MASARYKOVA UNIVERZITA

PV213 Enterprise Information Systems in Practice

03 – Architecture of the EIS in the cloud

MASARYKOVA UNIVERZITA

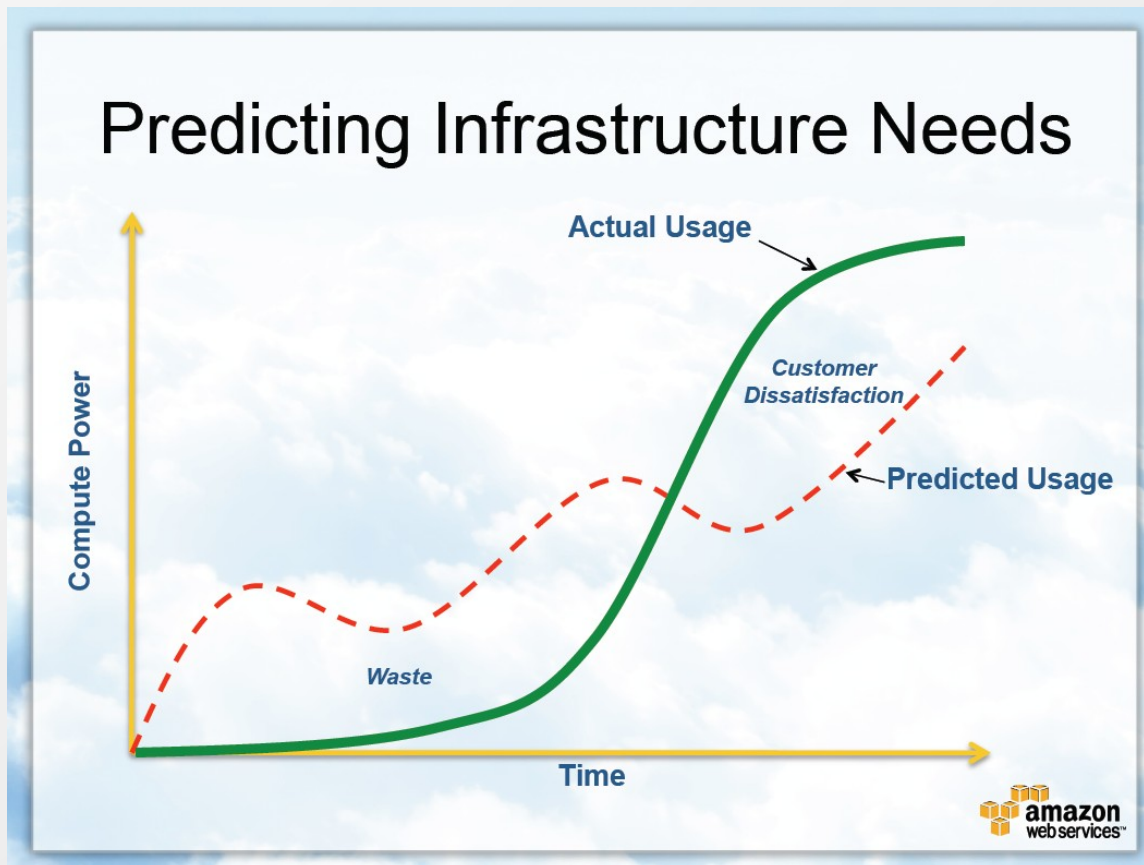**INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ**

PV213 EIS in Practice: 03 – Architecture of the EIS in the cloud

Tento projekt je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.

EVROPSKÁ UNIE

MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

OP Vzdělávání
pro konkurenceschopnost

UNIVERSITAS MASARYKIANA BRUNENSIS

I N V E S T I C E   D O   R O Z V O J E   V Z D Ě L Á V Á N Í

Cloud Computing motivation

How can my application grow with customers?

How can I pay only for what I use?

What is Cloud Computing?

Cloud computing is Internet-based computing, whereby shared resources, software and information are provided to computers and other devices on-demand, like the electricity grid.

Five Attributes of Cloud Computing

- Service based
- Scalable and elastic
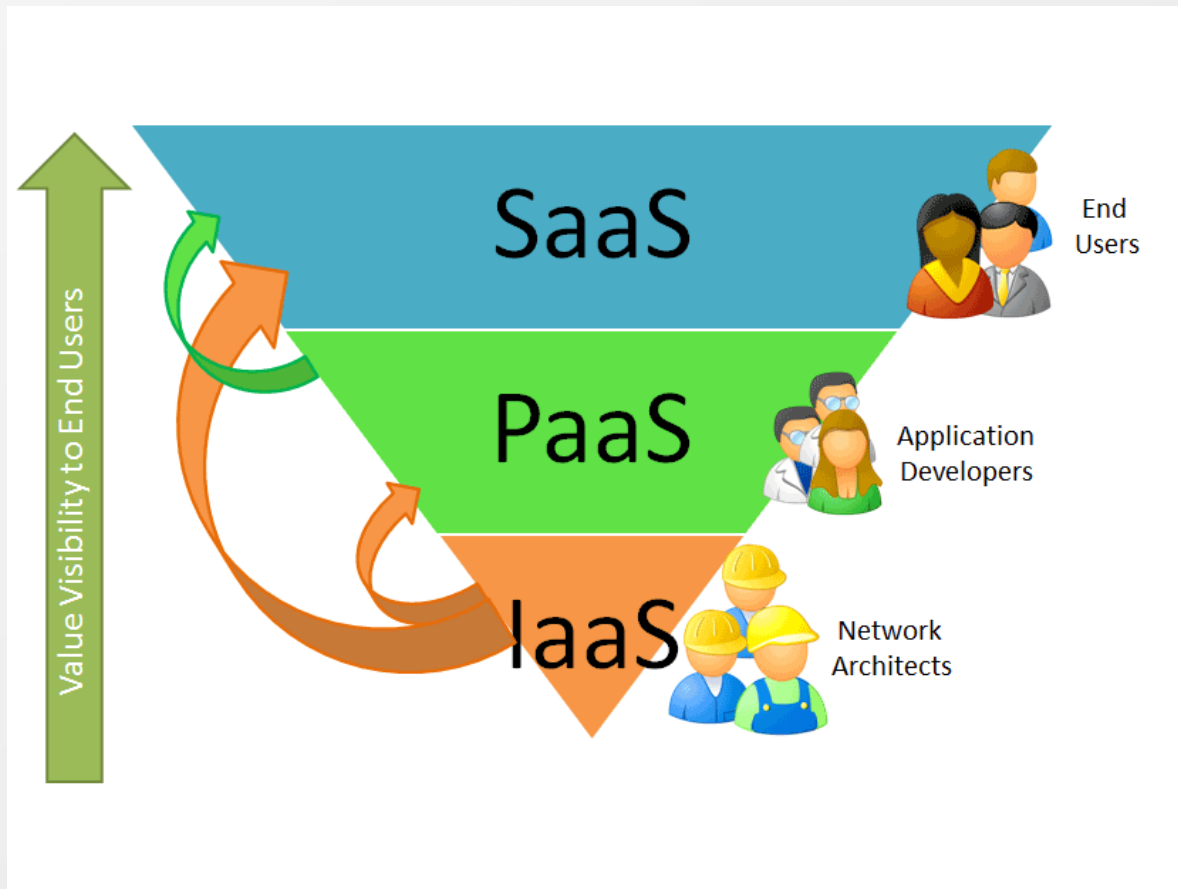- Shared
- Metered by use
- Uses internet technologies

History

1960s	First ideas
1999	Salesforces.com
2002	Google Search API Amazon Web Services
2006	Amazon Elastic Cloud
2008	Google App Engine Microsoft Azure Platform
2009	VMware vCloud Express
2010	Microsoft Office WebApps Google Services
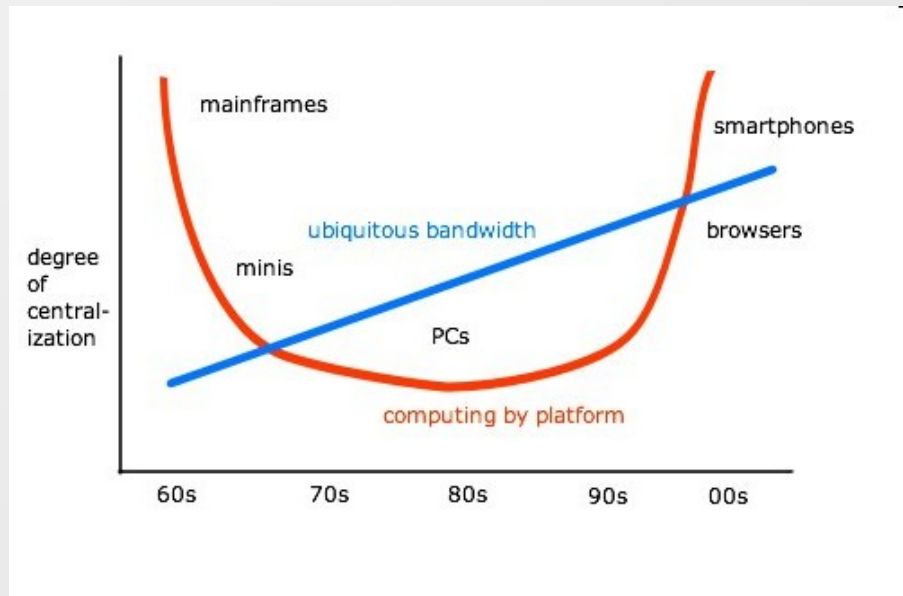
Cloud Services Stack I

Software
    as a Service

Platform
    as a Service

Infrastructure
    as a Service

Cloud Services Stack II

- **Virtualization** is a basis for all cloud services
  - Everything runs on virtualized HW
- **Infrastructure** defines basic available services
  - Operating system
  - Simple services
- **Platform** allows to build applications fast
  - Compatibility!
- **Application** is an only part visible for end users
 - Important is the value for end user

Types of cloud platforms

- **Public cloud**
    - "Classical" model
    - The same functionality for all users
- **Private cloud**
    - Used only by one organization
- **Community cloud**
    - Functionality is used by several parties which share the same interest (e.g. area of business)
- **Hybrid cloud**
    - Use of two or more clouds (public, private, community) together as a one platform

Amazon Web Services (AWS)

Amazon is not only a web shop
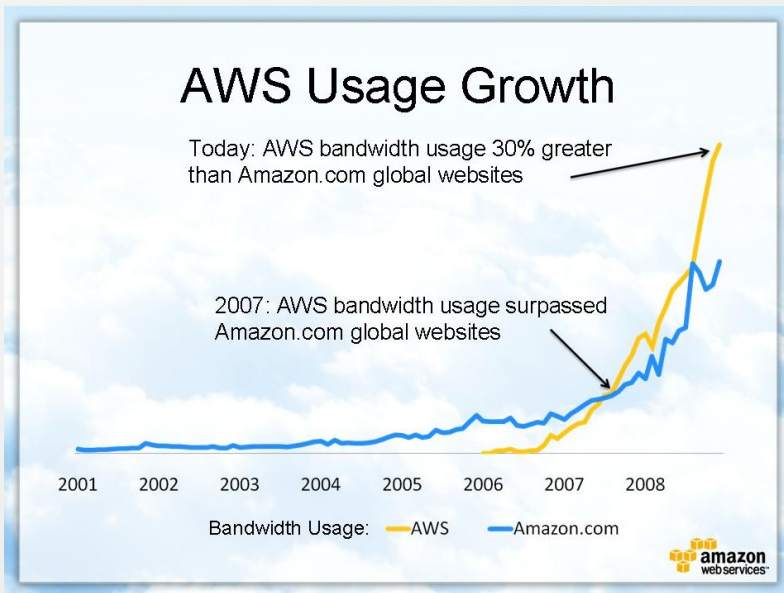
Amazon is a leader in **IaaS** and **PaaS**

AWS introduced in 2002

- Online services for other web sites
- Accessing web shop data
- Evolved into big amount of services from different areas
- Available worldwide
- Billed on usage
- Accessed via HTTP using REST and SOAP
- Number of services increase every year

Amazon Web Services – Regions and availability zones I

Available regions (February 2013)

- US East – Northern Virginia (Mar 2006)
- US West – Northern California (Dec 2009)
- US West – Oregon (Nov 2011)
- EU – Ireland (Nov 2007)
- Asia-Pacific – Singapore (Apr 2010)
- Asia-Pacific – Tokyo (Apr 2011)
- Asia-Pacific – Sydney (Nov 2012)
- South America – Sao Paulo (Dec 2011)
- AWS GovCloud – Special region for US government in NW of USA (Aug 2011)
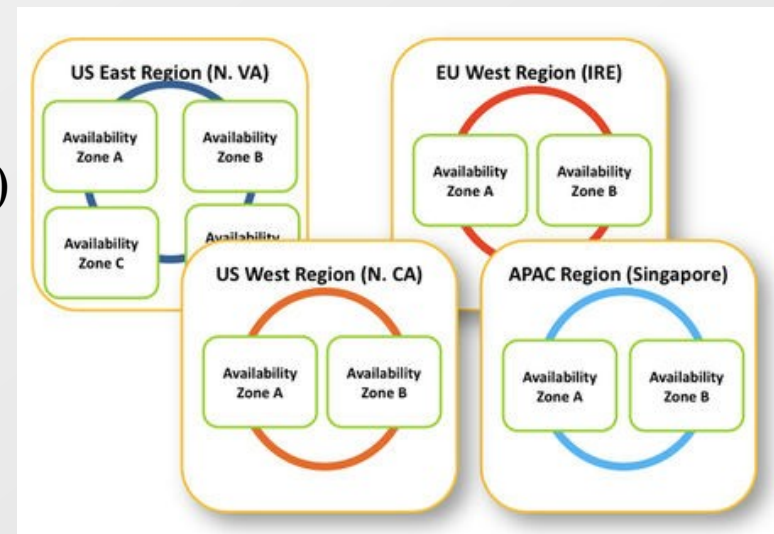
# Amazon Web Services – Regions and availability zones II

Availability zones

- Distinct locations within one region
- Availability zones in the region are physically separated
- Even disasters like flooding or tornados should affect only single availability zone
- Transfers between availability zones in the same region are cheap

AWS Edge Locations

- Content delivery network for distributing static content
- Distributed around the all world
    - North / South America – 13 locations
    - North America – 1 location
    - Europe – 8 locations
    - Asia – Pacific – 5 locations

Amazon Web Services – Regions and availability zones III

- Each region has two or more availability zones
- You put your application to the region which is "nearest" to your customers
- You can use more regions (but you have to pay for each)
- Not all services are available in all regions
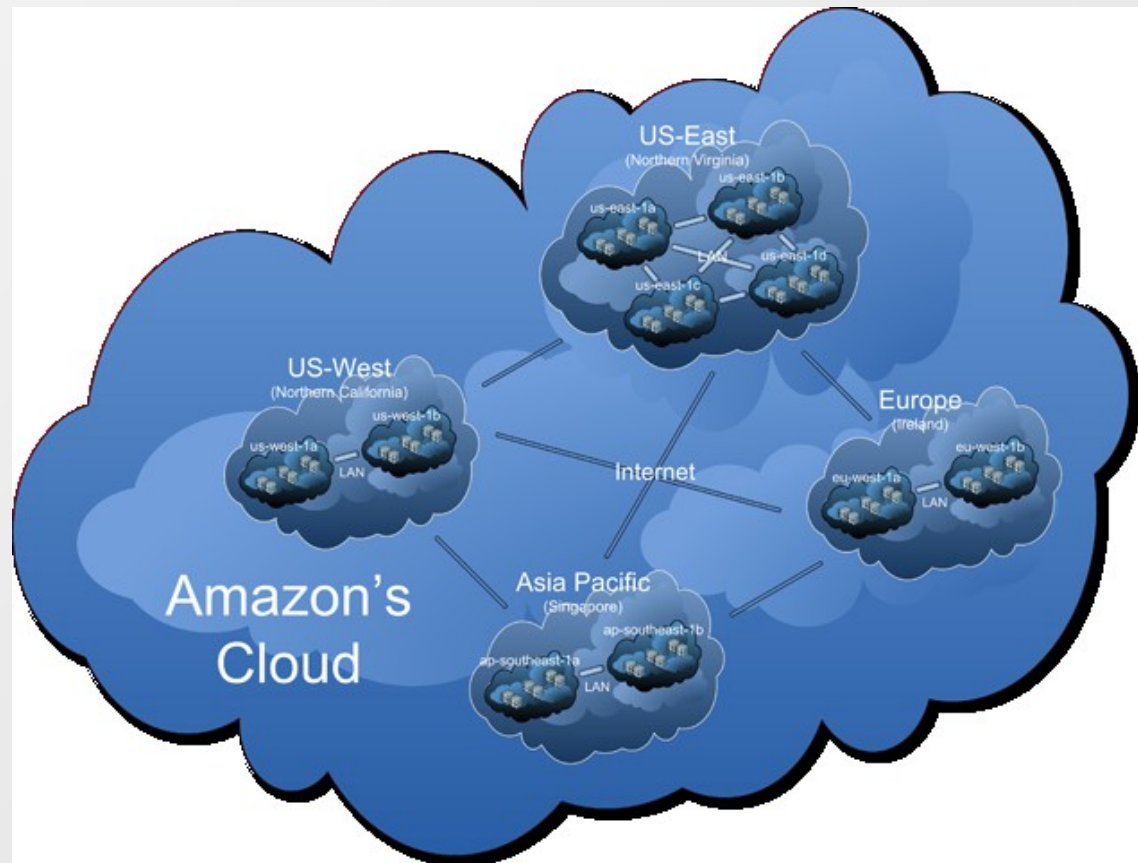
Amazon Web Services – Portfolio I

Computing

- Elastic Compute Cloud (EC2, August 2006)
  - Virtualized OS in Amazon's infrastructure
- Auto Scaling (May 2009)
  - Allows to automatically scale EC2 capacity up or down
- Elastic Load Balancing (May 2009)
  - Distributes requests to multiple EC2 instances
- Elastic MapReduce (April 2009)
  - Process vast amount of data (Hadoop framework)

Content Delivery

- CloudFront (November 2008)
  - Content delivery service

Amazon Web Services – Portfolio II

Storage

- Simple Storage Service (S3, March 2006)
  - Key-value storage for big objects (images, video, documents, etc.)
- Elastic Block Store (EBS, August 2008)
  - Block device mounted by EC2 instances
- Import/Export (May 2009)
  - Allows to import/export huge amount of data from/to portable device
- Storage Gateway (January 2012)
  - Integrates on-premises IT environments with Cloud storage
- Glacier (August 2012)
  - Low cost storage service for data archiving and backup

Amazon Web Services – Portfolio III

Database
- Relation Database Service (October 2009)
  - Relational database in the cloud (MySQL, Oracle)
- SimpleDB (December 2007)
  - Non-relational data store
- DynamoDB (January 2012)
  - NoSQL database
- ElastiCache (August 2011)
  - In-memory cache in the cloud
- Redshift
  - Petabyte-scale data warehouse service (data analysis)

Amazon Web Services – Portfolio IV

Application Services

- Simple Queue Service (SQS, July 2006)
  - Message queue in the cloud
- Simple Notification Service (SNS, April 2010)
  - Sends notifications from the cloud
- Simple Email Service (SES, January 2011)
  - Scalable e-mail sending service
- CloudSearch (April 2012)
  - Managed search service
- Simple Workflow Service (February 2012)
  - Workflow service for coordinating application components
- Elastic Transcoder (January 2013)
  - Easy-to-use scalable media transcoding

Amazon Web Services – Portfolio V

Networking

- Route 53 (December 2010)
  - Scalable domain name system (DNS)
- Virtual Private Cloud (August 2009)
  - Allows to create isolated services in virtual network
- Direct Connect (August 2011)
  - Connects your existing private data center with Amazon's cloud

Payments & Billing

- Flexible Payments Service (FPS, August 2007)
  - Payment service for developers
- DevPay (December 2007)
  - Simple to use payment and account management service

Amazon Web Services – Portfolio VI

Deployment and Management
- Management Console (January 2009)
  - Web-Based User Interface
- Identity and Access Management (IAM, September 2010)
  - Securely controls access to services and resources
- CloudWatch (May 2009)
  - Monitoring of cloud resources
- Elastic Beanstalk (January 2011)
  - Deployment and management tool
- CloudFormation (February 2011)
  - Tool for creating templates of related AWS resources
- Data pipeline (December 2012)
  - Orchestration Service for Periodic, Data-Driven Workflows
- AWS Marketplace (April 2012)
  - Find software that runs on EC2 and launch with 1-Click

Amazon Web Services – Portfolio VII

Web Traffic

- Alexa Web Information Service (October 2004)
    - Web statistics
- Alexa Top Sites (January 2006)
    - Top internet sitest statistics

Workforce

- Mechanical Turk (November 2005)
    - Manage tasks made by human via cloud

Amazon Web Services – Compute Services I

- Elastic Compute Cloud (EC2) is an infrastructure for running virtual machines
- EC2 is a core functionality for all customer specific computation
- EC2 works as IaaS
    - You can install everything what you want (licenses!)
- There exists pre-prepared AMIs (Amazon Machine Images) – virtual machines for "ready to use" (more than 2000 AMIs exists in February 2013)
- AMIs are available with following OSes:
    - Linux: Red Hat, OpenSUSE, Oracle, Ubuntu, Fedora, Gentoo, Debian, …
    - Windows server: 2003 R2, 2008, 2008 R2, 2012
    - Open Solaris

Amazon Web Services – Compute Services II

- AMIs also contain pre-prepared middleware or even applications
  - Databases, web servers, application development environments, application servers, video encoding & streaming, …
  - SAP HANA, Oracle ID Edwards, Oracle PeopleSoft, …
- Price of AMIs is AMI dependant
  - Some AMIs are free of charge
  - For other AMIs you pay per hours virtual machine (instance) is running

You can create your own EC2 instances with your required software!

Amazon Web Services – Compute Services III

- Amazon Elastic Compute Cloud represents core functionality for you app
- You deploy your application to given region
- For dynamic behavior use Auto Scaling functionality
    - You define rules for higher/lower workloads
        - Based on current statistics
        - Time scheduled
    - Good when you have peaks in usage
- Elastic Load Balancing allows to distribute requests to your app
    - Can increase availability of your application
    - Works together with Auto Scaling
- Elastic IP Addresses allows you to have virtual IP addresses
    - Public IP addresses mapped dynamically to given instance

Amazon Web Services – Compute Services IV

- ▣ You can monitor instances via CloudWatch
- ▣ Amazon guarantees 99.95% availability for region
- ▣ By default instances are not backed-up!
    - ▣ Don't store your data into the instance. Use e.g.
        - ▣ S3 for simple unstructured data
        - ▣ Database for structured data (keep scalability in mind)
        - ▣ Elastic Block Store (EBS - file system) – can be attached just to one EC2 instance
    - ▣ For storage you have to pay extra
- ▣ If you run several instances (e.g. in different regions) you pay for each
    - ▣ Prices in different regions are usually different
- ▣ You can order on-demand, reserved or spot instance

Amazon Web Services – Compute Services V

Amazon EC2 surroundings

- Around EC2 are management services
    - Auto scaling
    - Elastic load balancing
    - CloudWatch
    - Management console
    - ...

Amazon Web Services – Instance types I

You have to choose hardware on which you want to run, e.g.

- Standard first generation small instance (default) – 1.7 GiB memory, 1 ECU (1 virtual core with 1 ECU), 160 GB instance storage, 32 or 64 bit

- Standard first generation medium instance – 3.75 GiB memory, 2 ECUs (1 virtual core with 2 ECUs), 410 GB instance storage, 32 or 64 bit

- Standard first generation large instance – 7 GiB memory, 4 ECUs (2 virtual cores with 2 ECUs each), 850 GB instance storage, 64-bit

- Standard first generation extra large instance – 15 GiB memory, 8ECUs (4 virtual cores with 2 ECUs each), 1690 GB instance storage, 64 bit

- Standard second generation extra large instance – 15 GiB memory, 13 ECUs (4 virtual cores with 3.25 ECUs each), EBS storage only, 64 bit

ECU = EC2 Compute Unit. Equivalent of 1.0-1.2 GHz 2007 Opteron or Xeon processor

Amazon Web Services – Instance types II

- Standard second generation double extra large instance – 30 GiB memory, 26 ECUs (8 virtual cores with 3.25 ECUs each), EBS storage only, 64 bit

- Micro instance – 613 MiB memory, up to 2 ECUs (for short periodic burst), EBS storage only, 32 or 64 bit

- High memory extra large instance – 17.1 GiB memory, 6.5 ECUs (2 virtual cores with 3.25 ECUs each, 420 GB of instance storage), 64 bit

- High memory double extra large instance – 34.2 GiB memory, 13 ECUs (4 virtual cored with 3.25 ECUs), 850 GB instance storage, 64 bit

- High memory quadruple extra large instance – 68.4 GiB memory, 26 ECUs (8 virtual cores with 3.25 ECUs each), 1690 GB instance storage, 64 bit

- High CPU medium instance – 1.7 GiB memory, 5 ECUs (2 virtual cores with 2.5 ECUs each), 350 GB instance storage, 32 or 64 bit

Amazon Web Services – Instance types III

- High CPU extra large instance – 7GiB memory, 20 ECUs (8 virtual cores with 2.5 ECUs each), 1690 GB instance storage, 64 bit

- Cluster compute eight extra large instance – 60.5 GiB memory, 88 ECUs (2 x Intel Xeon E5-2670, eight core), 3370 GB instance storage, 64 bit

- High memory cluster eight extra large instance – 244 GiB memory, 88 ECUs (2 x Intel Xeon E5-2670, eight core), 240 GB SSD instance storage, 64 bit

- Cluster GPU quadruple extra large instance – 22 GiB memory, 33.5 ECUs (2 x Intel Xeon X5570, quad core), 2x NVIDIA Tesla "Fermi" M2050 GPUs, 1690 GB instance storage, 64 bit

- High I/O quadruple extra large instance – 60.5 GiB memory, 35 ECUs, 2 SSD-based volumes each with 1024 GB of instance storage

- High storage instance – 117 GiB memory, 35 ECUs, 24 hard disk drives each with 2 TB of instance storage, 64 bit

Amazon Web Services – Simple Storage Service (S3) I

- Purpose of Simple Storage Service (S3) is to store static unstructured data (pictures, videos, documents, …)
- Simple key-value storage with possible metadata organized into "buckets"
- Maximal size of each data (file) is 5 TB, unlimited number of objects
- Data accessible as http://s3.amazonaws.com/*bucket/key* (DNS CNAME supported - http://yourdomain/bucket/key)
- Data stored in regions, they never leave region if you don't require it
- Data can be versioned (different data with the same key)
- Supports protocols HTTP and BitTorrent
- 99.999999999% durability, 99.99% availability
- With Reduced Redundancy Storage (RRS) durability and availability 99.99%

Amazon Web Services – Simple Storage Service (S3) II

- With Amazon Glacier you can reduce storage costs (with some restrictions)
    - Glacier is low-cost storage with long retrieval times (several hours)
- "Unexpected" behavior due to distributed nature
    - Immediate read after write can return "not available"
    - Immediate list of bucket content after write can return list without currently written object
    - Immediate read after re-write can still return "old" data
    - Immediate read after delete can still return "deleted" data
- Some web sites uses S3 for hosting static content
- E.g. well known service DropBox is using Amazon S3 for storing data

Amazon Web Services – Database Services I

- **Relational Database Service**
  - Allows to easily migrate your existing applications
  - Support for MySQL or Oracle
  - By default you are allowed to run up to 20 instances (but you can ask for more)
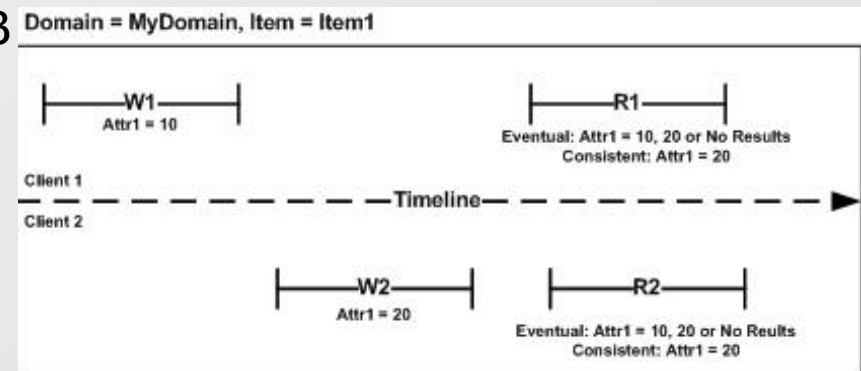 - Size limit per instance is 1TB
- **SimpleDB**
  - Non-relational data store, automatically indexes your data
  - Each row can have different attributes
  - Domain – item – attribute approach
  - Limit is 10GB per domain, you can have 250 domains by default (but you can ask for more)

Amazon Web Services – Database Services II

▣ **SimpleDB (continues)**

    ▣ 256 attributes per item, attribute length is limited to 1024 bytes

    ▣ Consistency

        ▣ Eventually consistent reads (default)

        ▣ Consistent reads

    ▣ Limited transactions (only conditional Put/Delete)

    ▣ Good for lower-scale workloads

    ▣ Older technology than DynamoDB

Domain = MyDomain, Item = Item1

Client 1

W1 — Attr1 = 10

R1 — Eventual: Attr1 = 10, 20 or No Results
Consistent: Attr1 = 20

Timeline

Client 2

W2 — Attr1 = 20

R2 — Eventual: Attr1 = 10, 20 or No Reults
Consistent: Attr1 = 20

Amazon Web Services – Database Services III

- **DynamoDB**
  - NoSQL (non-rational) database
  - Fast predictable performance, stored all data on SSD
  - No limits on amount of data, scales automatically
  - By default you can order guaranteed throughput rate up to 10000 reads or writes per second (in exceptional cases you can order higher)
  - Table – item – attribute approach
  - Requires primary key for item, only primary key is indexed
    - Primary key can be compound (hash and range type primary key)
  - Item size limit is 64kB (sum of all attribute names and values)
  - Not so flexible queries
  - Supports batch processing

Amazon Web Services – Database Services IV

☑ **DynamoDB (continues)**

- ☑ Supports scalar data types: number, string, binary
  - ☑ Plus multi-valued types: number set, string set, binary set
- ☑ Consistency
  - ☑ Eventually consistent reads (default, better performance)
  - ☑ Consistent reads
- ☑ Integrates Elastic MapReduce
  - ☑ MapReduce is programming model for processing large data
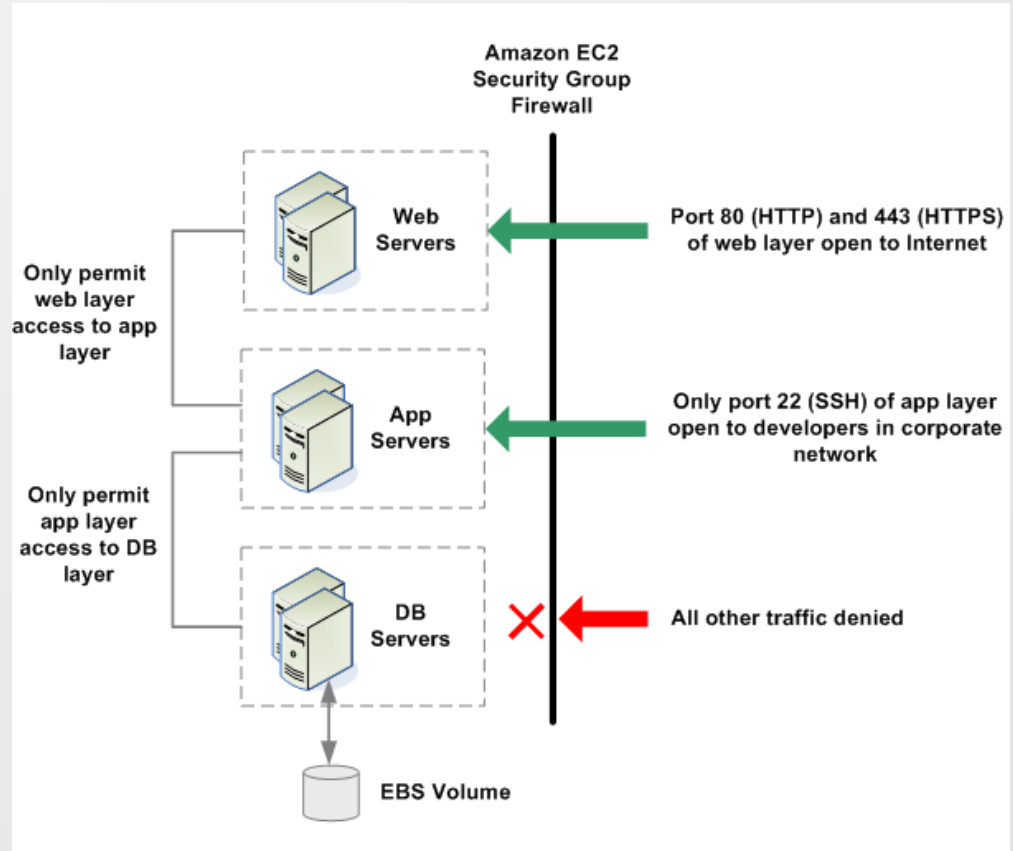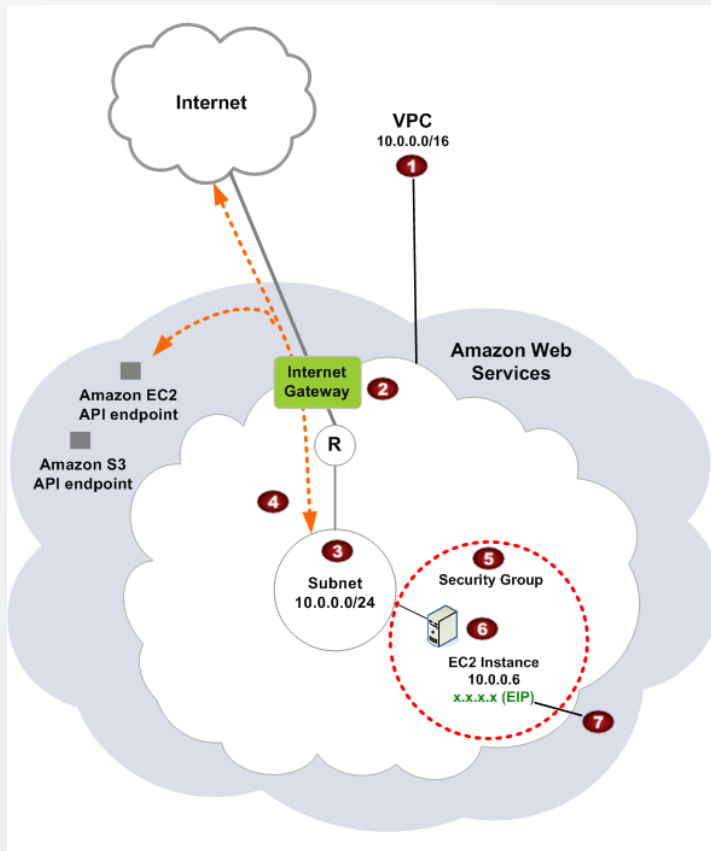- ☑ Still in beta phase (in February 2013)

You can run also other databases like Microsoft SQL Server, Oracle SQL, IBM DB2, Postgre SQL, Sybase, etc. but not all of them are officially supported.

Amazon Web Services – Virtual Private Cloud

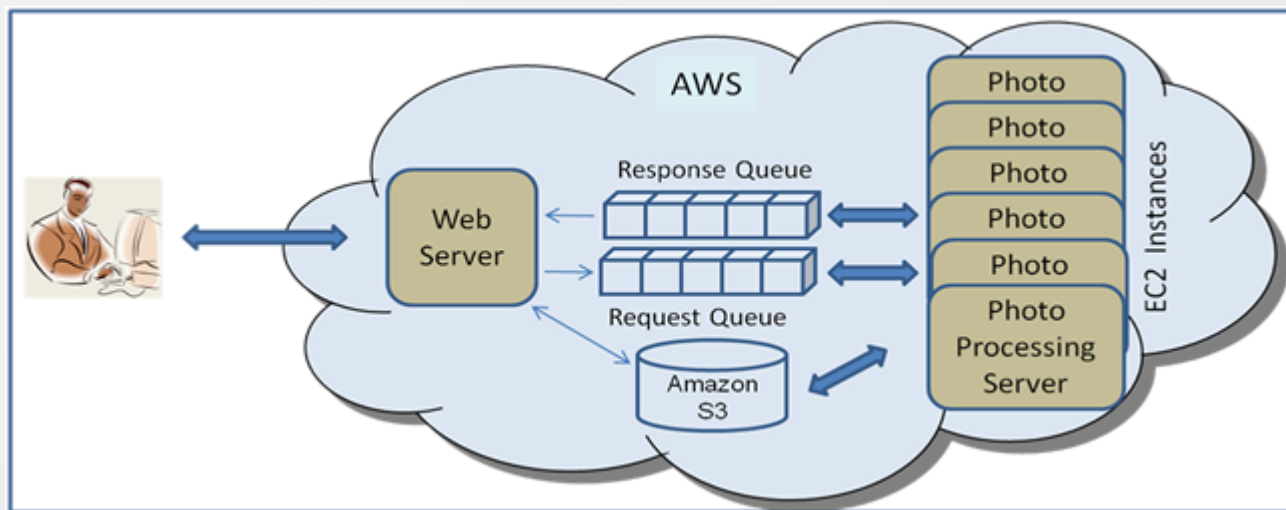⇨ Connection via VPN

⇨ Isolated network access

⇨ Subnets
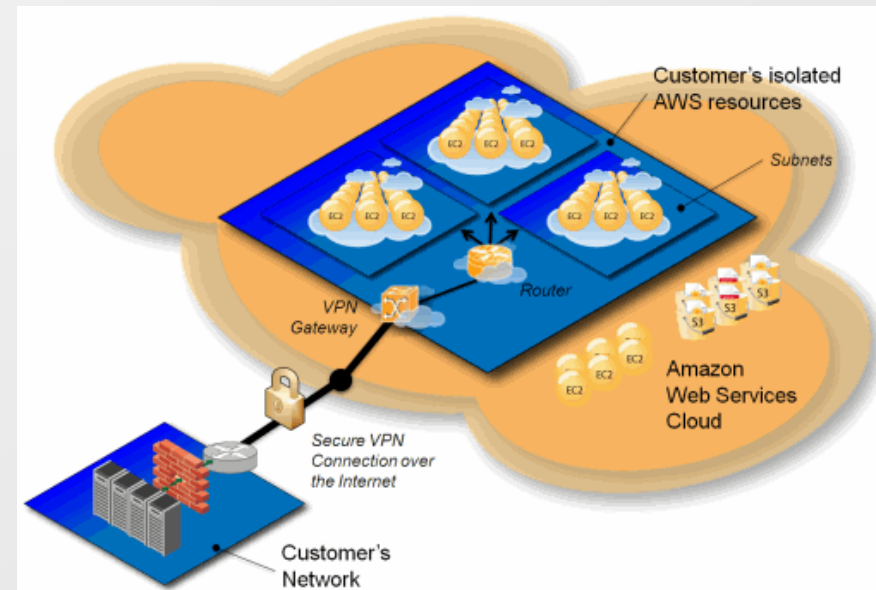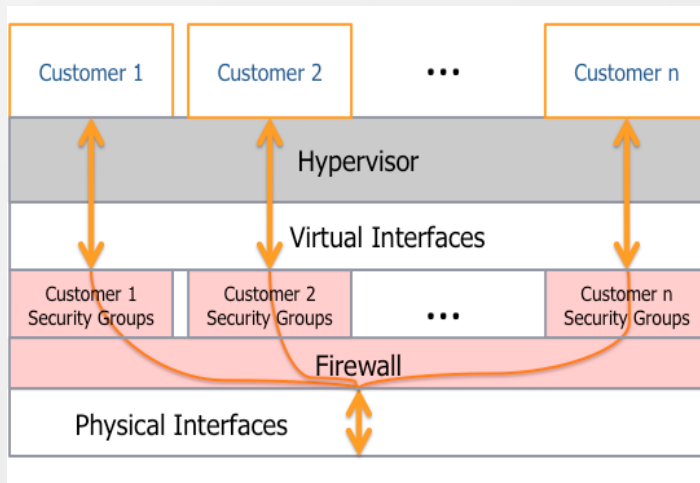
Amazon Web Services – Virtual Private Cloud, Security

Amazon Web Services – EC2 and S3 Example

Photo processing example

Amazon Web Services – Pricing I

Pricing

- Different for region, service, OS, ...
- Based on usage - you can have AWS account but you don't need to pay at all if you don't use Amazon's services
- Usually calculated per hour or per GB
- Some services are paid per item or per item and month (e.g. CloudWatch, Export/Import)
- Hours for EC2 instances means time when your instance is live!
- Prices are developing through years
- There are "action prices" all the time

Amazon Web Services – Pricing II

EC2 pricing depends on

- Operating system – Linux (Unix) / Windows (Linux is cheaper)
- Region
- Instance type (small / medium / large / ...)
- On-demand instance / reserved instance / spot instance
- Length of contract for reserved instances (1 year / 3 years)

S3 pricing depends on

- Region
- Amount of stored data
- Standard redundancy / reduced redundancy / glacier storage
- Number of GET, PUT, POST, ... requests
- Amount of data transferred

Amazon Web Services – Pricing III

New AWS customers get each month for one year (February 2013)

- 750 hours of EC2 running Linux/Unix Micro instance usageOperating system – Linux (Unix) / Windows (Linux is cheaper)
- 750 hours of EC2 running Microsoft Windows Server Micro instance usage
- 750 hours of Elastic Load Balancing plus 15 GB data processing
- 30 GB of Amazon Elastic Block Storage (EBS) plus 2 million IOs and 1 GB snapshot storage
- 15 GB of bandwidth out aggregated across all AWS services
- 1 GB of Regional Data Transfer

Amazon Web Services – Pricing IV

Amazon pricing overview

- http://aws.amazon.com/pricing/

Amazon pricing for specific products

- http://aws.amazon.com/ec2/pricing/
- http://aws.amazon.com/s3/pricing/
- http://aws.amazon.com/dynamodb/pricing/
- …

Simple Monthly Calculator and TCO calculator

- http://aws.amazon.com/calculator/
- http://aws.amazon.com/tco-calculator/

Amazon Web Services – Alternatives

Open source alternative – Eucalyptus

⇨ Runs some of Amazon Web Services under your Linux

⇨ http://www.eucalyptus.com
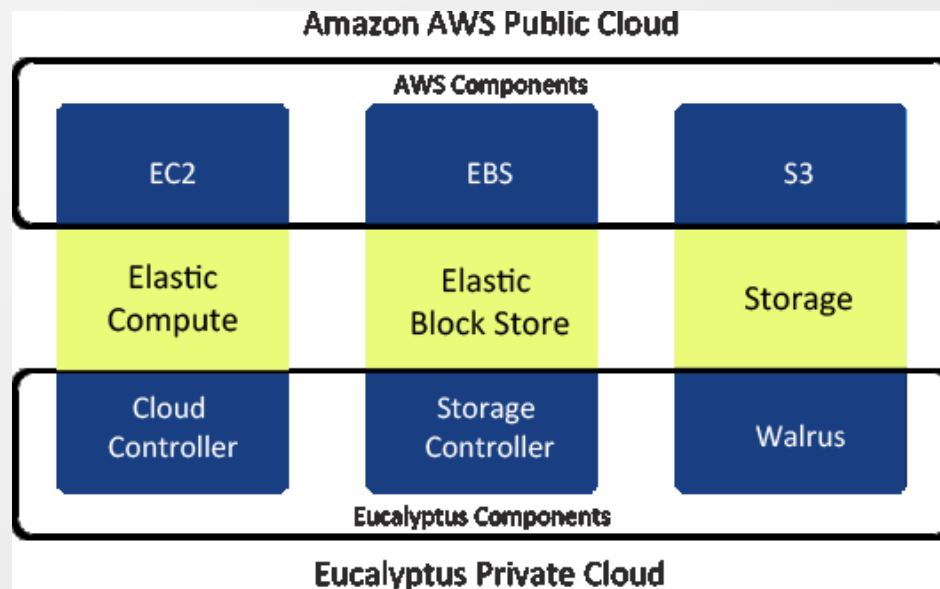
Microsoft Azure Platform

- Microsoft's response to the cloud computing
    - Announced in October 2008
    - Commercially available in February 2010
    - IaaS, PaaS and SaaS
- Windows Azure OS as a "base OS"
    - Windows Server 2008 SP for Guest OS 1.x
    - Windows Server 2008 R2 or R2 SP1 for Guest OS 2.x
    - Windows Server 2012 for Guest OS 3.x
- SLAs (monthly)
    - 99.95% for compute
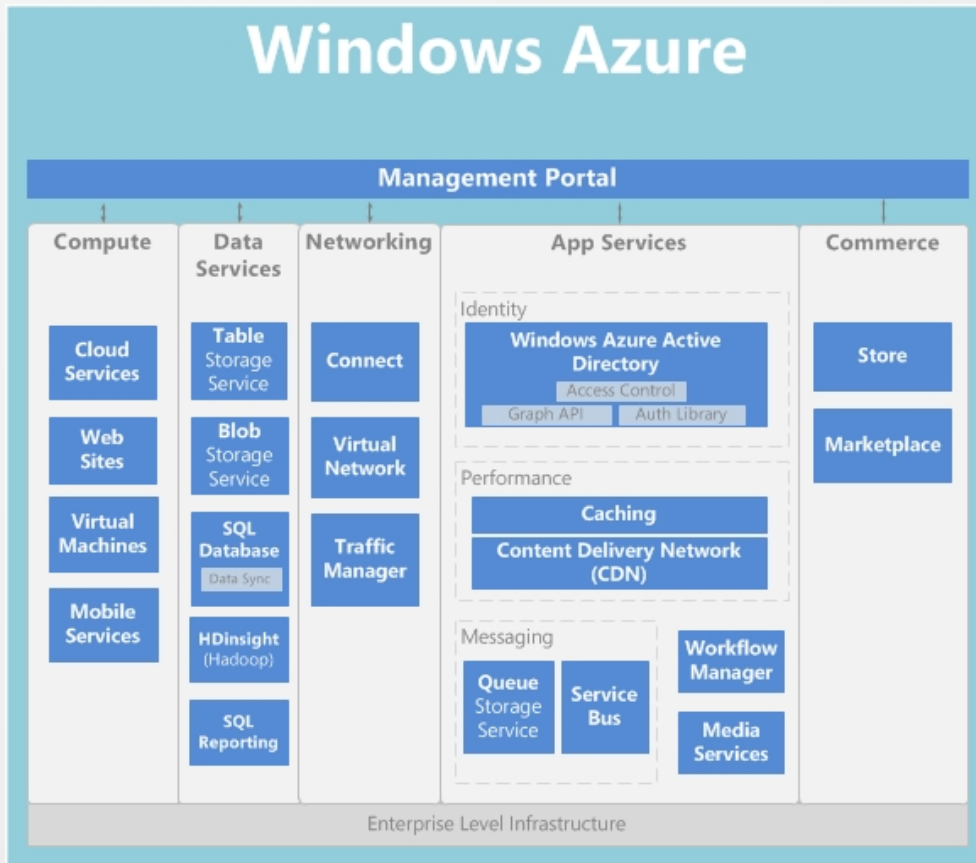   - 99.9% for storage, CDN, SQL Database, service bus, caching

Windows Azure Data Centers (2013)

- US North Central – Illinois
- US South Central – Texas
- US East – Virginia
- US West - California
- West Europe - Amsterdam
- North Europe - Dublin
- East Asia - Hong Kong
- South-East Asia – Singapore

- 24 CDN locations

Windows Azure - Services overview
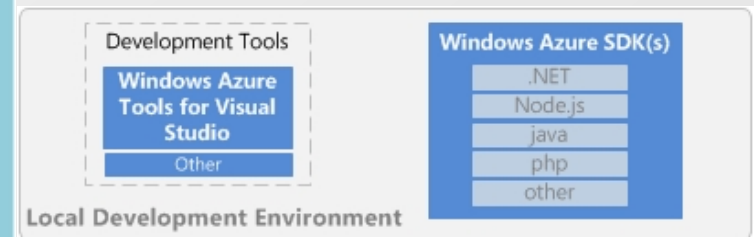
Windows Azure – Compute services I

- **Web Sites**
  - Hosting web applications written in ASP.NET, PHP or Node.js
- **Cloud Services**
  - Multi tier web applications ("advanced web sites")
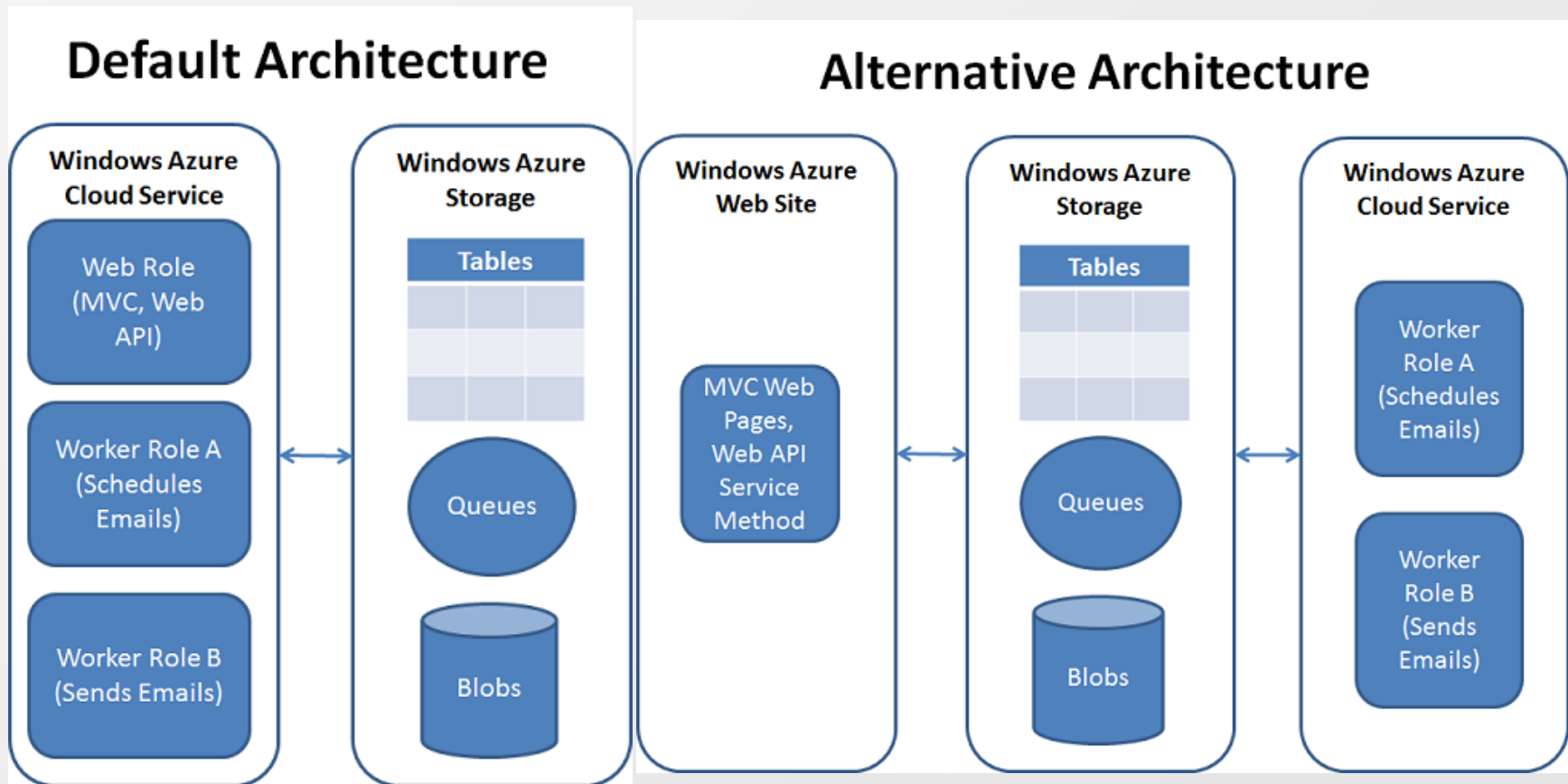   - Web role
    - Worker role
- **Virtual Machines**
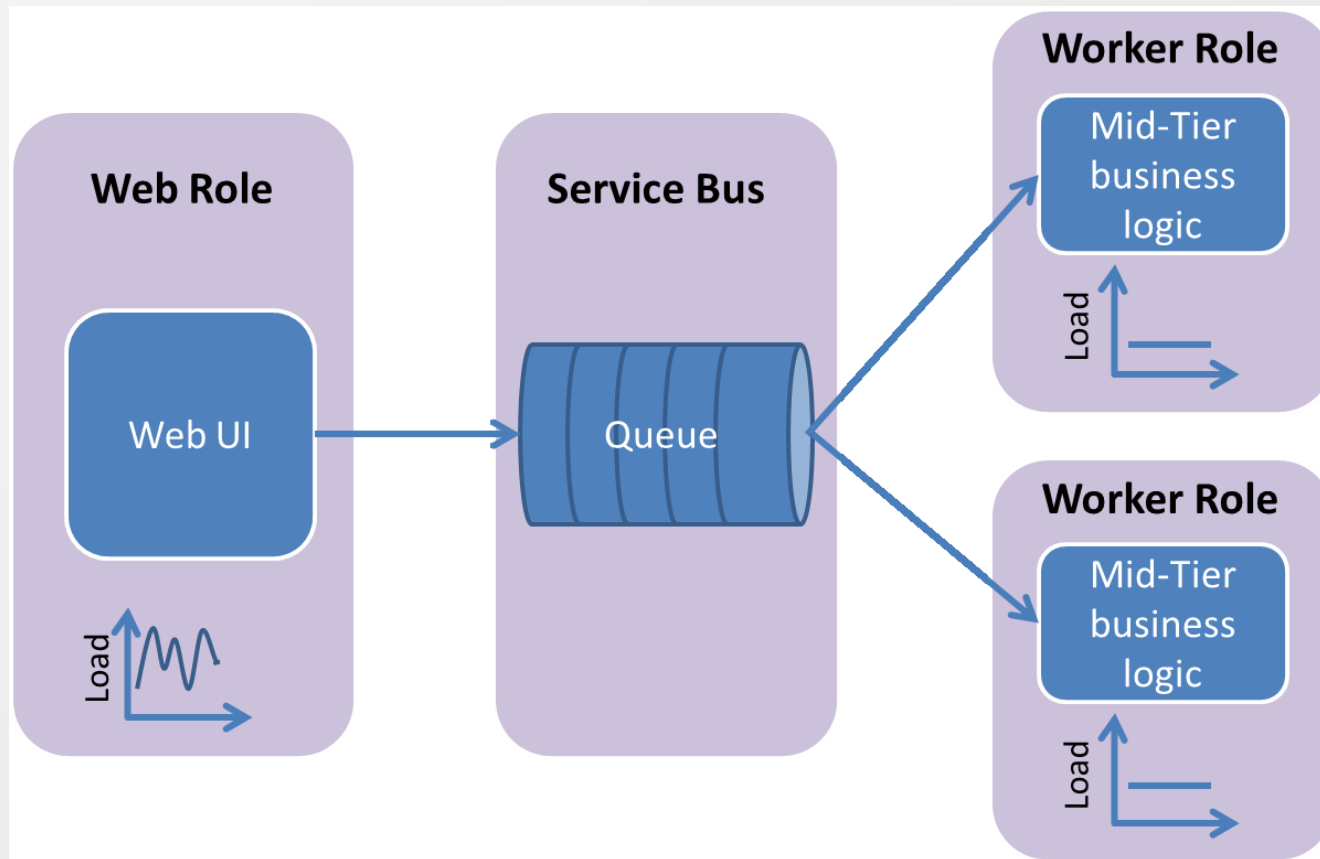 - IaaS – Run Windows or Linux, predefined or own image
- **Mobile Services**
  - PaaS for building mobile applications

Windows Azure – Compute services II

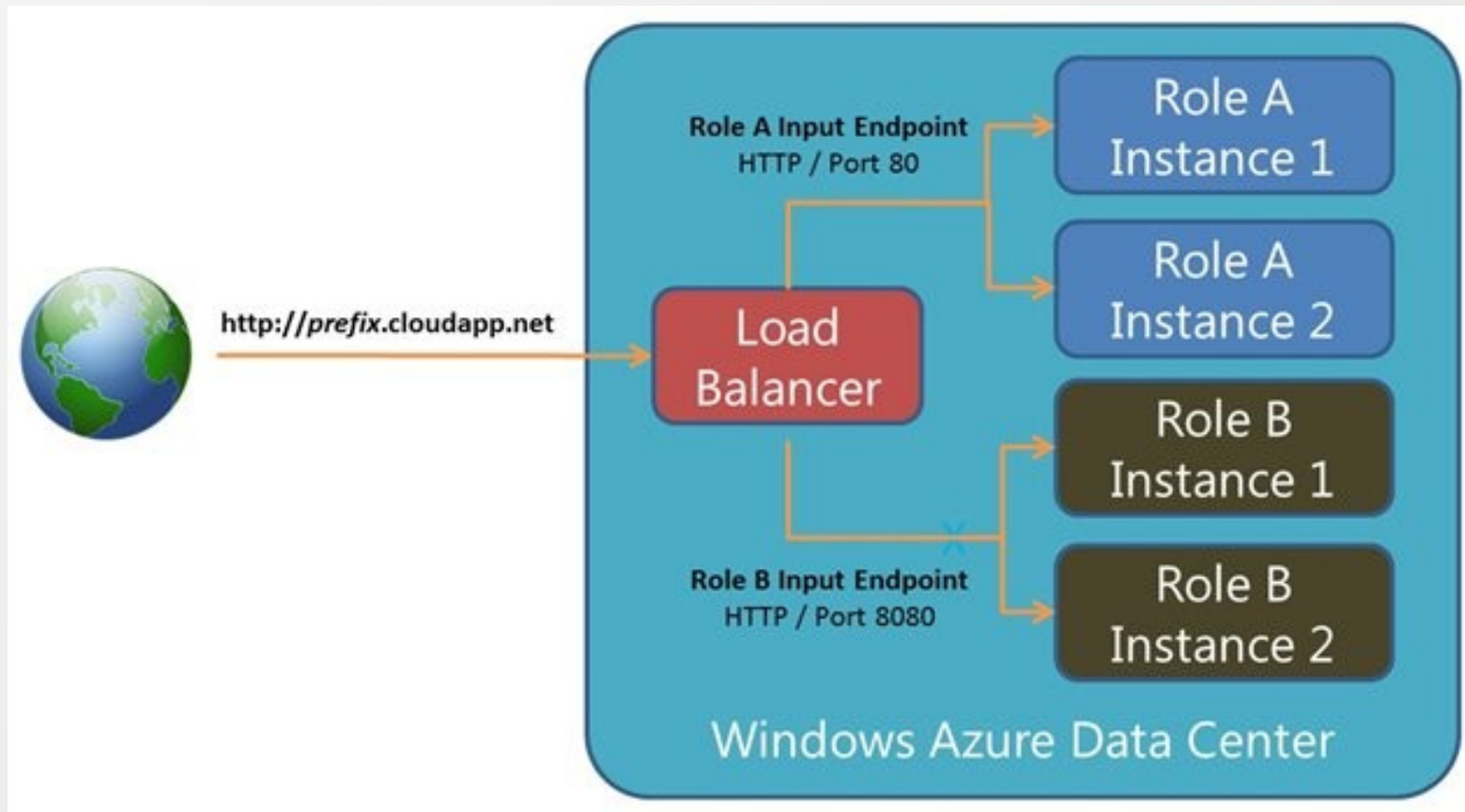# Windows Azure – Web role and worker role

Windows Azure – Load balancing explanation

Windows Azure – Data services I

- **BLOB (Binary Large Objects)**
  - Key-value storage for unstructured data (video, documents, etc.)
- **Table**
  - NoSQL database
- **Windows Azure Drive**
  - Standard NTFS drive which can be mounted to compute instance
- **SQL Database**
  - MS SQL Server database in the cloud
  - Doesn't support all features of standard MS SQL Server

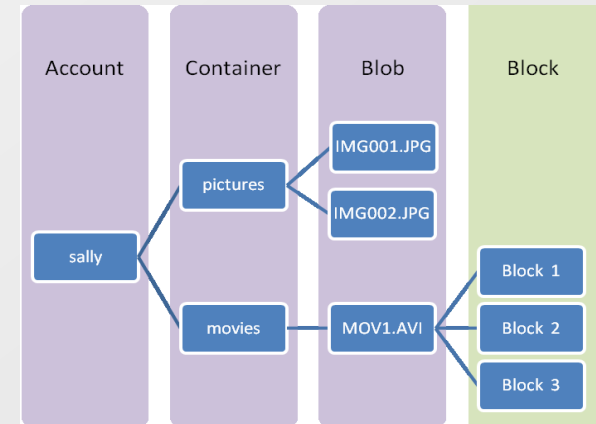# Windows Azure – Data services II

- **SQL Data Sync**
  - Service for synchronization of SQL databases
  - Either Azure <-> on-premises or Azure <-> Azure
- **SQL Reporting**
  - MS SQL Server reporting in the cloud for cloud or desktop applications
- **HDInsight**
  - Based on Apache Hadoop (framework for data-intensive distributed applications)
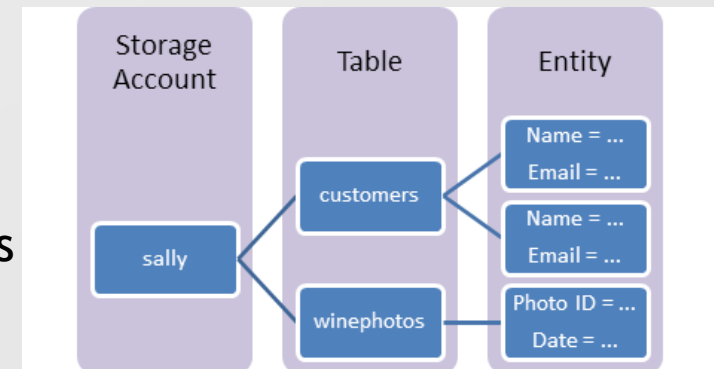
Microsoft Azure – Data Services – BLOB and Table

BLOB

- Maximal 100TB of data per account
- Blob – Key-value storage with "block" support
  - Up to 200GB for block blob
  - Up to 1TB for page blob
    - Optimized for random read/write
  - http://<storage account>.blob.core.windows.net/<container>/<blob>

Table

- NoSQL indexed database
- Maximal entity size is 1MB
- Entity can have up to 252 custom properties

Microsoft Azure – Data Services – Azure SQL Database

- Supports T-SQL
- Maximal 150 databases in each SQL Server
- Maximal limits per database
  - Web Edition – 5GB
  - Business Edition – 150GB
- For larger data you have to use sharding (split data into several databases)
- Every table must have clustered index
- Limits: no full text search, CLR types, replication, database mirroring, jobs, backup and restore, distributed transactions, …

Windows Azure – Network services

- **Windows Azure Connect**
  - IPsec connection between individual machines in your intranet and role instances running in Windows Azure
- **Windows Azure Virtual Network**
  - VPN between your intranet and Azure
- **Traffic Manager**
  - Controls distribution of user traffic to similar hosted services running in the same or different data center according to conditions like performance, price or compliance

Windows Azure – App Services I

- **Identity**
  - Access Control Service
    - Service for authentication and authorization of users
  - Active Directory Graph
    - API to Windows Azure Active Directory
- **Performance**
  - Windows Azure Caching
   - General caching service
  - Content Delivery Network
    - Hosting static content around the world

Windows Azure – App Services II

- **Messaging**
  - Storage Queues
    - Reliable and persistent messaging between application tiers running in Windows Azure
  - Service Bus
    - Platform for widespread communication
- **Workflow Manager**
  - Workflow service for authoring, hosting and managing workflows
- **Media Services**
  - Platform for building media intensive applications
- **Azure Store and Market Place**
  - Market place for buying or services

Microsoft Azure – Service Bus
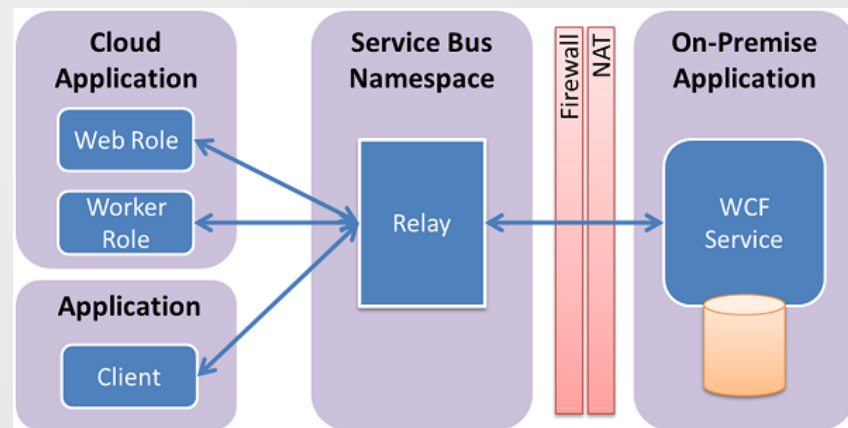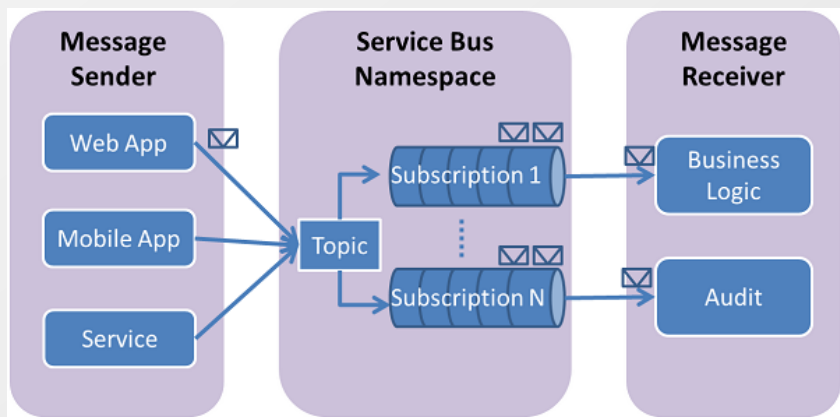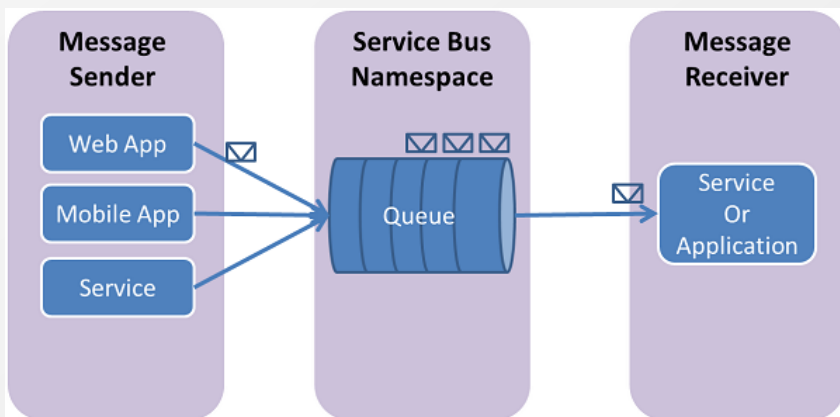
Windows Azure - Staging

- Allows to do application upgrade without downtime
- First new version is prepared into staging area
- You test the application in staging area
- Then staging area becomes production and vice versa
- Switch is done on the load balancer – it redirects all requests to the new version
- Old version is still running (currently in staging area). You are charged!
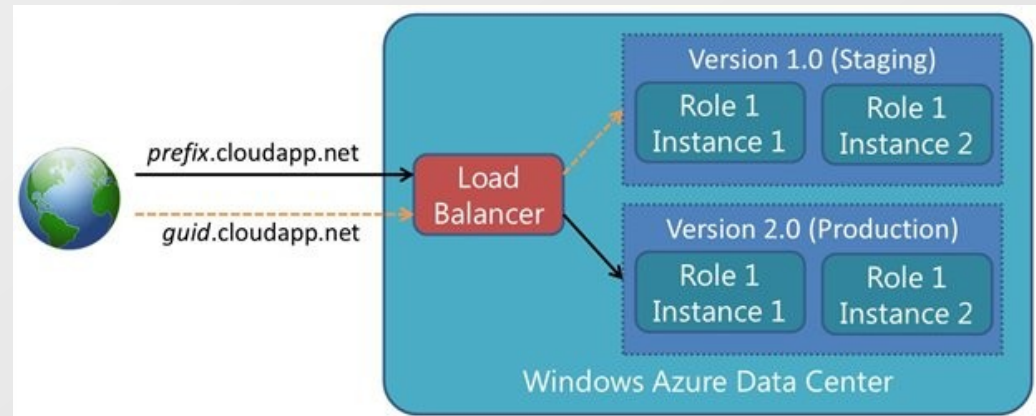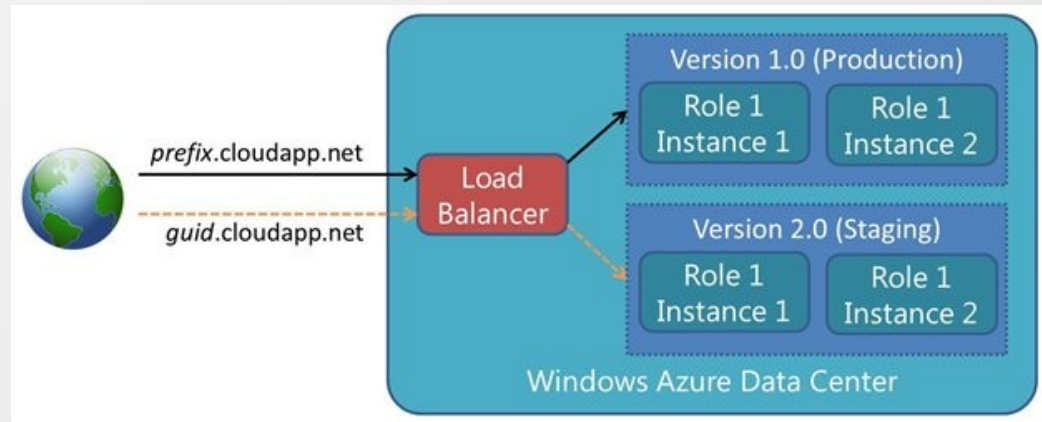
Microsoft Azure – Pricing

⊡ Prices are dependant on datacenter location (valid only for some services, some services are data center location independent)

⊡ Different virtual server sizes have different prices

⊡ Prices are based on usage (used storage, amount of send messages, amount of data transfers, etc.)

⊡ Prices are dependant on the used service (web, VM, database, ...)

⊡ You pay when your application is deployed even if it is not running!

⊡ Pricing details:

   ⊡ http://www.windowsazure.com/en-us/pricing/details

⊡ Pricing calculator:

   ⊡ http://www.windowsazure.com/en-us/pricing/calculator

Microsoft Azure – Pricing Trial

New customers will get each month for 90 days (February 2013)

- Compute: 750 small compute hours per month
- Web sites: 10 web sites
- Mobile services: 10 mobile services
- Relational database: 1 SQL database
- SQL reporting: 100 hours
- Storage: 70GB with 50,000,000 storage transactions
- Data transfer: Unlimited inbound and 25 GB outbound
- Media services encoding: 50 GB (input and output combined)
- CDN: 20 GB outbound with 500,000 transactions
- Service bus: 1,500 relay hours and 500,000 transactions

Microsoft Azure – Billing Explanation
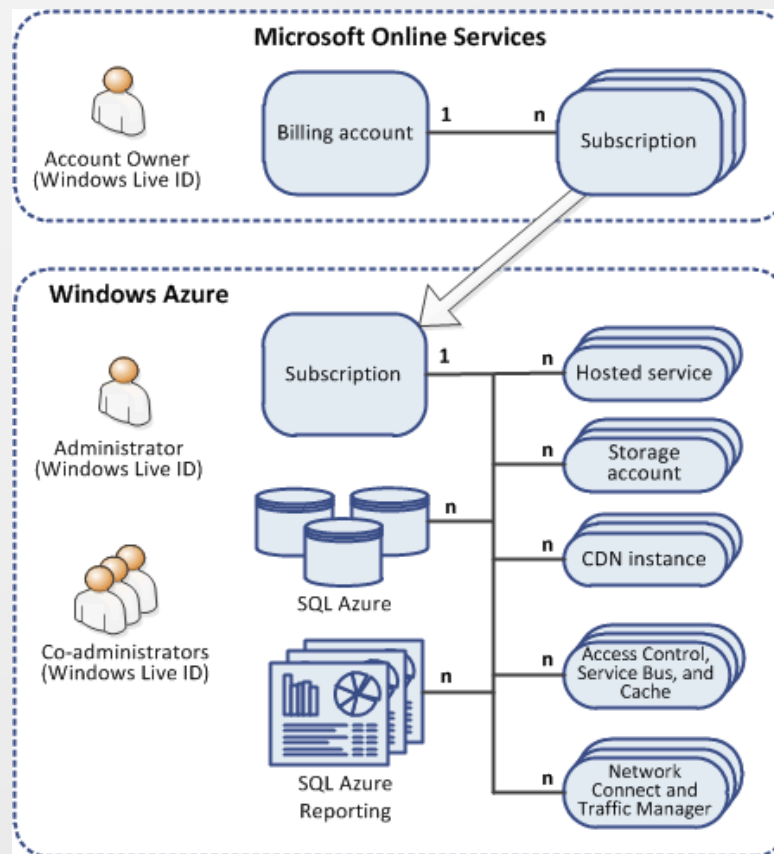
Google Cloud Platform I

- **App Engine**
  - Run web applications in the Google infrastructure
- **Compute Engine**
  - Large-scale computing workloads on Linux virtual machines
- **Cloud Storage**
  - Key-value storage
- **Big Query**
  - Data analysis service

Google Cloud Platform II

- **Cloud SQL**
  - MySQL database in Google cloud
- **Prediction API**
  - Pattern-matching and machine learning capabilities
- **Translate API**
  - Machine translation between different languages

- **SLA**
  - 99.95% for App Engine, Compute Engine, Cloud SQL
  - 99.9% for Cloud Storage, Prediction API, BigQuery

Google Cloud Platform – App Engine I

Offers three environments where your application

- Java
  - Java 5 or 6 JVM in safe sandboxed environment
  - You can use standard JavaServer Pages (JSP)
  - Some limits exist. You cannot:
    - Write to the file system
    - Open a socket or access another host dire
    - Create processes or threads directly or indirectly
    - Directly communicate with system via JNI
  - There are limits for web application in general e.g.
    - Maximal response time 60 seconds
    - Maximal request and response size 32MB

Google Cloud Platform – App Engine II

Supported environments (continue)

- Java
  - You can use also other languages compatible with JVM: Ruby, Scala, Jython, …
- Python
  - Python 2.5 or 2.7
  - Similar limits like for Java environment (write to file system, use sockets, create threads, size of web application, …)
- Go
  - Go is a new Google's language
  - Similar limits like for Java or Python environment

Google Cloud Platform – App Engine III

App Engine Datastore

- Schemaless object data store with some features of "standard DBs" (e.g. transactions and references to other entities)
- Objects in data store are entities one or more properties
- Properties have types like integers, strings, reference to entity, etc.
- Entities are of given kind (like person)
- Entities can be hierarchically structured
- For querying objects is used GQL (variant of SQL)
- Consistency
  - Strong consistency for reads and ancestor queries
  - Eventual consistency for other queries
- Maximum entity size is 1 MB

Google Cloud Services – App Engine IV

- Several instance types (with different memory and CPU) available
- **Frontend instance**
    - Runs code based on some request
    - Time how long execution of the request can run is limited
    - Scales up and down automatically as number of requests increase and decrease
- **Backend instance**
    - Runs code in the background (queues, cron tasks, etc.)
    - Doesn't scale automatically – you specify number of instances you want to use
    - Resident (runs continuously) or dynamic (started and stopped dynamically based on required work)

Google Cloud Platform – Cloud Storage I

- Key-value storage for big objects ("bucket" based)
- Maximum file size is 5TB
- Consistency
  - Strong read after write consistency for objects (it is guaranteed that immediate read after write will return object)
  - Strong read after delete consistency for objects
  - "List" operations are eventually consistent
- Different versions of the same object possible
- You can specify location of bucket: US or Europe

Google Cloud Platform – Cloud Storage II

- Data accessible as
    - https://storage.cloud.google.com/bucket/object
- Redirects via DNS CNAME possible - http://yourdomain/object
- Support access control
- Supports resumable uploads
- Supports streaming

Google Cloud Platform – Cloud SQL

- Based on MySQL 5.5
- Data stored in US or EU region
- Maximal database size is 100GB
- Maximal request and response size is 16 MB
- All database requests must finish within 60 seconds
- Some features are not supported against standard MySQL
    - User defined functions
    - MySql Replication
    - Some MySQL statements and functions

Google Cloud Platform – Pricing

- **App Engine**
  - You pay fixed price per application per month plus additional price above free quota for used instance hours, storage, used bandwidth, number of transactions, …
- **Compute Engine**
  - Price depends on virtual machine size, used bandwidth, disk space, location (US / EU)
- **Cloud Storage**
  - Price depends on amount of stored data, used bandwidth, location (US / EU)
- **Cloud SQL**
  - Price depends on virtual machines size, amount of stored data, used bandwidth

Google Cloud Platform – Free trial quota I

- **App Engine**
    - 28 on-demand frontend instance hours per day
    - 9 backend instance hours per day
    - 1 GB Datastore per day
    - 1 GB outgoing / incoming bandwidth per day
- **Cloud Storage (quota valid until June 30, 2013 for first project)**
    - 5GB of storage
    - 25GB download and upload of data
    - 30000 GET, HEAD requests
    - 3000 PUT, POST, list objects, list buckets

Google Cloud Platform – Free trial quota II

- **BigQuery**
  - 100 GB of data processes per month
- **Cloud SQL (quota valid until June 30, 2013)**
  - One Google Cloud SQL instance with a small amount of RAM
  - 0.5 GB of database storage

Pros of cloud computing

- You don't need to maintain your infrastructure
  - You don't need to have unused (reserved) hardware
  - You don't need appropriate specialists
- Pay only for what you use
- Costs can be lower
  - Especially in cases when you expect usage peaks
- Easy scalability for increased number of users
  - But application must be prepared as well

Cons of cloud computing

- Security and privacy
- Strong dependency on the network availability
- Vendor lock-in
    - There aren't standards to migrate between vendors
- Migration costs from existing infrastructure
- Legal aspects
    - Provider and user can have different law
- Cloud platforms provide less functionality than existing platforms
- Requires new know-how for the whole team

Děkuji za pozornost.

EVROPSKÁ UNIE

MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

OP Vzdělávání
pro konkurenceschopnost

2007-13

UNIVERSITAS · MASARYKIANA BRUNENSIS

I N V E S T I C E   D O   R O Z V O J E   V Z D Ě L Á V Á N Í