

云数据存储与管理

柴云鹏

IDC 的统计表明，2007 年一年的数据增量就达到了 281EB，最近几年的数据增长率在 60%左右，相当于每过 18 个月翻一番。在数据量如此迅速增长的背景下，传统的存储系统和数据管理系统在扩展性、效率和成本等方面遇到了巨大的挑战，无法满足需求。而云数据管理具有扩展性强、性价比高、容错性好等优势，可以承担起超大规模数据存储的重任。

我们目前在云数据管理系统方面与 IBM、NSN 等公司进行合作，研究和开发适合企业级应用的云数据管理系统。具体来说，在云数据库方面，我们基于 Map-Reduce 模型，并行实现 SQL 操作，并通过多维索引结构来优化系统的性能；在云存储系统方面，我们在云存储中引入了闪存缓存层，一方面能够提高系统的访问性能，另一方面通过合理的数据分布策略来减少整个系统的能耗。

Cloud Data Management (CDM)

Yunpeng Chai

Outline

- ◆ **Motivation of CDM**
- ◆ Survey of CDM
- ◆ IBM SUR Cloud DB Project
- ◆ NSN Cloud Project
- ◆ Future Research Work

Motivation of CDM

- ◆ Rapid Data Increase
 - ◆ 2007: 281EB (1EB = 1000PB)
 - ◆ IDC: 60% per year (double - 18month)
 - ◆ Facebook: 850 million photos
& 8 million videos ONE day
 - ◆ Web pages, logs, media contents



3/26

Motivation of CDM

- ◆ Challenges
 - ◆ Scalability
 - ◆ Load Balance
 - ◆ Congestion / Delay
 - ◆ Efficiency
 - ◆ Cost
 - ◆ Fault Tolerance
 - ◆ Energy Conservation



4/26

Motivation of CDM

- ◆ Advantages of CDM
 - ◆ Scalability
 - ◆ Fault Tolerance
 - ◆ Performance/Efficiency
 - ◆ Performance-cost
 - ◆ Living and evolving



5/26

Motivation of CDM

- ◆ Comparison
 - ◆ Traditional Storage and DB
 - ◆ RAID/NAS/SAN
 - ◆ Hundreds of nodes
 - ◆ Cloud Data Management
 - ◆ Commodity Computers + SATA Disk
 - ◆ Tens of thousands of storage nodes



6/26

Motivation of CDM

- ◆ Cloud Data Management Systems
 - ◆ Scalability
 - ◆ 100 → 10,000 nodes
 - ◆ Efficiency
 - ◆ Application-driven
 - ◆ Cost
 - ◆ Commodity Computers + SATA Disks
 - ◆ Flash / Disk Energy Conservation



7/26

Motivation of CDM

- ◆ CDM Industry
 - ◆ Internet Service Cloud Platform
 - ◆ Google/Amazon
 - ◆ Cloud Products
 - ◆ IBM – Blue Cloud
 - ◆ MS - Azure
 - ◆ Enterprise-class Private Cloud
 - ◆ China Mobile
 - ◆ National Health Care



8/26

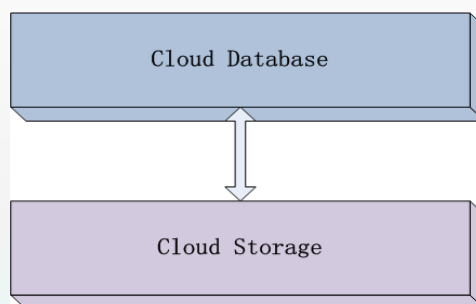
Outline

- ◆ Motivation of CDM
- ◆ **Survey of CDM**
- ◆ IBM SUR Cloud DB Project
- ◆ NSN Cloud Project
- ◆ Future Research Work

9/26

Survey of CDM

- ◆ CDM System Survey
 - ◆ Commercial Systems
 - ◆ Google- GFS + BigTable
 - ◆ MS- Azure
 - ◆ Yahoo!- PNUTS
 - ◆ Amazon- Dynamo
 - ◆ Open Source Systems
 - ◆ HDFS
 - ◆ KFS
 - ◆ Hbase
 - ◆ HyperTable
 - ◆ Cassandra
 - ◆ CouchDB
 - ◆ Voldemort



10/26

Survey of CDM

- ◆ Cloud Database:
 - ◆ Key-value data model
 - ◆ R/W performance / Parallelism
 - ◆ No/ Simple SQL operations



11/26

Survey of CDM

- ◆ Cloud Storage:
- ◆ Architecture:
 - ◆ Master-Slave
 - ◆ GFS, HDFS
 - ◆ Baidu Cloud
 - ◆ P2P
 - ◆ Cassandra



12/26

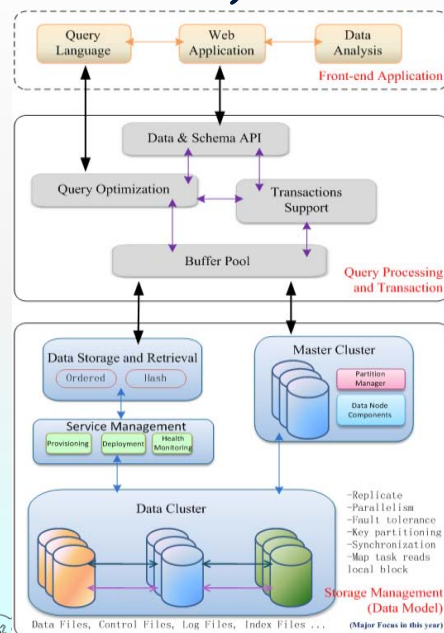
Outline

- ◆ Motivation of CDM
- ◆ Survey of CDM
- ◆ **IBM SUR Cloud DB Project**
- ◆ NSN Cloud Project
- ◆ Future Research Work

13/26

IBM SUR Cloud DB Project

- ◆ Cloud Data Management Systems
 - ◆ Based on Cassandra (Facebook)



14/26

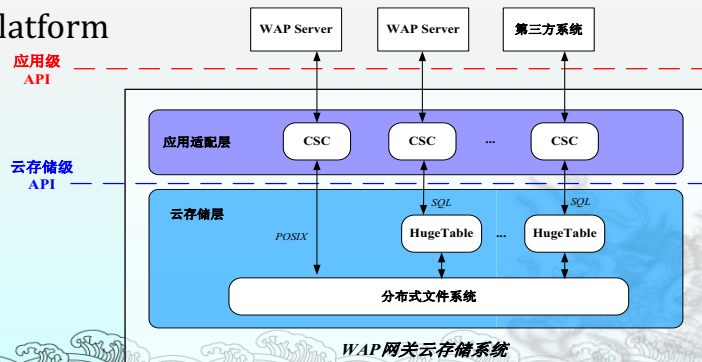
Outline

- ◆ Motivation of CDM
- ◆ Survey of CDM
- ◆ IBM SUR Cloud DB Project
- ◆ **NSN Cloud Project**
- ◆ Future Research Work

15/26

NSN Cloud Project

- ◆ Background
 - ◆ Telecom Cloud
 - ◆ Scalability
 - ◆ Cut costs
 - ◆ Unified platform



16/26

NSN Cloud Project

◆ Deliverables

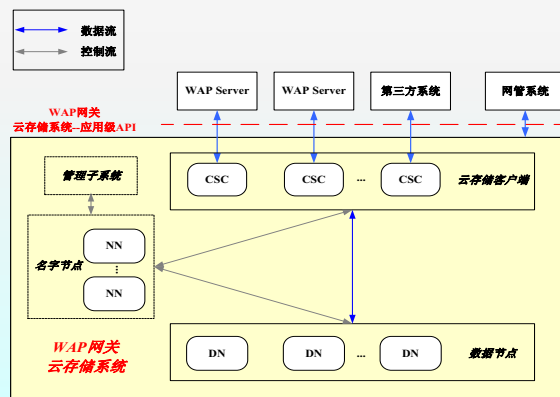
- ◆ Package of source code for CDM
 - ◆ CMRI WAP specification as input
 - ◆ Based on open source cloud project
 - ◆ Suit telecom app/workload
- ◆ Demo of CDM*
 - ◆ Integrated with NSN WAP Server
- ◆ Relevant investigation report and technical documentation

17/26

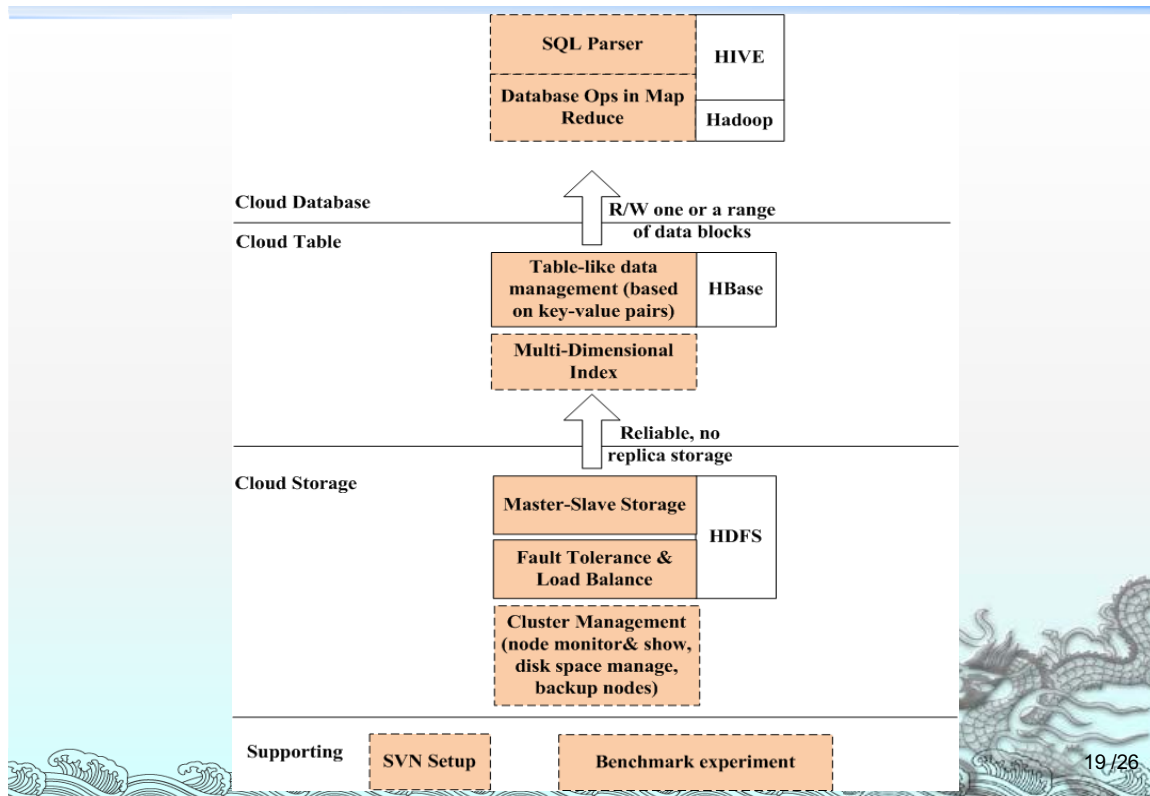
NSN Cloud Project

◆ Solution

- ◆ Cloud Storage
 - ◆ Based on HDFS (Java, M-S arch)
 - ◆ Data dist strategy
- ◆ Cloud Database
 - ◆ Hbase (Java)
 - ◆ Data type
 - ◆ DB Func
 - ◆ SQL support
 - ◆ Performance opt



18/26



Outline

- ◆ Motivation of CDM
- ◆ Survey of CDM
- ◆ IBM SUR Cloud DB Project
- ◆ NSN Cloud Project
- ◆ **Future Research Work**

Future Research Work

- ◆ Cloud Database
 - ◆ Data model
 - ◆ DB operation extension
 - ◆ Efficiency
 - ◆ Multi steps
 - ◆ Distributed operations



21/26

Future Research Work

- ◆ Cloud Storage
 - ◆ Architecture: Master-Slave vs P2P
 - ◆ Flash
 - ◆ Energy Conservation



22/26

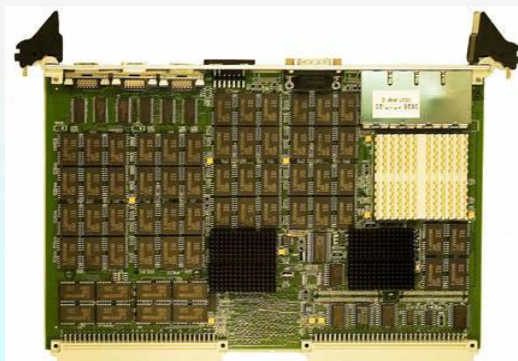
Future Research Work

- Architecture: Master-Slave vs P2P
 - ◇ P2P
 - ◆ Easy to manage, scalable
 - ◆ Fault tolerant
 - ◇ Master-Slave
 - ◆ High efficiency

23/26

Future Research Work

- ◇ Mixed Storage (Flash + Disk)
 - ◇ Flash Database
 - ◆ Commercial SSD based optimization
 - ◆ Flash storage board



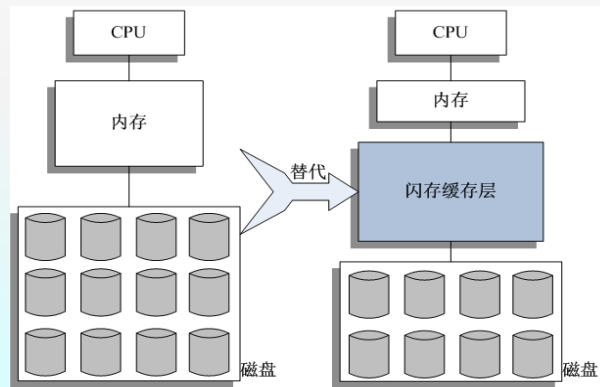
24/26

Future Research Work

- ◆ Mixed Storage (Flash + Disk)

- ◆ Flash's Drawbacks:

- ◆ Costs
 - ◆ Erase



Future Research Work

- ◆ Energy Conservation

- ◆ More Energy-Efficient SSD
 - ◆ Dynamic data distribution

