

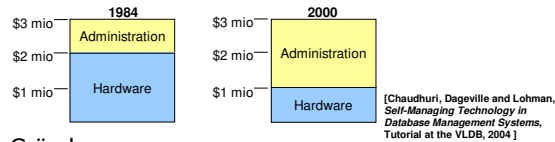
Seminar Database-as-a-Service

Einführung und Themenvergabe



Kosten für Datenverwaltung

- Problem: Hohe DBS-Kosten



- Gründe
 - Funktionsvielfalt und Komplexität
 - Teures Expertenwissen
- Lösungsansätze
 - Autonome DBS
 - Database-as-a-Service

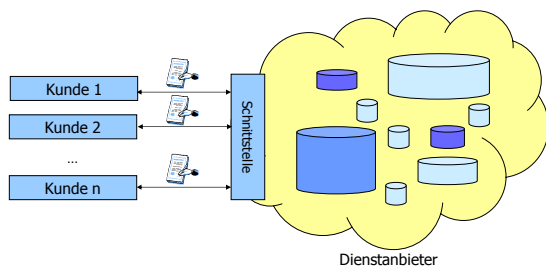


22.10.2009



2

Database-as-a-Service



22.10.2009

3



Herausforderungen

- Mandantenfähigkeit
- Verschlüsselung
- Überwachung der Dienstqualität
- Selbst-Verwaltung
- Verwaltung der Infrastruktur
- Vorhersage der Leistungskennzahlen
- Elastizität
- Betrieb eines „Service Ecosystem“
- Caching von Datenobjekten
- Autorisierung
- Portierung
- Datenanalyse
- Vertrauen



22.10.2009



4

Organisation: Prüfungsleistungen

- Alle Teilnehmer
 - Aktive Teilnahme
 - Referat
- Bachelor-Studiengang Informatik/
Diplom-Studiengang Wirtschaftsinformatik
 - Erstellung einer Seminararbeit
 - Bewertung: Referat und Seminararbeit (zu gleichen Teilen)
- Diplom-Studiengang Informatik
 - Nur für Leistungsbestätigung zusätzliche Erstellung einer Seminararbeit



22.10.2009

5



Organisation: Vortrag

- Vortragsdauer: 40 Minuten
- Anschließende Diskussion
- Bewertung
 - Inhaltlich
 - Inhaltliche Richtigkeit
 - Tiefgang
 - Nutzung von Beispielen
 - Kompetenz in der Fragenbeantwortung
 - Präsentation
 - Vortragsstil
 - Foliengestaltung
 - Verständlichkeit



22.10.2009



6

Organisation: Seminararbeit

- Umfang: 8-10 Seiten (netto)
- Bewertung
 - Inhaltlich
 - Inhaltliche Richtigkeit
 - Tiefgang
 - Präsentation
 - Aufbau der Arbeit
 - Illustrationen
 - Nutzung von Beispielen
 - Arbeitstechnik
 - Selbständige wissenschaftliche Arbeitsweise
 - Quellensuche
 - Korrekte Zitierung

Organisation: Termine

20.10.2009	Einführung und Themenvergabe
nach Absprache	Besprechung der Seminararbeit
05.01.2010	Abgabe erste Version Seminararbeit
19.01.2010	Besprechung erste Version Seminararbeit
26.01.2010	Abgabe finale Version Seminararbeit
02.03.2010	Abgabe Folien
09.03.2010	Besprechung Folien
15.03.2010	Blockseminar Tag 1
16.03.2010	Blockseminar Tag 2

- Abgaben an: dbaas09@informatik.uni-hamburg.de

Organisation: Weitere Informationen



... gibt es auf unserer Webseite:

<http://vis-www.informatik.uni-hamburg.de/teaching/ws-09.10/dbaas/>

Richtlinien: Schein/Prüfungskriterien

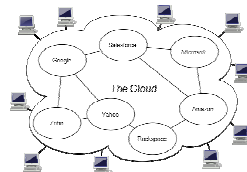
- Prüfung gilt als nicht bestanden / Leistungsschein wird nicht erteilt wenn:
 - Präsentation am vereinbarten Termin nicht erfolgt
 - Seminararbeit nicht angefertigt wird
 - Zwei der folgenden Termine nicht eingehalten werden
 - Erste Abgabe Seminararbeit
 - Endgültige Abgabe Seminararbeit
 - Abgabe Folien
 - Einer der Termine um mehr als eine Woche überschritten wird
- Teilnahmechein wird nicht erteilt wenn:
 - Präsentation am vereinbarten Termin nicht erfolgt
 - Termin "Abgabe Folien" um mehr als eine Woche überschritten wird

Richtlinien: Folien und Seminararbeit

- Vortragsfolien
 - Verbindliche Verwendung der Folien-Vorlage
 - PowerPoint oder OpenOffice Impress
 - Verbindliche Folienanzahl: 15-25
 - Anteil Text-Folien: Max. 50%
 - Ausschließliche Verwendung von Vektorgraphiken
- Seminararbeit
 - Verbindliche Verwendung der Seminararbeits-Vorlage
 - MS Word oder OpenOffice Writer oder LaTeX
 - Verbindlicher Umfang: 8-10 Seiten

Grundlagen: Cloud Computing

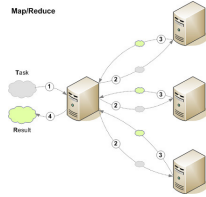
- Cloud Computing
 - (Versuch einer) Definition
 - Charakteristika
 - Beispiele
- Cloud Architektur
 - Eucalyptus



- Allgemein
 - Michael Armbrust, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy Katz, Andy Konwinski, Gunho Lee, David Patterson, Ariel Rabkin, Ion Stoica, and Matei Zaharia: Above the Clouds: A Berkeley View of Cloud Computing <http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.html>
 - Neal Leavitt: Is Cloud Computing Really Ready for Prime Time? <http://blogs.bva.com/analyst/leavitt/2009/08/10/is-cloud-computing-really-ready-for-prime-time/>
- Infrastrukturen / Eucalyptus
 - The Eucalyptus Open-source Cloud-computing System. http://eucalyptus.cs.ucsb.edu/documents/num1_et_al-eucalyptus_open_source_cloud_computing_system-cca_2008.pdf
 - Eucalyptus: A Technical Report on an Elastic Utility Computing Architecture Linking Your Programs to Useful Systems. http://eucalyptus.cs.ucsb.edu/documents/num1_et_al-eucalyptus_tech_report-august_2008.pdf
 - Eucalyptus Walrus
 - http://open.eucalyptus.com/wiki/EucalyptusStorage_v1.4
 - Cloud-Infrastrukturen mit Eucalyptus selbst aufbauen. IX-4/2009. S.128-130

Grundlagen: Cloud Processing

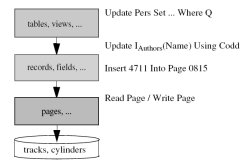
- Programmier-Modell: Map/Reduce
- Implementierungen von Map/Reduce
- Einsatzbeispiel in der Datenanalyse



- Jeffrey Dean, Sanjay Ghemawat: MapReduce: Simplified Data Processing on Large Clusters <http://labs.google.com/papers/mapreduce.html>
- Hadoop MapReduce <http://hadoop.apache.org/mapreduce/>
- Scope <http://research.microsoft.com/en-us/um/people/jrzhou/pub/Scope.pdf>

Grundlagen: DBMS Architektur

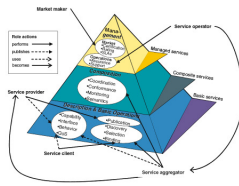
- Schichtenmodell
- Transaktionale Verarbeitung
- Spezialisierte DBMS-Architekturen
- Herausforderungen für DBaaS



- Härder, T.: DBMS Architecture - The Layer Model and its Evolution. Datenbank Spektrum 13 (2005) 45-57
- Härder, T.: DBMS Architecture - New Challenges Ahead. Datenbank Spektrum 14(5) (2005) 38-48
- Holze, M., Krebs, K., Panse, F., Ritter, N.: Data-Management-as-a-Service: Safe Ground or Terra Incognita?, Local Proceedings of the of the 13th East European Conference on Advances in Databases and Information Systems. Riga Technical University, 2009, 2-17
- Härder, T., Rahm, E.: Datenbanksysteme. Springer Verlag, 2001, ISBN 3540421335

Grundlagen: SOA

- Grundlagen
 - nachrichtenbasierte Interaktion
 - lose Kopplung
 - Autonomie
- Web-Services
- Was hat das ganze mit Cloud und SaaS zu tun?



- Papazoglou, M. P.: Service-Oriented Computing: Concepts, Characteristics and Directions <http://infobab.uvt.ro/pub/papazogloump-2003-51.pdf>
- M.P. Papazoglou and D. Georgakopoulos, Service Oriented Computing, Communications of the ACM <http://infobab.uvt.ro/pub/papazogloump-2003-52.pdf>
- Webber, J. and Parasitidis, S., Demystifying Service Oriented Architectures <http://webservicessys-on.com/read/3908.htm>
- Kuroopka, D. What does Service-oriented Computing really mean? <http://itrops.dgk.stuhl.de/opus/willertu/2006/522/pdf/05-462.KuroopkaDominik.Paper.522.pdf>
- Sanjiva Weerawarana, F. Curbera, F. Leymann: Web Services Platform Architecture <https://hhas21.rzr.uni-hamburg.de/DB=1.18/LNG=DU/CM?ACT=SRCH&IKT=54&SRT=YOP&TRM=AWE837597>

DBaaS: Skalierbarkeit

- Daten-Management in Cloud-Umgebungen
 - Probleme
 - Anforderungen
- Architekturen
 - Shared-Nothing
 - Shared-Disk
- Datenbanken auf Amazon S3



- Daniel J. Abadi: Data Management in the Cloud: Limitations and Opportunities <http://sites.computer.org/debul/A09mar/abadi.pdf>
- Ashraf Aboulnaga, Kenneth Salem, Ahmed A. Soror, Umar Farooq Minhas, Peter Kokosieli, Sunil Kamath: Deploying Database Appliances in the Cloud <http://sites.computer.org/debul/A09mar/aboulnaga.pdf>
- Matthias Brantner, Daniela Florescu, David Graf, Donald Kossmann, Tim Kraska: Building a Database on S3 <http://portal.acm.org/citation.cfm?id=1376645>

DBaaS: SLA / QoS für Services

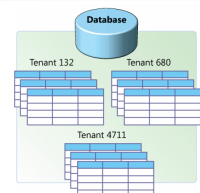
- Definition und Durchsetzung von Dienstgüte (im Bereich von Datenbanken)
- Mechanismen, um definierte Dienstgüten für einen Dienst zu erreichen



- Stefan Krompass, Daniel Gmach, Andreas Scholz, Stefan Seltzmann, Alfons Kemper: Quality of Service Enabled Database Applications
- Daniel Gmach, Stefan Krompass, Andreas Scholz, Martin Wimmer, Alfons Kemper: Adaptive Quality of Service Management for Enterprise Services
- D. McWhorter, B. Schroeder, N. Allamaki, M. Harchol-Balter: Priority Mechanisms for OLTP and Transactional Web Applications
- A. Krauss, F. Schön, G. Weikum, U. Deppisch: With HEART Towards Response Time Guarantees for Message-Based E-Services
- S. Elnikety, E. Nahum, J. Tracey, W. Zwaenepoel: A Method for Transparent Admission Control and Request Scheduling

DBaaS: Mandantenfähigkeit

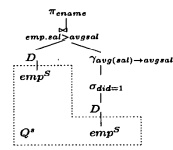
- Gemeinsame Nutzung zur Kostenreduktion
 - Server
 - DBS
 - Schema
- Problem: Datenschutz



- Aulbach, S., Grust, T., Jacobs, D., Kemper, A., Rittinger, J.: Multi-tenant databases for software as a service: schema-mapping techniques. Proc. of the International Conference on Management of Data. ACM Press, 2008 1195-1206
- Wong C., Musker, S.: DB2 Label-Based Access Control, a practical guide, Part 1: Understand the basics of LBAC in DB2, IBM Corporation, 2006
- Oracle: Oracle Label Security with Oracle Database 11g Release 2, Oracle White Paper, 2009
- Chong, F., Carraro, G., Wolter, R.: Multi-Tenant Data Architecture. <http://msdn.microsoft.com/en-us/library/aa479086.aspx> (2006)
- Oracle: SaaS Data Architecture, Oracle White Paper, 2008

DBaaS: Verschlüsselung

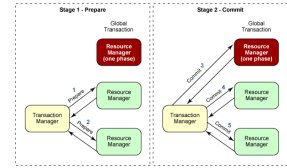
- Problem: Schutz der Daten vor Einsicht durch den Dienstanbieter
- Verschlüsselte Speicherung
- Verschlüsselte Anfrageverarbeitung



- Hacıgümüş, H., Iyer, B., Li, C., Mehrotra, S.: Executing SQL over Encrypted Data in the Database Service Provider Model. Proc. of the ACM SIGMOD International Conference on Management of Data, ACM Press (2002) 216-227
- Agrawal, R., Kiernan, J., Srikant, R., Xu, Y.: Order Preserving Encryption for Numeric Data. Proc. of the ACM SIGMOD International Conference on Management of Data, ACM Press (2004) 563-574

DBaaS: Transaktionen

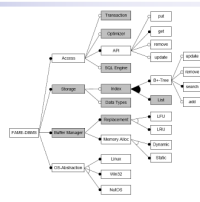
- Grundlagen von Transaktionen
- Probleme und Anforderungen in der Cloud
- Aktuelle Ansätze



- Zhou Wei, Guillaume Pierre, Chi-Hung Chi: Scalable Transactions for Web Applications in the Cloud http://www.globule.org/publi/STWAC_europar2009.html
- Sudipto Das, Divyakant Agrawal, Amr El Abbadi: ElasTraS: An Elastic Transactional Data Store in the Cloud

DBfromS: Fame-DBMS

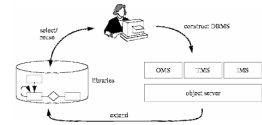
- Ziel: Bedarfsgerechte DBS
- Kostenreduktion
- DBS from Services
- Fame DBSM
 - Software-Produktlinien-Ansatz



- Subasu, I.E., Ziegler, P., Dittrich, K.R.: Towards Service-Based Database Management Systems. Workshop Proceedings der 12. Fachtagung zu Datenbanksysteme in Business, Technologie und Web, Verlagshaus Mainz, 2007, 296-306
- Ionut Emanuel Subasu, Patrick Ziegler, Klaus R. Dittrich, Harald Gall: Architectural Concerns for Flexible Data Management. Software Engineering for Tailor-made Data Management, 2008, 35-40
- M. Rosenmüller, N. Siegmund, H. Schirmeier, J. Sincero, S. Apel, T. Leich, O. Spinczyk, and G. Saake. FAME-DBMS: Tailor-made Data Management Solutions for Embedded Systems. In Workshop on Software Engineering for Tailor-made Data Management, 2008.
- M. Rosenmüller, T. Leich, and S. Apel. Konfigurierbarkeit für ressource

DBfromS: Komponenten DBMS

- Aufbau von DBS aus SW-Komponenten
- Beispiel: KIDS



- Geppert, A., Scherrer, S., and Dittrich, K. R. KIDS: Construction of Database Management Systems Based on Reuse. Technical Report. University of Zurich, 1997
- Dittrich, K., Geppert, A.: Component Database Systems, Morgan Kaufman, 2001, ISBN 1558606424

Industrielle Ansätze: Amazon

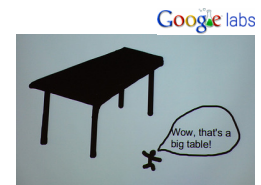
- Amazon Web Services
 - Simple Storage Service
 - Dynamo
 - SimpleDB



- Amazon Web Services LLC: Amazon SimpleDB. Developer guide, Amazon Web Services LLC (2009)
- Amazon Web Services LLC: Amazon Simple Storage Service. Developer guide, Amazon Web Services LLC (2009)
- DeCandia, G., Hastorun, D., Jampani, M., Kakulapati, G., Lakshman, A., Pilchin, A., Sivasubramanian, S., Vosshall, P., and Vogels, W. 2007. Dynamo: amazon's highly available key-value store. SIGOPS Oper. Syst. Rev. 41, 2007, 205-220.

Industrielle Ansätze: Google

- Google Bigtable
- Implementierung: Hadoop HBase



- Fay Chang, Jeffrey Dean, Sanjay Ghemawat, Wilson C. Hsieh, Deborah A. Wallach, Mike Burrows, Tushar Chandra, Andrew Fikes, Robert E. Gruber: Bigtable: A Distributed Storage System for Structured Data <http://labs.google.com/papers/bigtable.html>
- Hadoop HBase <http://hadoop.apache.org/hbase/>

Industrielle Ansätze: Yahoo! & Microsoft

- Microsoft Azure
 - SQL Data Services

- Yahoo PNUTS



- Lee, J.: Microsoft SQL Data Services - Under the Hood. Technical white paper, Microsoft Corporation, 2008
- <http://channel9.msdn.com/pdc/2008/BB03/>
- Cooper, B. F., Ramakrishnan, R., Srivastava, U., Silberstein, A., Bohannon, P., Jacobsen, H., Puz, N., Weaver, D., and Yerneni, R.: PNUTS: Yahoo!'s hosted data serving platform. Proc. VLDB Endow. 1, 2 (Aug. 2008), 1277-1288.



22.10.2009



25

Themenüberblick

01	15.03.2010	Grundlagen: Cloud Computing	Teil 1: Wetzel Teil 2: Frei
02	15.03.2010	Grundlagen: Cloud Processing	Zenlake
03	15.03.2010	Grundlagen: DBMS Architektur	Teil 1: Klutt Teil 2: Keskin
04	15.03.2010	Grundlagen: SOA	Twiefel
05	15.03.2010	DBaaS: Skalierbarkeit	Zhang
06	15.03.2010	DBaaS: SLA / QoS für Services	Linke
07	15.03.2010	DBaaS: Mandantenfähigkeit	Resemann
08	16.03.2010	DBaaS: Verschlüsselung	Detbarn
09	16.03.2010	DBaaS: Transaktionen	Meder
10	16.03.2010	DBfromS: Fame-DBMS	Kerkloh
11	16.03.2010	DBfromS: Komponenten DBMS	Eins
12	16.03.2010	Industrielle Ansätze: Amazon	Teil 1: Kuhlmann Teil 2: Bücklers
13	16.03.2010	Industrielle Ansätze: Google	Zimmermann
14	16.03.2010	Industrielle Ansätze: Yahoo! & Microsoft	Teil 1: Azure Gessert Teil 2: Pnuts Slotke



22.10.2009



26