



# XData: Sugon all-in-one BigData Machine

Huaiming Song  
[songhm@sugon.com](mailto:songhm@sugon.com)  
R&D Center, Sugon

# Outline



**Big Data Background**



**XData: Sugon Bigdata Machine**



**Application Cases**

# Big Data Anywhere



**420/564 Million**  
Mobile/Internet headcounts  
2012.12 China

**\$10 Billion**  
LBS worth in 2016



**209 Billion**  
# of FRID in 2021  
12M in 2011



**1.8ZB in 2011**  
Data Vol. of 2 days  
> Total by 2003



**200PB/Quarter**  
Intelligent City Data  
One big City in CN



**5PB/Year**  
Health Record  
One big City in CN

"Data are becoming the new raw material of business: an economic input almost on a par with capital and labour."  
--2010 , The Economist  
"Information will be the oil of the 21<sup>st</sup> Century"  
--2010 , Gartner

# Application Scenarios



## 1. Telecommunications

- Phone call Records
- Text&speech Retrieval
- Local Based Service



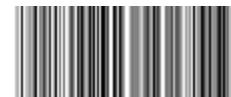
## 2. Internet&Mobile Internet

- Netflow Stats
- Social Network Service
- Text&video Retrieval



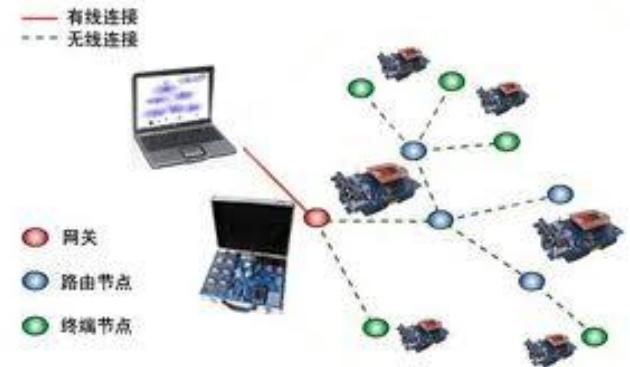
## 3. Transaction Data

- History Transaction Stats
- Anomaly Detection
- Business Intelligence

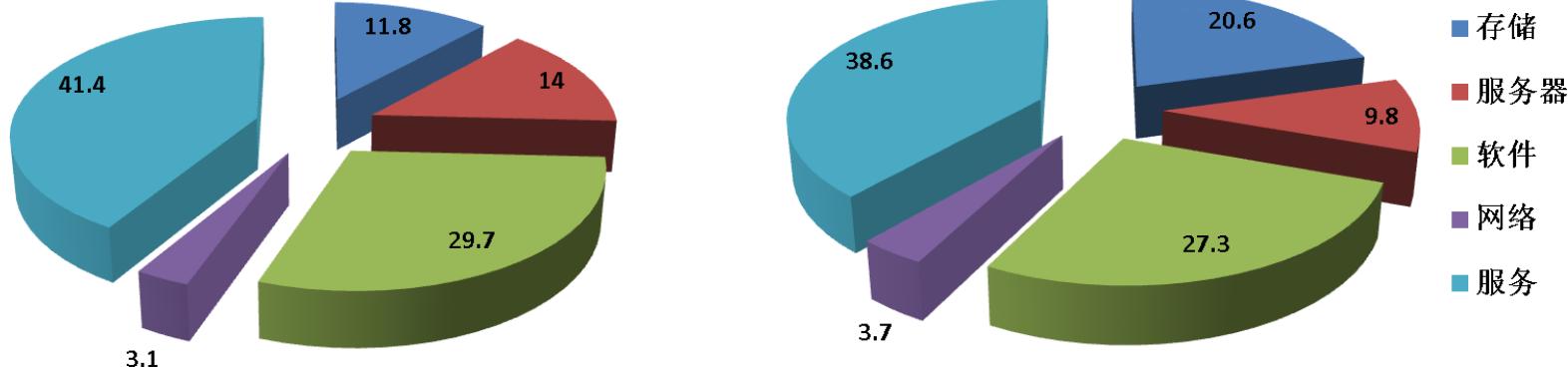


## 4. Sensors& The Internet of Things

- Streaming Processing
- Offline Data Statistics
- Streaming Mining



# Changes by Big Data



The Big Data market share in the world in 2011 & 2015. Annual growth rate is more than 50% next 5 years. IDC Big Data report (2012)

- Industry: business merges & acquisitions, e.g. EMC, IBM, HP, Teradata, Intel...
- Government: Big Data R&D Plan (Mar. 2012, US; 2012, CN...)
- Scientific research: committee of experts (CCF, CIC...)

Big Data is one of the core technologies of Cloud Computing.

# Opportunities & Challenges

The Big Data market is in the ascendant, bringing new opportunities and challenges

## OPPORTUNITY

- ◆ Lots of systems have already been very big
  - Fast growth of Mob., Int. ...
- ◆ Technology in fast development, not become standard
  - Still in an early age
- ◆ Requirement is not clear, needs technique to promote
  - Intelligent City, Camera video analysis, etc.

## CHALLENGE(4v)

- ◆ Volume
  - TBs to PBs
- ◆ Variety
  - Structured & unstructured
- ◆ Velocity
  - Require fast processing
- ◆ Value
  - Valuable information

# Three Stages of Big Data

## First : Storage of Big Data

- Sys. prop. : Data acquisition and Storage
- Comp. prop. : Simple analysis of unstructured data (filtering and statistics)

## Second : Integrated storage & processing platform

- Sys. prop. : integrated big data storage and processing platform
- Comp. prop. : form the calculation model ,structured and unstructured data uniformly processing, support complex problem solving (cascade and associative processing)

## Third : Analysis & Mining , Service Integration

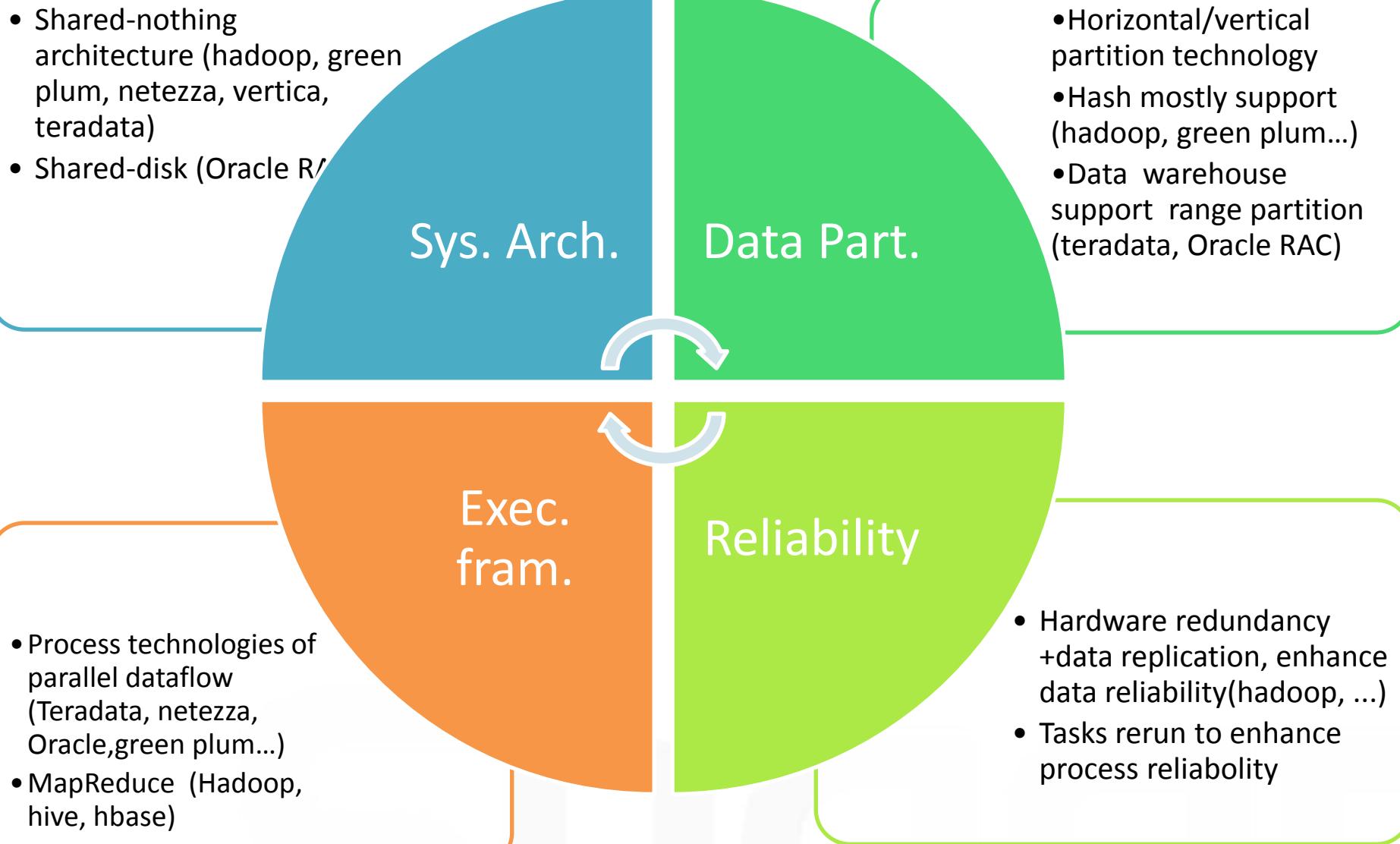
- Sys. prop. : lots of packaged application systems , data trans. to value
- Comp. prop. : integrated business process, complicated analysis & mining, approximation analysis

# System Requirements



- Functionality: structured & unstructured data processing
- Performance: quick response
- Throughput: high parallel degree
- Manageability: graphics interface , easy to use
- Availability: runtime fault-tolerant
- Scalability: easy to scale-out

# Summary of Related Works



# Outline

1

**Big Data Background**



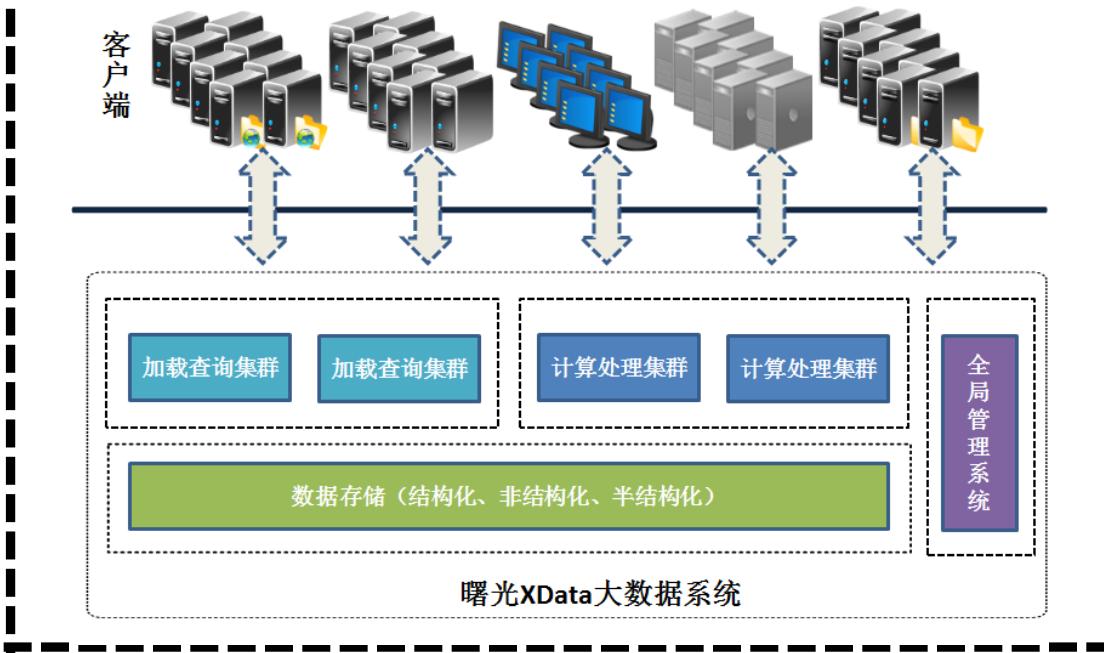
2

**XData: Sugon Bigdata Machine**

3

**Application Cases**

# XData: Sugon BigData Machine



XData is a general big data platform developed by R&D Center, Sugon. XData supports both structured & unstructured data processing, and it is widely used in communication data statistics, Internet/mobile Internet log and user behavior analysis, Internet of things/sensor network data analysis, and transaction data offline statistics and mining, etc.

# Why XData?

## □ XData = X + Data

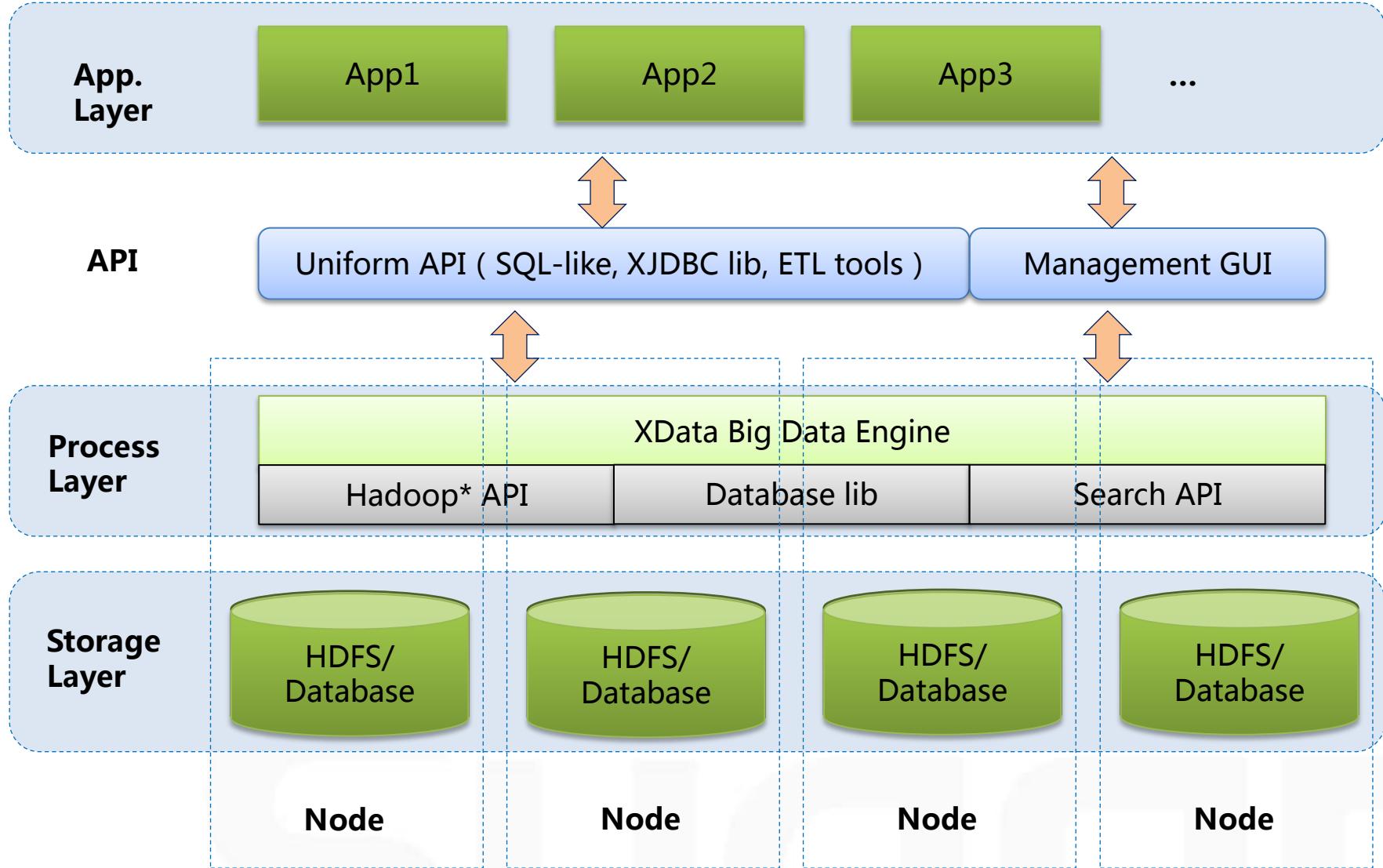
## □ X: Three Meanings

- Large
  - Exascale, means data **volume** is very big
- Unknown
  - X usually stands for an unknown number in functions, means **variety** data types
- Execution
  - Verb, means to process data

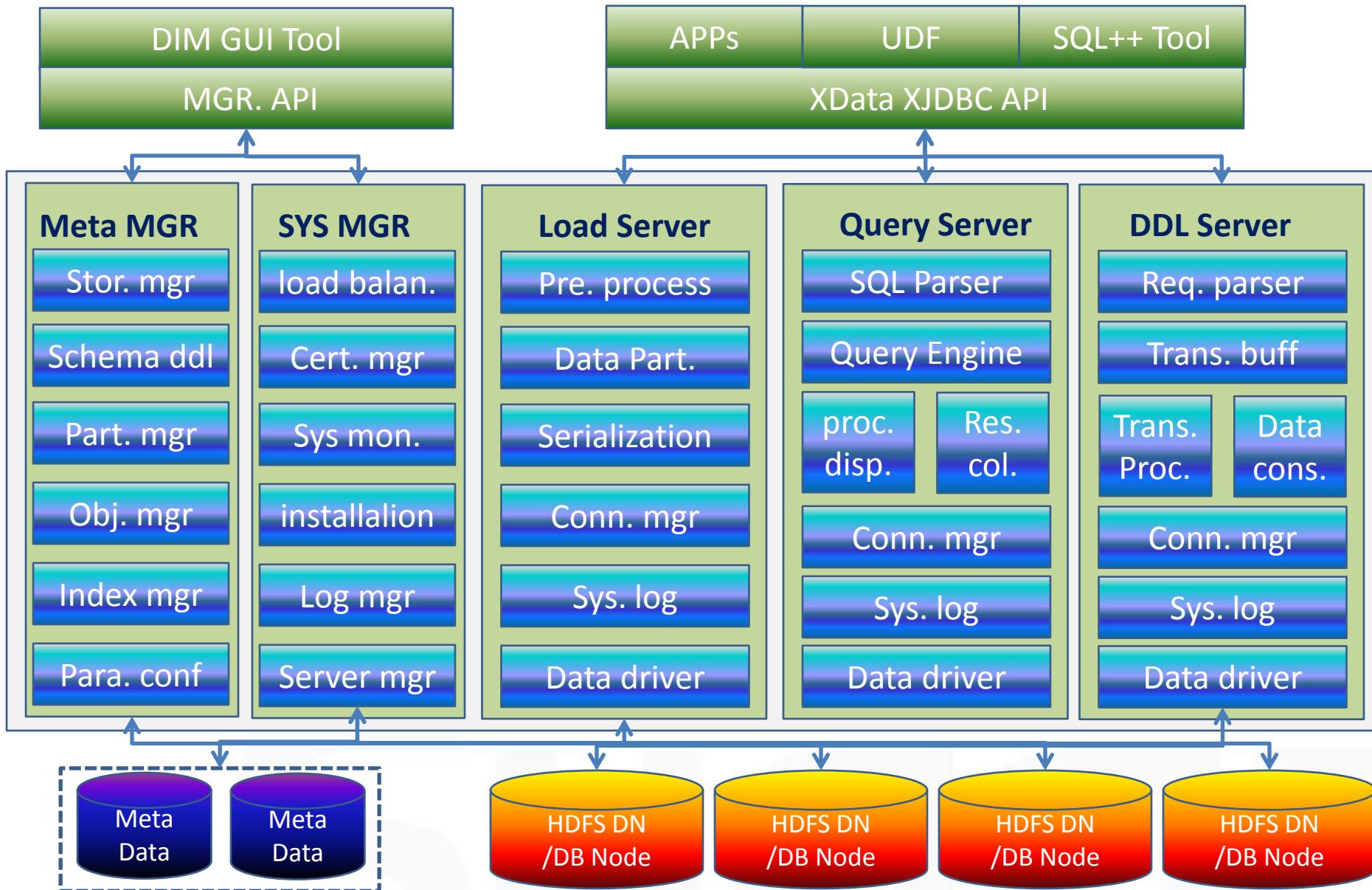
## □ Data

- Data platform: storage & process platform

# XData System Layer



# XData Software Architecture



# Dedicated Hardware (storage node)



处理器	最大支持2颗Intel E5 2600系列处理器 ( Cache : 4M/8M, 12M )
芯片组	Intel C606
内存/Max	支持DDR3 ECC内存，16根DIMM 最大支持512GB 内存
存储	最大支持24个热插拔SAS/SATA硬盘位，最高支持72TB存储 ( 内嵌2.5寸槽位*2 ) ； 升级可支持：36块3.5寸硬盘 ( 升级后无内置2.5寸硬盘槽位 ) 48块2.5寸硬盘 ( 内置两颗2.5寸磁盘 )
网卡	集成两个Intel 千兆网卡 ( RJ45接口 i350芯片 )
扩展性能	支持6根PCI-E 3.0扩展插槽 ( 3个x16,3个x8 )
电源	1200W 金牌高效冗余电源
管理	支持IPMI与KVM over IP；服务器
产品特点	高性价比大容量存储服务器、具有
应用领域	Cache应用、搜索引擎、游戏应用 多媒体、数据存储、VOD、媒体渲染系统、气象数据存储、视频监控、医疗影像、网 络托管



# Core Technologies

-  1 Data partition and layout
  - Mapping logical to physical layer
  - Data schema definition and explain
-  2 Computing/request definition
  - UDF pre-definition
  - SQL/MR cascading and association
  - Hybrid job flow define & parser
-  3 Parallel job flow execution control
  - Hybrid engine: MPI + MapReduce
  - Data Shard and dispatch
  - BigData join and nested query
-  4 Data transform and migration
  - Data conversion between platforms
  - Implicit & explicit conversion
  - Import/export, migration tools
-  5 Data associative loading
  - Data scheme mapping & APP migration
  - Multi-source join loading to HDFS\*
-  6 XData management system
  - System install deploy and config
  - HA and fault-tolerant

# 1. Data Partition & Layout

- **Physical layout of data object**

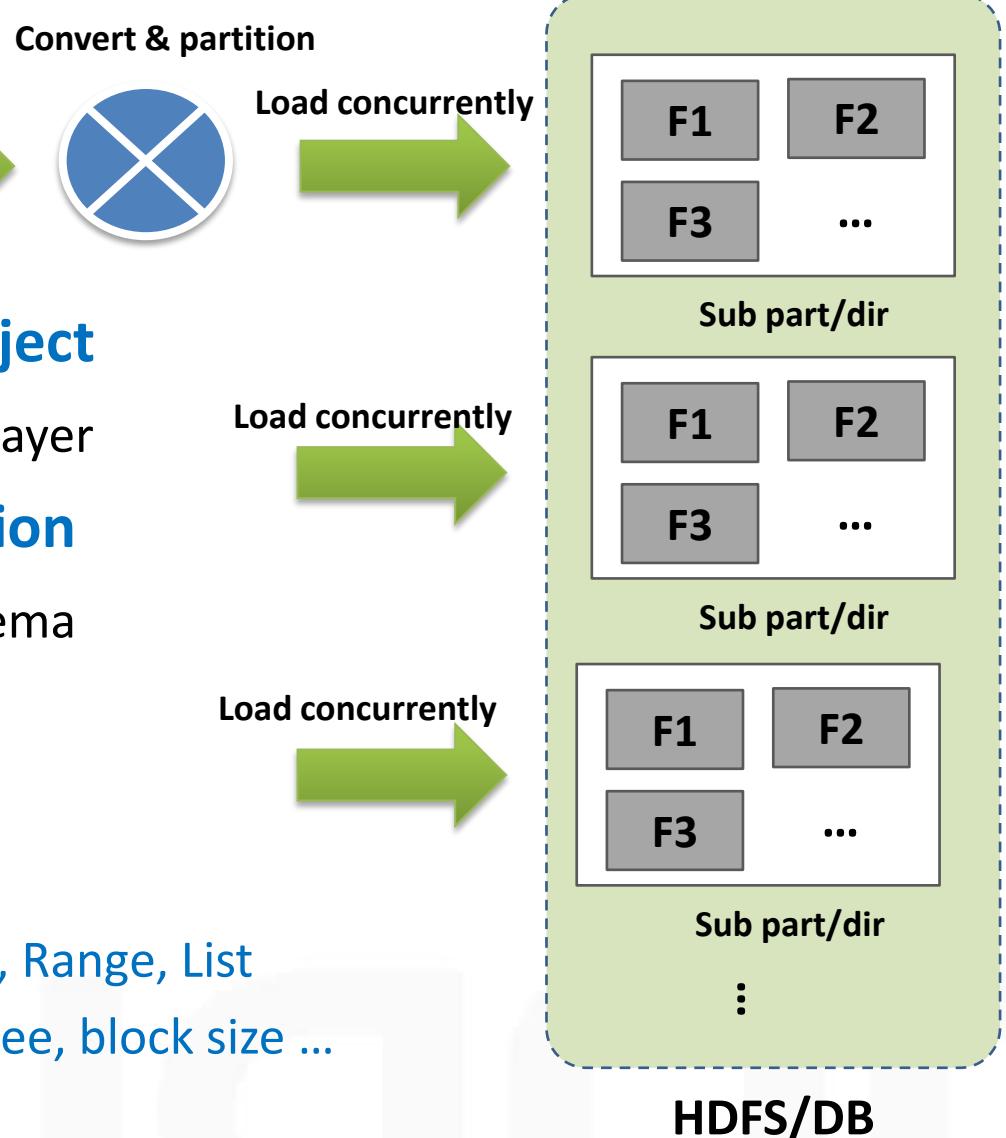
- Mapping logical to physical layer

- **Data object inner description**

- Mapping logical layer to schema

- **Data partitioning**

- partition by key
  - Partition = sub part/dir
  - Partition strategy : Hash, Range, List
  - Parameter: parallel degree, block size ...



# 2.Computing semantics definition

## ● Computing semantics

- Mapping of computing semantics to Job Flow
- Basic Job: MapReduce(MR), SQL, UDF

## ● Job cascading

- S(S), M(M), M(S), S(M)

## ● Job correlation (join)

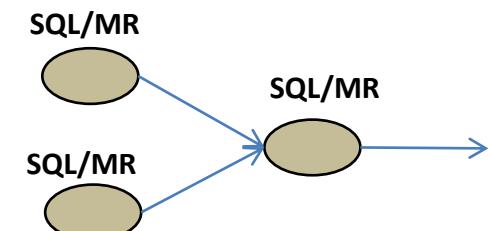
- SxS, MxM, SxM, MxS

## ● Formalizing description

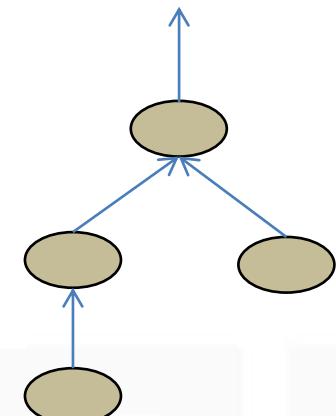
- $OP = \{\emptyset, S, M\}; OP=OP(OP) ; OP=OP \times OP$
- Job flow diagram to describe any computing semantics



**Cascading**



**Correlation(join)**



**Job flow**

# Samples

## ● Mapreduce job defination

- define wordcount(\$1) WordCount.jar input [runtime parameters];
- define sort(\$1) Sort.jar input [runtime parameters];

## ● Mapreduce job cascading

- sort(wordcount(file1));

## ● Mapreduce&SQL cascading

- wordcount(select name, age from table1);
- select count(key) from (wordcount(file1));
- select sum(value) from (wordcount(file1));

## ● Mapreduce/SQL Correlation (join)

- select name, age from table1 A join wordcount(dir1) B where A.name = B.key and B.value > 100;

# 3. Parallel Job Flow Description

## ● Query operation

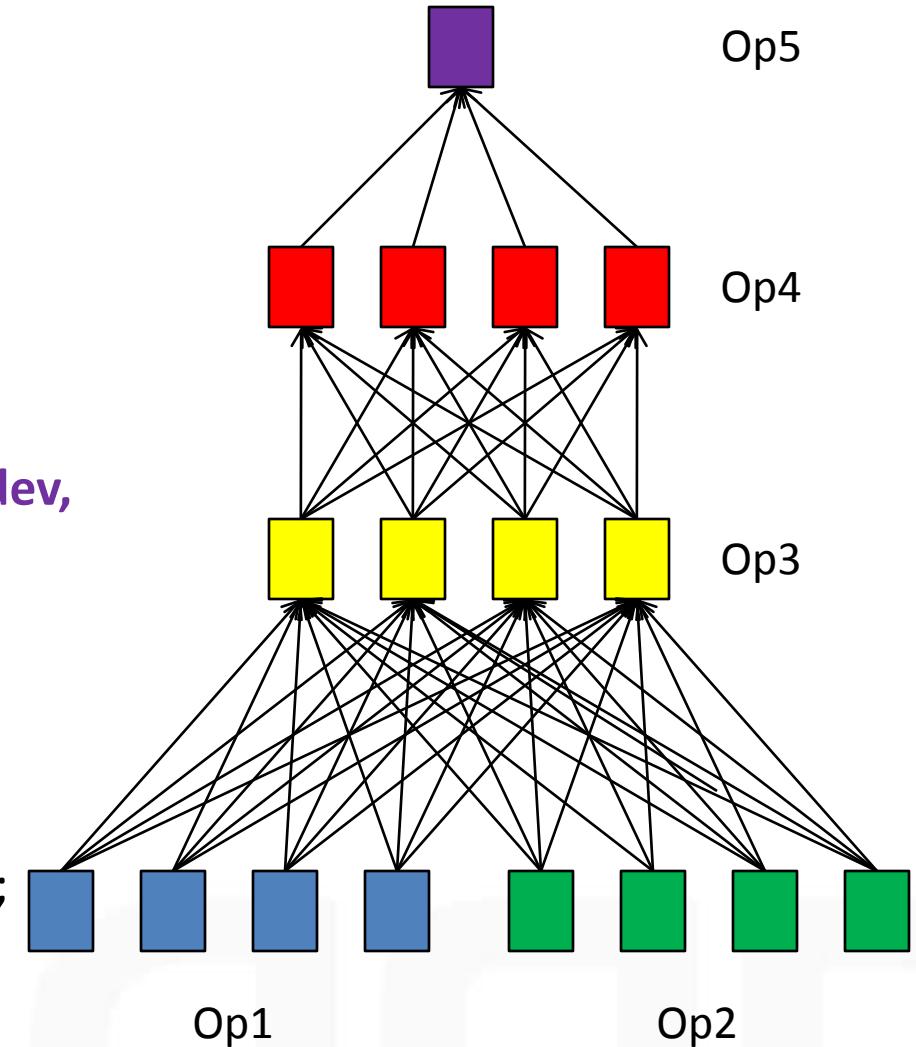
- Basic unit can be executed at each node
- Structured Data: SQL
- Unstructured: MR or UDF

## ● Query Job

- Four tuples<OP, src\_nodev, dst\_nodev, dd\_func>
- Dd\_func: hash, range

## ● Parallel Job Flow

- A directed acyclic graph(DAG)
- E.g. <T1, T2, [S], T3, [S], T4, [S], T5>;  
[S] is a synchronization point



# Job Flow Execution Control

## ● Request analysis

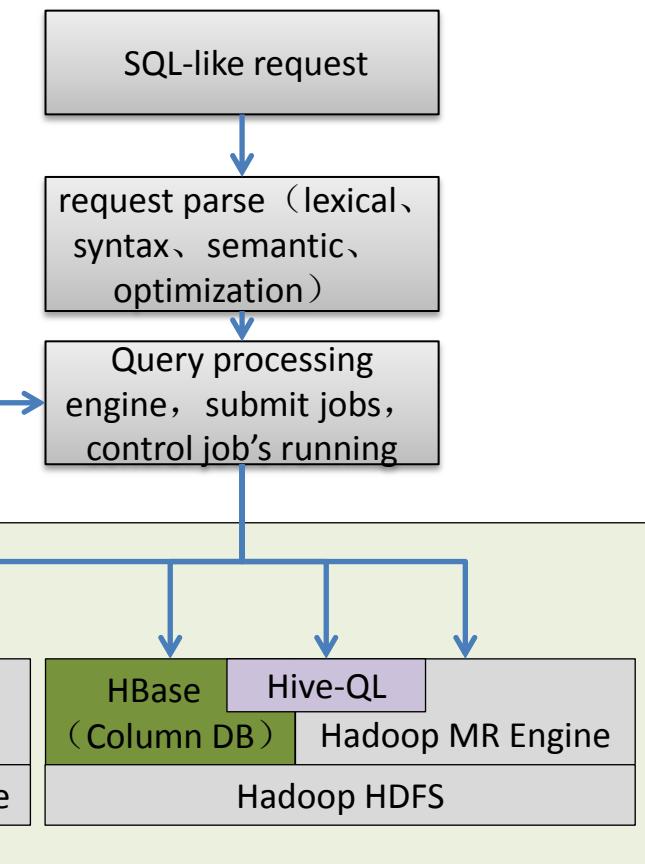
- SQL-like language=> Job flow
- Parser: Lex & bison

## ● Execute engine

- MR job: run on Hadoop
- SQL job: execute in DB
- HBase job: HBase driver
- Hive query: Hive driver

## ● Job Synchronous control

- Hybrid engine: MR + MPI
- Query server: submit command, control execution, gather results
- Data Node: run Jobs/tasks



XData Query Engine

## ● Data distribution

- Data format transform
- Data partition and distribution

# 4. Data Transforming & Migration

## ● DB/HDFS data conversion

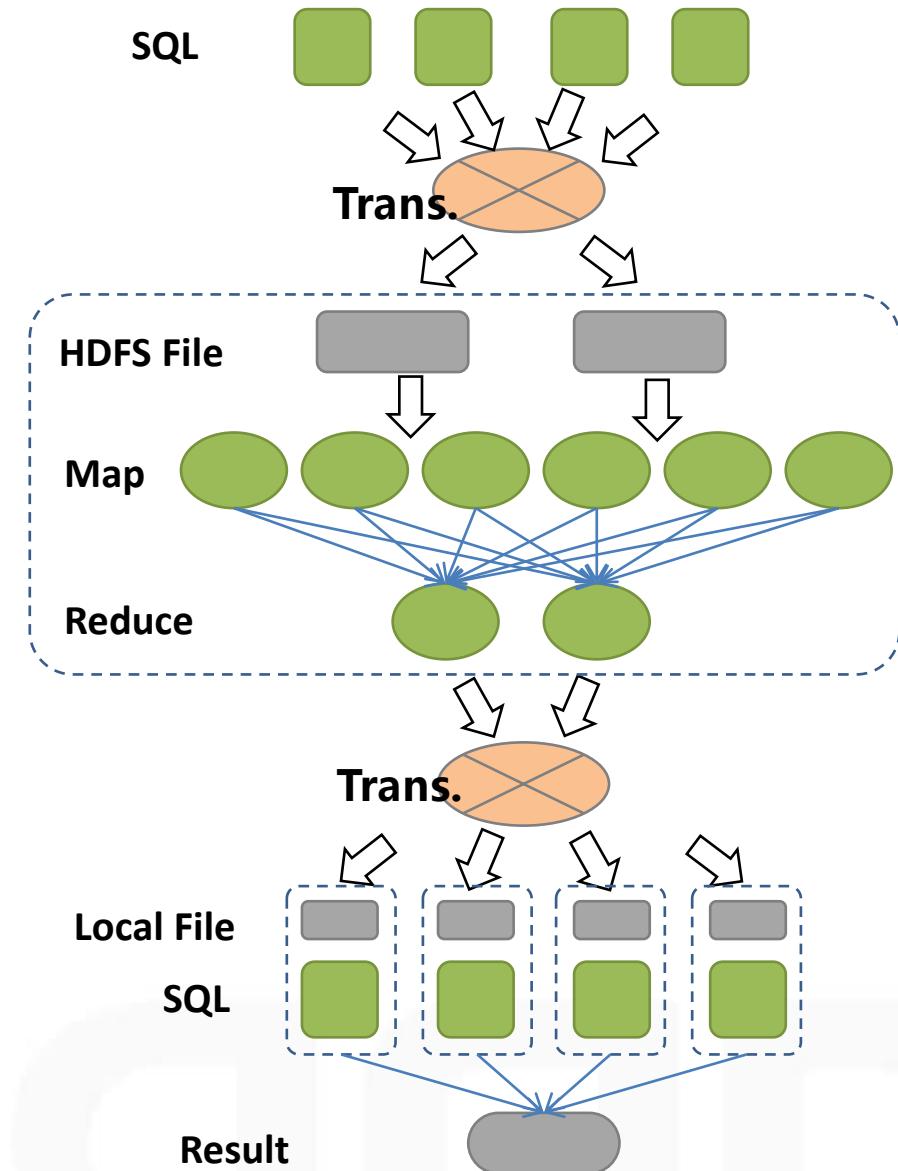
- ETL between DB/HDFS
- Data interaction between jobs

## ● Explicit transform

- DB <=> HBase/Hive
- Data import/export through sqoop
- Conditional import/export
- Data replication/recovery

## ● Implicit conversion

- MR/SQL hybrid jobs'  
cascading/correlation
- Modify data schema/layout  
(DB+HDFS)

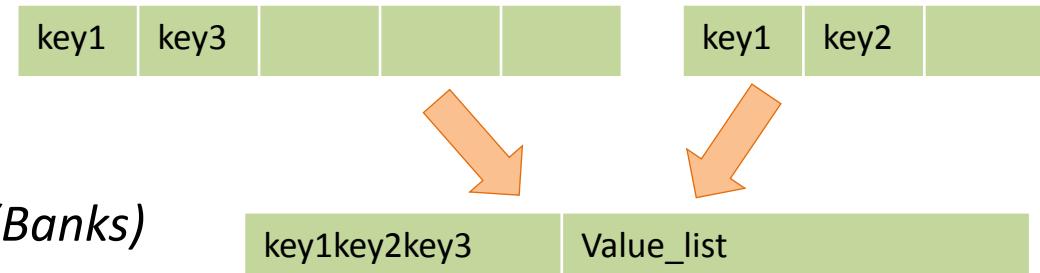


# 5. Associative Loading

## ● Application migration

- DB => HBase

*Example: POC APP migration(Banks)*



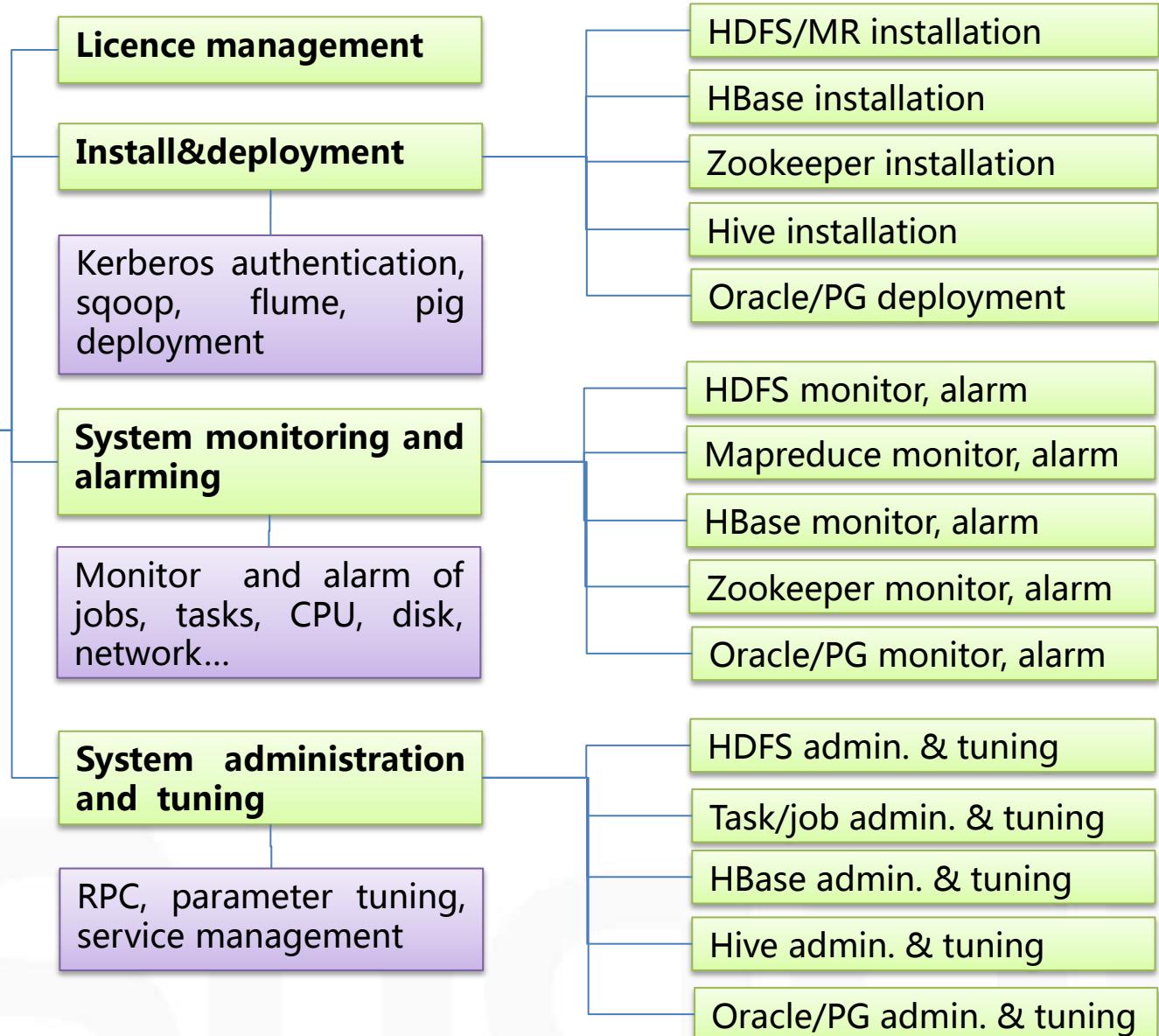
## ● Technical difficulty

- Hadoop/HBase only support data of simple schema <key , value>
- Limited operation: access data only by key, range scan, full table scan
- The design of row key and column family
- Data merge associatively introduces heavy overhead of disk-IO, network, computing

## ● Implement method

- Run through MapReduce, Hash Partition, parallel associative loading
- General schema definition of associative loading
- Let HBase support more applications

# 6.XData Management GUI



# XData Management GUI

XData 曙光通用海量数据处理平台

 监控中心
 集群配置
 系统管理

欢迎 admin




**监控中心**

- [Hadoop监控](#)
- [资源监控](#)
- [告警](#)
- [报表](#)

**Hadoop监控**

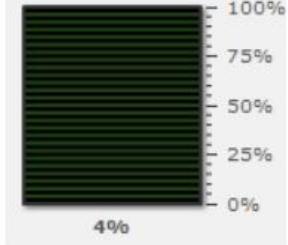
组件信息

名称	状态	角色配置
 Zookeeper	<span style="color: green;">运行中</span>	3个Zookeeper
 HDFS	<span style="color: green;">运行中</span>	1 个SecondaryNameNode, 1一个Namenode, 3个DataNode
 MapReduce	<span style="color: green;">运行中</span>	1 个JobTracker, 3个TaskTracker
 Hbase	<span style="color: pink;">未运行</span>	3 个HMaster, 3个RegionServer
 Hive	<span style="color: grey;">未配置</span>	未配置

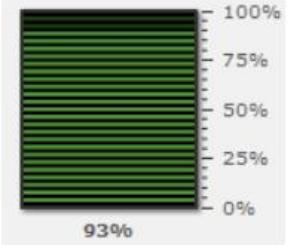
**集群性能**



平均负载  
0.45



CPU使用情况  
4%

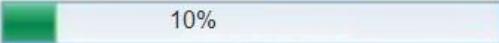


内存使用情况  
93%

**HDFS存储信息**

**Namenode**

总容量: 100TB 已用容量: 10%



**Datanode**

总容量: 100TB 已用容量: 10%



**MapReduce信息**

**Slots**

总数: 100 已用: 10



 5
 20
 20
 20

# Outline

1

**Big Data Background**

2

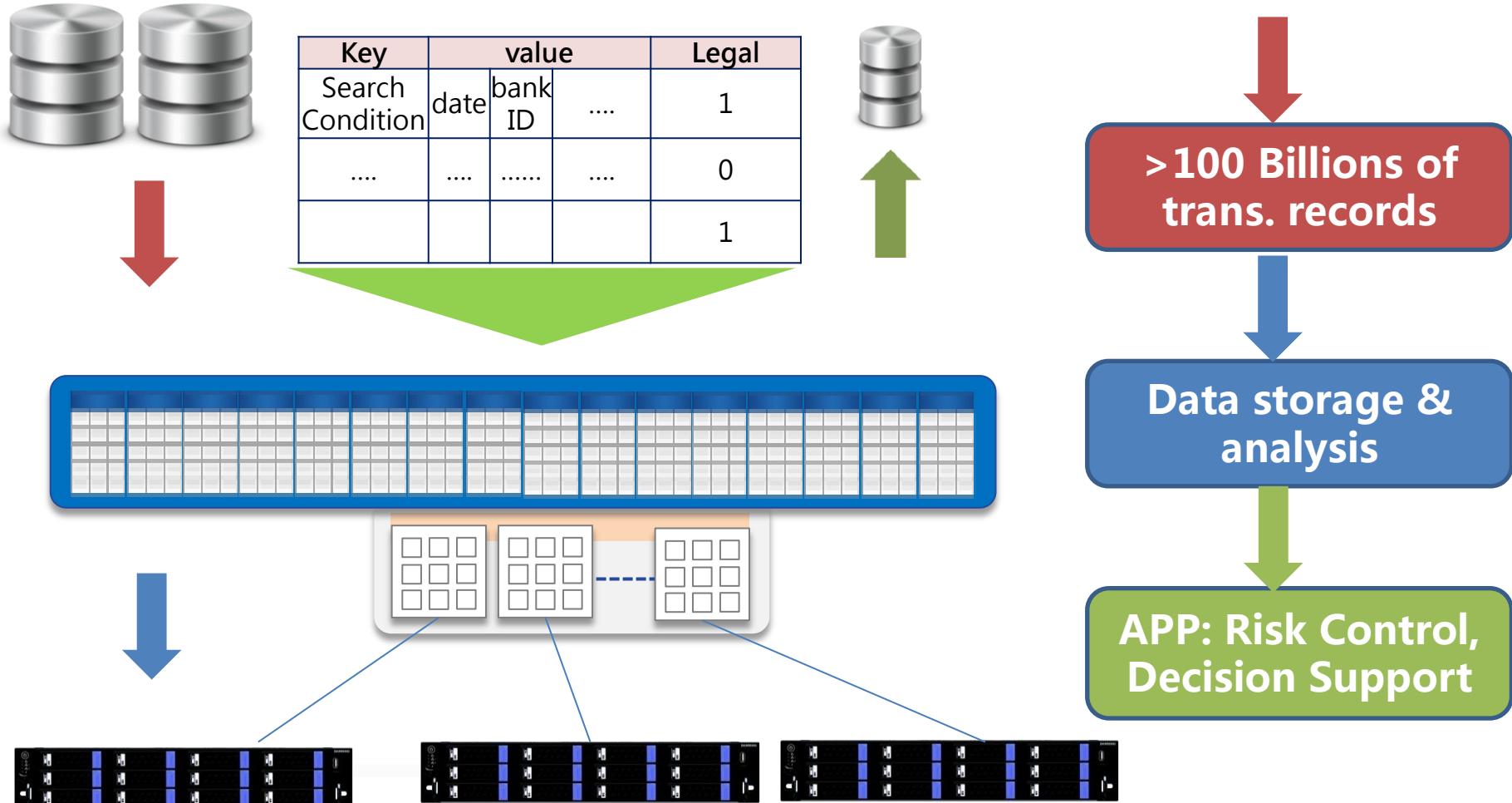
**XData: Sugon Bigdata Machine**

3

**Application Cases**

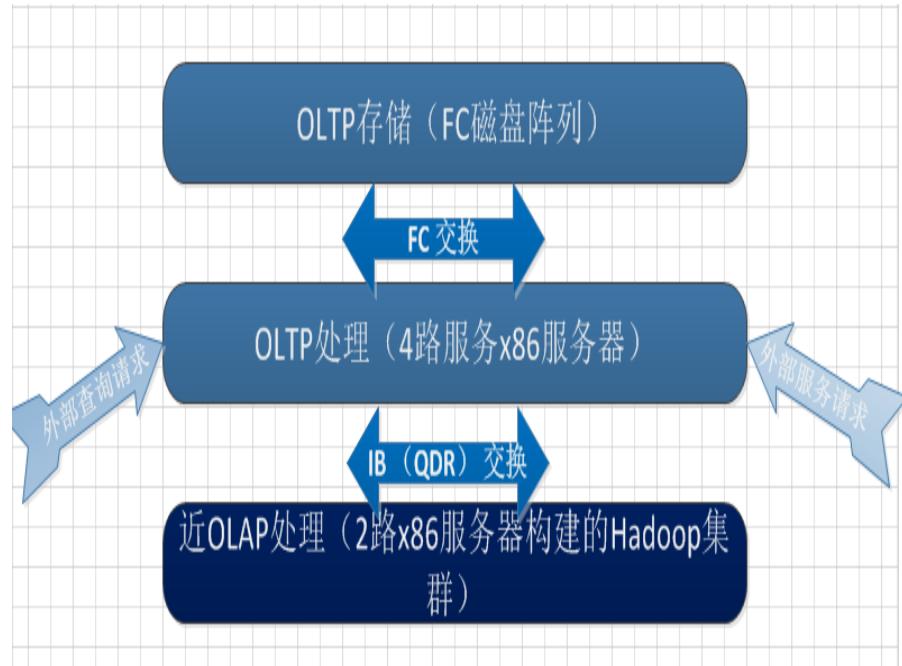
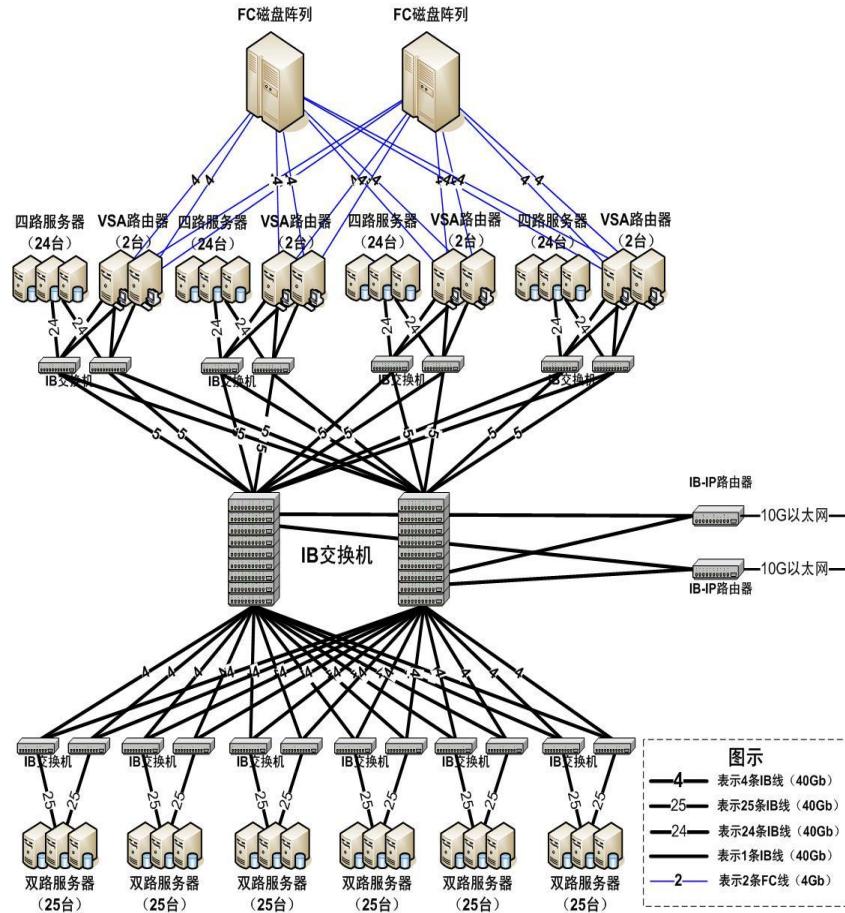


# Trans. Data Offline Analysis



Trans. data offline analysis system for China UnionPay

# Traffic accounting (Henan)



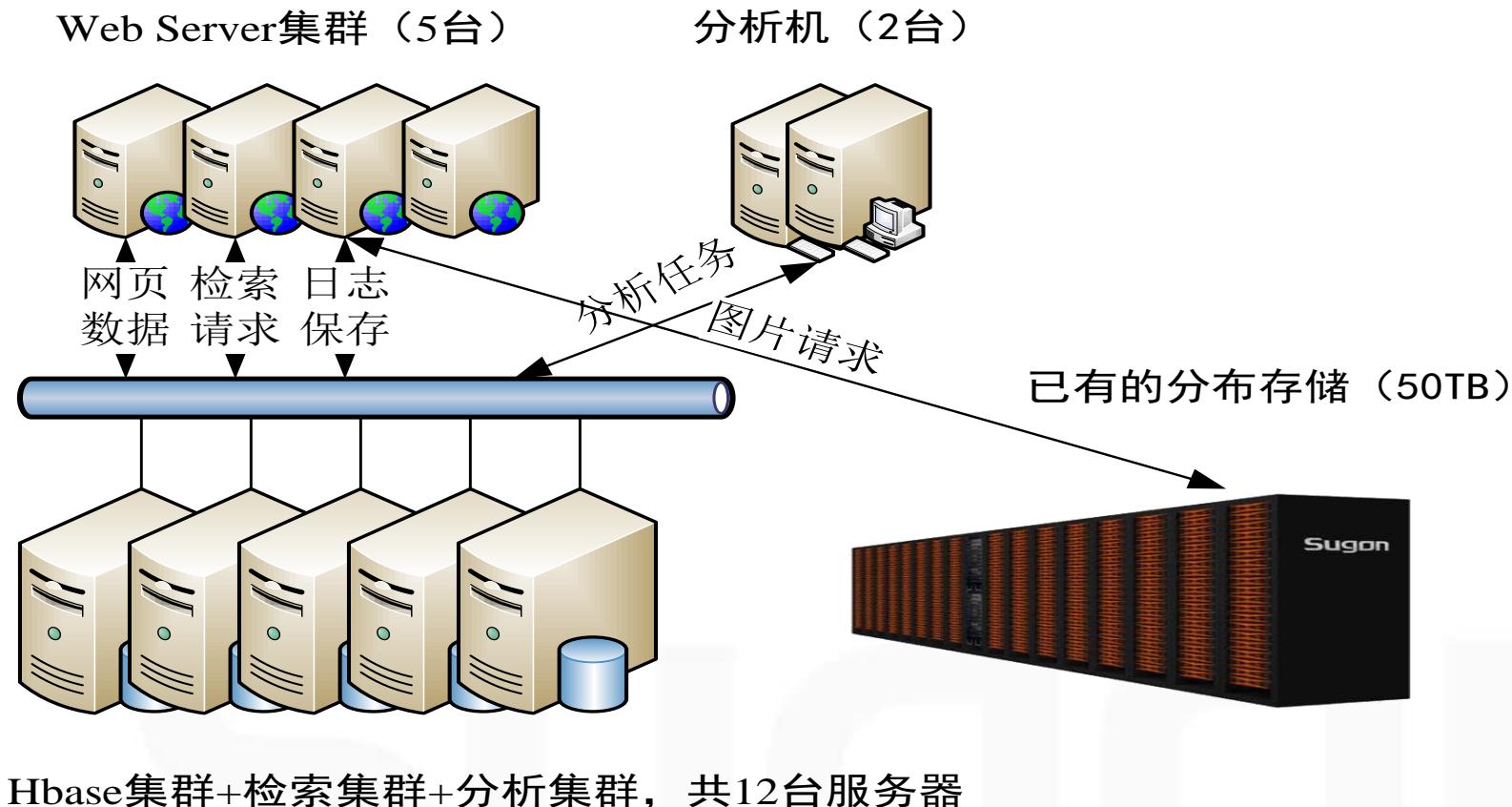
Sugon XData platform for traffic accounting app. in CMCC, Henan

# Photograph Forum



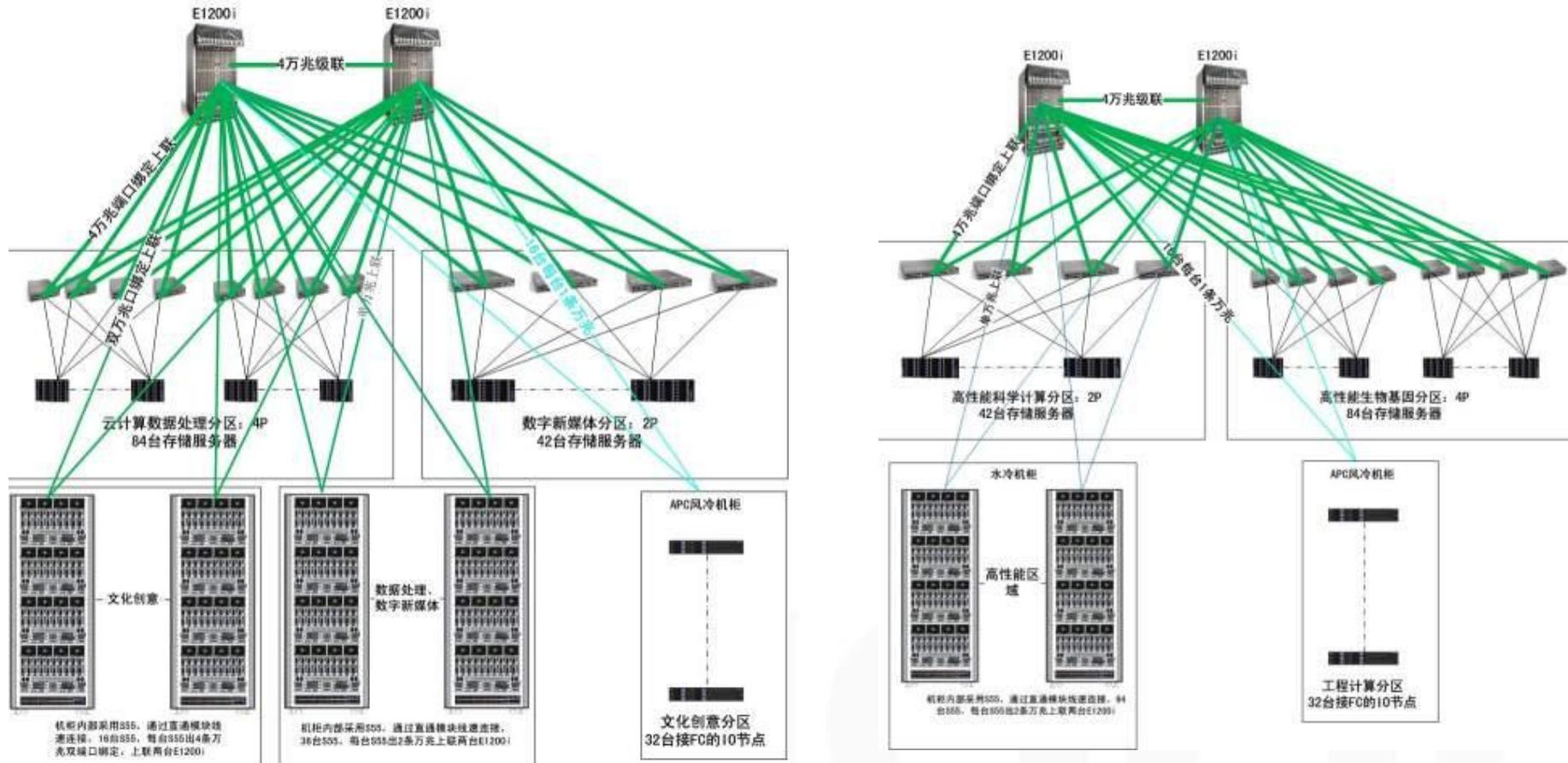
## ● System Architecture

- Vm to build PHPBB platform
- Storage server to build up Hadoop platform
- Sugon Parastor200 parallel file system

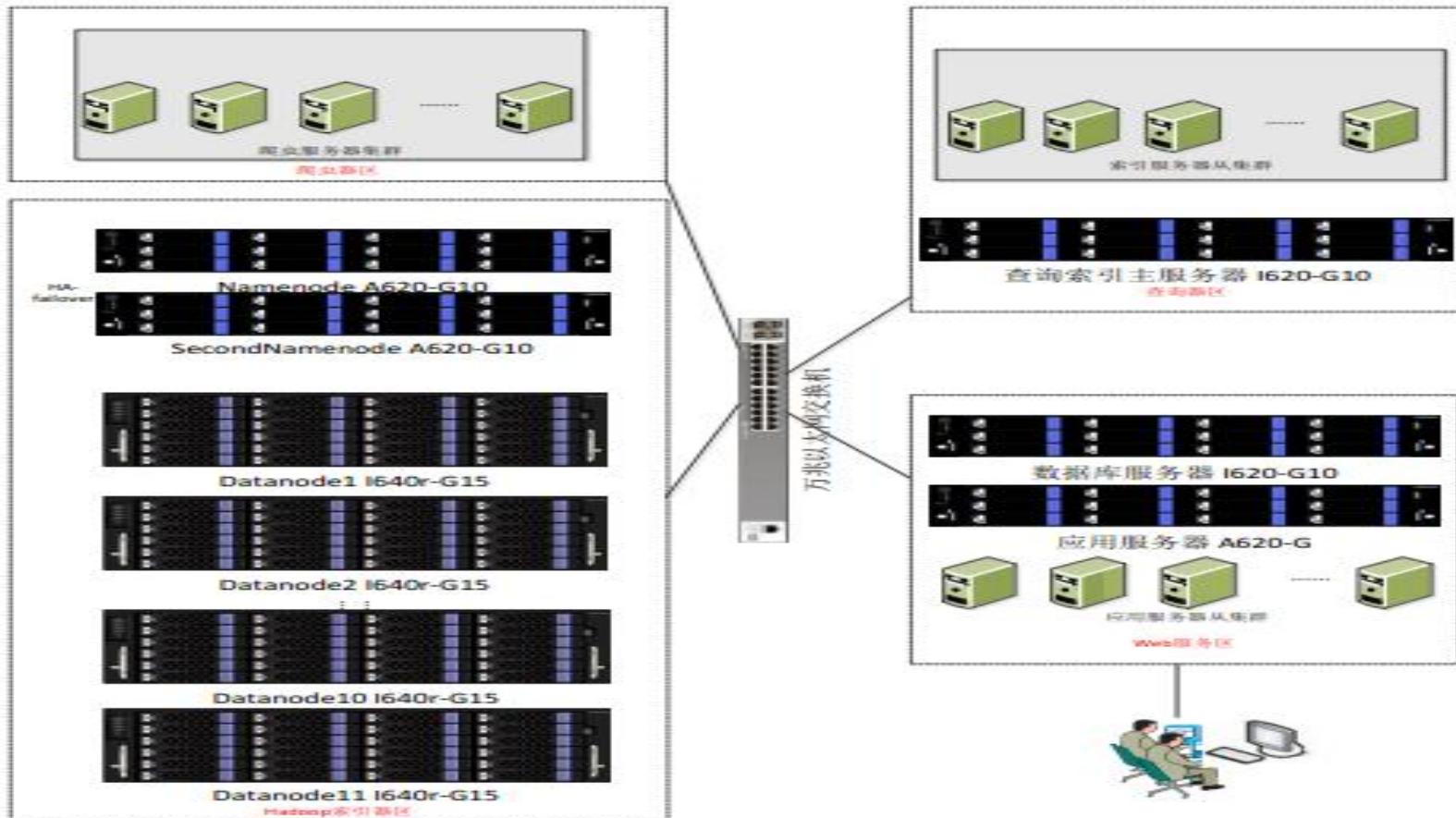


# PFS in Huanan SC. Center

- Deploy with Sugon Nebula, for HPC applications
- 16PB capacity, largest single parallel file system in China

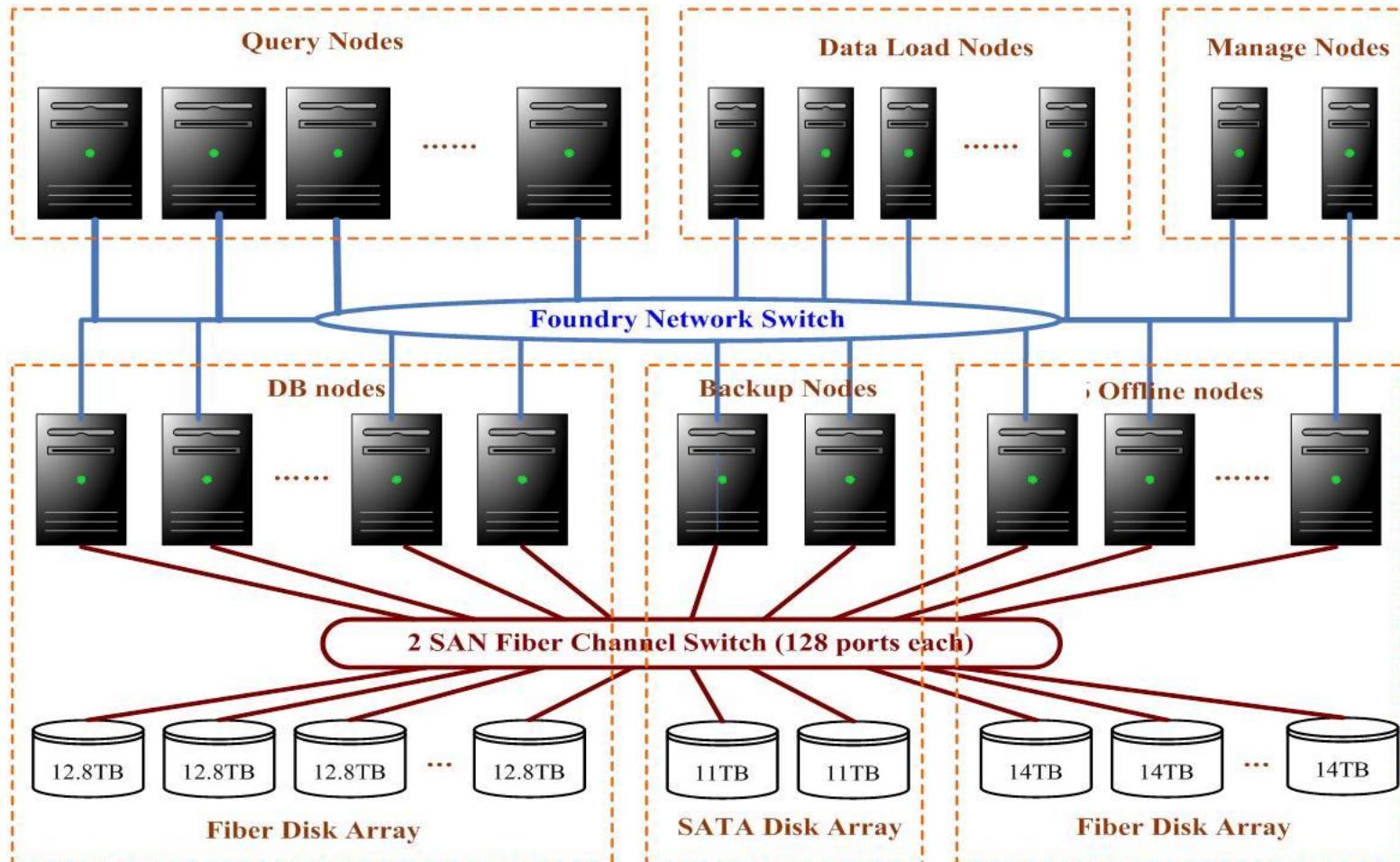


# BigData platform(Tongji university) Sugon



Flexible data processing platform: PaaS layer, VM + PM, easy to reconfiguration and allocation, supporting all scientific research in the university

# Info. Secur. Ana. Sys. (gov)



Data volume 600TB, hierarchical storage, for information retrieval, statistics, and log analysis



谢谢！

Thanks!