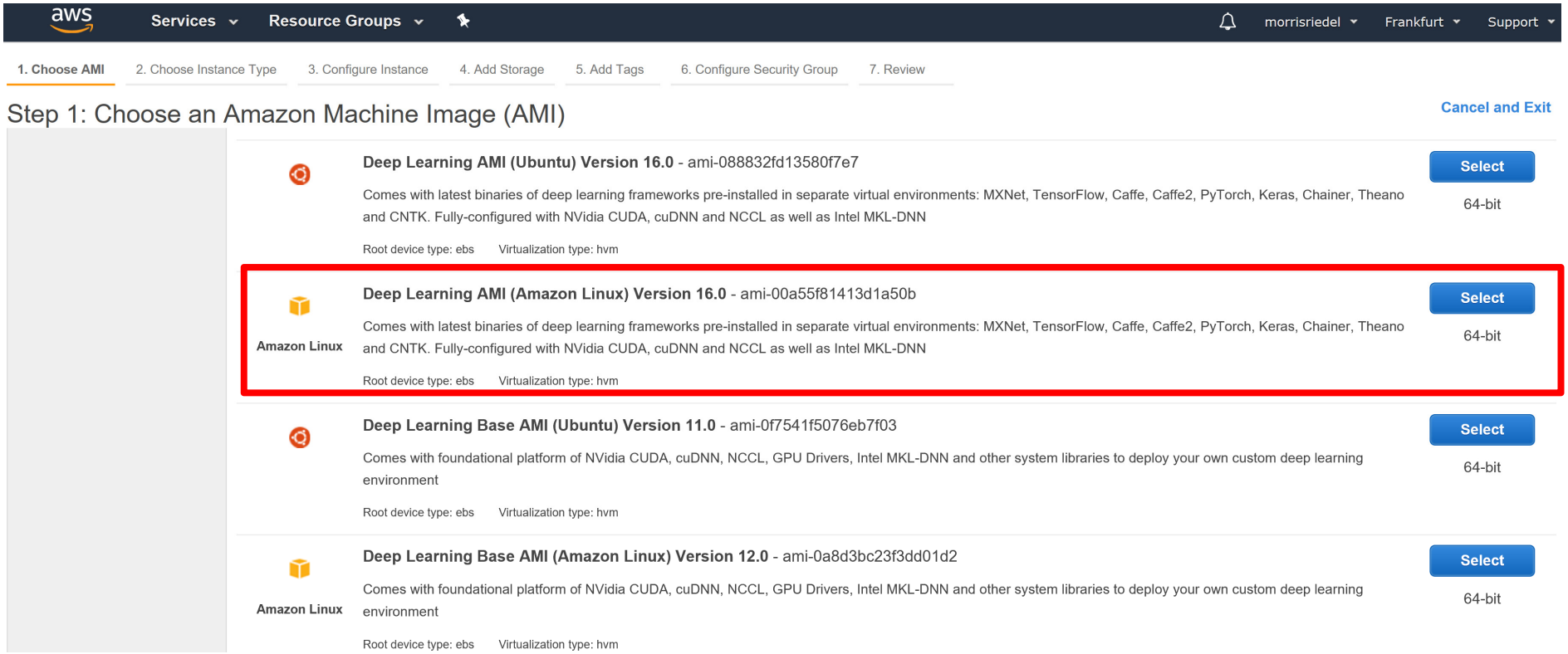


  3. Select EC2 instance type
  4. Create key pair: download private key
  5. Create particular EC2 instance



- **Amazon EC2 provides virtual machines (cf. Lecture 4) in the AWS cloud known as EC instances**

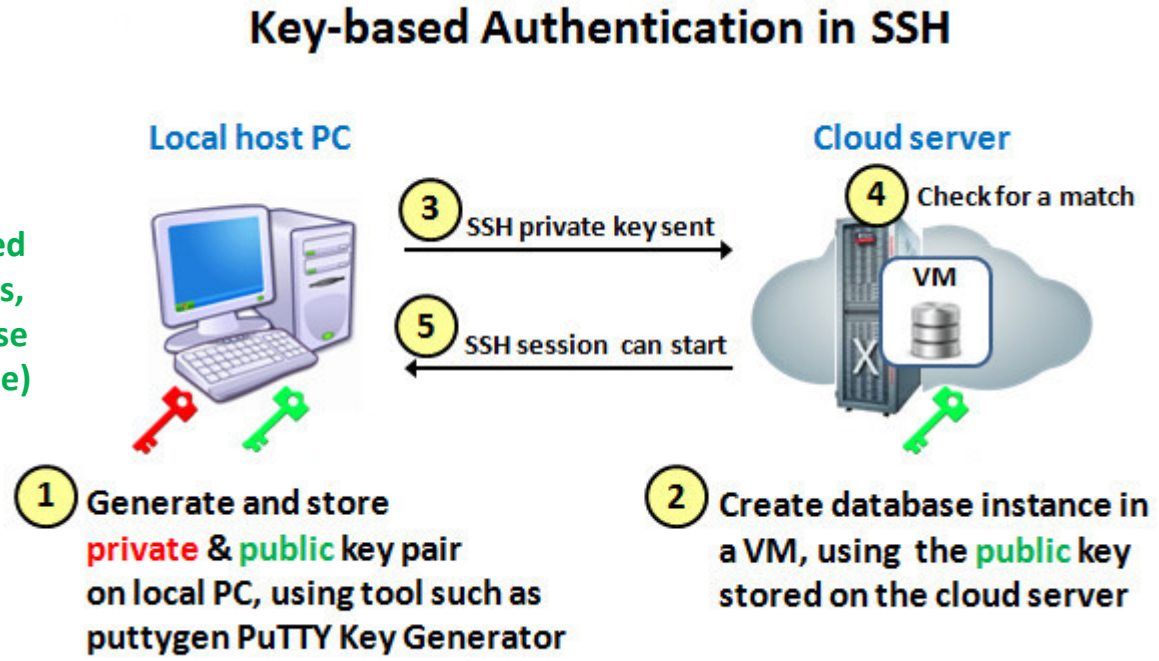
AWS EC2 – AMIs for Deep Learning (cf. Lecture 7)



- **Amazon offers pre-configured AMIs for deep learning consisting of preinstalled deep learning packages such as MXNet, TensorFlow, Caffee, Caffe2, PyTorch, Keras, Chainer, Theano, or CNTK**
- **Pre-configured AMIs for deep learning feature preinstalled GPGPU Nvidia CUDA & cuDNN libraries that usually requires a lot of efforts in installation and version checks with deep learning packages**

Secure IAAS Access – Private / Public Key Pairs in Detail

- Used for IAAS cloud users
  - Enables authentication (identity check w/o passwords)
  - Used with secure shell (SSH) connections between laptops & servers



(putty is an often used SSH tool for Windows, Linux distributions use SSH via command line)

[20] Key-Based Authentication

- **Private/Public key pairs are often used to enable the authentication of users without passwords**
- **Private keys remain on the laptops of the user and the public key is stored at the cloud server side**

AWS Amazon EC2 Instances – Launching EC2 Instances

- Launching
  - Might take a couple of minutes depending on selected options

cloudserverone

Status: In progress...

Your Instance is Launching!

Amazon EC2 is launching your instance. This process should only take a few minutes.
You can proceed to the EC2 Console while this process takes place.

cloudserverone

Status: Completed!

Quick Start Steps

- ✓ **Name**
- ✓ **Operating System**
- ✓ **Instance Type**
- ✓ **Private Key**

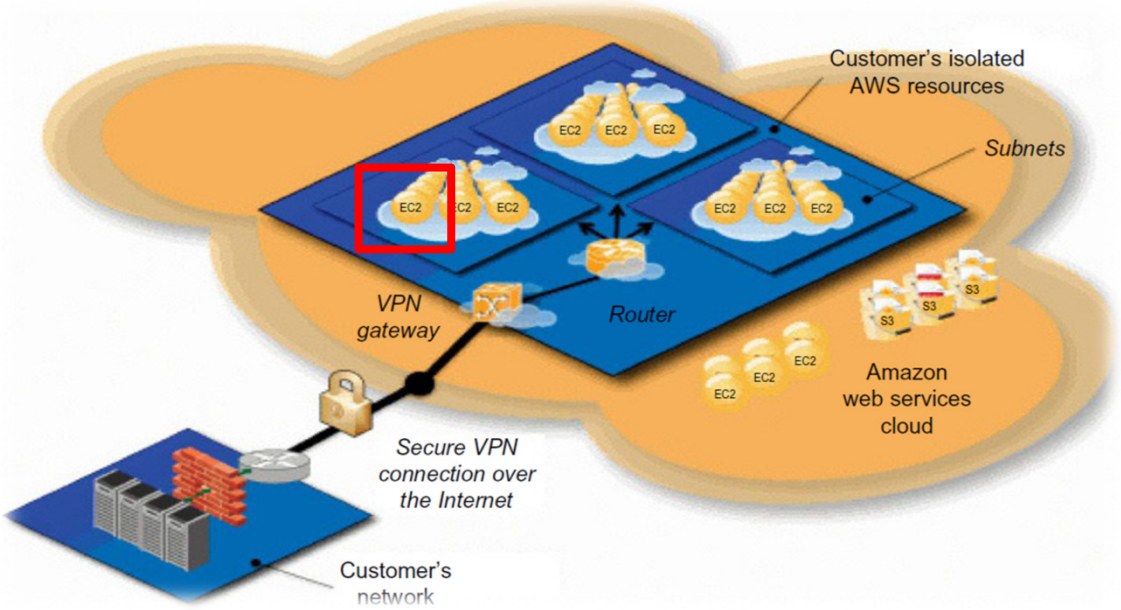
What we're launching for you

Your virtual machine will be set up using the following AWS services:

Virtual Machine
EC2 Instance

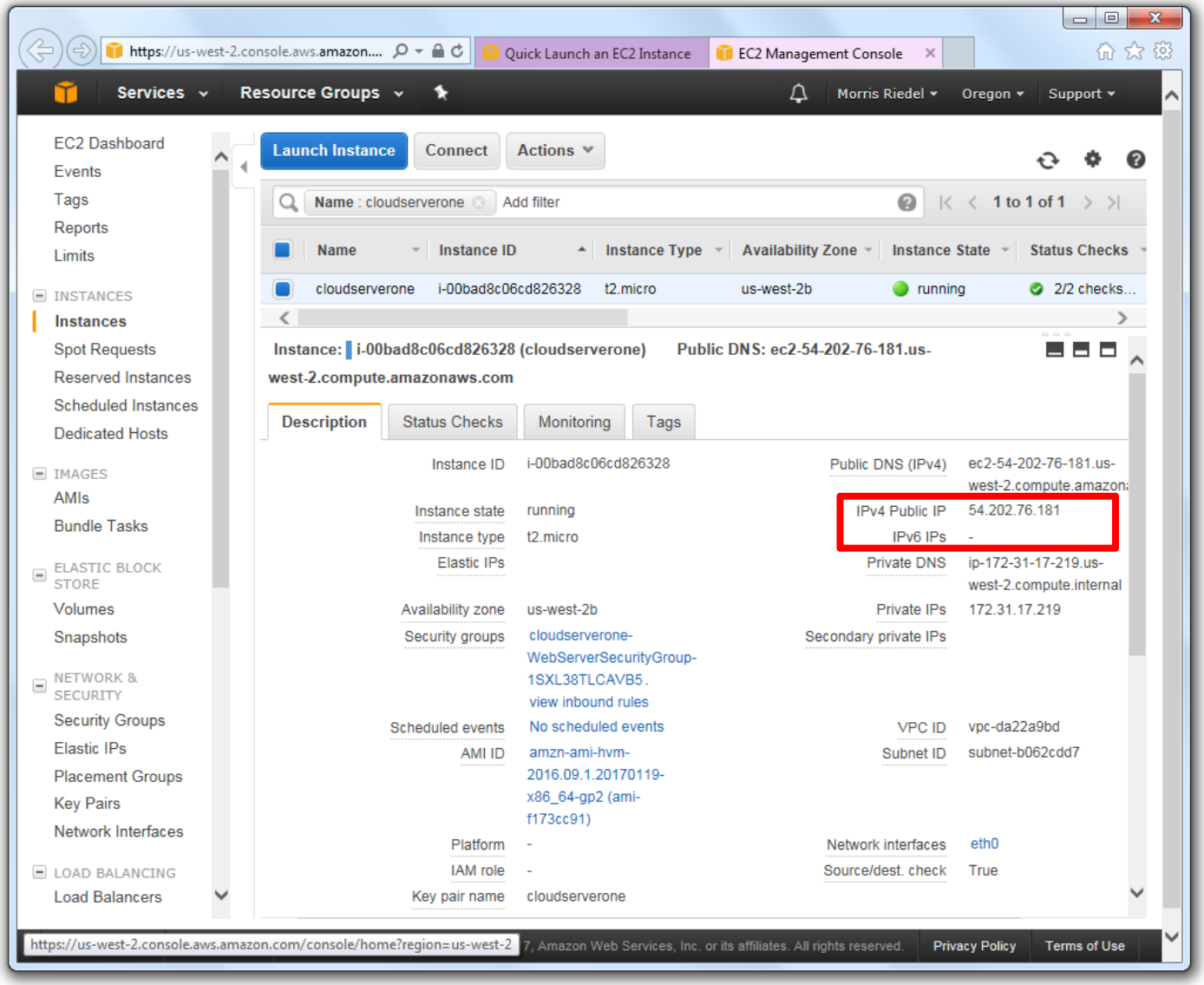
Storage
EBS Volume

Firewall
Security Group



[1] Distributed & Cloud Computing Book

AWS Amazon EC2 Instances – Manage EC2 Instances



Different Payment Models – Revisited

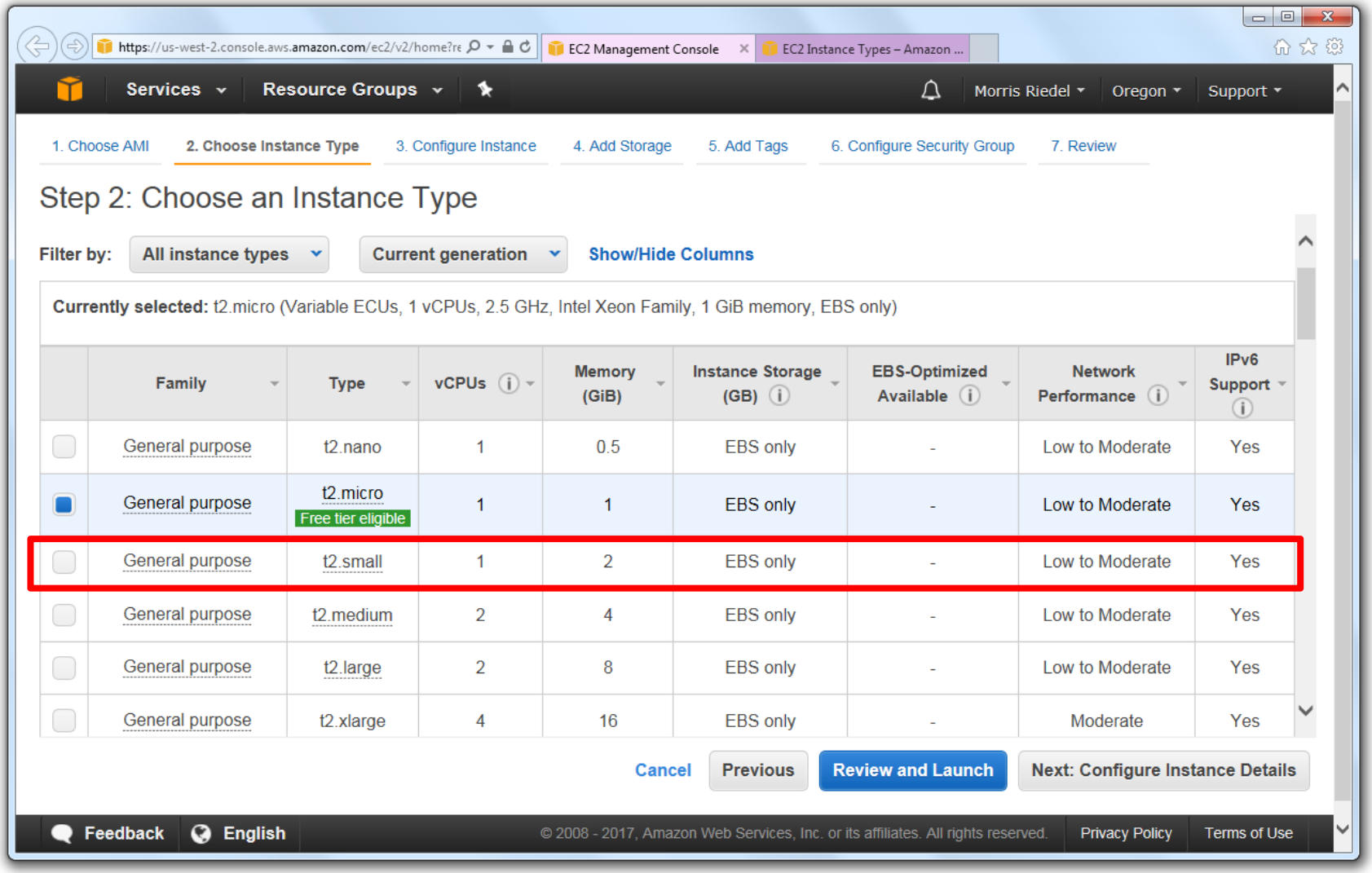
- Pay-as-you-go / pay-per-use model **(informally known as 'swipe the credit card')**
  - Cost is often significantly reduced vs. cost of ownership (not always)
  - All hardware and software resources are leased from cloud provider
  - Without capital investment on the part of the users
  - Example: Amazon Web Services (AWS) offers a pay-as-you-go model for pricing ~70 cloud services

Classical Computing	**Cloud Computing**
(Repeat the following cycle every 18 months)	(Pay as you go per each service provided)
Buy and own	**Subscribe**
Hardware, system software, applications to meet peak needs	- - - -
Install, configure, test, verify, evaluate, manage	**Use** (Save about 80-95% of the total cost)
- - - -	- - - -
Use	(Finally)
- - - -	**$ - Pay for what you use**
Pay $$$$$ (High cost)	based on the QoS

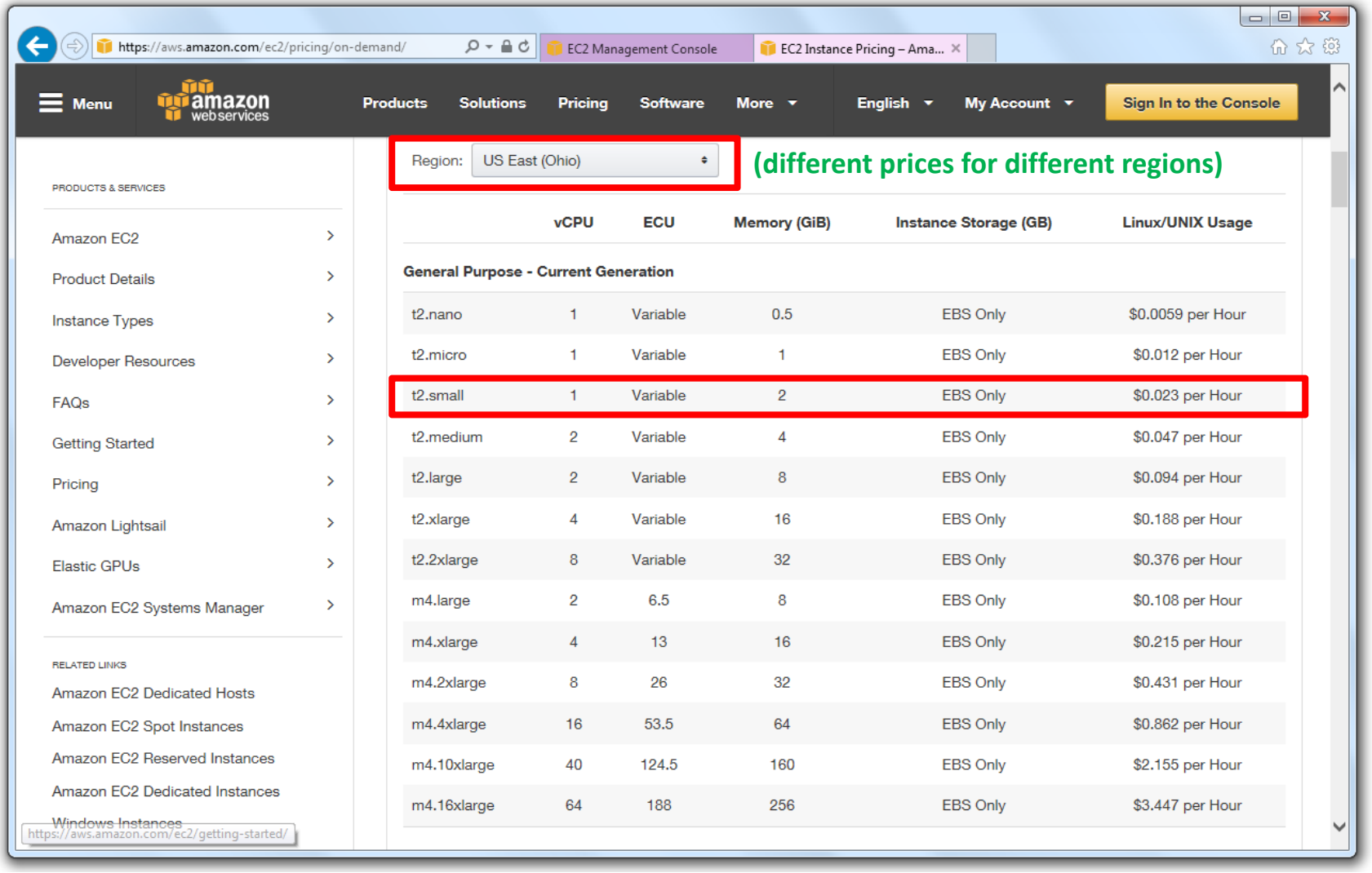
[1] Distributed & Cloud Computing Book

- **AWS uses four pricing models: 'on-demand' to pay for compute capacity by the hour, EC2 spot instances to bid on spare Amazon EC2 computing 90% off, reserved instances, dedicated hosts**

AWS EC2 Compute Payment Models – vCPUs/Memory



AWS EC2 Compute Payment Models – ‘On-Demand‘ Costs



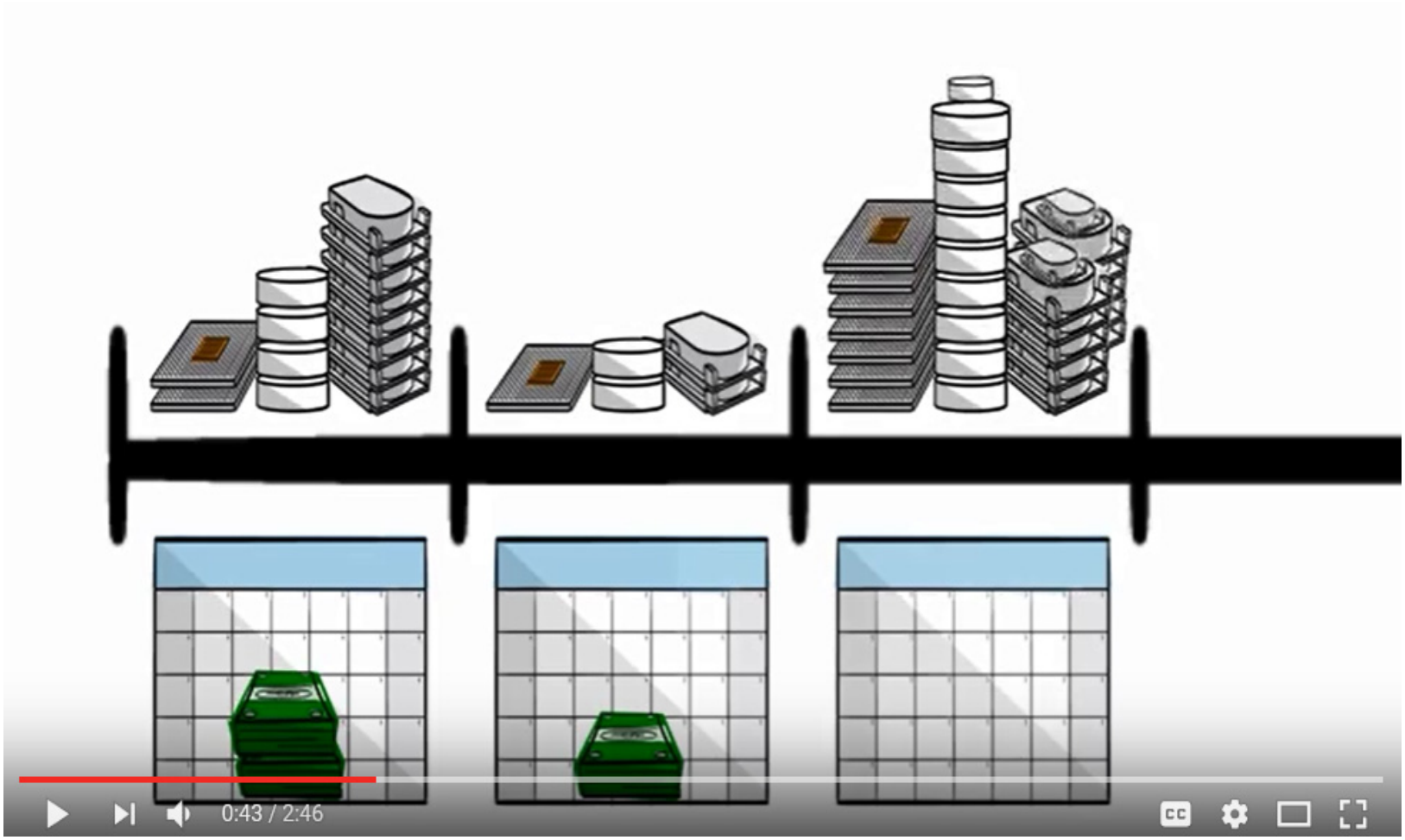
	vCPU	ECU	Memory (GiB)	Instance Storage (GB)	Linux/UNIX Usage
General Purpose - Current Generation					
t2.nano	1	Variable	0.5	EBS Only	$0.0059 per Hour
t2.micro	1	Variable	1	EBS Only	$0.012 per Hour
t2.small	1	Variable	2	EBS Only	$0.023 per Hour
t2.medium	2	Variable	4	EBS Only	$0.047 per Hour
t2.large	2	Variable	8	EBS Only	$0.094 per Hour
t2.xlarge	4	Variable	16	EBS Only	$0.188 per Hour
t2.2xlarge	8	Variable	32	EBS Only	$0.376 per Hour
m4.large	2	6.5	8	EBS Only	$0.108 per Hour
m4.xlarge	4	13	16	EBS Only	$0.215 per Hour
m4.2xlarge	8	26	32	EBS Only	$0.431 per Hour
m4.4xlarge	16	53.5	64	EBS Only	$0.862 per Hour
m4.10xlarge	40	124.5	160	EBS Only	$2.155 per Hour
m4.16xlarge	64	188	256	EBS Only	$3.447 per Hour

AWS EC2 Compute Payment Models – Elastic Capability

- General Purpose T2
  - ‘Pay for what you use‘ → t2.SIZES
  - Burstable performance Instances
  - Provide a baseline level of CPU performance with the ability to burst above the baseline
  - T2 instances are for workloads that don’t use the full CPU often, but occasionally need to burst
  - Accrue CPU Credits when they are idle
  - Use CPU credits when they are active
  - A CPU Credit provides the performance of a full CPU core for one minute
  - E.g. t2.small instance receives credits continuously at a rate of 12 CPU Credits per hour (baseline performance equivalent to 20% of a CPU core)
  - Usage below 20% / core stored in credit balance (used when above 20%)

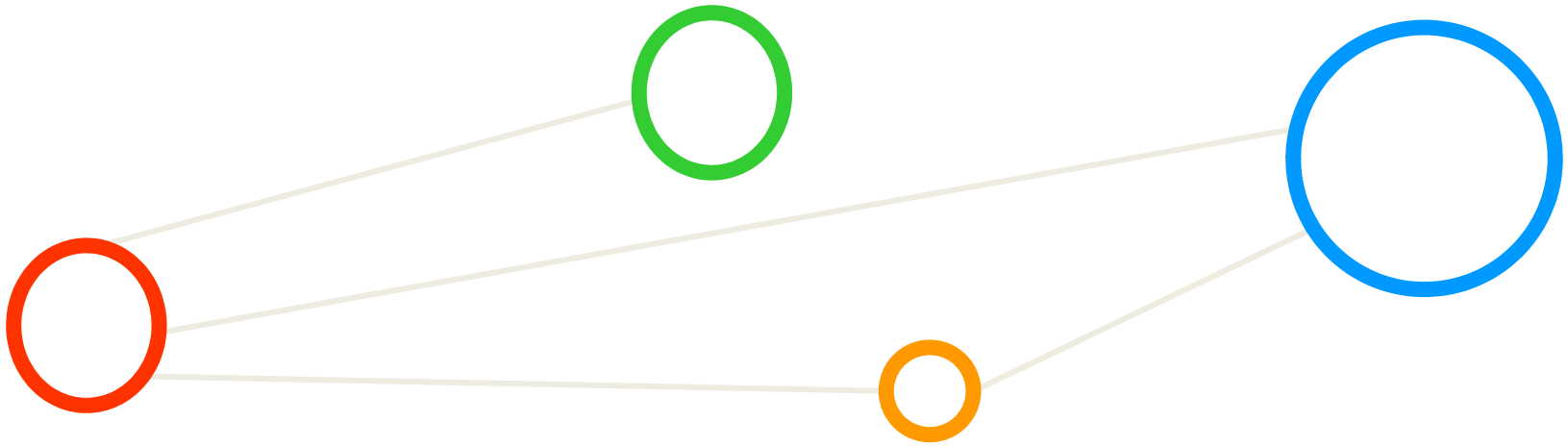
Model	vCPU	CPU Credits / hour	Mem (GiB)
t2.nano	1	3	0.5
t2.micro	1	6	1
t2.small	1	12	2
t2.medium	2	24	4
t2.large	2	36	8
t2.xlarge	4	54	16
t2.2xlarge	8	81	32

[Video] Amazon Web Services



[11] YouTube, 'What is Cloud Computing - Amazon Web Services '

Advanced IAAS Topics



IAAS Provide Example – AWS Services Overview

Compute
- Amazon EC2
- Amazon EC2 Container Registry
- Amazon EC2 Container Service
- Amazon Lightsail
- Amazon VPC
- AWS Batch
- AWS Elastic Beanstalk
- AWS Lambda
- Auto Scaling
- Elastic Load Balancing

Storage
- Amazon Simple Storage Service (S3)
- Amazon Elastic Block Storage (EBS)
- Amazon Elastic File System (EFS)
- Amazon Glacier
- AWS Storage Gateway
- AWS Snowball
- AWS Snowball Edge
- AWS Snowmobile

Database
- Amazon Aurora
- Amazon RDS
- Amazon DynamoDB
- Amazon ElastiCache
- Amazon Redshift
- AWS Database Migration Service

Migration
- AWS Discovery Services
- AWS Database Migration Service
- AWS Server Migration Service
- AWS Snowball
- AWS Snowball Edge
- AWS Snowmobile

Networking & Content Delivery
- Amazon VPC
- Amazon CloudFront
- Amazon Route 53
- AWS Direct Connect
- Elastic Load Balancing

Developer Tools
- AWS CodeCommit
- AWS CodeBuild
- AWS CodeDeploy
- AWS CodePipeline
- AWS X-Ray
- AWS Command Line Interface

Management Tools
- Amazon CloudWatch
- Amazon EC2 Systems Manager
- AWS CloudFormation
- AWS CloudTrail
- AWS Config
- AWS OpsWorks
- AWS Service Catalog
- AWS Trusted Advisor
- AWS Personal Health Dashboard
- AWS Command Line Interface
- AWS Management Console
- AWS Managed Services

Security, Identity & Compliance
- AWS Identity and Access Management (IAM)
- Amazon Inspector
- AWS Certificate Manager
- AWS CloudHSM
- AWS Directory Service
- Amazon Cloud Directory
- AWS Key Management Service
- AWS Organizations
- AWS Shield
- AWS WAF
- AWS Artifact

Analytics
- Amazon Athena
- Amazon EMR
- Amazon CloudSearch
- Amazon Elasticsearch Service
- Amazon Kinesis
- Amazon Redshift
- Amazon QuickSight
- AWS Data Pipeline
- AWS Glue

Artificial Intelligence
- Amazon Lex
- Amazon Polly
- Amazon Rekognition
- Amazon Machine Learning

Mobile Services
- AWS Mobile Hub
- Amazon API Gateway
- Amazon Cognito
- Amazon Pinpoint
- AWS Device Farm
- AWS Mobile SDK

Application Services
- AWS Step Functions
- Amazon API Gateway
- Amazon Elastic Transcoder
- Amazon AppStream

Messaging
- Amazon SQS
- Amazon Pinpoint
- Amazon SES
- Amazon SNS

Business Productivity
- Amazon WorkDocs
- Amazon WorkMail

Desktop & App Streaming
- Amazon WorkSpaces
- Amazon AppStream 2.0

Software
- AWS Marketplace

Internet of Things
- AWS IoT Platform
- AWS Greengrass
- AWS IoT Button

Game Development
- Amazon Lumberyard

[9] AWS Web page

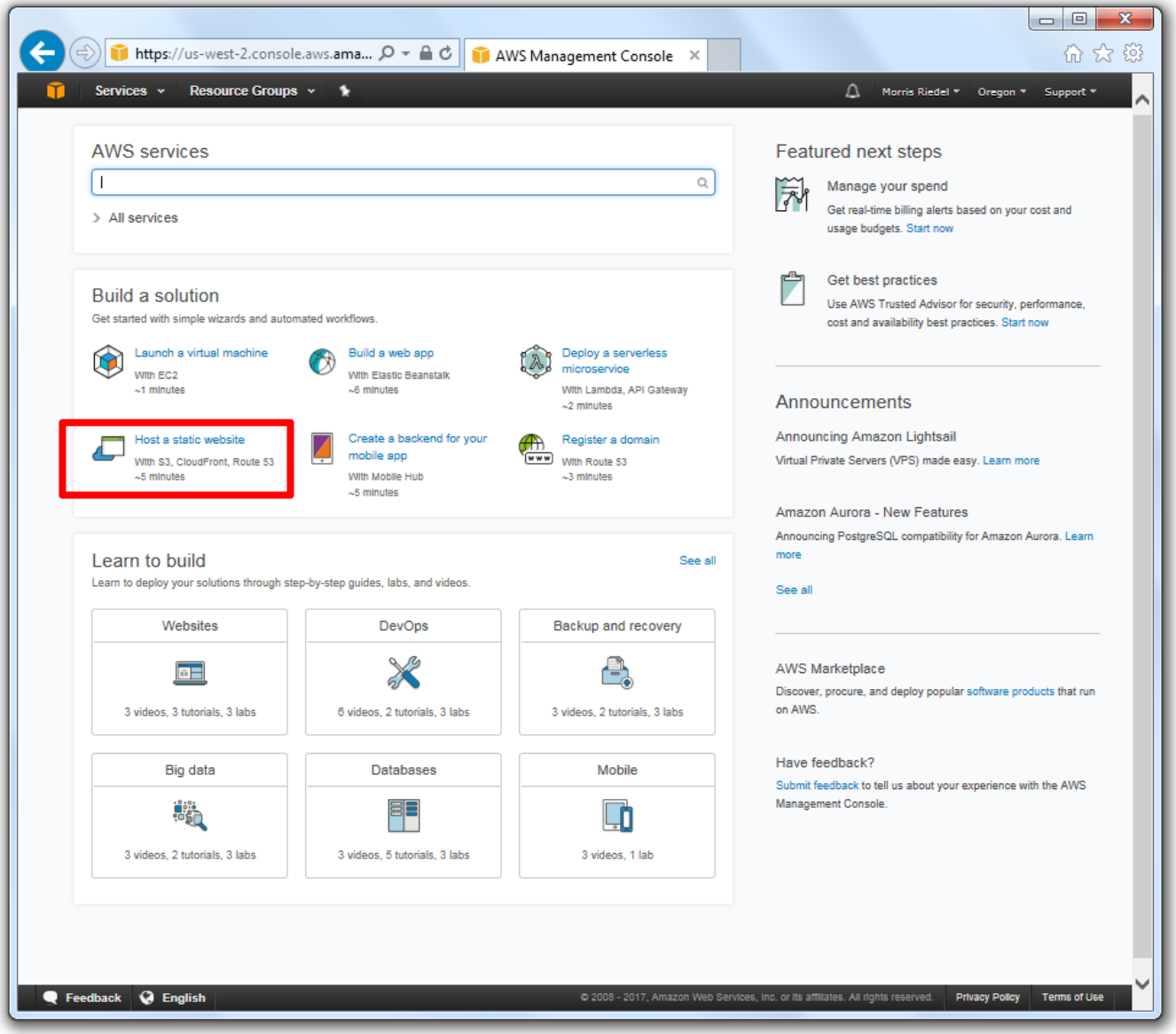
Understand AWS Cloud Service Portfolio – Storage

- Multiple storage-related products
  - A reliable, scalable, and secure place for data
- Products & Usage
  - Amazon S3: Scalable storage in the cloud
  - Amazon EBS: Elastic block storage for EC2
  - Amazon Elastic File System: Managed file storage for EC2
  - Amazon Glacier: Low-cost archive storage in the Cloud
  - AWS Storage Gateway: Hybrid storage integration
  - AWS Snowball: Petabyte-scale data transport
  - AWS Snowball Edge: Petabyte-scale data transport (+ on-board compute)
  - AWS Snowmobile: Exabyte-scale data transport

[9] AWS Web page

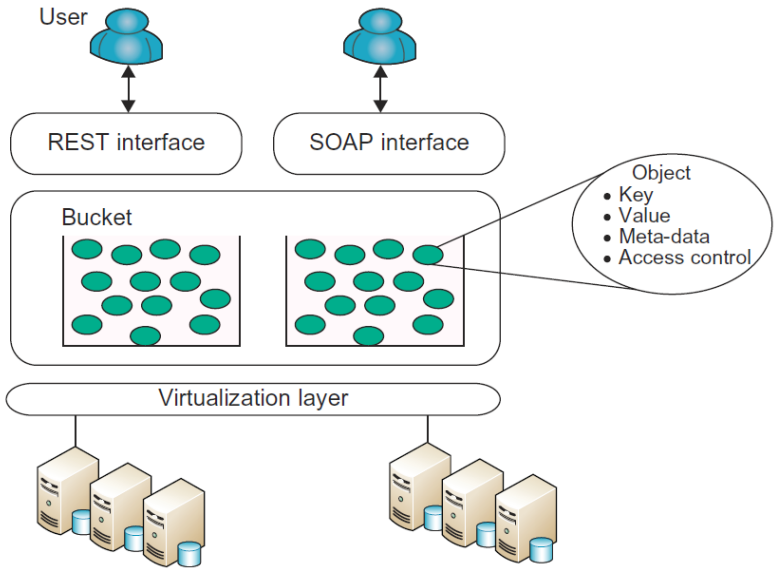
➢ **Lectures 10 & 11 provide details on data-intensive computing & storage: data transfers & archives**

AWS Console


https://us-west-2.console.aws.ama...
AWS Management Console
Services
Resource Groups
Morris Riedel
Oregon
Support
AWS services
All services
Build a solution
Get started with simple wizards and automated workflows.
Launch a virtual machine
With EC2
~1 minutes
Build a web app
With Elastic Beanstalk
~6 minutes
Deploy a serverless microservice
With Lambda, API Gateway
~2 minutes
Host a static website
With S3, CloudFront, Route 53
~5 minutes
Create a backend for your mobile app
With Mobile Hub
~5 minutes
Register a domain
With Route 53
~3 minutes
Learn to build
See all
Learn to deploy your solutions through step-by-step guides, labs, and videos.
Websites
3 videos, 3 tutorials, 3 labs
DevOps
6 videos, 2 tutorials, 3 labs
Backup and recovery
3 videos, 2 tutorials, 3 labs
Big data
3 videos, 2 tutorials, 3 labs
Databases
3 videos, 5 tutorials, 3 labs
Mobile
3 videos, 1 lab
Featured next steps
Manage your spend
Get real-time billing alerts based on your cost and usage budgets. Start now
Get best practices
Use AWS Trusted Advisor for security, performance, cost and availability best practices. Start now
Announcements
Announcing Amazon Lightsail
Virtual Private Servers (VPS) made easy. Learn more
Amazon Aurora - New Features
Announcing PostgreSQL compatibility for Amazon Aurora. Learn more
See all
AWS Marketplace
Discover, procure, and deploy popular software products that run on AWS.
Have feedback?
Submit feedback to tell us about your experience with the AWS Management Console.
Feedback
English
© 2008 - 2017, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Privacy Policy
Terms of Use

IAAS Storage Provisioning – AWS Amazon S3 Example

- S3 is ‘storage as a service’ with a Web messaging interface
  - Using API with Representational State Transfer (REST)
  - Using API with Simple Object Access Protocol (SOAP)
- Remote object storage
  - Used with trillions of objects
  - 99,99 % durability
  - Data considered objects to be named by end users
  - Objects alongside metadata are stored in bucket containers
  - Buckets enable namespaces for user identification & accounting

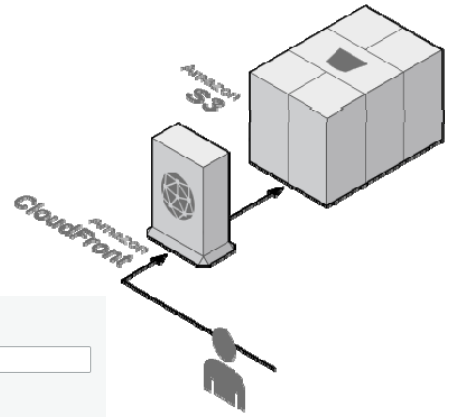


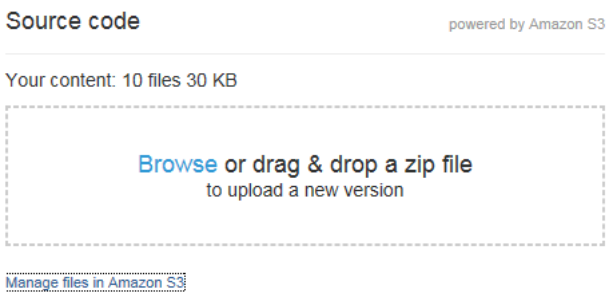
[1] Distributed & Cloud Computing Book

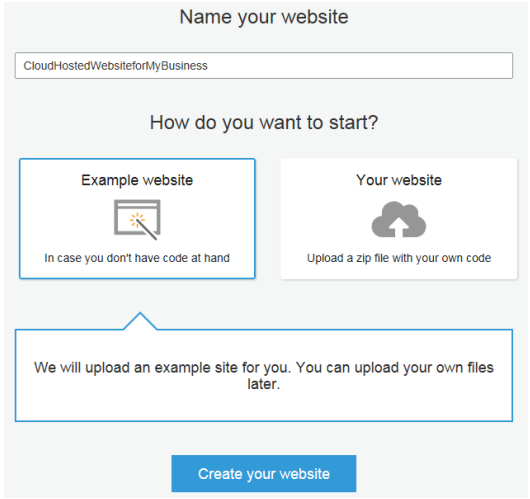
- **AWS Amazon Simple Storage Service (Amazon S3) is a scalable object storage to store/retrieve data**
- **S3 Web service interfaces (REST, SOAP) enable remote data access from anywhere on the Web**

AWS Amazon S3 Instances – Host a static Website

- Launch a S3-based service via console
  - E.g. Host a static Website lets one easily create S3 storage for a Website
- Combined with other AWS Services
  - Amazon Cloud Front: assigns a URL that is used to access a Website ('buy domain') and managed content delivery
  - Amazon Route 53: Scalable domain name system
  - Amazon S3: data, code, etc.





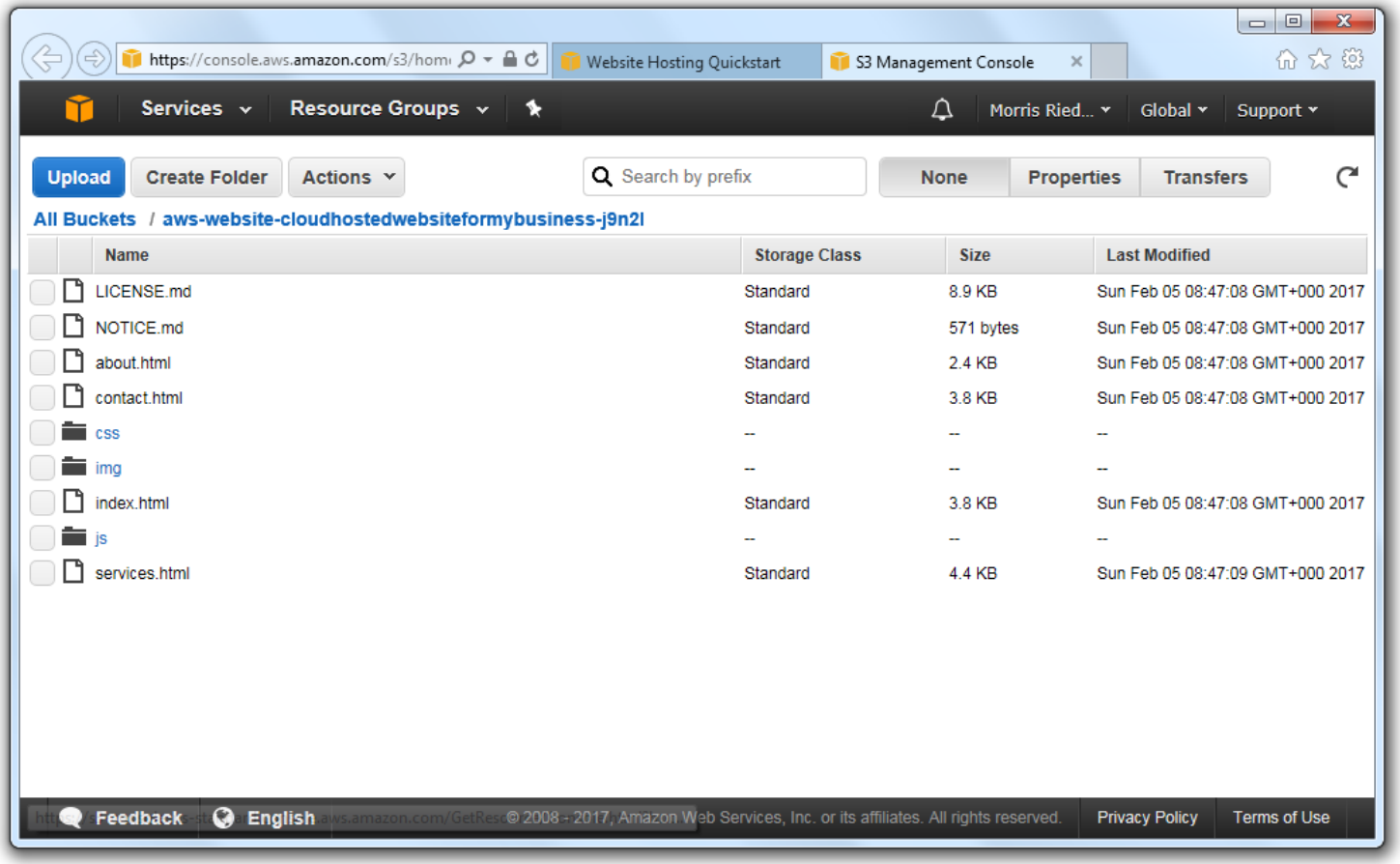


[9] AWS Web page

- **Amazon S3 provides replicated storage at multiple data centers (cf. Lecture 4) in the AWS cloud**

AWS Amazon S3 Instances – Manage Files in S3 Example

- Using the Web console to upload/download content
  - Static Website example is mainly based on ‘files with content‘



AWS Amazon S3 Instances – Cost Models

- S3 (standard) storage prices including long-term archive (glacier)

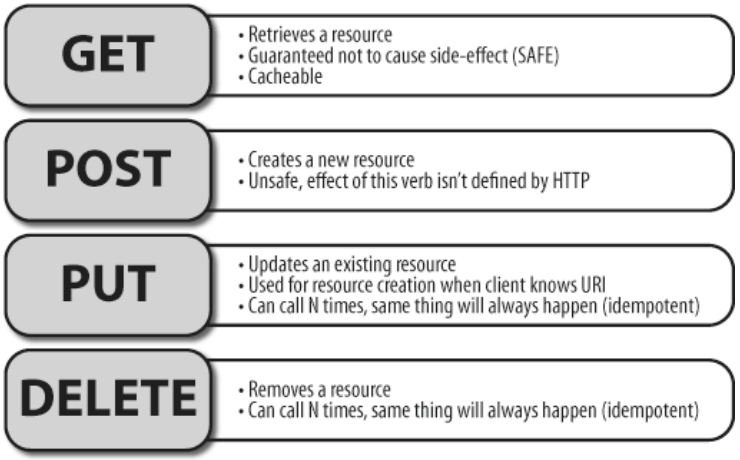
	Standard Storage	Standard - Infrequent Access Storage †	Glacier Storage
First 50 TB / month	$0.024 per GB	$0.0131 per GB	$0.0045 per GB
Next 450 TB / month	$0.023 per GB	$0.0131 per GB	$0.0045 per GB
Over 500 TB / month	$0.022 per GB	$0.0131 per GB	$0.0045 per GB

(Amazon Glacier is a low-cost archive storage in the cloud)

- Transfer in/out prices (!)

For Requests Not Otherwise Specified Below	
PUT, COPY, POST, or LIST Requests	$0.0053 per 1,000 requests
GET and all other Requests	$0.0042 per 10,000 requests
Delete Requests	Free †
For Standard – Infrequent Access Requests	
PUT, COPY, or POST Requests	$0.01 per 1,000 requests
GET and all other Requests	$0.01 per 10,000 requests
Lifecycle Transition Requests into Standard – Infrequent Access	$0.01 per 1,000 requests
Data Retrievals	$0.01 per GB

(GET, POST, PUT, DELETE are all Web service message exchanges used by REST/SOAP APIs)



[12] RESTful.NET

- **Amazon S3 pricing includes cost models per GB data and for data transfer & Web message requests**

AWS Amazon Elastic Block Storage (EBS)

- Block storage
  - Manages data as blocks (compared to object storage with data as objects)
  - Split files into raw blocks of data (each block controlled like a hard drive)
- EBS Service
  - EBS volumes are automatically replicated
  - Scales usage up/down (~min)

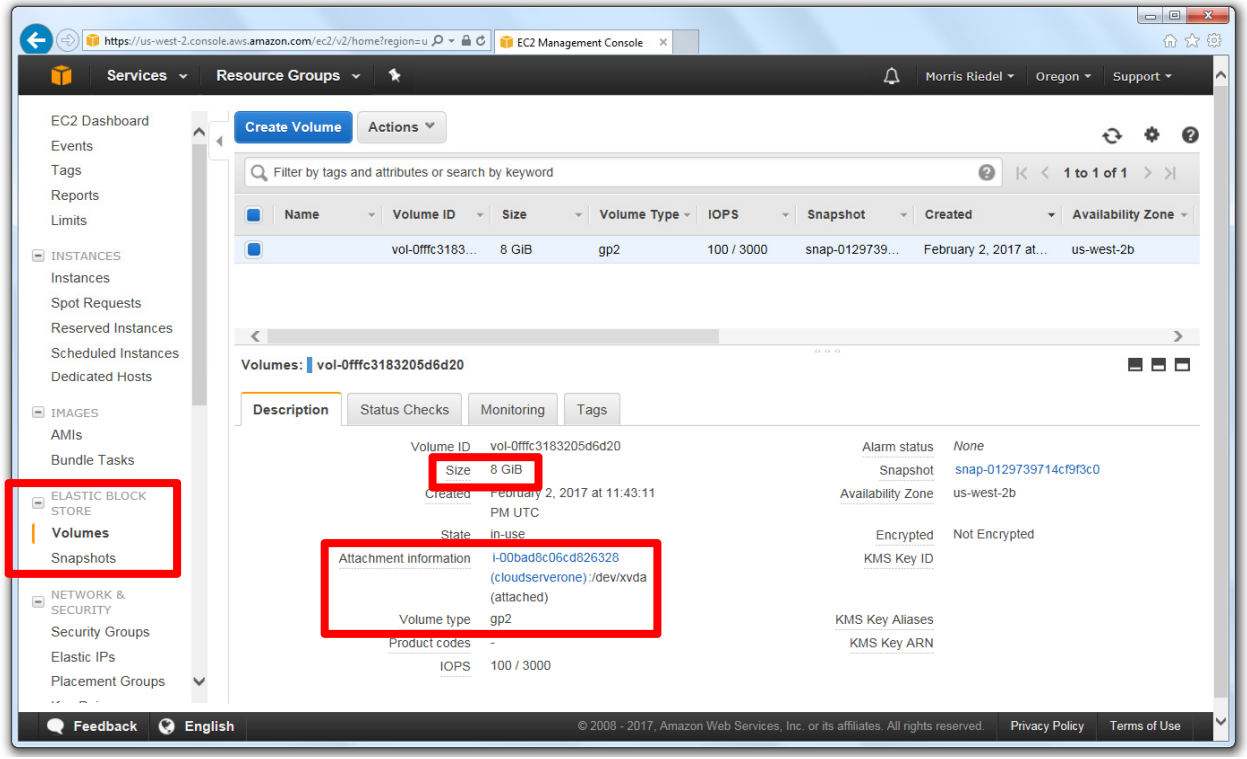
Amazon EBS General Purpose SSD (gp2) volumes
- $0.116 per GB-month of provisioned storage

Amazon EBS Provisioned IOPS SSD (io1) volumes
- $0.145 per GB-month of provisioned storage
- $0.076 per provisioned IOPS-month

Amazon EBS Throughput Optimized HDD (st1) volumes
- $0.053 per GB-month of provisioned storage

Amazon EBS Cold HDD (sc1) volumes
- $0.029 per GB-month of provisioned storage

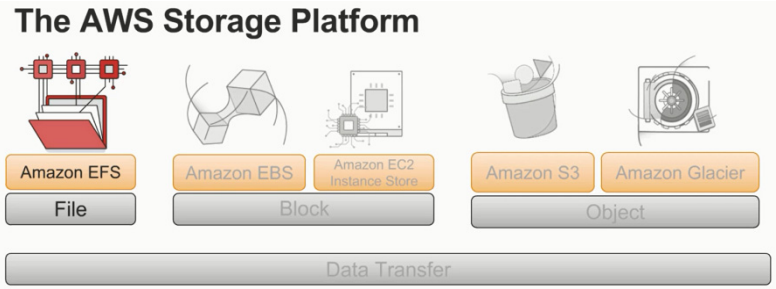
Amazon EBS Snapshots to Amazon S3
- $0.053 per GB-month of data stored



- **Amazon Elastic Block Storage (EBS) provides persistent block storage volume used with Amazon EC2**

AWS Amazon Elastic File System (EFS)

- EFS for working with traditional files
  - Simple, scalable & reliable file storage for Amazon EC2 instances
  - Storage capacity is elastic, growing and shrinking automatically



  - Add and remove files and applications have the storage as needed
  - File system object (i.e. directory, file, and link) is redundantly stored
- Typical usage
  - Amazon EFS file system provides a standard file system interface
  - Typical file system access semantics for existing applications and tools
  - Multiple Amazon EC2 instances access an Amazon EFS at the same time
  - Using standard operating system mount commands (e.g. UNIX) (e.g. 'mount' / connect Amazon EFS file systems via the NFSv4.1 protocol)

- **Amazon Elastic File System (EFS) is a scalable and reliable elastic file storage sytem for AWS EC2**

AWS Marketplace and Users

- AWS Marketplace
  - E.g. collection of community and Amazon created pre-installed images
  - Software infrastructure, developer tools, business & desktop software
  - User success stories and details of how AWS was adopted in solutions
- ‘Startup company‘ Airbnb (travel)

[13] AWS Marketplace

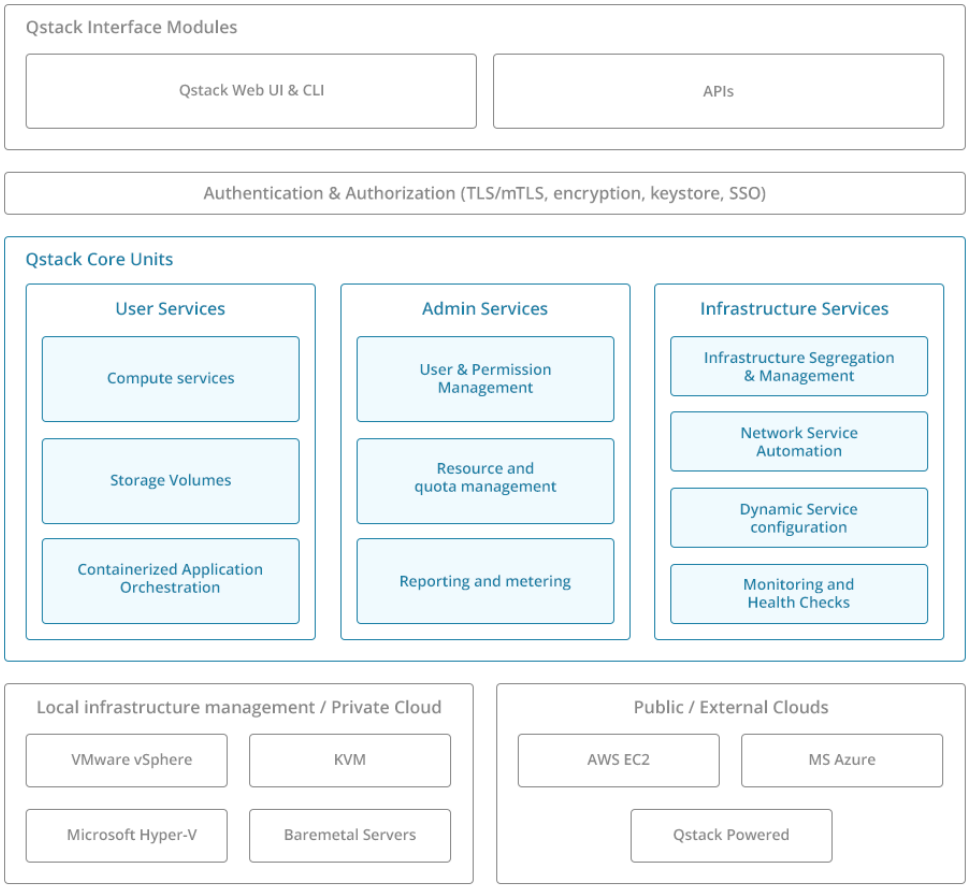
  - Scales infrastructure automatically using AWS
  - Uses 200 Amazon EC2 instances for its application
  - Uses elastic load balancing with Amazon EC2 instances
  - Analyzes 50 GB of data daily via Amazon Elastic MapReduce (Amazon EMR)
- ‘Startup company‘ Spotify (music)

  - Instant access to over 16 million licensed songs
  - Stores its huge volume of content in Amazon S3

- **The AWS Marketplace offers a great opportunity to start with AWS services since it offers a wide range of user stories how AWS was adopted in practice and gives access to pre-installed images**

Qstack Cloud Management Platform

- Software features
  - User-friendly User Interface
  - Provisioning bare metal infrastructure & servers
  - Works with different types of hypervisors (cf. Lecture 4)
  - Interoperable with public clouds (AWS EC2, MS Azure)
- Adoption
  - E.g. Advania Cloud Services

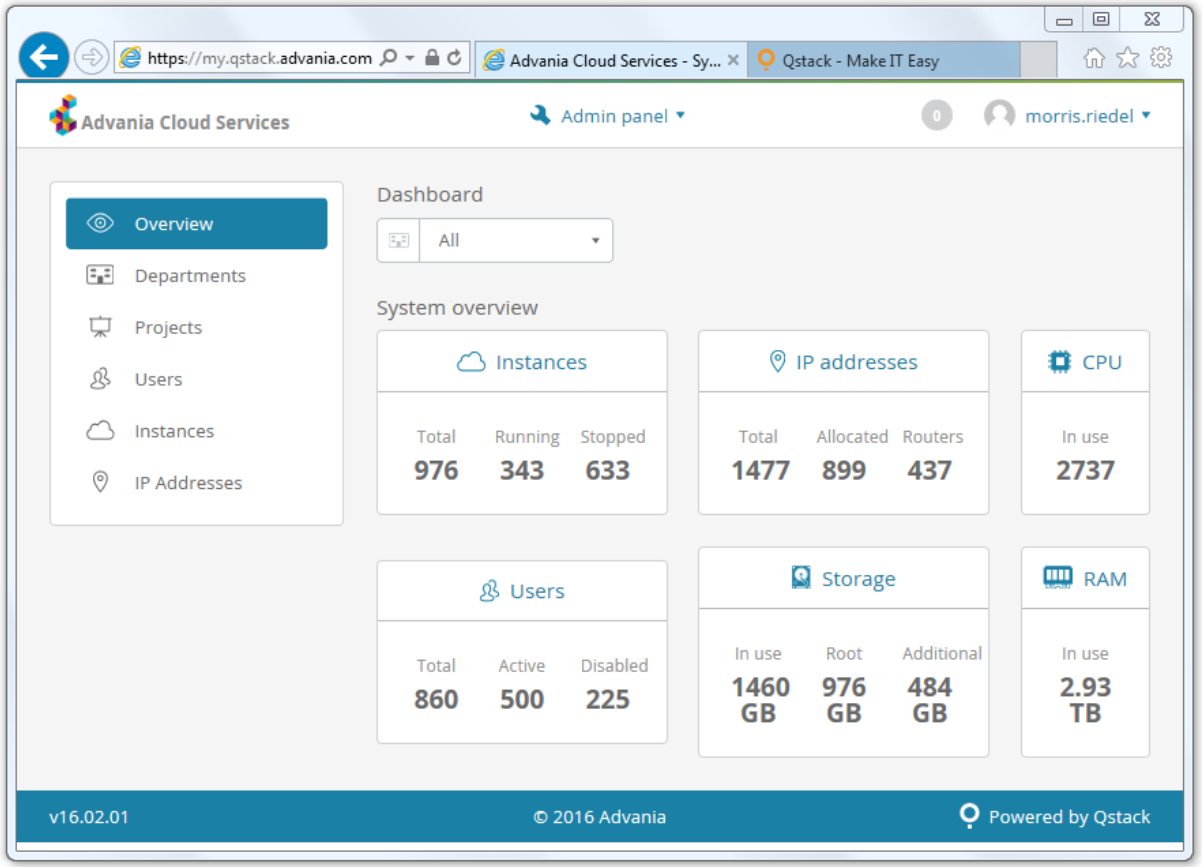


[14] Qstack Web Page

- **Qstack is a cloud management platform that makes it easy to deploy private/public/hybrid clouds**
- **Qstack runs on commodity hardware and is hypervisor-agnostic in order to avoid vendor locks**

Advania Cloud Services

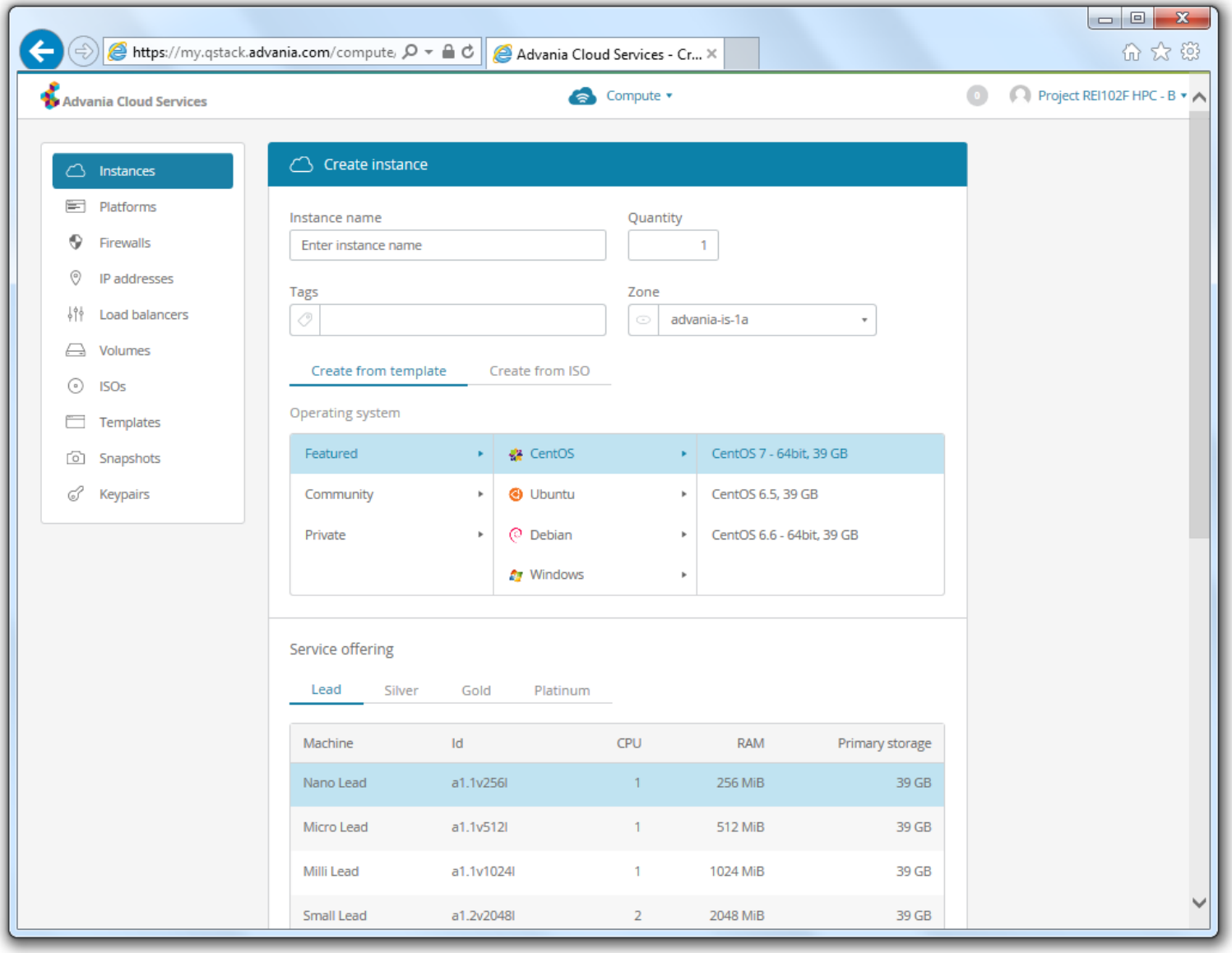
- Selected features
  - Uses templates for deploying virtual instances
  - Web console to manage well computing and storage servers
  - IT resources hosted in 100% green data center (cf. Lecture 4)



[15] Advania Cloud Services Web page

- **Advania Cloud Services provides an open cloud as scalable IAAS platform that is managed with a robust user interface and offers access to scalable virtual servers and storage billed by the hour**

Advania Cloud Services – Create Compute Instance Example

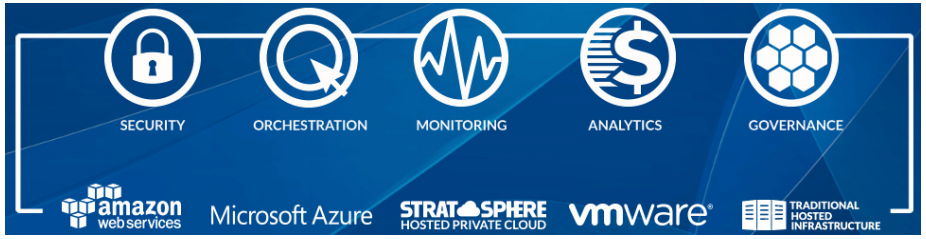


IAAS Providers & Users: GoGrid/DataPipe & Enomaly

- GoGrid/DataPipe
  - [Datapipe acquired big data multi-Cloud pioneer GoGrid]
  - Architects deploy & manage multi-platform solutions tailored to client needs

DATAPIPE

[17] Datapipe Web page



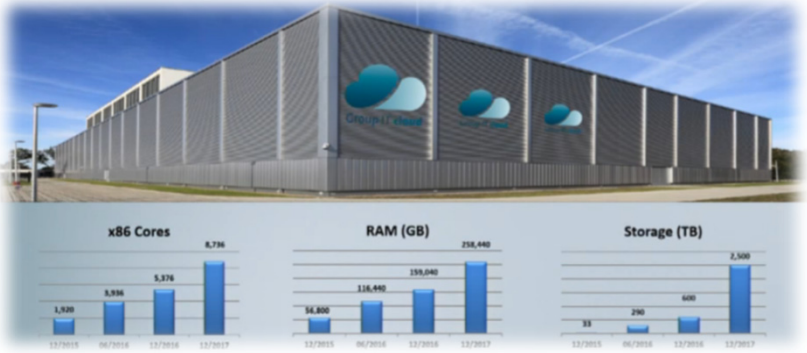
  - Full 'pipe': traditional IT, public, private & hybrid clouds (cf. Lecture 2)
- Enomaly/Virtustream/Dell
  - ['hype' business: one company buys another]
  - Enomaly Elastic Computing Platform (ECP) is an IAAS provider
  - 'Cloud in a box' solution that enables telcos and hosting providers to deliver IAAS cloud computing services to their customers
  - Offer SpotCloud market platform that provides public cloud service providers a mechanism to sell unused cloud capacity
  - SpotCloud & ECP are now integrated into Virtustream portfolio

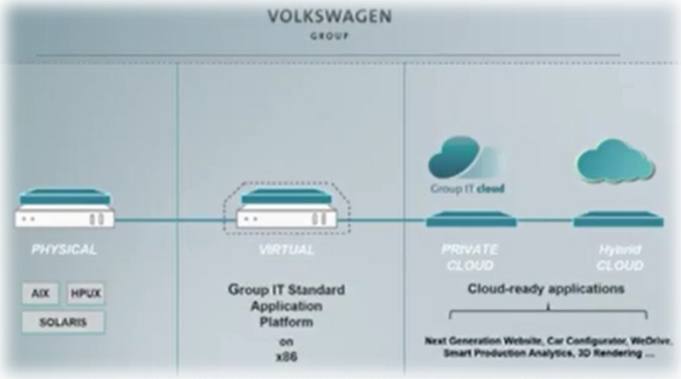
enomaly elastic computing

[18] Enomaly Web page

OpenStack and IAAS Solutions using Private/Public Clouds

- OpenStack Cloud Operating System
  - Provides an architecture to build own IAAS clouds
  - One can create a massive product portfolio (e.g. AWS, but on small level)
  - Underlying virtual IT resources like storage, compute, networking needed
  - Openstack software & hardware provide IAAS basis for many applications
  - Compatible with public clouds via APIs: Amazon EC2 and Google Compute Engine





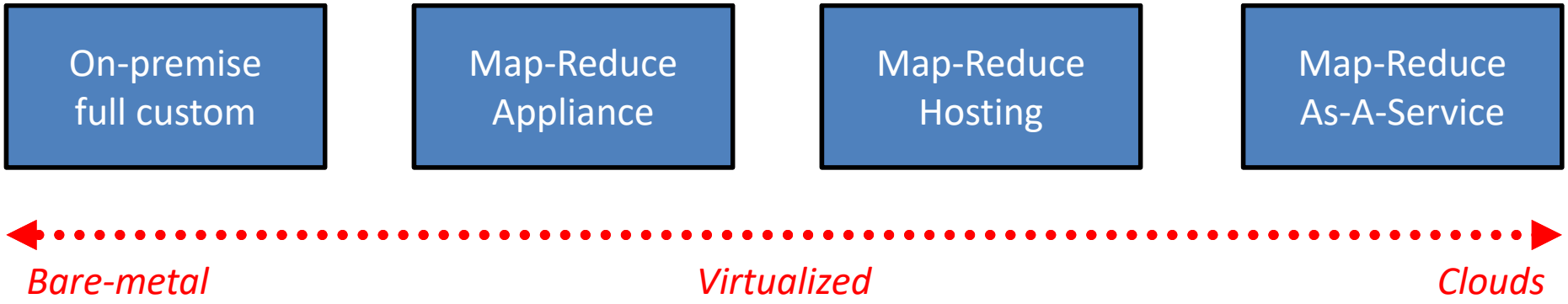
[16] VW & Openstack

- **The Volkswagen Group is using OpenStack for their IAAS-based private and public cloud solutions**

➢ **Lectures 13 offers more insights into the various building blocks of the OpenStack Cloud Software**

IAAS & Map-Reduce Potential Deployment Models Example

- **'Deployment Models' are different ways of how software can be installed, configured, provisioned, or used by end-users in context of underlying hardware**
- **Map-Reduce deployment models can be categorized in 'on-premise', 'map-reduce appliance, 'map-reduce hosting', or 'map-Reduce-As-A-Service'**



modified from [21] Accenture Deployment Models

- **Traditional approaches to store and process data was around the approach of isolated 'data warehouses' with customer data**
- **The products Hadoop Map-Reduce and HDFS store and process very large datasets in a scalable and cost effective way**

[22] Apache Hadoop Web Page

IAAS & Map-Reduce – Choosing Deployment Models

- **Infrastructure in IAAS is different per users: for some pure hardware for others maybe Map-Reduce**
- **The five key-areas that influence a deployment model decision are 'price-performance ratio', 'data privacy','data gravity', 'data enrichment', 'productivity of developers and data scientists'**

- Price-Performance Ratio
  - Virtualized environments safe costs vs. slower with I/O operations
- Data privacy (cf. Lecture 4 'Private Clouds' vs. 'Public Clouds')
  - Concern when storing data outside of a 'self-owned' infrastructure
- Data gravity
  - Data increases tremendously → lock in data infrastructure, migration costs
- Data enrichment
  - Leveraging multiple datasets to uncover new insights, needs now vs. future
- (Less) Productivity of data scientists (and developers/admins)
  - Explore the needs, create solution, deploy vs. in clouds already 'production'

modified from [21] Accenture Deployment Models

On-premise Full Custom Deployment

On-premise full custom

- **Purchase commodity hardware and employ administrator(s)**
- **Install and operate the map-reduce software on own costs**
- **Enabling full control of the map-reduce cluster and configuration**

- A bare-metal map-reduce cluster
  - Deployed on top of physical servers without any virtualization layer
  - E.g. plain Hadoop installation on a physical cluster

- 'Total Cost of Onwership (TCO)' matters
  - Compare the performance at the matched budget (not hw specifications)
- Bare-metal with 'On-premise full custom' deployment
  - Costs for 'staff for operation', 'technical support', 'data center facility', 'server hardware'
  - Don't forget: 'training the staff to use the deployment' → 'Learning curve'

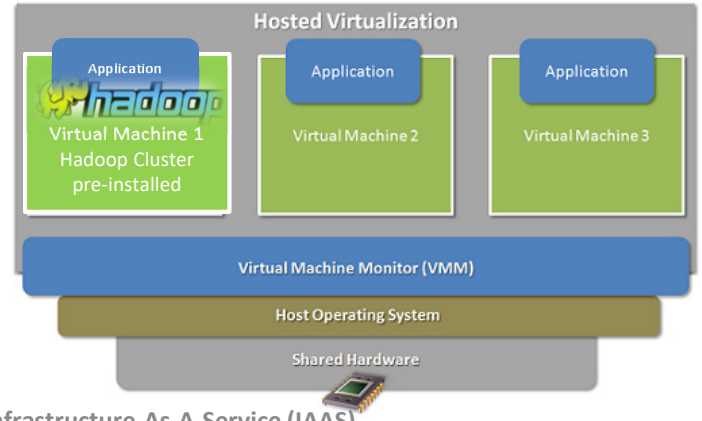
modified from [21] Accenture Deployment Models

Map-Reduce Appliance Deployment

Map-Reduce Appliance

- **'Appliance' means pre-configured map-reduce cluster**
- **Allows for bypassing detailed technical configuration decisions**
- **Enables an early start of data analysis based on map-reduce**

- Take advantage of advancements in 'virtualization technology'
  - cf. Lecture 3 (Advancements of Virtualization Technologies)
  - Not a cloud offering by a service provider, just a pre-installed image
  - Can be used to run different 'map-reduce' appliances
- E.g. Pre-installed and configured Hadoop cluster



modified from [23] National Instruments 'Virtualization'

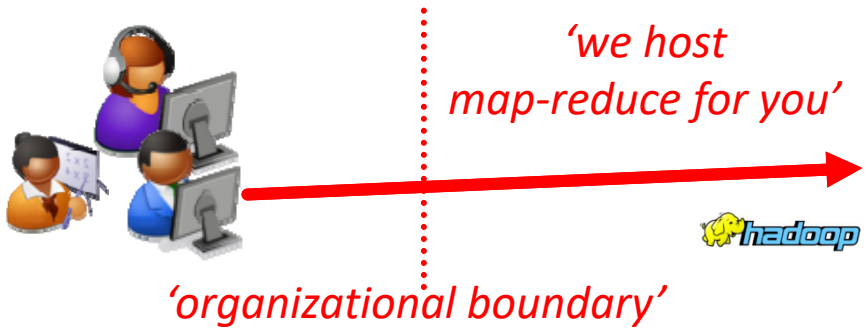
modified from [21] Accenture Deployment Models

Map-Reduce Hosting Deployment

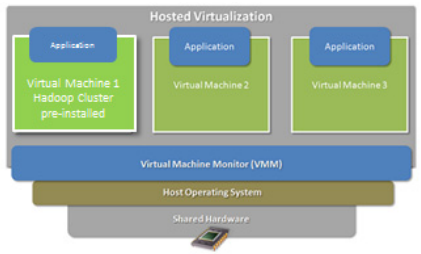
Map-Reduce Hosting

- **Take advantage of 'Internet Service Provider (ISP)' model**
- **Buy-in a service provider to deploy and operate Map-Reduce cluster**
- **Enabling automating the deployment and operation processes**

- Take advantage of 'map-reduce hosting' by another organization
  - 'Outsourcing' of the deployment and oepration of a map-reduce cluster
  - Use 'map-reduce hosting' based on a 'Service Level Agreement (SLA)'
- E.g. Small-Medium Enterprise (SMS) offers 'map-reduce hosting'
  - E.g. Provide access of fractions of their own Hadoop deployment

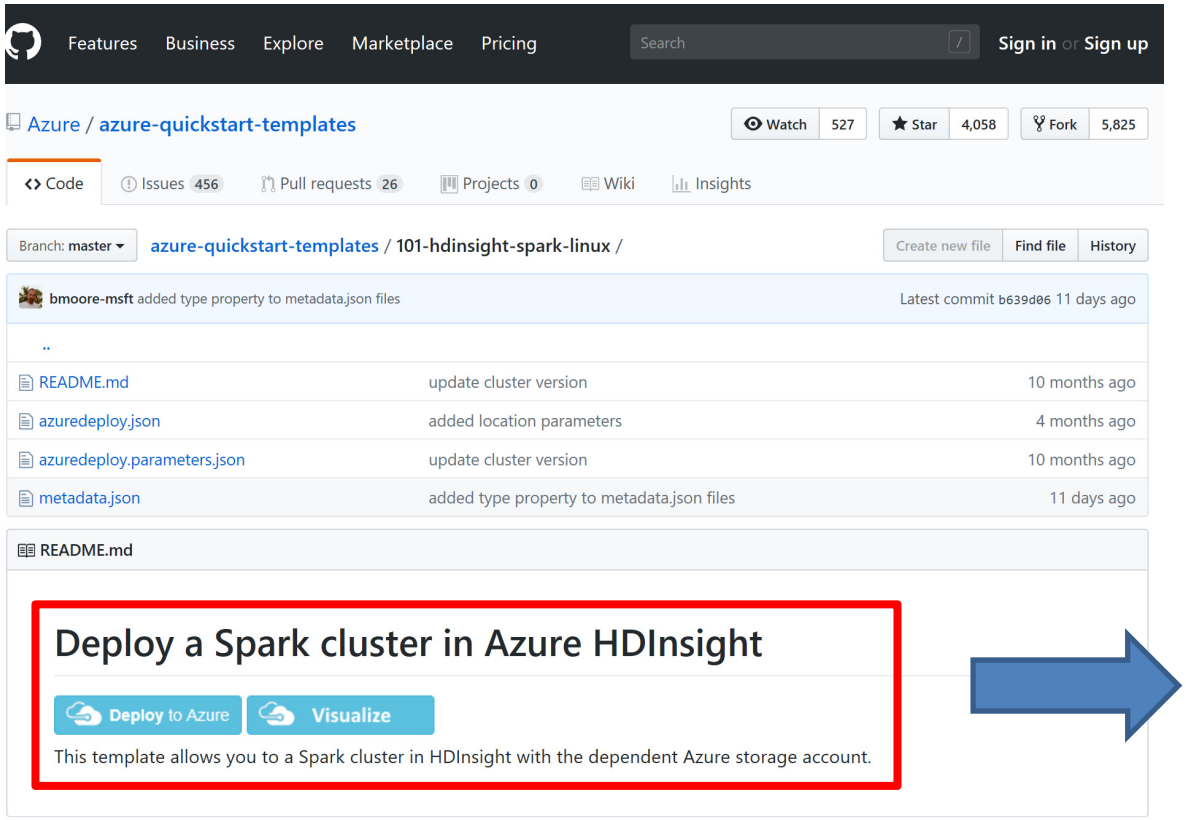


modified from [21] Accenture Deployment Models



modified from [23] National Instruments 'Virtualization'

Deploy Spark Cluster in MS Azure HDInsight (cf. Lecture 3)



[25] Github template Web page

[26] MS Azure HDInsight Web Page

Microsoft Azure

HDInsight

hadoop

APACHE Spark

- **Microsoft Azure HDInsight Spark Clusters can be deployed via pre-configured resource manager templates that use Azure computing resources & Azure Storage Blobs as cluster storage**
- **A wide variety of templates are available on Github pages for various general cloud services**

Map-Reduce Hosting Deployment

Map-Reduce As-A-Service

- **Provides instant access to a map-reduce cluster as a 'cloud service'**
- **Take advantage of the 'pay-per-use' model (aka 'swipe credit card')**
- **Offers dynamic usage of the service according to workload peaks**

- A 'cloud' that offers a Map-Reduce-As-A-Service capability
  - Ideally connected to to a 'storage-as-a-service' capability ('data locality')
  - E.g. 'Hadoop-as-a-Service' as cloud service directly to use
- No 'TCO', but also costs with 'pay-per-use' model
- E.g. Amazon Elastic MapReduce (EMR), cf. Lecture 5
  - Enables users to deploy Hadoop clusters on demand
  - Fastest way to run MapReduce jobs
  - Tear down the clusters when the jobs are completed
  - Dynamic usage (perfect for workload peaks)
  - Based on Apache Hadoop & Spark

modified from [21] Accenture Deployment Models

AWS Cloud Service Portfolio – Analytics (cf. Lecture 5)

- Multiple analytics products
  - Extracting insights and actionable information from data requires technologies like analytics & machine learning

[9] AWS Web page

- Products & Usage
  - Amazon Athena: Serverless Query Service
  - Amazon ElasticMapReduce: Hadoop
  - Amazon ElasticSearch Service: Elasticsearch on AWS
  - Amazon Kinesis: Streaming Data
  - Amazon QuickSight: Business Analytics
  - Amazon Redshift: Data Warehouse
  - ...

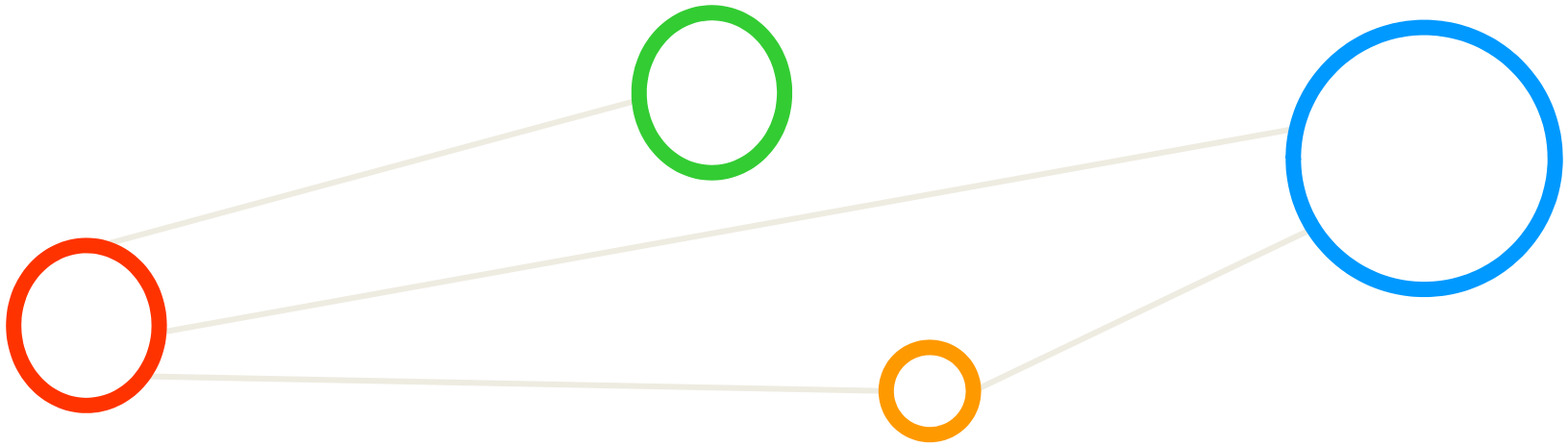
[24] Apache Spark on Amazon EMR

- **Apache Spark on Hadoop YARN is natively supported in Amazon ElasticMapReduce (EMR)**
- **The AWS management console offers to easily create and manage Apache Spark clusters**

[Video] VMware and Rackspace

[19] YouTube, VMware and Rackspace'

Lecture Bibliography



Lecture Bibliography (1)

- [1] K. Hwang, G. C. Fox, J. J. Dongarra, 'Distributed and Cloud Computing', Book, Online: http://store.elsevier.com/product.jsp?locale=en_EU&isbn=9780128002049
- [2] Keras Python Deep Learning Library, Online: https://keras.io/
- [3] Tensorflow Deep Learning Framework, Online: https://www.tensorflow.org/
- [4] Jupyter Web page, Online: http://jupyter.org/
- [5] Amazon Web Services EC2 On-Demand Pricing models, Online: https://aws.amazon.com/ec2/pricing/on-demand/
- [6] Coursera, 'Deep Learning Specialization', Online: http://www.coursera.org
- [7] Guide to Iceland, Mount Esja Online: https://guidetoiceland.is/connect-with-locals/1625/mount-esja-the-city-mountain-of-reykjavik
- [8] Wikipedia, ‚Cloud Computing', Online: http://de.wikipedia.org/wiki/Cloud-Computing
- [9] Amazon Web Services Web Page, Online: https://aws.amazon.com
- [10] Big Data Tips 'Amazon Web Services', Online: http://www.big-data.tips/amazon-web-services
- [11] YouTube Video, 'What is Cloud Computing', Online: https://www.youtube.com/watch?v=jOhbTAU4OPI

Lecture Bibliography (2)

- [12] Safari book online, ‘RESTful.NET‘,
  Online: https://www.safaribooksonline.com/library/view/restful-net/9780596155025/ch04.html
- [13] AWS Marketplace,
  Online: https://aws.amazon.com/marketplace/
- [14] Qstack,
  Online: https://www.qstack.com/
- [15] Advania Cloud Services,
  Online: https://advania.com/cloud/open-cloud/
- [16] YouTube Video, ‘Driving the Future of IT Infrastructure at Volkswagen Group‘,
  Online: https://www.youtube.com/watch?v=HL_pzkDnaI4
- [17] Datapipe Web page,
  Online: https://www.datapipe.com/
- [18] Enomaly/Virtustream Web page,
  Online: http://www.virtustream.com/
- [19] YouTube Video, ‘VMware and Rackspace - Delivering Value to Our Customers Through A Strategic Partnership’,
  Online: https://www.youtube.com/watch?v=cdBxtjfYAy8
- [20] Key-based Authentication in SSH, Online:
  http://www.oracle.com/webfolder/technetwork/tutorials/obe/cloud/dbaas/obe_dbaas_connecting_to_an_instance/images/key%20based%20authentication.jpg

Lecture Bibliography (3)

- [21] Accenture, Hadoop Deployment Comparison Study,
  Online: http://www.accenture.com/SiteCollectionDocuments/PDF/Accenture-Hadoop-Deployment-Comparison-Study.pdf
- [22] Apache Hadoop Wep page,
  Online: http://hadoop.apache.org/
- [23] Virtualization, National Instruments White Paper,
  Online: http://www.ni.com/white-paper/8709/en/
- [24] Apache Spark on Amazon EMR,
  Online: https://aws.amazon.com/emr/details/spark/
- [25] Github template: Deploy a Spark cluster in Azure HDInsight,
  Online: https://github.com/Azure/azure-quickstart-templates/tree/master/101-hdinsight-spark-linux
- [26] Microsoft Azure HDInsight Cluster Web page,
  Online: https://azure.microsoft.com/en-us/services/hdinsight/


services
measurement
within
fusion
new
funding
Policy-based
concepts
device
Analysis
forms
cross-disciplinary
resources
Centre
climate
Computer
expertise
year
like
Computational
Test
E.g
Computing
disciplines
results
Approaches
construction
Fusion
Enable
Cross-Disciplinary
methods
High
big
nodes
PB
community
Environment
Modelling
different
EU
Science
centers
Key
scientific
important
files
etc
DKRZ
Use
storage
basis
Infrastructure
model
international
SMAQ
engineering
References
formats
simulations
ENES
technologies
physical
brain
Big
Project
Data
increasing
via
EUDAT
processing
IPCC
Common
TB
Methods
MapReduce
create
computing
Using
European
databases
GB
Supercomputing
Work
images
codes
analysis
DLCL
Level
stored
Resources
analyze
Large
Summary
data-intensive
Service
Scientific
levels
often
use
Matters
Concrete
external
performance
JSC
computational
Research
research
e.g
science
HPC
Juelich
system
datasets
acquisition
Landscape
compute
System
Climate
modelling
AR5
Hardware
large-scale
cores
access
hundreds
Nutshell
Services
management
network
Technologies
Earth
Structure
current
DLCLs
Understanding
structures
distributed
directory
project
General
Germany
Simulation
Processing
Provide
NASA
Energy
systems
day
Health