

# 百度云计算介绍

百度云计算架构师 肖伟

xiaowei@baidu.com



# 百度内部系统架构

搜索

广告

社区

业务

云平台

数据智能

海量数据存储与处理

数据中心、网络、服务器

软件基础架构

硬件基础架构

云计算



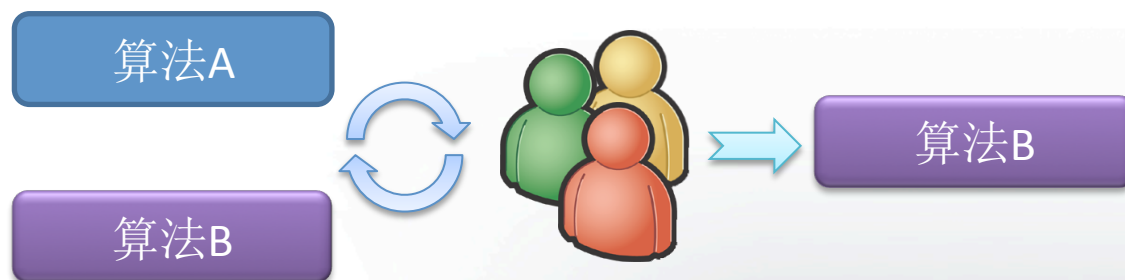
# 百度内部云特点

快速迭代

超大规模



# 快速迭代特征



通过反馈来验证算法优劣



离线分析与在线实验相结合



# 互联网对快速迭代的需求

数据收集  
A/B测试

数据智能



- ✓ 产品开发全过程
- ✓ 方法论+流程+平台



# 超大规模特征

- 海量数据存储 百P
- 大规模数据计算





# 大规模计算

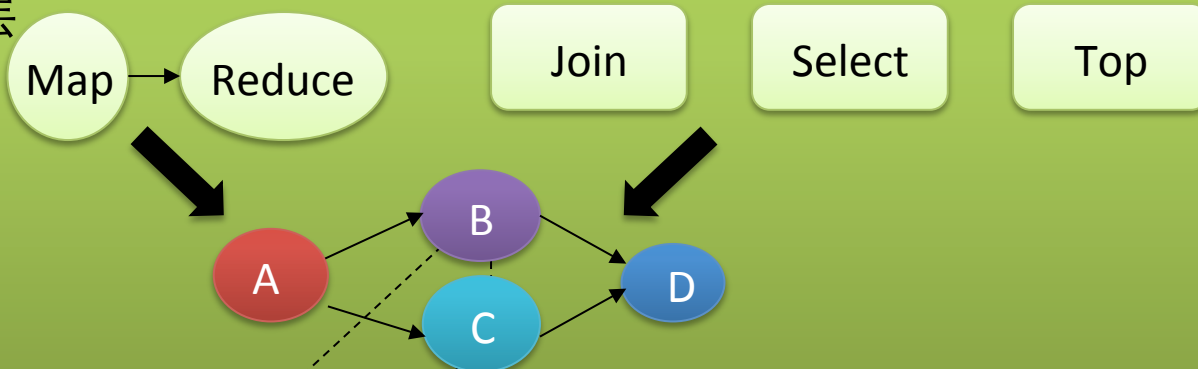
表示层

SQL-like

描述能力

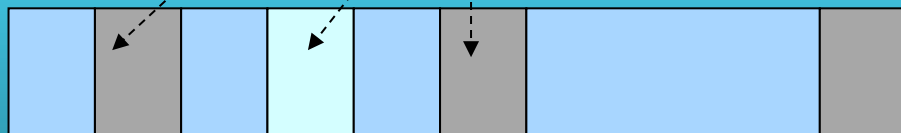
翻译

模型层



数据流优化  
控制流管理

执行层



资源分配  
优先级、并发控制  
隔离、安全



# 计算类型

## Map Reduce

- 通用批量数据处理

## HPC

- 大规模机器学习

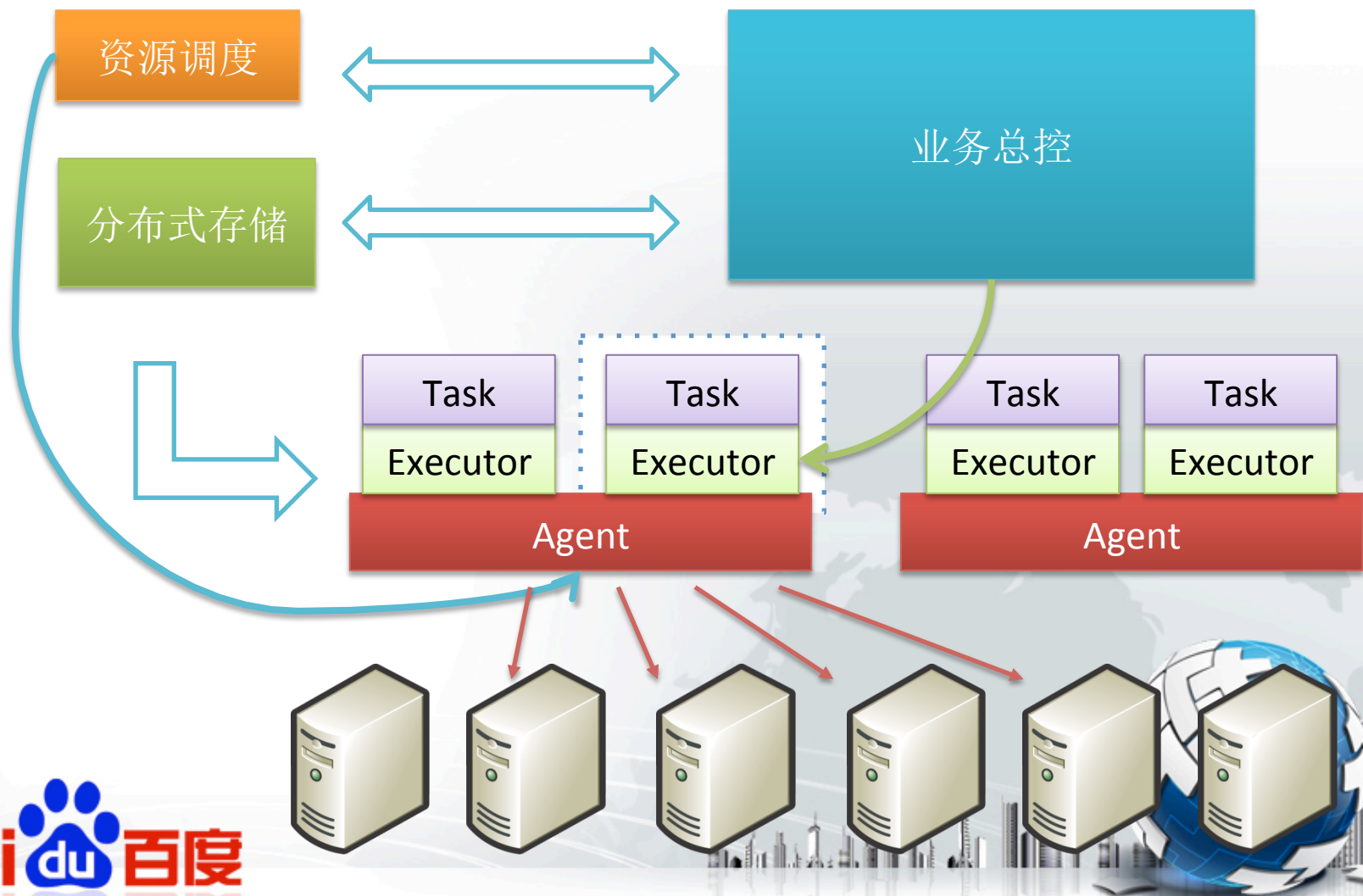
## Volunteer Computing

- 计算密集型上下文无关数据处理





# 资源调度系统



# 百度存储体系



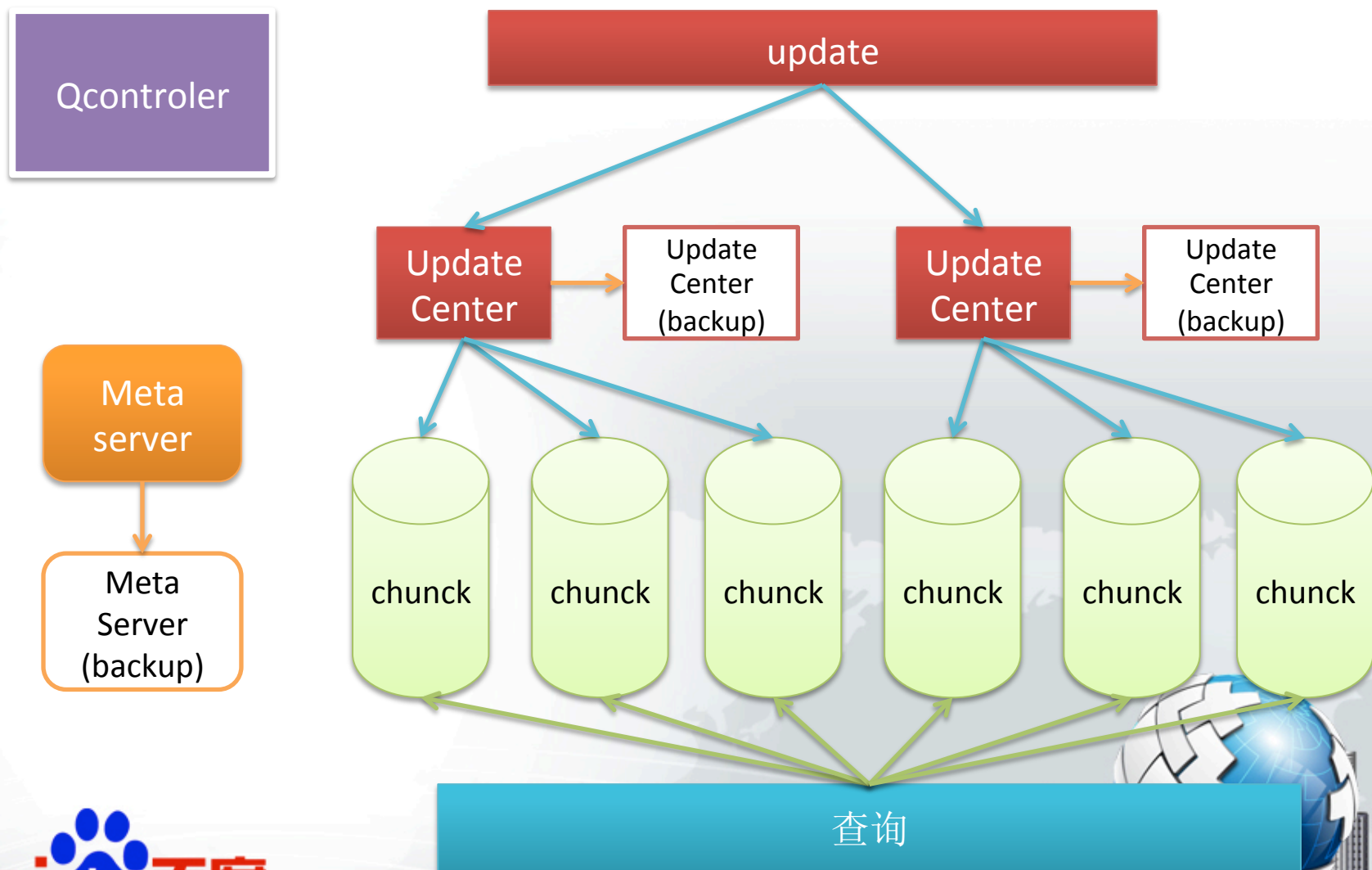
## ✓ 统一存储体系

- 平衡大容量、高并发、低延迟
- 不同访问模式通过组合满足

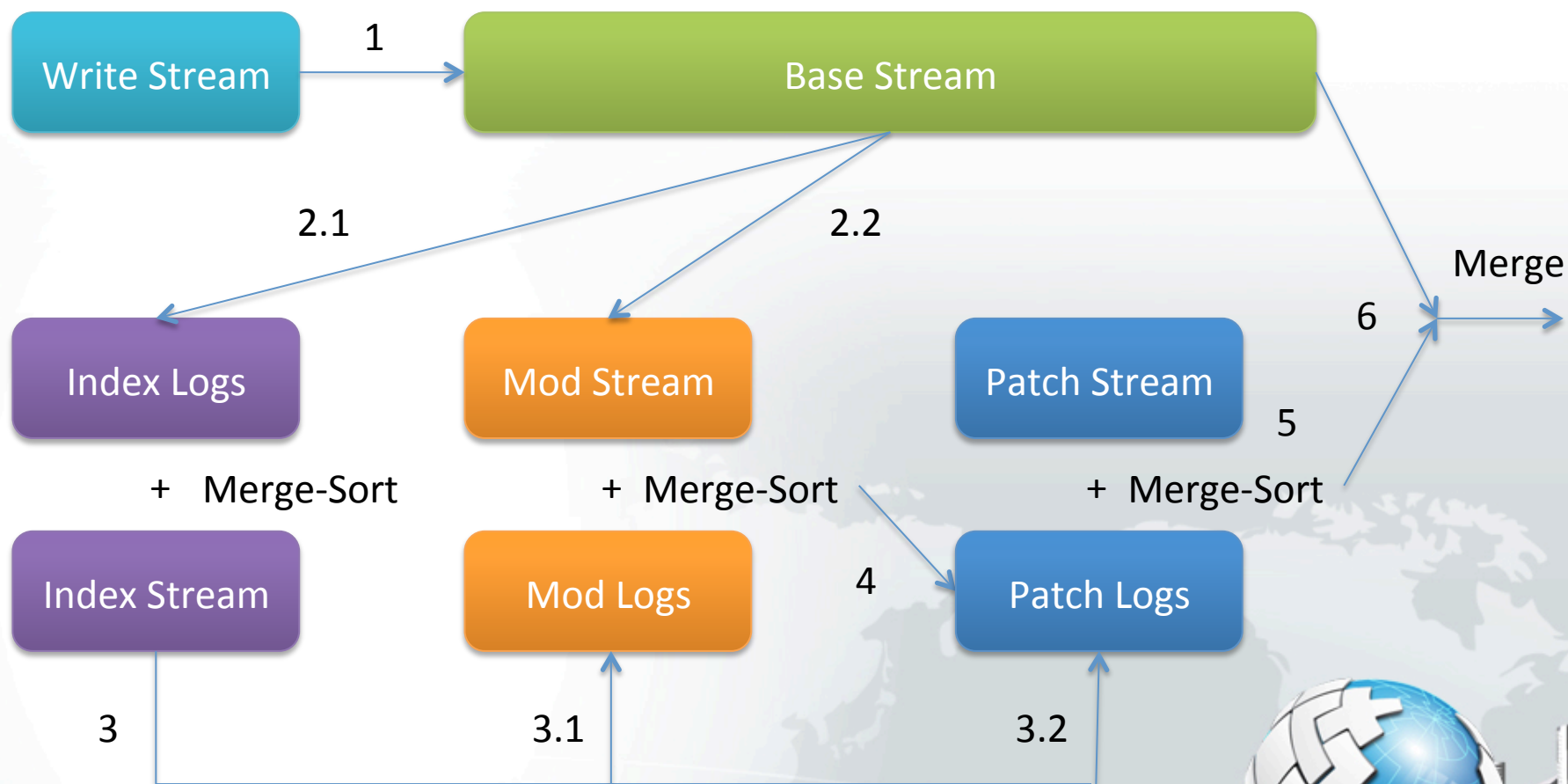
## ✓ 统一访问与传输



# 典型实时存储系统：Mola



# 典型的离线存储系统：Bailing



# 百度开放云预览



# 私有云到开放云

## 云化机制

- 私有云到公有云的转化机制
- 让云接口更易用
- 使更贴近云应用开发

## 云安全保障

- 统一的权限管理
- 虚拟化独立隔离环境





# 开放云逐步开放

- <http://yun.baidu.com/>



**成功验证!!!**  
**支持PV亿级别的应用**

**开放云平台**  
beta  
在百度的基础架构上运行您的产品，更加高效、安全、可靠  
[进入管理后台](#)

**产品与服务**

- 云存储**  
云存储为用户提供大容量、高可靠、高可用的网络存储服务。  
[查看详情](#)
- 云环境**  
云环境为所有基于开放云平台的所有云服务开发的云应用提供执行环境，同时也帮助开发者构建和维护云应用。  
[查看详情](#)
- 关系服务**  
关系服务为开放云平台提供了大规模、高可用关系型数据存储的支持。  
[查看详情](#)
- 消息服务**  
消息服务提供了消息在云端传输以及从云端到终端传输的一揽子解决方案。  
[查看详情](#)

**实例**

- ting!**  
ting! 是百度旗下音乐平台，共有广场、电台、排行榜、我的ting!四大栏目。由于数据量庞大，用户访问量巨大，普通的存储系统难以支撑，云存储为ting!提供了大容量、高可靠、高可用性的数据存储服务，确保ting!安全可靠地运行。
- hao123**  
hao123网址之家——最专业的上网导航，提供最简单便捷的网上导航服务，是数千万网民的上网主页。Hao123网址之家是搭建在云环境之上的应用，利用云环境良好的伸缩性，无需考虑集群扩容、负载均衡等问题，即能承载日益增长的访问量。
- Baidu 应用**  
百度应用是一个面向用户的应用分享平台，其应用资源主要由第三方开发者或提供商通过百度应用开放平台提供。百度应用为了支持第三方开发者，向其提供了云环境。第三方开发者不用考虑部署等细节，即能实现应用的快速开发上线。
- Baidu 团购**  
百度团购是国内最好的团购导航，全国团购网站大全。汇聚全国多家大型知名团购网站如拉手团购、美团、糯米团购、团宝、点评团等。百度团购是搭建在云环境之上，利用云环境提供多种服务，实现了快速开发和上线。





# 开放云的公益性质

## 开放程度

- 已经开放云存储
- 预计Q2能基本功能全开放

## 诚意

- 对于百万PV的网站，免费额度基本够用
- 资源不够，协商后可以免费继续提高，无规定上限
- 目前提供1T的分布式存储（多副本、多机房、CDN支持）

## 开放资质

- 面向有经验的中小开发团队

## 开放流程

- 查看：<http://dev.baidu.com/wiki/yun>
- 发邮件到 [yun@baidu.com](mailto:yun@baidu.com) 详细介绍自己



# 云环境

网络 & 动态负载均衡

云接入与管理

Network & 调度

云环境

调度器

业务开发流程

镜像服务

CPU & Memory

资源池( 大规模计算结点)

数据  
分析

消息  
服务

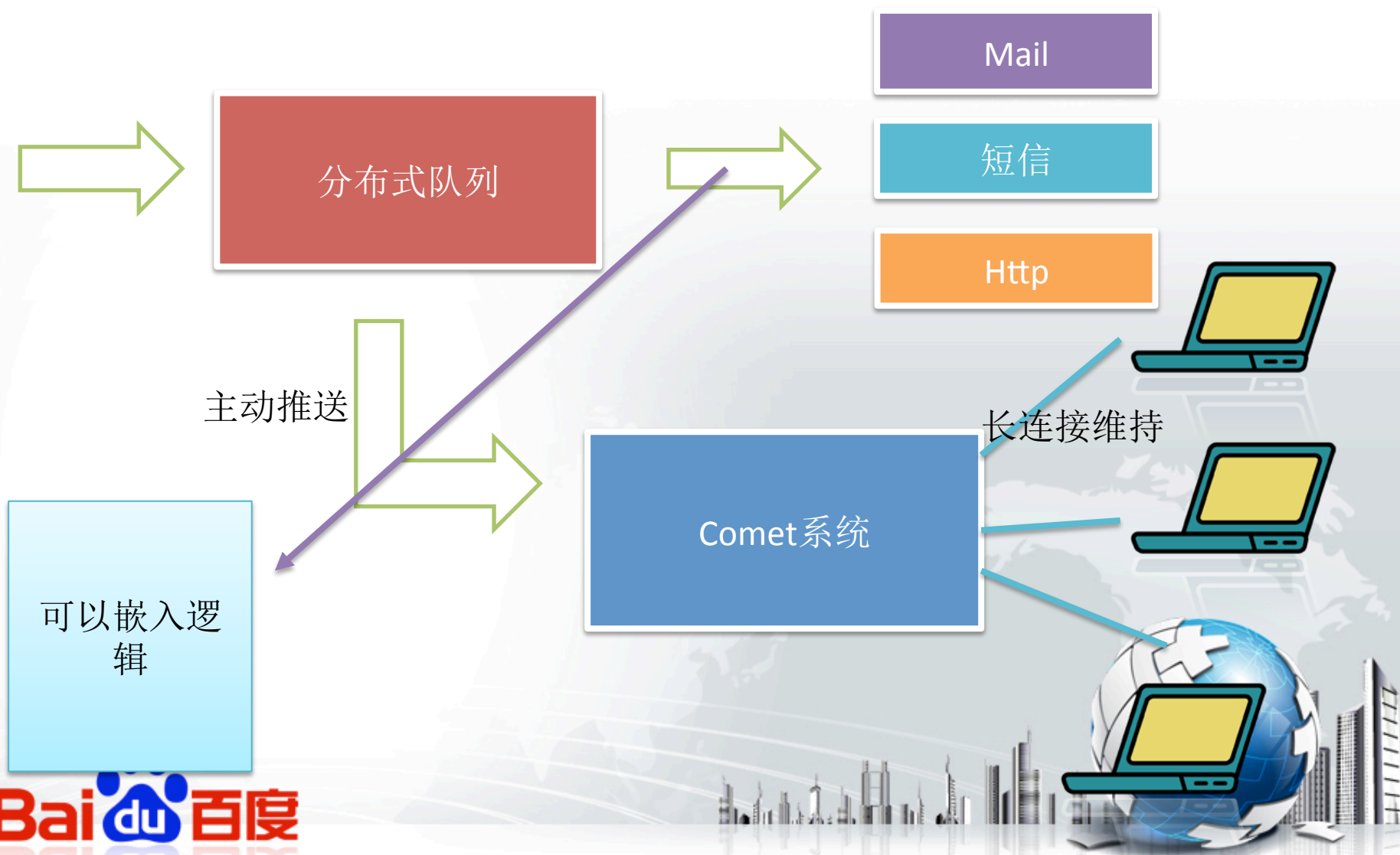
IO

分布式存储

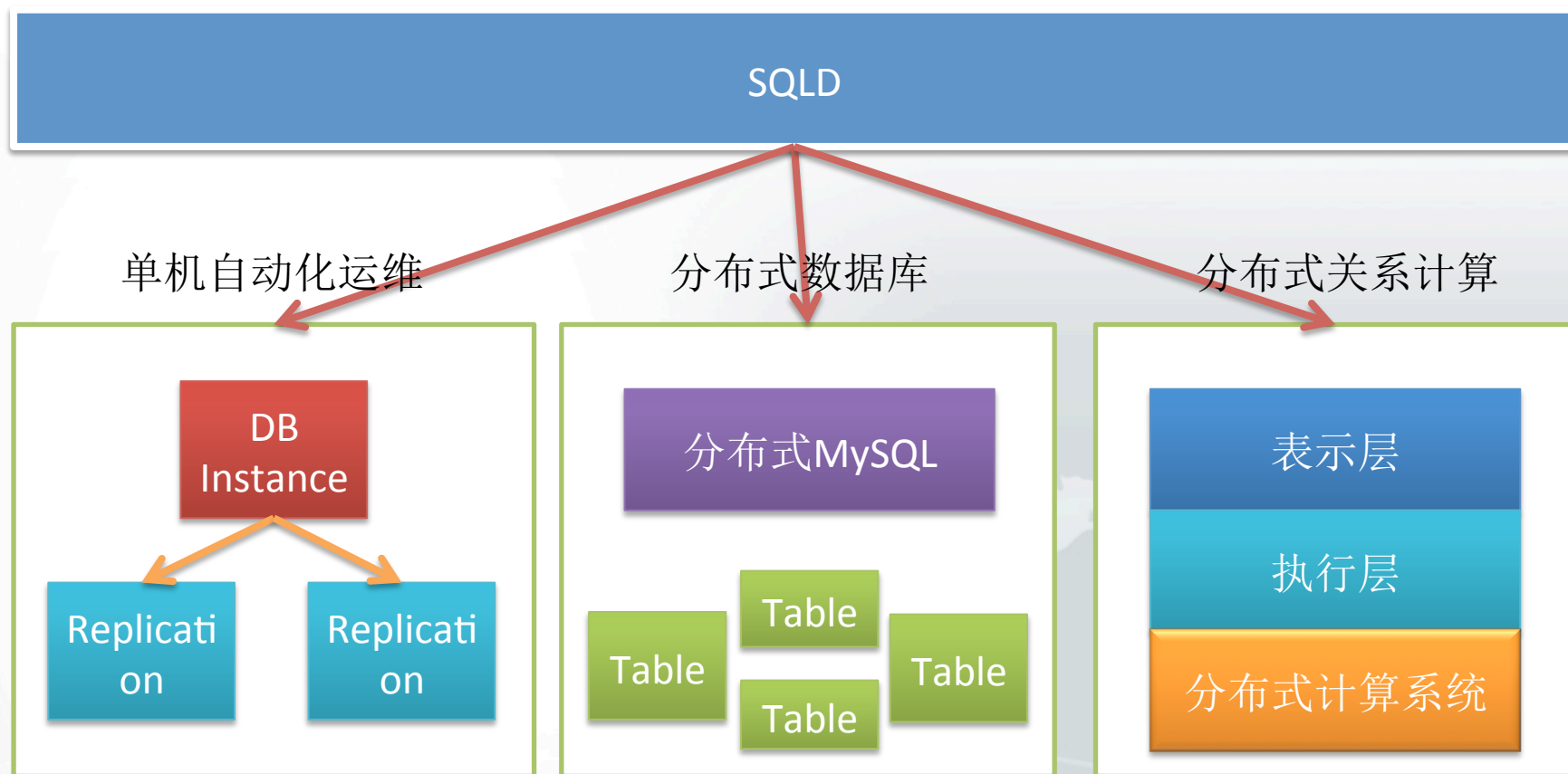
分布式计算

本地存储

# 消息服务

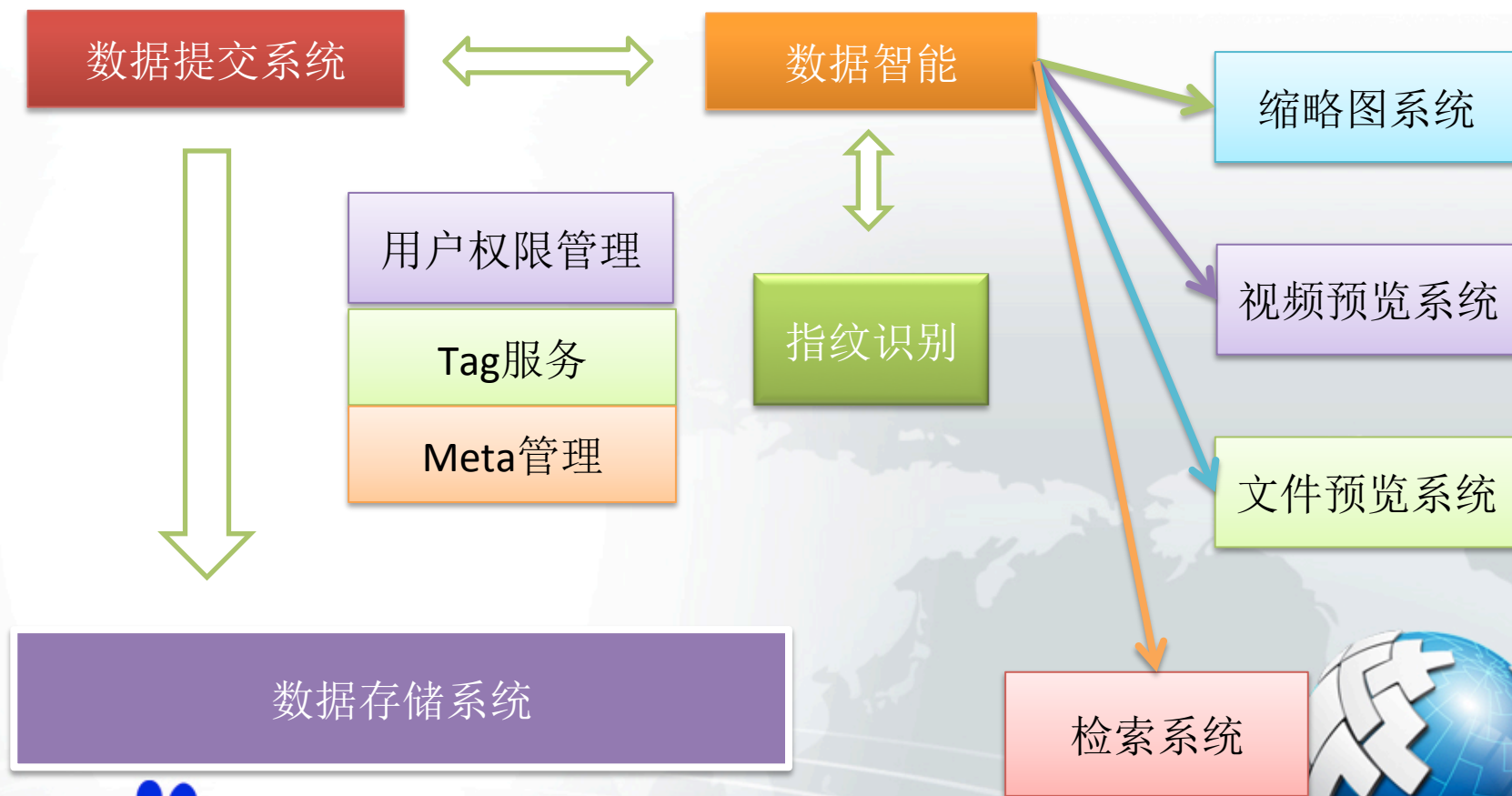


# 关系计算服务

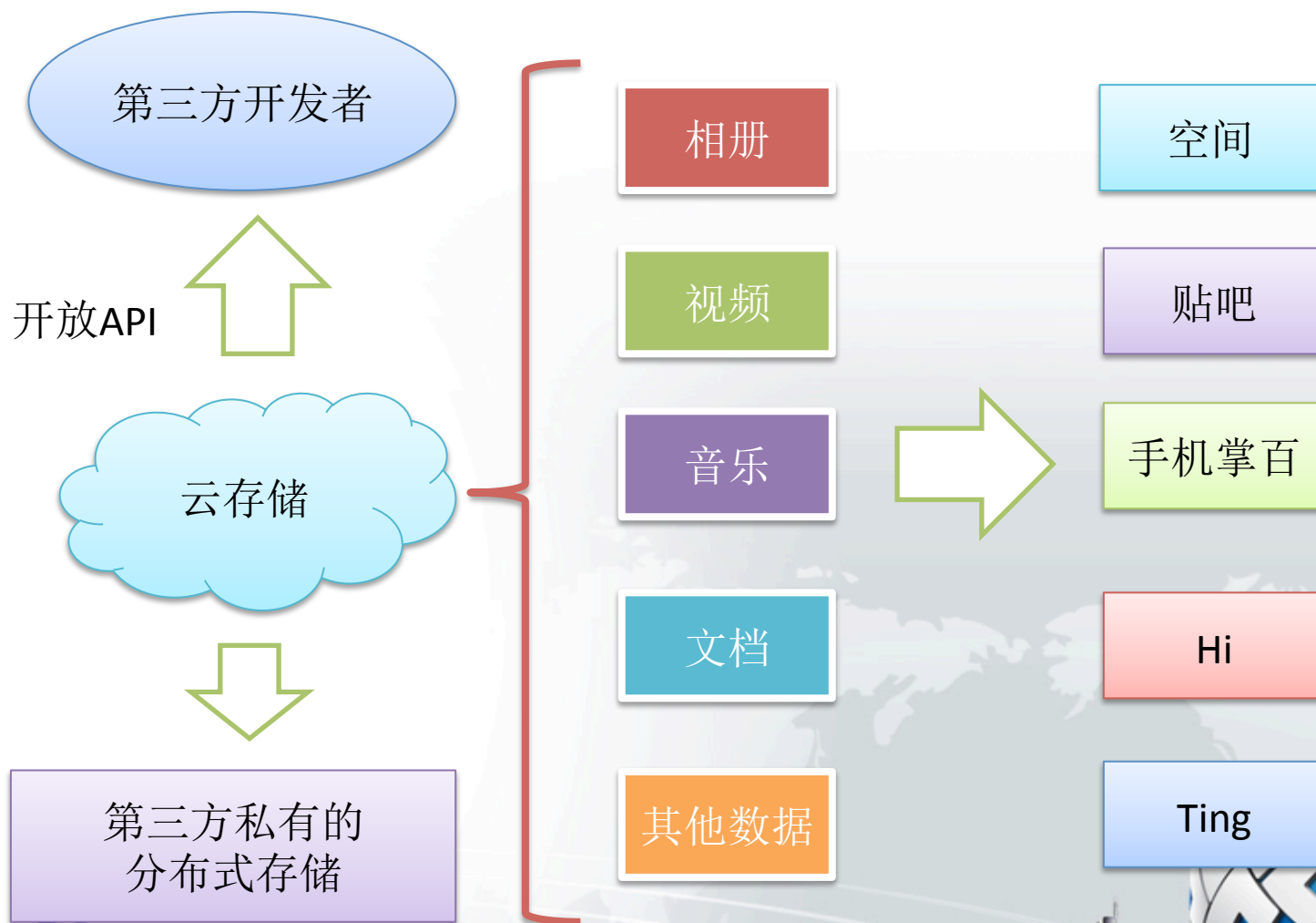


# 云存储

个人为中心数据存储系统



# 云存储



# 百度获发改委 项**最高激励**

国家

- 百度持续投入得到领域专家和政府部门高度认可和持续鼓励
- 百度云计算获三部委最高激励

百度

- 将大力发展智能数据服务、典型行业应用和公共云计算平台





# 联系方式

- 开发者运营平台: [yun@baidu.com](mailto:yun@baidu.com)
- 招聘渠道: [yun@baidu.com](mailto:yun@baidu.com)      Title [应聘]





北京站 · 2012年4月18~20日  
[www.qconbeijing.com](http://www.qconbeijing.com) (11月启动)

QCon杭州站官网和资料  
[www.qconhangzhou.com](http://www.qconhangzhou.com)

全球企业开发大会

INTERNATIONAL  
SOFTWARE DEVELOPMENT  
CONFERENCE