Cloud Computing

Chapter 1: Overview and Introduction



Summer Term 2017

Complex and Distributed IT Systems
TU Berlin



Who are we? (1/2)



- Research group "Complex and Distributed IT Systems"
 - Prof. Dr. Odej Kao
 - Homepage: http://www.cit.tu-berlin.de
 - Secretary: Jana Bechstein, E-N 157
 - ◆ E-mail: jana.bechstein@tu-berlin.de



Who are we? (2/2)

- Prof. Dr. Odej Kao
 - Responsible for lecture
 - odej.kao@tu-berlin.de



- Research assistant at CIT
- Responsible for tutorials
- anton.gulenko@tu-berlin.de







Research at CIT

Three main research categories

Connected, Embedded Devices

- How to ensure seamless integration of embedded devices in everyday's life?
- How to meet security/privacy demands of modern services?

Scalable, Adaptive, Data-Intensive Cloud Middleware

- How to write a program for a computer with thousands of CPUs?
- How to adapt to changing resource requirements?

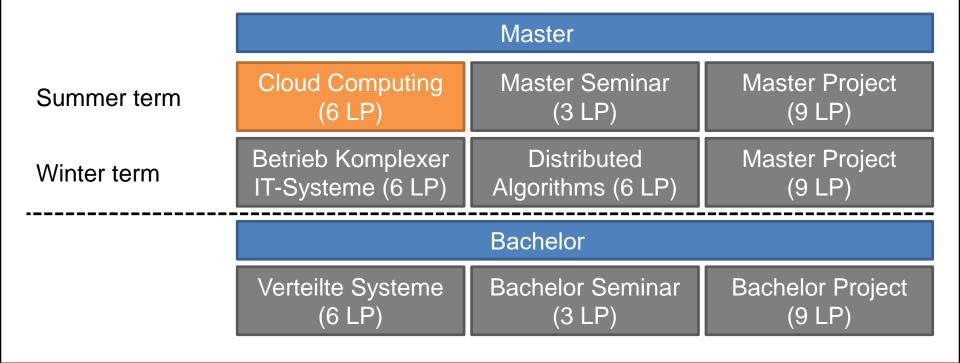
QoS-Aware Operation of Virtualized Infrastructures

- How to operate a complex IT infrastructure?
- How to guarantee the quality of service to the customer?



Teaching at CIT

- Research topics are reflected in the CIT curriculum
 - Bachelor: Basic education in Distributed Systems
 - Master: Advanced topics with stronger research focus





The Cloud Computing Module

- Credits (according to ECTS): 6
- No upfront registration necessary
- By registering for the exam through QISPOS, you implicitly sign up for the module
 - Deadlines will be announced



Structure of the Cloud Computing Course

	Mondays	Fridays
Time	14-16 h	14-16 h
Room	EMH 225	HFT-TA 251

Lecture: 21.04.-19.05.

■ Dates: 21.04., 24.+28.04., 05.05., 08.+12.05., 15.+19.05.

Tutorial: starting 22.05.

■ Dates (subject to change): 22.05., 09.06., 19.06., 03.07.



Programming Projects Details

- Goal: Hands-on experience with major Cloud platforms!
- Kindly supported by Amazon



 AWS Educate grants students AWS credits each semester



- Feel free to play around and solve project,
 but remember to shut down machines to avoid charges
- Each of the 4 projects is tackled by group of 3-4 students
 - Two weeks for completing the project (mandatory)
 - Each project counts for 10% of the final grade



Lecture Resources

- Lecture resources available on ISIS
 - ISIS also used for announcements!
 - Please sign on as soon as possible!
- There will be no script
 - Slides will be online before the lecture
 - References to additional material for further reading







Overview

- What is Cloud Computing?
- Definition of Cloud Computing
- Structure of the lecture

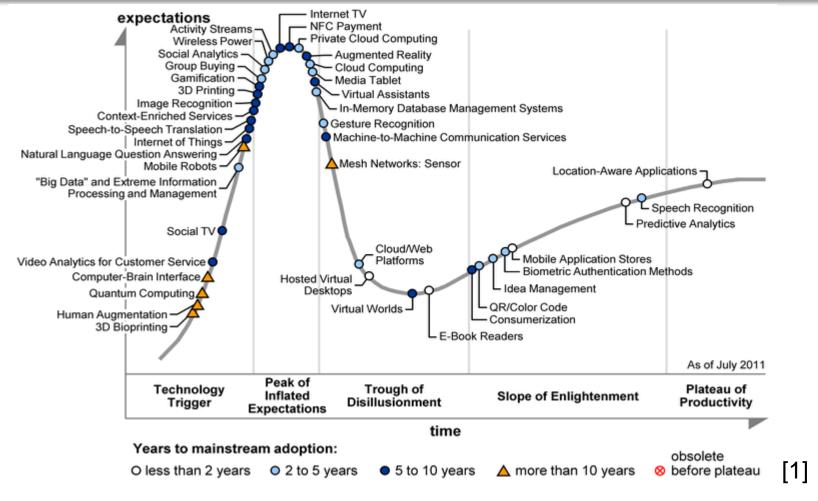


What is Cloud Computing?





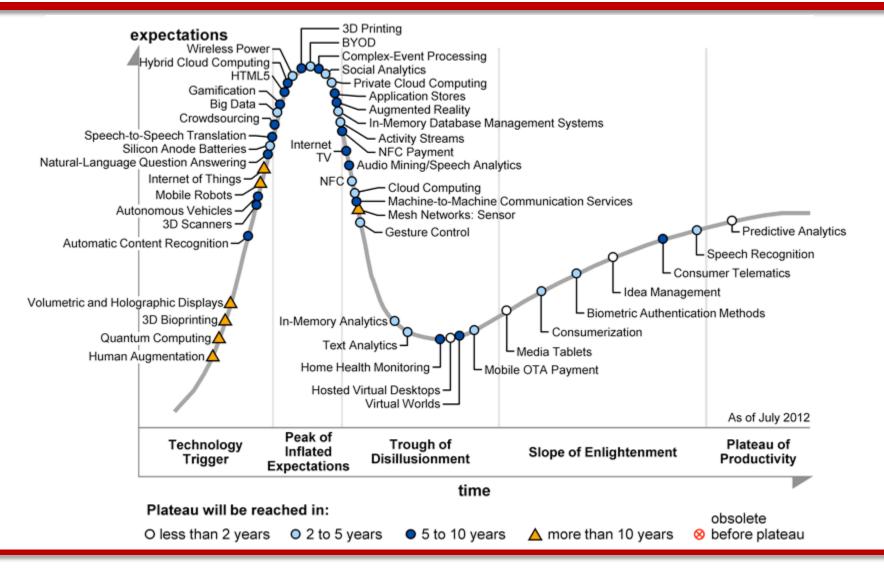
Gartner Hype Cycle for Cloud Computing (2011)



Great expectations for Cloud Computing

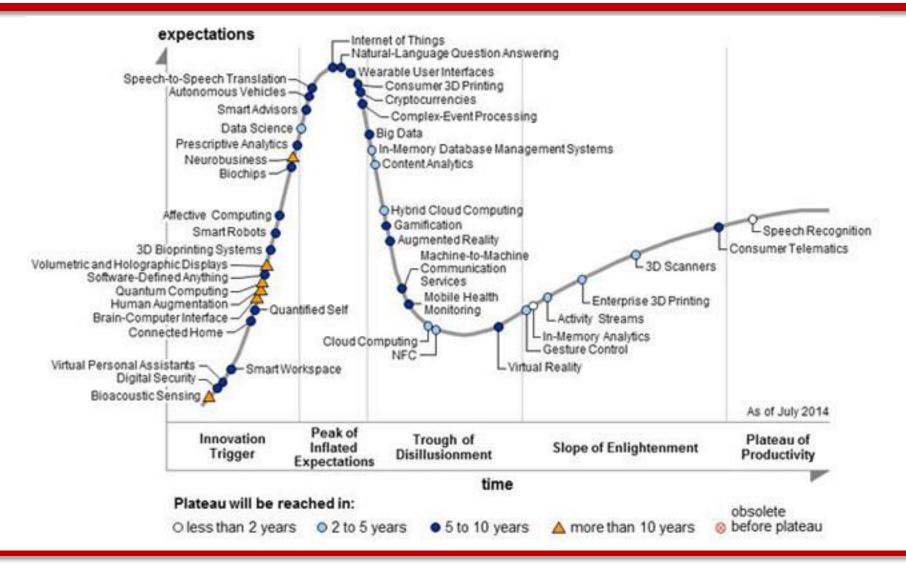


Gartner Hype Cycle for Cloud Computing (2012)



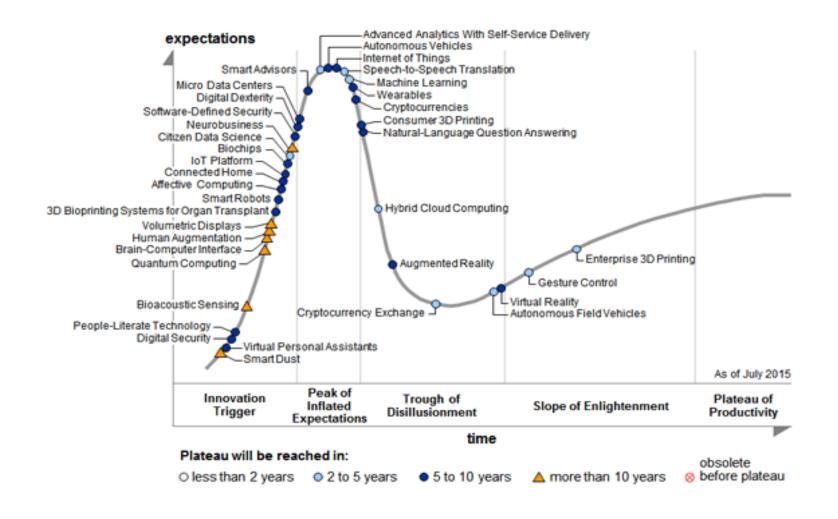


Gartner Hype Cycle for Cloud Computing (2014)





Gartner Hype Cycle for Cloud Computing (2015)





Definitions of Cloud Computing

- There are dozens of Cloud definitions out there...
 - Here are two (random) examples, both from 2008

"A large-scale distributed computing paradigm that is driven by economies of scale, in which a pool of abstracted, virtualized, dynamically-scalable, managed computing power, storage, platforms, and services are delivered on demand to external customers over the Internet."[2]

"Clouds are a large pool of easily usable and accessible virtualized resources (such as hardware, development platforms and/or services). These resources can be dynamically reconfigured to adjust to a variable load (scale), allowing also for an optimum resource utilization. This pool of resources is typically exploited by a pay-per-use model in which guarantees are offered by the Infrastructure Provider by means of customized SLAs."[3]



NIST Definition of Cloud Computing

- NIST: National Institute of Standards and Technology
 - Agency of U.S. Department of Commerce
 - Responsible for standardization processes
- Definition of Cloud Computing according to NIST[4]:

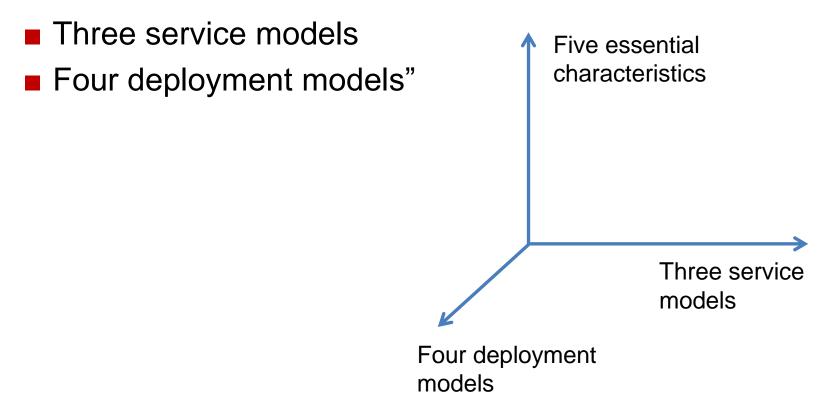
"Cloud computing is a model for enabling ubiquitous, convenient, ondemand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.

This cloud model is composed of five essential characteristics, three service models, and four deployment models."



Dimensions of Cloud Computing (NIST)

- NIST: "Cloud model is composed of
 - Five essential characteristics





Five Characteristics of Cloud Computing (NIST)

- 1. On-demand self-service
 - No human interaction required for resource provisioning
- 2. Broad network access
 - Accessible over network with standard mechanisms
- 3. Resource pooling
 - Pooled resources dynamically shared among several consumers, location independence
- 4. Rapid elasticity
 - Capabilities can be provisioned/released on demand
- Measured service
 - Resource usage is monitored, controlled, and reported



Three Service Models of Cloud Computing (NIST)

- Software as a Service (SaaS)
 - Provider's application runs on cloud infrastructure
 - Consumer can access application over the network
 - Consumer does not control/manage underlying infrastructure
- Platform as a Service (PaaS)
 - Consumer can deploy custom application onto cloud infrastructure using programming languages, libraries, services and tools supported by provider
 - Consumer does not control/manage underlying infrastructure
- Infrastructure as a Service (laaS)
 - Provider provisions processing, storage, network resources to consumer
 - Consumer does not control/manage underlying infrastructure but has control over operating systems, storage, and deployed applications



Four Deployment Models of Cloud Computing (NIST)

	Private Cloud	Community Cloud	Public Cloud	Hybrid Cloud
User of the cloud infrastructure?	Single organization	Organizations with shared concerns	Open for the general public	Composition of private/community/public cloud: • Remain distinct entities • Bound together by standard mechanisms • Goal: Enable data /application portability
Owner of the cloud infrastructure?	Organization, third party, combination thereof	Organizations, third party, combination thereof	Business, academic, government organization, combination thereof	
Location of the cloud infrastructure?	On premise, off premise	On premise, off premise	On premise of cloud provider	



What the Lecture will Discuss

- Goal: Understand what Cloud Computing entails on a technical level
 - Understanding the different levels of abstractions
 - Understanding the implications of resource sharing
 - Learning to take advantage of Cloud platforms
- Lecture will be split in two major parts
 - Discussion of IaaS and PaaS
 - Understanding building blocks, performance characteristics
 - Writing scalable and fault-tolerant applications
 - "How to run an application of hundreds of CPUs?"



What the Lecture will not Discuss

- Economical implications
 - "CAPEX vs. OPEX"
 - Consequences for accounting, tax deductions, etc.
- Legal issues
 - Disclosure of confident data (c.f. "US Patriot Act")
 - Liability in terms of outages, data loss, etc.
- Cloud data center certification
 - Methods for risk assessment
 - Cloud certification bodies, audit procedures (SAS 70, ...)



References

- [1] J. Fenn, H. LeLong: "Hype Cycle for Emerging Technologies", Gartner, 2011
- [2] I. Foster, Y. Zhou, I. Raicu, S. Lu: "Cloud Computing and Grid Computing 360-Degree Compared", Proceedings of the Grid Computing Environment Workshop, 2008
- [3] L.M. Vaquero, L: Rodero-Merino, J. Caceres, M. Lindner: "A Break in the Clouds: Towards a Cloud Definition", ACM SIGCOMM Computer Communication Review, 39(1), 2008
- [4] P. Mell, T. Grance: "The NIST Definition of Cloud Computing", Technical Report, National Institute of Standards and Technology, 2011, http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf