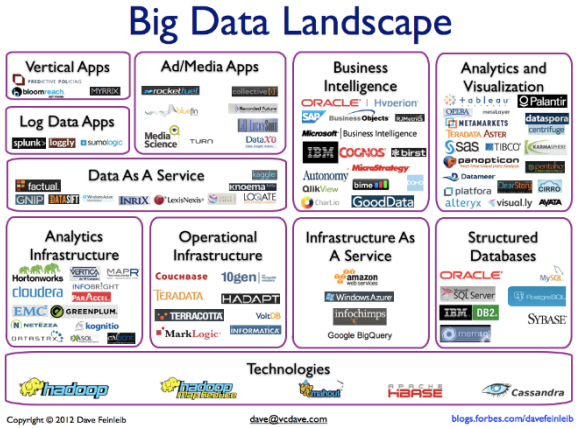


New trends of big data management

Seminar on big data management

Lecturer: Jiaheng Lu

Spring 2017

Outline

- New trends of big data management
- 1. Move to cloud for more companies
- 2. Multi-model databases

Trend 1: Move to the cloud

Why we use cloud computing?

Why we use cloud computing?

Case 1:

Write a file

Save

Computer down, file is lost

Files are always stored in cloud, never lost

Why we use cloud computing?

Case 2:

Use IE --- download, install, use

Use QQ --- download, install, use

Use C++ --- download, install, use

……

Get the serve from the cloud

What is cloud and cloud computing?

Cloud

Demand resources or services over Internet

scale and reliability of a data center.

What is cloud and cloud computing?

Cloud computing is a style of computing in which **dynamically scalable** and often **virtualized** resources are provided as a serve over the Internet.

Users need not have knowledge of, expertise in, or control over the technology infrastructure in the "cloud" that supports them.

Characteristics of cloud computing

- **Virtual.**

  software, databases, Web servers, operating systems, storage and networking as virtual servers.

- **On demand.**

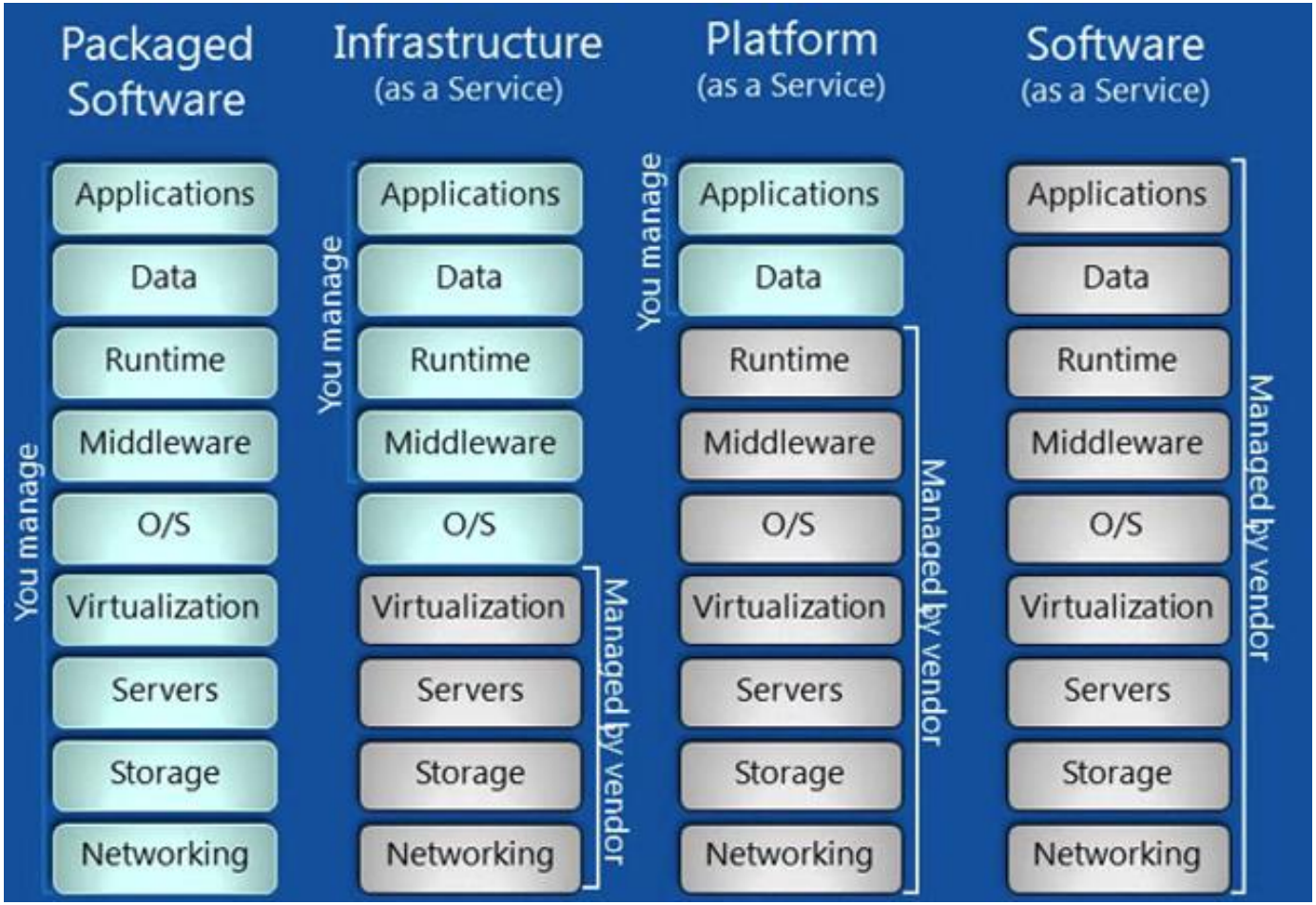
  add and subtract processors, memory, network bandwidth, storage.

Types of cloud service

SaaS
Software as a Service

PaaS
Platform as a Service

IaaS
Infrastructure as a Service


Packaged Software
Infrastructure (as a Service)
Platform (as a Service)
Software (as a Service)
Applications
Data
Runtime
Middleware
O/S
Virtualization
Servers
Storage
Networking
You manage
Managed by vendor

Software delivery model

SaaS

- No hardware or software to manage
- Service delivered through a browser
- Customers use the service on demand
- Instant Scalability

Examples

- Your current CRM package is not managing the load or you simply don't want to host it in-house. Use a SaaS provider such as Salesforce.com

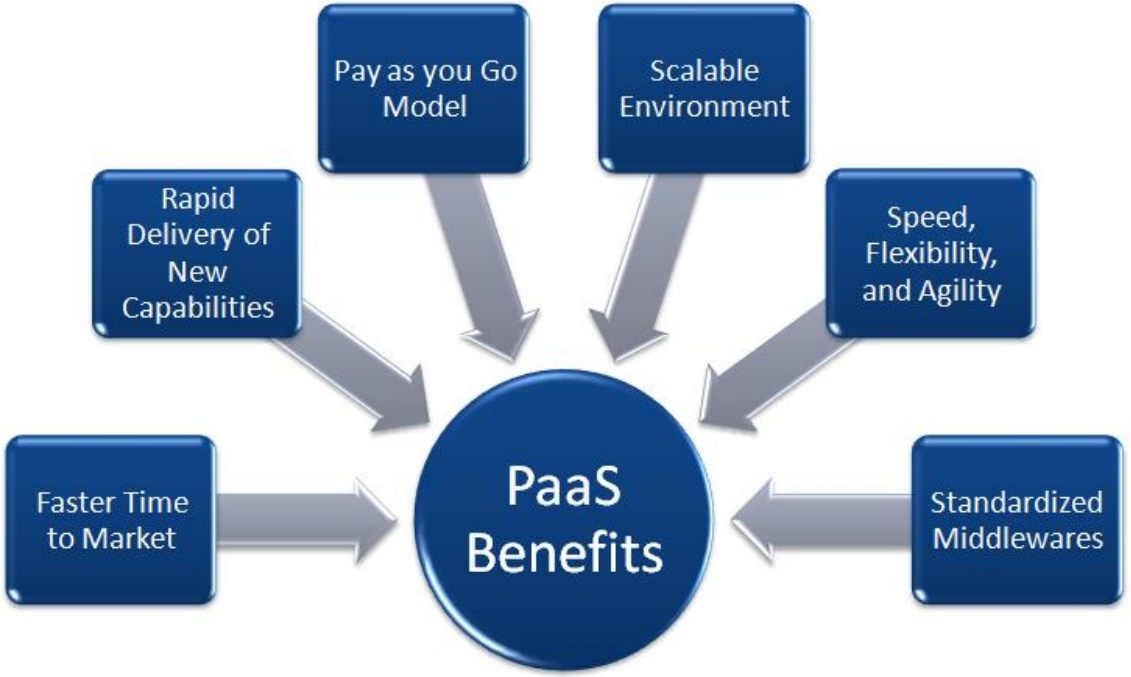
Platform delivery model

PaaS

- Platforms are built upon Infrastructure, which is expensive
- Estimating demand is not a science!
- Platform management is not fun!

Examples

- You need to host a large file (5Mb) on your website and make it available for 35,000 users for only two months duration. Use Cloud Front from Amazon.



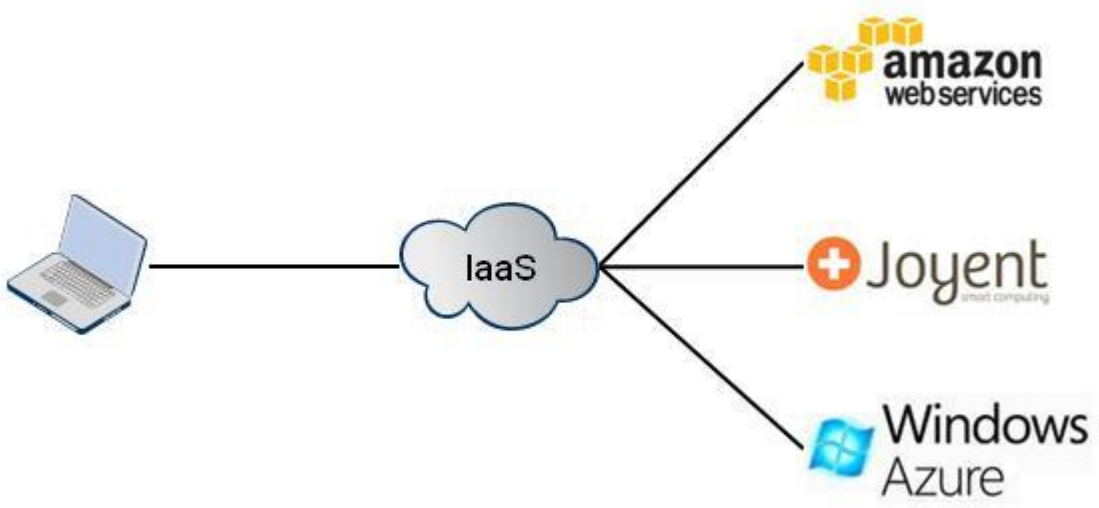
Computer infrastructure delivery model

IaaS

- A platform virtualization environment
- Computing resources, such as storing and processing capacity.
- Virtualization taken a step further

Examples

- You want to run a batch job but you don't have the infrastructure necessary to run it in a timely manner. Use Amazon EC2.



Why many companies move to the cloud

- Reduced spend in their data centers
- Greater flexibility in terms of plugging into and out of solutions.
- No big data analysis experts in the company

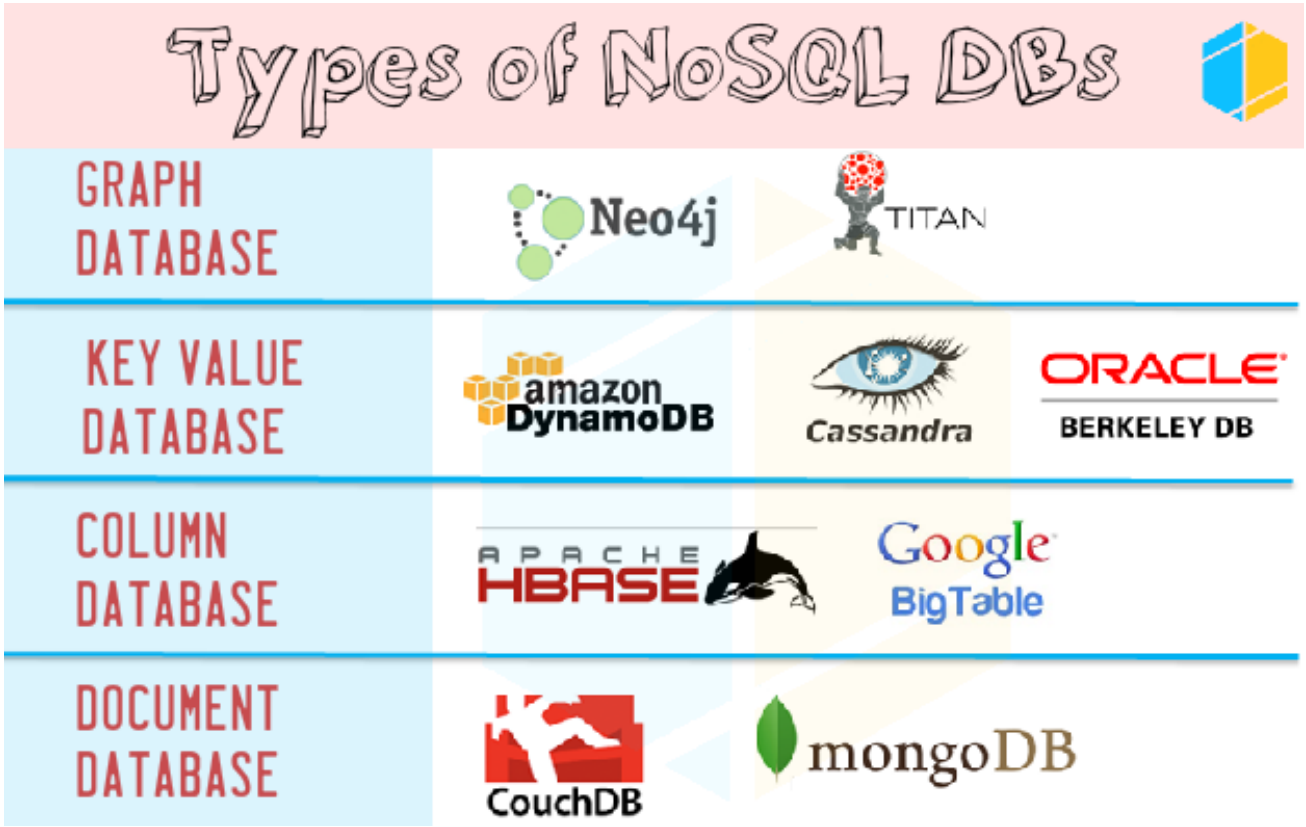
Watch video and discussion

- Watch two videos on cloud computing
- https://www.youtube.com/watch?v=uYGQcmZUTaw
- https://www.youtube.com/watch?v=DGDtujmOBKc

- Discuss with your neighbors about the learning objectives.

Trend 2: most operational DBMSs will offer multiple data models, relational and NoSQL, in a single DBMS platform.

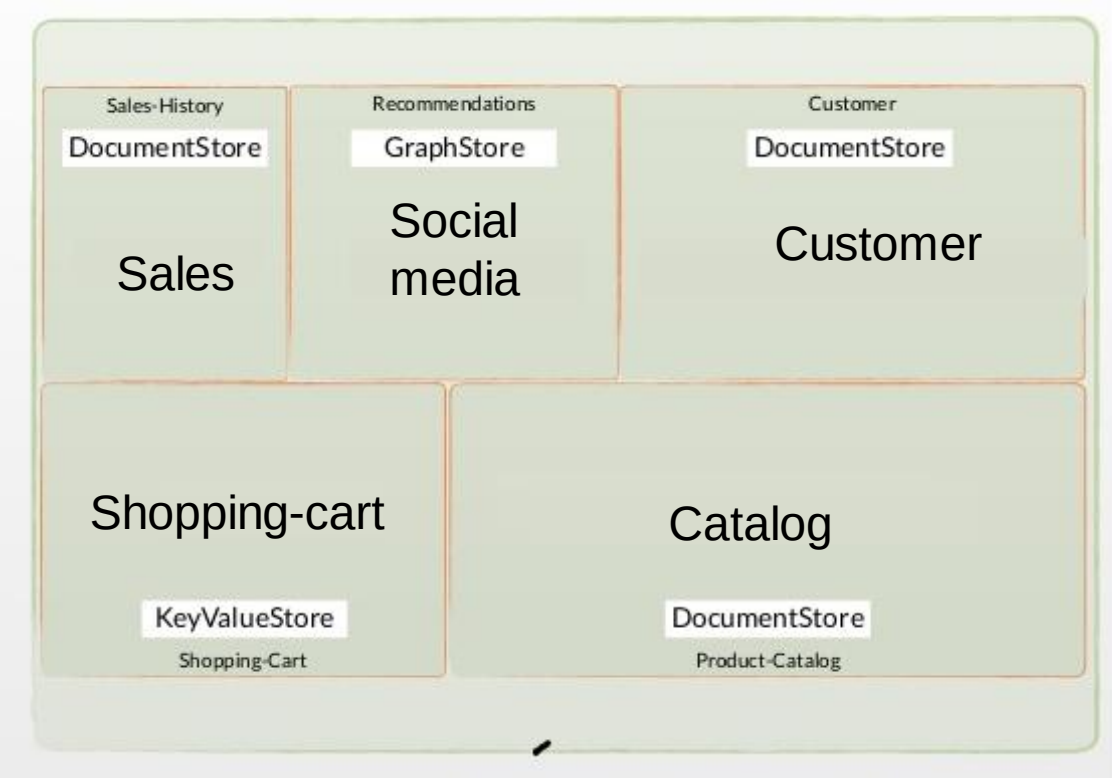
NoSQL database types



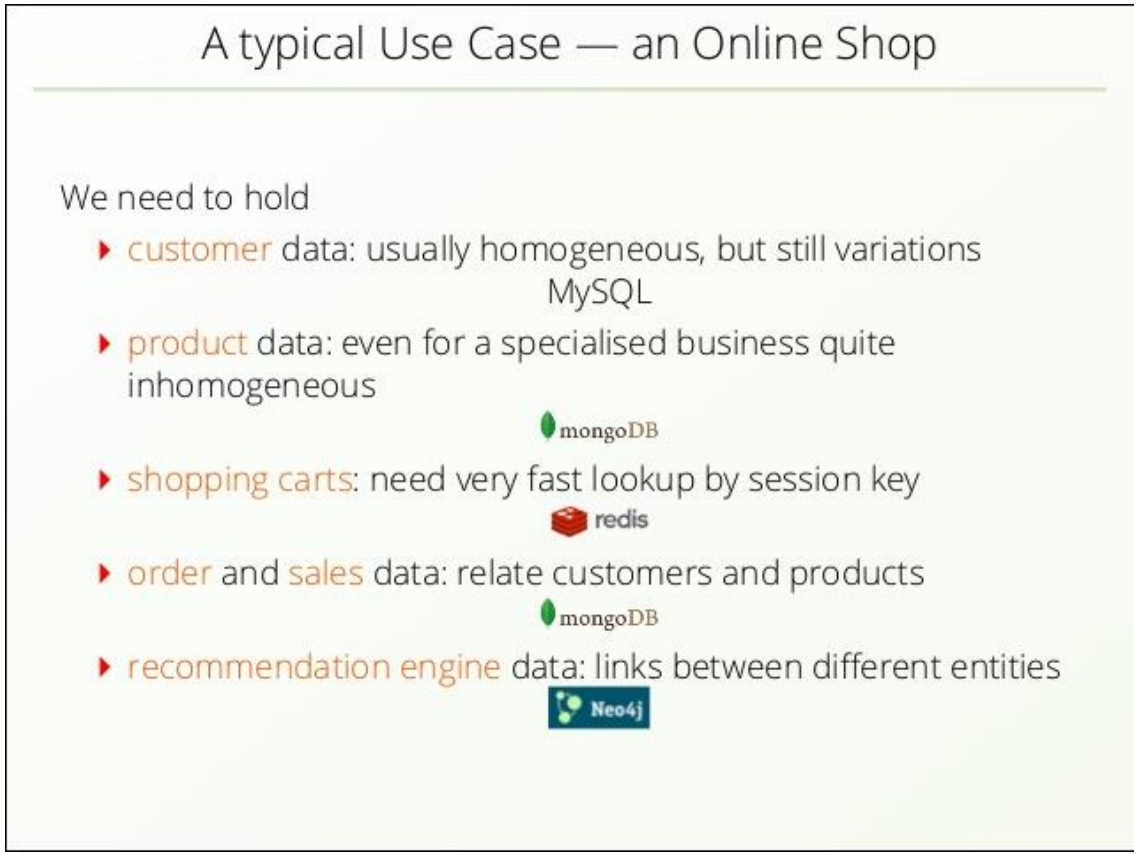
Polyglot modeling- Multiple tools

- “Use the right tool for the job”
- If you have structured data with some differences
  - Use a document store
- If you have relations between entities and want to efficiently query them
  - Use a graph database
- If you manage the data structure yourself and do not need complex queries
  - Use a key-value store

But one application for multi-model databases

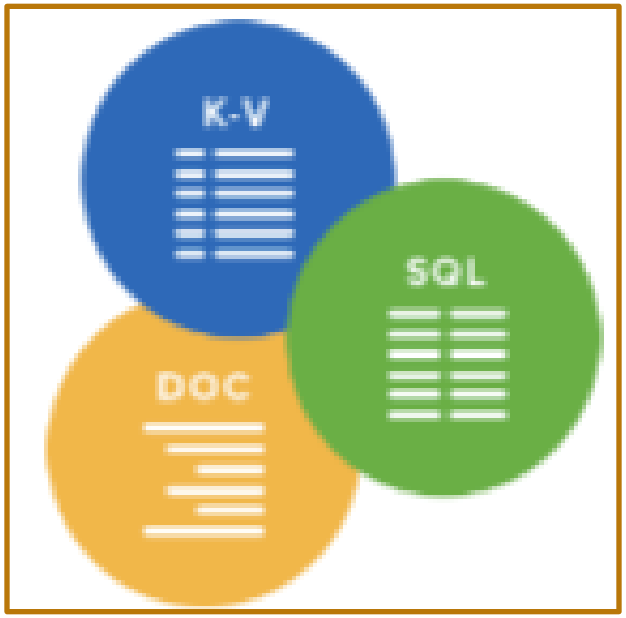


An online-shop use case



Multi-model database

- One unified database for multi-model data



One single database system

Multi-model databases

- A multi-model database is designed to support multiple data models against a single, integrated backend.

- Document, graph, relational, and key-value models are examples of data models that may be supported by a multi-model database.

- Watch a video on a multi-model database: ArangoDB

- Presentation slides:
- http://www.slideshare.net/MichaelHackstein/multi-modeldatabases

- Discuss the questions with your neighbors

Summary

- Two trends for big data management:

1. Move to the cloud
2. One Multi-model DBMS will host NoSQL and SQL

No meeting next week and the first presentation will be given on 20.02.2017.