Distributed Data Management
(Summer Semester 2015)

Cloud Computing

Guest Lecture by Matthias Bender
Dr.-Ing. Matthias Bender
matthias.bender@dataone.de

• Business Development Manager (Data One GmbH, Saarbruecken)

• PhD Computer Science 2007
  Distributed Information Retrieval / P2P
  (Max-Planck-Institute for Computer Science)
# The Evolution of Computing Technology

<table>
<thead>
<tr>
<th>Mainframe (1950’s – 1970’s)</th>
<th>IT has little impact. It is a curiosity.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Client-Server (1980’s-1990’s)</td>
<td>IT invades the home and starts to change behaviors.</td>
</tr>
<tr>
<td>Web (2000’s)</td>
<td>IT enables seamless communication around the globe. Growth is explosive.</td>
</tr>
<tr>
<td>Cloud (2010’s)</td>
<td>IT becomes fully integrated into our daily lives.</td>
</tr>
</tbody>
</table>

[http://history-computer.com/ModernComputer/Software/Windows.html](http://history-computer.com/ModernComputer/Software/Windows.html)


What is Cloud Computing?

- Cloud computing refers to the practice of transitioning computer services such as computation or data storage to multiple redundant offsite locations available on the Internet, which allows application software to be operated using internet-enabled devices. (Wikipedia)

- 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 – National Institute of Standards and Technology)

- ... no commonly accepted definition ...

Cloud computing is a general term for anything that involves delivering hosted services over the Internet.
Some Popular Cloud Computing Services

- Facebook
- Gmail
- Spotify
- Dropbox
- Amazon Web Services
Gartner Hype Cycle for Emerging Technologies

http://www.gartner.com/newsroom/id/2819918
Cloud Computing is here to stay

**Cloud applications will account for 90% of total mobile data traffic by 2018, compared to 82% at the end of 2013. Mobile cloud traffic will grow 12-fold from 2013 to 2018, attaining a compound annual growth rate of 64%.


**42% of C-level executives are increasing their investments in cloud computing over the next twelve months.**

Computerworld 2014 State of the Enterprise Survey

**Cloud software market will surpass $75B by 2017 attaining a five year compound annual growth rate of 22% in the forecast period.**

IDC, Midsize Enterprises Leading the Way with Cloud Adoption

**SaaS and cloud-based business application services revenue will have grown from $13.5b in 2011 to $32.8b in 2016.**

Centaur Partners SaaS Market Overview
Why Cloud Computing?

- **Example 1:** You are a start-up and want to go viral with a cool new app.
  - How many servers do you need?

- **Example 2:** You are an online florist.
  - How do you handle peaks on mother’s day or valentines day?

- **Example 3:** You offer an adult online game.
  - User demand peaks late evening between 10pm and midnight.
  - You want to internationalize (Europe, US, China).
  - Servers should be „close“ to users (to minimize latencies)
  - How to give best performance to your users with as little invest as possible?

**Common:** Focus on core business – not on IT administration!
Typical Characteristics of Cloud Computing

- **On-demand self-service**
  - Consumer provisions resources automatically, without human interaction

- **Broad network access**
  - Resources available over the network, accessed via standard mechanisms

- **Resource pooling**
  - Provider’s computing resources are pooled to serve multiple consumers using a multi-tenant model

- **Rapid elasticity**
  - Capabilities can be elastically provisioned and released to scale rapidly

- **Measured service**
  - Automatic control and optimization of resource use by leveraging an appropriate metering abstraction
Cloud Computing Pricing Paradigm

- Typically „Pay-as-you-go“
  - Costs scale up and down with usage dynamically
  - little to no upfront invest necessary (HW / SW!)
  - Capital expenses (CapEx) vs. Operating expenses (OpEx)

- Popular metering abstractions
  - Storage
  - Bandwidth
  - User accounts
  - ...
Pricing Example – Microsoft Azure VM

General purpose compute: Basic tier

An economical option for development workloads, test servers, and other applications that don’t require load balancing, auto-scaling, or mem virtual machines.

<table>
<thead>
<tr>
<th>INSTANCE</th>
<th>CORES</th>
<th>RAM</th>
<th>DISK SIZES</th>
<th>PRICE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A0</td>
<td>1</td>
<td>0.75 GB</td>
<td>20 GB</td>
<td>€0.0135/hr (~€10/mo)</td>
</tr>
<tr>
<td>A1</td>
<td>1</td>
<td>1.75 GB</td>
<td>40 GB</td>
<td>€0.0574/hr (~€43/mo)</td>
</tr>
<tr>
<td>A2</td>
<td>2</td>
<td>3.5 GB</td>
<td>60 GB</td>
<td>€0.1147/hr (~€86/mo)</td>
</tr>
<tr>
<td>A3</td>
<td>4</td>
<td>7 GB</td>
<td>120 GB</td>
<td>€0.2294/hr (~€171/mo)</td>
</tr>
<tr>
<td>A4</td>
<td>8</td>
<td>14 GB</td>
<td>240 GB</td>
<td>€0.4588/hr (~€342/mo)</td>
</tr>
</tbody>
</table>

Prices include Windows Server licensing fee but not sales tax. Monthly Pay-As-You-Go estimates based on 744 hours of continuous use.
Statistical Multiplexing

- Suppose resource has capacity $C$, shared by $N$ identical tasks. Each task requires capacity $c$.
  - If $C \geq Nc$, the resource is underloaded
  - But: If at most 10% of tasks are active at any given time, $C \geq Nc/10$ is enough

Cloud Computing provider can deliver the demands of multiple customers with less resources by means of statistical multiplexing („sharing“), because resource demands are distributed over time.

- Example: How many external phone lines does a hotel with 40 rooms need?
  - We need good statistics
  - Statistics may change over time (for this example: mobile phone usage...)

- Further examples: Car Sharing, WLAN, ...
Cloud Economics

- Real world estimates of server utilization in datacenters range from 5% to 20%
- Peak workload exceeds the average by factors of 2 to 10
  → Provisioning for peaks “wastes” capacity

Illustration taken from [Egli]
Some Key Technologies Enabling Cloud Computing

- **Virtualization**
  - Running virtual machines (VM) decoupled from physical hardware helps to achieve higher availability & utilization

- **Pervasive network access**
  - Increasing bandwidths and availability of mobile internet connections allows moving applications to the cloud

- **Web Services / Service-Oriented Architectures**
  - Many cloud-based applications offer Web-based APIs and allow other applications to orchestrate new services

- **Web Technologies**
  - Progress of technologies like HTML, CSS and Javascript allows feature-rich, interactive user interfaces on different devices
Approaches for enabling multi-tenancy

1. Single Application Instance
2. Multiple Application Instances, Shared Address Space
3. Multiple Application Instances, Separate Address Spaces
4. Multiple Virtual Images
5. Multiple Instances on Separate Hardware

Scale to larger number of tenants and lower operational costs
Lower application development effort and time to market

Cloud Computing vs. Outsourcing

- **(Traditional) Outsourcing / Application Service Provider**
  - Infrastructure / service dedicated to one single customer („single-tenant“)
  - Typically long-term, static contracts
  - Changes to infrastructure typically require manual interaction by service provider

- **(Public) Cloud Computing**
  - Shared infrastructure / service with large pool of users to achieve economies of scale („multi-tenant“)
  - Dynamically scalable – adapts quickly to changing demands.
  - „Pay-per-use“ & „subscription fee“ pricing models
  - Management typically „self-service“, little interaction between service provider and customer
Common Categorizations of Cloud Computing

...by service model [NIST]

- **Software-as-a-Service (SaaS)**
  - Ready-to-use application software

- **Platform-as-a-Service (PaaS)**
  - Computing platform (operating system, application server, database, web server, ...) to build & run applications

- **Infrastructure-as-a-Service (IaaS)**
  - Physical or (more often) virtual machines, storage, or other kinds of computing resources
Cloud Computing Service Models

**Traditional IT**
- Applications
- Runtimes
- Security & Integration
- Databases
- Servers
- Virtualization
- Server HW
- Storage
- Networking

**IaaS**
- Applications
- Runtimes
- Security & Integration
- Databases
- Servers
- Virtualization
- Server HW
- Storage
- Networking

**PaaS**
- Applications
- Runtimes
- Security & Integration
- Databases
- Servers
- Virtualization
- Server HW
- Storage
- Networking

**SaaS**
- Applications
- Runtimes
- Security & Integration
- Databases
- Servers
- Virtualization
- Server HW
- Storage
- Networking

DATA ONE
Cloud Computing – Examples Revisited

Software-as-a-Service
- Doodle
- Facebook
- Gmail
- Dropbox
- SAP Business ByDesign
- Salesforce
- Spotify
- Microsoft Office 365

Platform-as-a-Service
- Apprenda
- Pivotal CF
- Google Cloud Platform
- Windows Azure
- Amazon Web Services

Infrastructure-as-a-Service
- Google Cloud Platform
- Windows Azure
- Rack Space
- Akamai
Common Categorizations of Cloud Computing

...by delivery model [NIST]

- **Public Cloud**
  - service rendered over a network that is open for public use

- **Private Cloud / Community Cloud**
  - service operated solely for a single organization / group of organizations

- **Hybrid Cloud**
  - composition of two or more clouds (private, community or public)
  - Sometimes also used for services combining on-premise and cloud services

Potential Threats

- Security & privacy
- Dependency on cloud provider
  - „Vendor-Lock-In“
- Limited control
- Limited flexibility
- Offline-usage
Security and Privacy

Before using a cloud service, you should ask yourself

- Is it allowed?
  - Legal perspective

- Is it safe?
  - Technical and organizational perspective

- Is it trustworthy?
Definitions and Legal Foundations (Germany)

- **Compliance**
  - General term for „observe all valid rules“

- **Commissioned data processing („Auftragsdatenverarbeitung“) (§11 BDSG)**
  - Purchaser is responsible for ensuring compliance when person-related data is stored „on-behalf“
  - Providers have to be „carefully selected“
  - Requires written contract with certain minimum requirements
  - Purchaser remains owner of data; provider is not allowed to use data for own purposes

- **Technical and Organizational Measures (§9 BDSG)**
  - Provider must fulfill eight specific requirements (physical access control, ...)
  - Purchaser is responsible to verify (→ e.g., ISO 27001)
Definitions and Legal Foundations (Germany)

- Data transfer only to countries with „appropriate data security“
  - List includes all European Union countries, but not US

- „Safe Harbor“
  - US companies can self-register and guarantee certain principles
  - Safe-harbor certified providers are generally accepted by German authorities

- Conclusions
  - German and EU legislation put requirements on the way into the Cloud
  - Compliance concerns prevent companies from moving data into the Cloud
    - Based on „psychological insecureness“ rather than hard facts
Thank you very much for your attention

- ... and good luck for the exams!

Selected open positions at Data One (Saarbruecken)

- Working Student „Software Development SAP Cloud Solutions“
- Working Student „Consulting & Development Business Intelligence“
- Working Student „Development Microsoft SharePoint / Office 365“

More positions and further information at http://www.dataone.de/karriere
Further Reading and References

Above the Clouds: A Berkeley View of Cloud Computing

Overview of Cloud Computing – Principles and Technologies

The NIST definition of Cloud Computing


2014 Future of Cloud Computing - 4th Annual Survey Results