

ERBD – VII Escola Regional de Banco de Dados

Sistemas de Gerenciamento de Dados na Nuvem

Carmem Satie Hara
Universidade Federal do Paraná

ERBD 2011

Computação na Nuvem

- Por que?
 - A WEB está substituindo o desktop
Google Gmail, Google Docs, Amazon, Flickr, Facebook, Twitter, YouTube
- Mudança de Paradigma:
 - Amazon Web Services
 - Windows Azure Platform
 - Google App Engine

ERBD 2011

Computação na Nuvem

É um modelo que proporciona acesso, através da rede, a um conjunto de recursos configuráveis (rede, servidores, armazenamento, aplicações e serviços), que são gerenciados pelo provedor do recurso, e que podem ser utilizados por clientes através de uma interface de serviço.

ERBD 2011

Computação na Nuvem

- Evolução dos conceitos de:
 - Virtualização**: encapsulamento de características físicas do recurso e visão de múltiplos recursos lógicos sobre um mesmo recurso físico
 - Arquitetura orientada a serviço (SOA)**: baixo acoplamento entre o provedor e consumidor
 - Computação autônoma**: auto-gerenciável, self-service e sob demanda
 - Computação como serviço público**

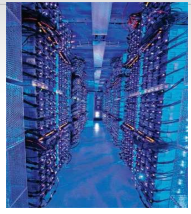
ERBD 2011

Computação na Nuvem

- Compartilhamento de recursos: CPU, armazenamento, banda de rede
- Disponibilidade, escalabilidade, elasticidade
- Gerenciamento, transferência de riscos
- Tolerância a falhas
- Baseado em computadores simples
- Pagamento pelo uso

ERBD 2011

Elasticidade



Data Center da Microsoft

- 1000 computadores usados por 1 hora custa o mesmo que 1 computador usado por 1000 horas
- Com paralelismo: resultados 1000 vezes mais rápido
- Exemplo: Animato – carga de trabalho dobrou a cada 12 horas por 3 dias.

ERBD 2011

Utilização de Recursos



Fonte: M. Ambrose et al., CACM 2010

- Provisionamento pela carga máxima pode causar subutilização.
- Provisionamento pela média pode aumentar o tempo de espera e perda de clientes.
- Dificuldade de prever a carga e variação da carga no tempo.

ERBD 2011

Modelos de Operação

- Nuvem Pública**
computação como serviço público
- Nuvem Comunitária**
compartilhamento de recursos entre membros de uma comunidade com interesses comuns
- Nuvem Privada**
virtualização de serviços em servidores locais
- Nuvem Híbrida**
combinação de abordagens

ERBD 2011

Paradigmas

Cliente	
Aplicação	SaaS (Software as a Service) : oferece software como serviço, eliminando a necessidade de instalação e execução na máquina do cliente.
Plataforma	PaaS (Platform as a Service) : oferece um conjunto de soluções como serviço para dar suporte a aplicações na nuvem. Ex: Google App Engine, MS Azure
Infraestrutura	IaaS (Infrastructure as a Service) : oferece uma plataforma computacional, em geral um ambiente virtualizado, como serviço. Ex: Amazon EC2
Servidor	

ERBD 2011

Banco de Dados como SaaS

- Um serviço de armazenamento e busca de dados na Internet com:
 - Ilusão de recursos infinitos: *escalabilidade*
 - Custo mínimo de instalação
 - Pagamento pela utilização do serviço (volume de dados e de acessos)
 - Disponibilidade

ERBD 2011



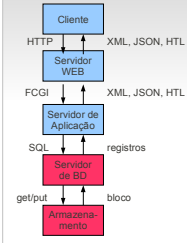
Carga de Trabalho

- OLAP (*Online Analytical Processing*)
 - Análise de um grande volume de dados
 - Poucas atualizações
 - Processamento pesado
- OLTP (*Online Transaction Processing*)
 - Transações curtas que envolvem poucos dados
 - Grande volume de atualizações

ERBD 2011



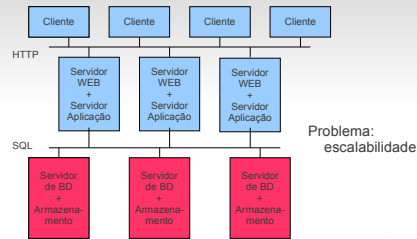
Aplicação WEB Tradicional



ERBD 2011



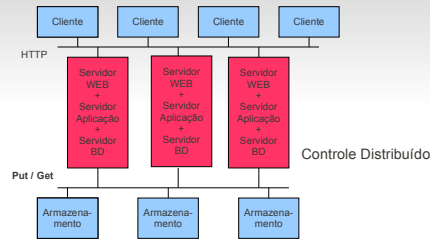
Arquitetura de Base Replicada ou Particionada



ERBD 2011



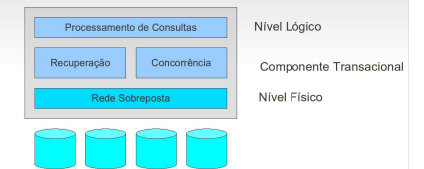
Arquitetura de Disco Compartilhado



ERBD 2011



OLTP - Arquitetura do Sistema

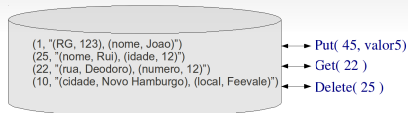


ERBD 2011



Nível Lógico Modelo Chave-Valor

A chave identifica um par e o valor associado é um BLOB.



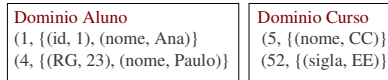
Exemplo: Voldemort, Scalaris

ERBD 2011



Nível Lógico Modelo de Documento

Similar ao Chave-Valor, mas o valor tem estrutura, como por exemplo um conjunto de (atributo, valor de atributo). São organizados em domínios.



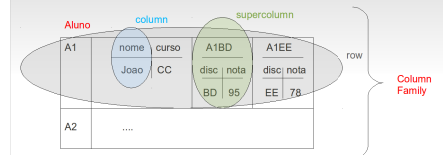
Select nome from Aluno where id = 1

Exemplo: SimpleDB, CouchDB

ERBD 2011



Nível Lógico - Modelo de Registro Extensível



Interface procedural: Get: key, range; Insert; Remove

Ex: Cassandra, HBase

ERBD 2011



Nível Lógico - Modelos

- Relacional
 - Interface: SQL
 - Ex: Amazon Relational Database Service (RDS)
- XML
 - Interface: Xquery
 - Ex: Sausalito

ERBDY 2011



Nível Físico

Envolve:

- a localização do dado
- mapeamento de operações atômicas com o armazenamento físico

métodos de acesso

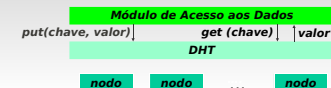
Pode ser implementado por:

- Tabela de Espalhamento Distribuída (DHT)
- Sistema de Arquivos Distribuído (DFS)
- dentre outras possibilidades

ERBDY 2011



DHT



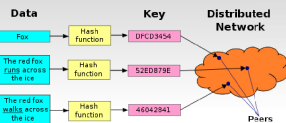
DHT é um serviço de busca de dados armazenados de forma distribuída em um conjunto de nodos.

- O espaço de chaves é uniformemente distribuído pelos nodos.
- Os nodos formam uma rede sobreposta.
- Os nodos mantêm uma lista de vizinhos em sua tabela de roteamento.
- É independente da topologia da rede.

ERBDY 2011



DHT



Propriedades:

- Descentralização
- Tolerância a falhas
- Escalabilidade

ERBDY 2011



Sistema de Arquivos Distribuído (DFS)

- Arquivo: unidade lógica de armazenamento
- Permite o acesso a arquivos armazenados em servidores (nodos) remotos
- Implementa controle de concorrência
- Pode fazer replicação para maior disponibilidade e tolerância a falhas

ERBDY 2011



Google DFS

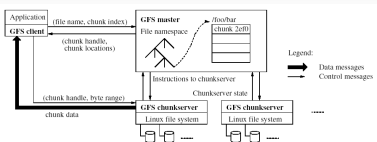
- Objetivo: armazenamento de um enorme volume de dados de forma redundante sobre servidores simples.
- Maioria das operações de leitura.
- Escrita – na maioria inserção de novos valores (sem sobrescrever valores existentes)
- Arquivos: são divididos em "pedaços" (*chunks*)
- Interface: create, delete, open, close, read, write, append, snapshot sobre arquivos

ERBDY 2011



Google DFS - Arquitetura

- Servidor mestre: coordena acessos e mantém metadados – com mestres *shadow*
- Múltiplos servidores de *chunks*



ERBDY 2011



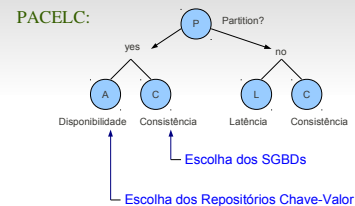
Componente Transacional

- Em SGBDs tradicionais: transações com propriedades ACID: Atomicidade, Consistência, Isolamento, Durabilidade
 - Em Sistemas Distribuídos: Teorema CAP
 - Consistência: cópias consistentes
 - Disponibilidade: tolerante a falhas em nodos
 - Tolerante a Particionamentos na rede
- só é possível garantir 2 características.

ERBDY 2011



CAP - Tradeoffs



ERBDY 2011



Propriedade BASE

- Basically Available, Soft state, Eventually consistent
 - Sempre disponível
 - Nem sempre consistente (inconsistência entre réplicas)
 - Torna-se consistente em um determinado momento
- Resolução de Conflitos
 - Durante Leitura / Escrita / Independente
 - Timestamp, Vector clocks
- Nível de Consistência
 - Quorum, Um, ~~Todos~~

ERBD 2011



Suporte a Transações

- Estender os repositórios chave-valor
 - MegaStore [CIDR '11], G-Store [SoCC '11]
- Arquitetura híbrida: componentes de SGBDs e repositórios / serviços
 - Deutoronomy [CIDR '09, '11], ElasTra5 [HotCloud '09], DB on S3 [SIGMOD '08]
- Migrar o SGBD para a Nuvem
 - RelationalCloud [CIDR '11], SQL Azure [ICDE '11]

ERBD 2011



Movimento NoSQL

- NoSQL: "Not Only SQL" ou "Not Relational" ?
- Armazenamento distribuído e escalável
- Replicação de dados → tolerância a falhas
- Sem esquema ou com esquema extensível
- Interface simples, baseada em chamadas de operações simples
- Consistência fraca

ERBD 2011



Sistemas - Amazon WS

- Serviços pagos por volume de armazenamento e quantidade de acessos
 - Simple Storage Service (S3)
 - SimpleDB
 - Relational Database Service (RDS)
- Dynamo: repositório baseado em DHT

ERBD 2011



S3

- Modelo: chave-valor, onde valor pode ter de 1B a 10GB, armazenados em "buckets". O bucket pode ser criado em uma das 5 regiões (EUA Leste, EUA Oeste, Europa, Japão, ou Singapura)
- Interface: put, get, scan(bucket), scan(bucket, prefix)
- Segurança: privilégios no nível de bucket (leitura, gravação)

ERBD 2011



SimpleDB

- Modelo: documento
- Interface: put, get, "select target from dominio where condição"
- Consistência: eventual ou forte (definida por operação)
- Para a implementação de transações pela aplicação: put/get condicionais (atualizam dependendo do valor prévio coincidir com o que foi lido)

ERBD 2011



Relational Database Service

- Modelo: relacional (com réplicas)
- Fornece funcionalidade e compatibilidade com o SGBD MySQL, mas com armazenamento na nuvem
- Capacidade: cada base de dados com 1.7GB a 68GB

ERBD 2011



Cassandra

- Modelo: registro extensível
- Repositório baseado em uma DHT, mantém os dados armazenados em ordem lexicográfica pela *row key*
- Fator de replicação configurável
- Consistência por operação: zero, one, quorum, all, ou any
- Write: escrito primeiro na *log* do nodo que contém o dado

ERBD 2011



Sistemas - Comparativo

Característica	S3	SimpleDB	Cassandra	SQL Azure	Rel. Cloud
Modelo	Chave	Documento	Reg. Ext.	Relac.	Relac.
Armaz.	DHT	DHT	DHT	Tabela	Tabela
Linguagem	API Simples	API Simples	API Simples	SQL	SQL
Transações	Não	Não	Não	Sim local	Sim distribuída
Consistência	Eventual	Eventual	Eventual	Forte	Forte
Escalabil.	Alta	Alta	Alta	Média	Baixa
Disponibil.	Alta	Alta	Alta	Alta	Média

ERBD 2011

[Souza, SBBD'2010]



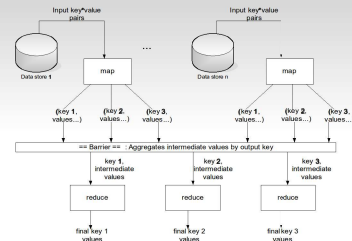
OLAP - Plataformas

- BD Paralelos
 - Existem desde a década de 80
 - É uma tecnologia madura e cara
- Map-Reduce
 - É um modelo de programação paralela
 - Está associada a uma implementação paralela e distribuída em um cluster de computadores
 - Introduzido pela Google e popularizado pelo Yahoo (Hadoop)

ERBD 2011



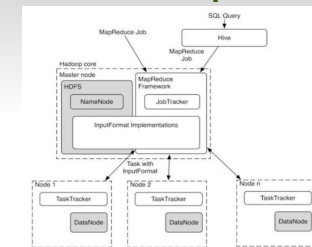
MapReduce



ERBD 2011



HadoopDB



ERBD 2011

HADOOP + SGBD



Desafios

- Segurança e privacidade dos dados
 - Modelo
 - Nível em que deve ser implementado
 - Criptografia, políticas de segurança
- Modelos de dados distintos para aplicações distintas?
 - independência física e lógica
 - Qual o modelo / arquitetura de compartilhamento?
 - OLAP / OLTP

ERBD 2011



Desafios (cont.)

- Mapeamento entre o modelo lógico e físico
 - Particionamento e Localização dos dados
- Processamento de consultas e métodos de acesso (Indexação, otimização e tuning)
- Modelo de suporte a transações
 - Pelo repositório chave-valor
 - Componente do sistema gerenciador de banco de dados

ERBD 2011



Desafios (cont.)

- SGBDs multi-inquilinos
 - Grande quantidade de inquilinos
 - Carga variável ao longo do tempo – balanceamento de carga
 - Migração dinâmica de dados

ERBD 2011



Gerenciamento de Dados na Nuvem

Um novo modelo de armazenamento de dados que apresenta diversos desafios para prover:

- Escalabilidade
- Elasticidade
- Consistência
- Facilidade de desenvolvimento de aplicações

ERBD 2011



Referências

- Brantner, M., Florescu, D., Graf, D., Kossmann, D., Kraska, T., "Building a Database on S3", SIGMOD'2008
- "Amazon Web Services" <http://aws.amazon.com>
- Agrawal, D., Das, S., Abadi, A.E., "Big Data and Cloud Computing: Current State and Future Opportunities", EDBT'2011
- Fox, A. "Cloud Computing—What's in It for Me as a Scientist?", Science 28 January 2011: Vol. 331 no. 6016 pp. 406-407

ERBD 2011



- Michael Armbrust, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy Katz, Andy Konwinski, Gunho Lee, David Patterson, Ariel Rabkin, Ion Stoica, Matei Zaharia, "A View of Cloud Computing", CACM, Vol. 53 No. 4, Pages 50-58
- Levandoski, J.J., Lomet, D., Mokbel, M.F., Zhao, K.K., "Deuteronomy: Transaction Support for Cloud Data", CIDR'2011
- Cattell, R., "Scalable SQL and NoSQL Data Stores", Relatório Técnico, 2011
- Ghemawat S., Gobioff H., Leung, S-T, "The Google File System", SOSP'2003

ERBD 2011

