推动云计算应用：Hadoop 开源平台

郑皓
首席架构师
雅虎全球软件研究和开发中心 北京

Agenda

- 雅虎Yahoo!公司的**云计算**战略
- Hadoop 开源社区和用户团体
- Hadoop 和 Yahoo!
- Hadoop 在 Yahoo! 的应用

为什么云计算对Yahoo!至关重要

我们的使命：成为亿万网民在线活动的中心

HUNDREDS

资产 / 平台

600M

独立用户 / 月

300M+

MAIL 用户 / 月

HUNDREDS

PB量级存储

BILLIONS

对象存储

PETABYTES

每日流量

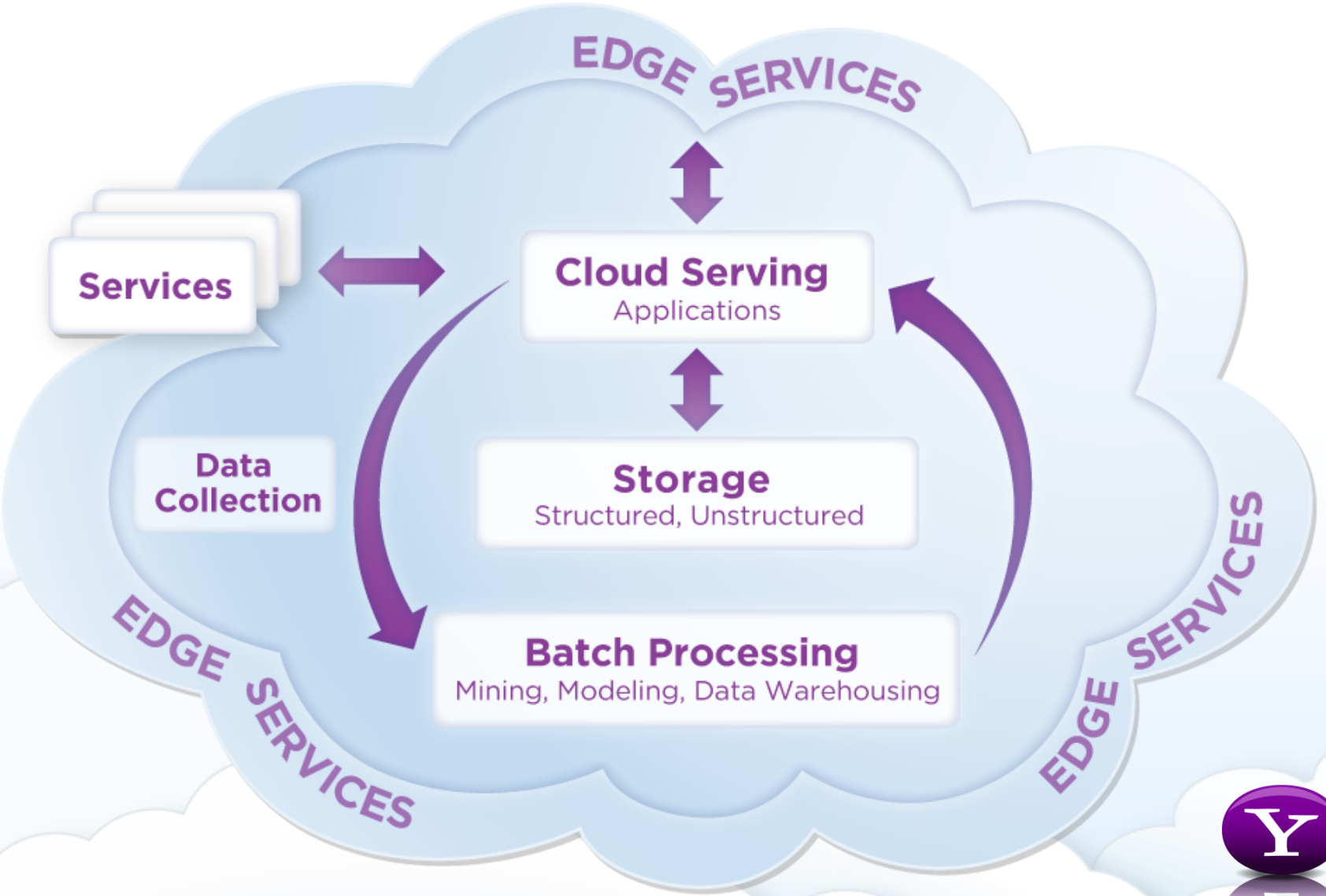
Yahoo!如何使用云计算平台

- 依赖云计算来提升用户和广告主的体验:
  1) 处理和分析海量数据
  2) 在全球范围内加速内容投放
  3) 灵活存储
- 大规模促进创新和科学研究
- 和开源社区合作共享我们的云计算平台

Yahoo!的云战略

- 创建Yahoo!自己的私有云
- 优化Yahoo!全球资产
- 数据处理和服务环境
- 核心收益：驱动创新
- 成熟之后将核心技术开源
- 付出了多年努力

Yahoo!云内部架构


EDGE SERVICES
Services
Cloud Serving
Applications
Data
Collection
Storage
Structured, Unstructured
Batch Processing
Mining, Modeling, Data Warehousing
EDGE SERVICES
EDGE SERVICES
EDGE SERVICES

业界挑战：日积月累的海量数据

1. 大量优质数据积累在相对少数人身上

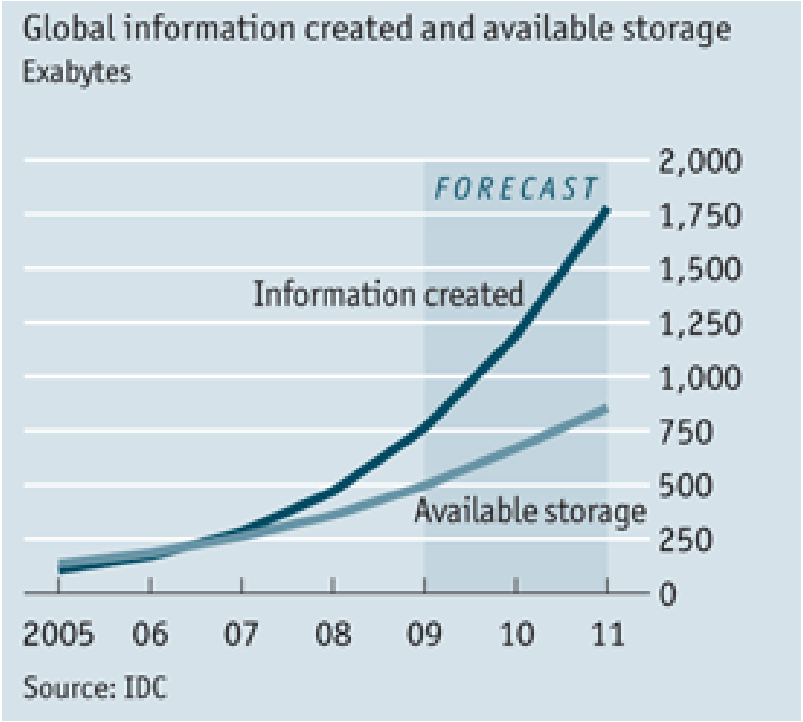
2. 从大量数据中挖掘有用信息依然是一个挑战

3. 数据成为竞争力的分水岭：
   搜索日志，广告点计击数据，浏览历史，社交图，等等

Yahoo!的技术解决方案：Hadoop

“现在是数据的工业革命。” – Joe Hellerstein, 加州伯可莱大学

“到2020年，全球数据规模会比2009年扩大44倍。” [IDC]



Hadoop是什么?

- 一个分布式文件系统和并行执行环境
- 让用户便捷地处理海量数据
- Apache软件基金会下面的一个开源项目
- 目前Yahoo!是最主要的贡献者

快速增长的用户群

2007

YAHOO! Powerset last·fm

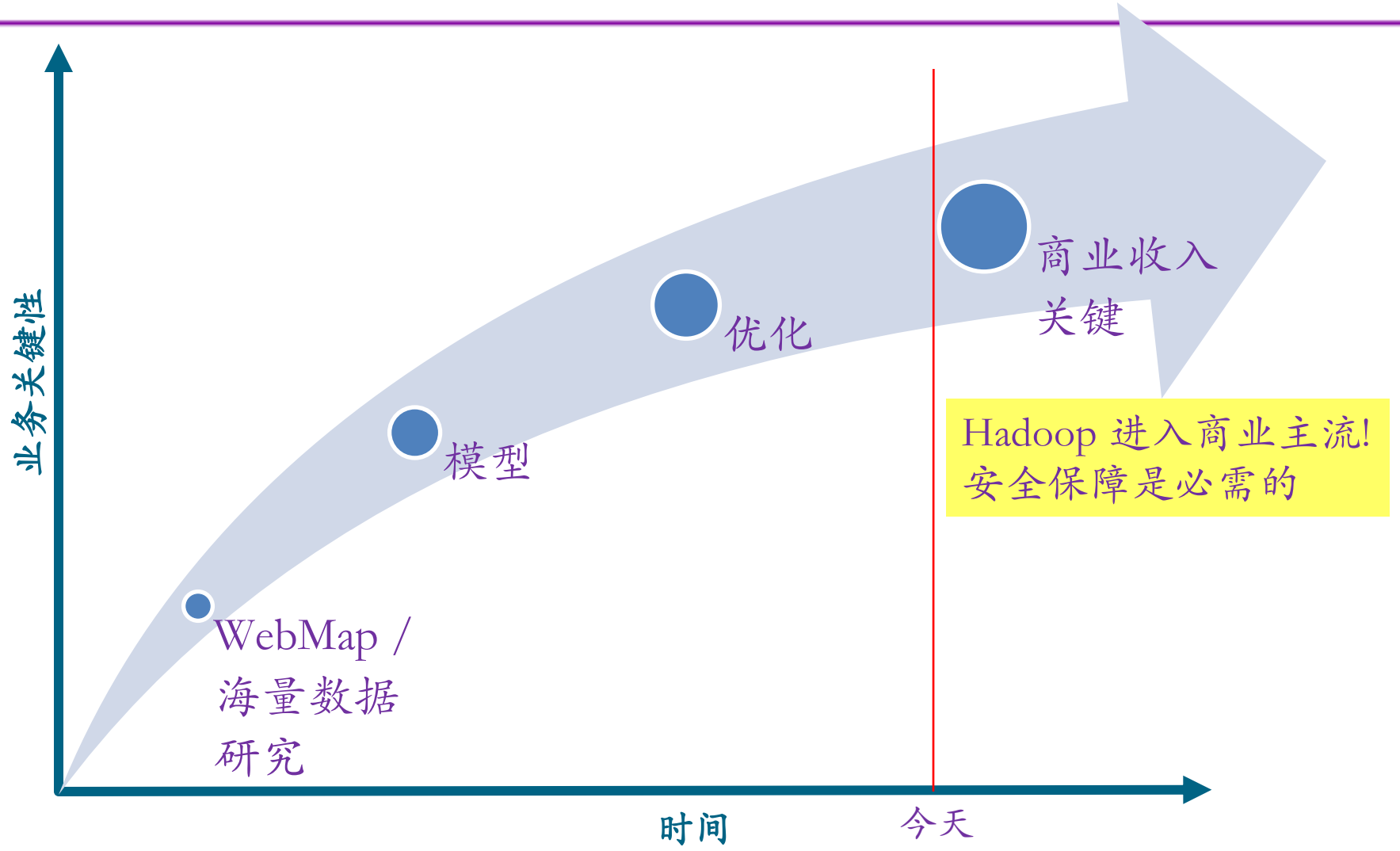
2008

Google ableGrape Cascading INFORMATION SCIENCES INSTITUTE Zvents Discover Things To Do Cornell University Computing and Information Science Visible MEASURES
ImageShack online media hosting A9 IBM ENORMO THE UNIVERSITY OF EDINBURGH Joost News Corporation
Lookery facebook krugle The New York Times LOTAME Locate, Target, & Message with Social Media veoh
SECURITY ENHANCED DOMAIN NAME SYSTEM NetSeer parc Palo Alto Research Center rackspace HOSTING

2009 - 2010

AOL WorldLingo cloudera cooliris deepdyve BETA TEXTMAP THE ENTITY SEARCH ENGINE Monitoring the World So PSG College of Technology
eyealike iterend tailsweep adknowledge hulu RapLeaf Ning quantcast USCMS
pressflip HOSTING HABITAT detikSearch stampede beta VK SOLUTIONS Global Solutions Provider
Microsoft Terrier CbIR HOLA SERVERS Systems@ETH Zürich NETFLIX
amazon web services twitter LinkedIn ebay TARAGANA Innovation + Quality + Simplicity

Hadoop逐渐成为主流云计算平台



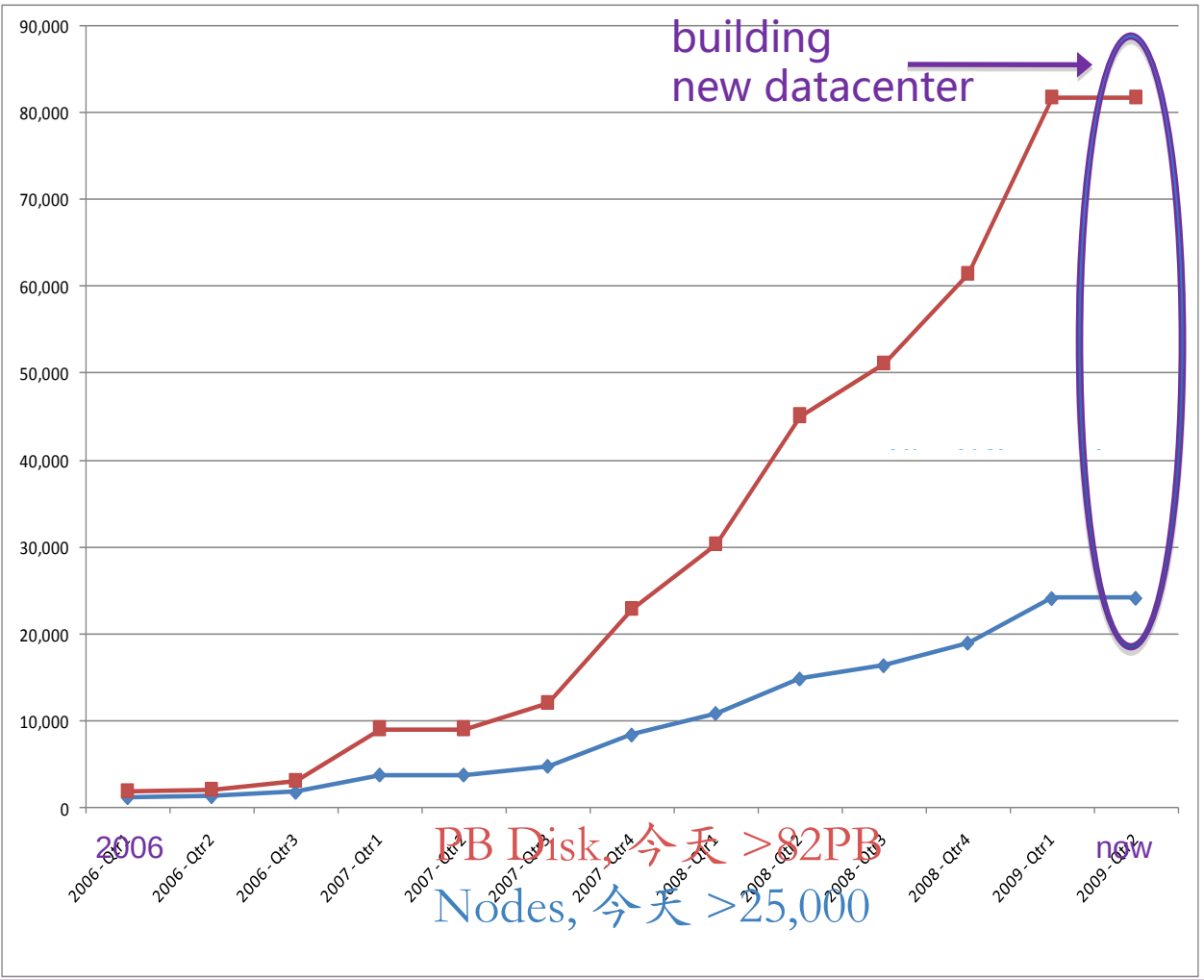
Yahoo!和Hadoop密切关系

- Yahoo!是:
  - Hadoop最大的用户
  - Hadoop最大的测试者
  - Hadoop最大的贡献者

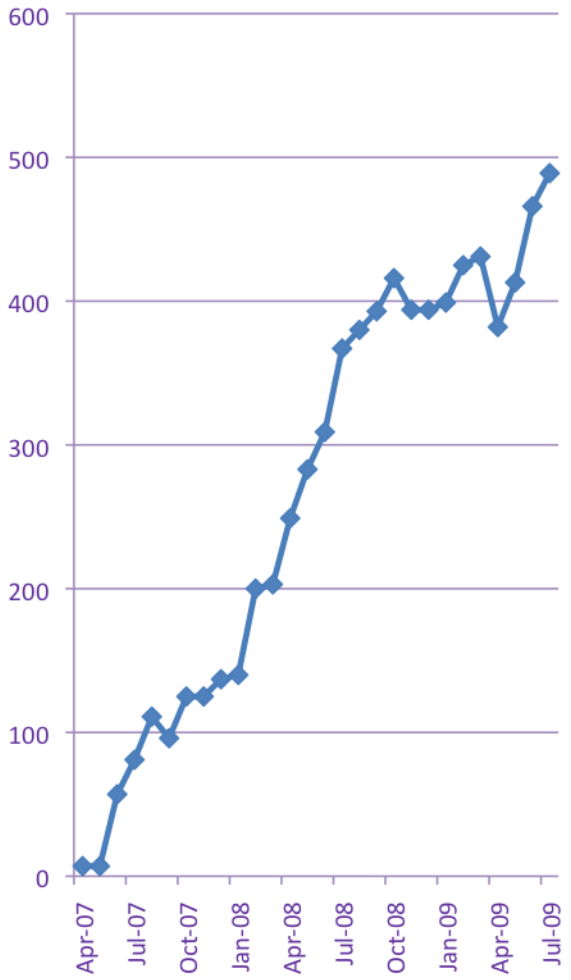
- 还有:
  - 我们发布了“*Yahoo! Hadoop 公开版*”
  - 我们贡献了所有在Hadoop上的工作给Apache软件基金会
  - 我们持续积极地投入到Hadoop开发中
  - 我们不以Hadoop的服务和支持盈利!
    - 我们使用Hadoop来驱动整个Yahoo!

Hadoop最大的用户

硬件规模



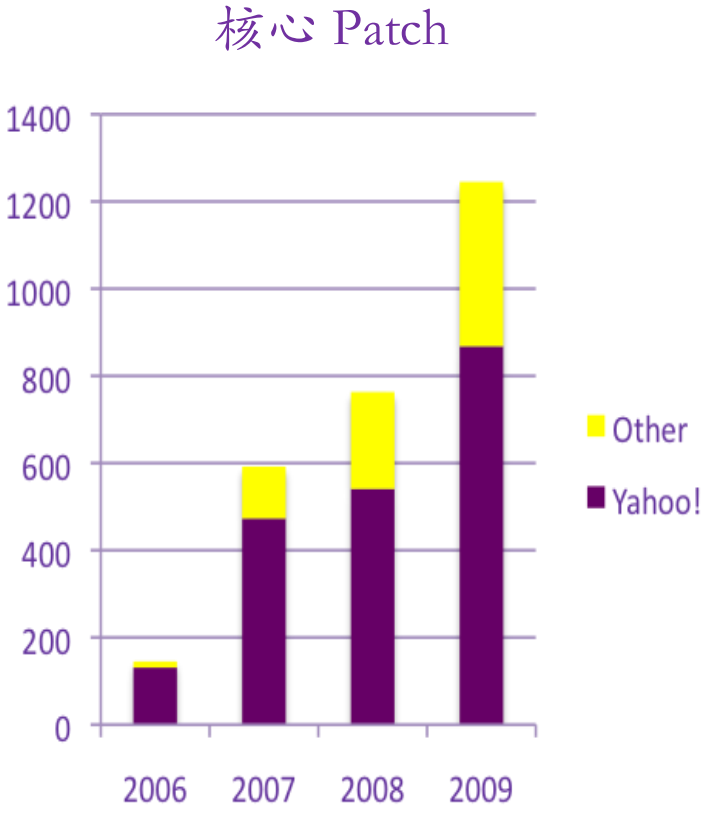
内部Hadoop用户数



Hadoop最大的测试者

- *Yahoo! Hadoop公开版* 的每一次发布都经过了很多级别的严格测试
- Hadoop集群的4个层级:
  - 开发，测试，和质检 *(~10%的硬件)*
    - 持续集成 / 测试最新代码
  - 概念证明和特设工作 *(~10%的硬件)*
    - 运行最新版本
  - 科学与研究 *(~60%的硬件)*
    - 运行比较稳定的版本
  - 产品 *(~20%的硬件)*
    - 运行最稳定的版本

Hadoop最大的贡献者



- 主要的patch都来自Yahoo!
  - 72% 的核心 patch
    - Core = HDFS, Map-reduce, Common
  - 存在低估现象
    - 一些雅虎员工使用apache.org的账户
    - 全职contributors提交大的patch
- 截止目前，Yahoo!是Hadoop贡献者的最大雇主
  - 我们将在Hadoop上做的所有工作回馈给Apache社区
  - 团队: 美国, 班加罗尔, 北京
  - http://careers.yahoo.com

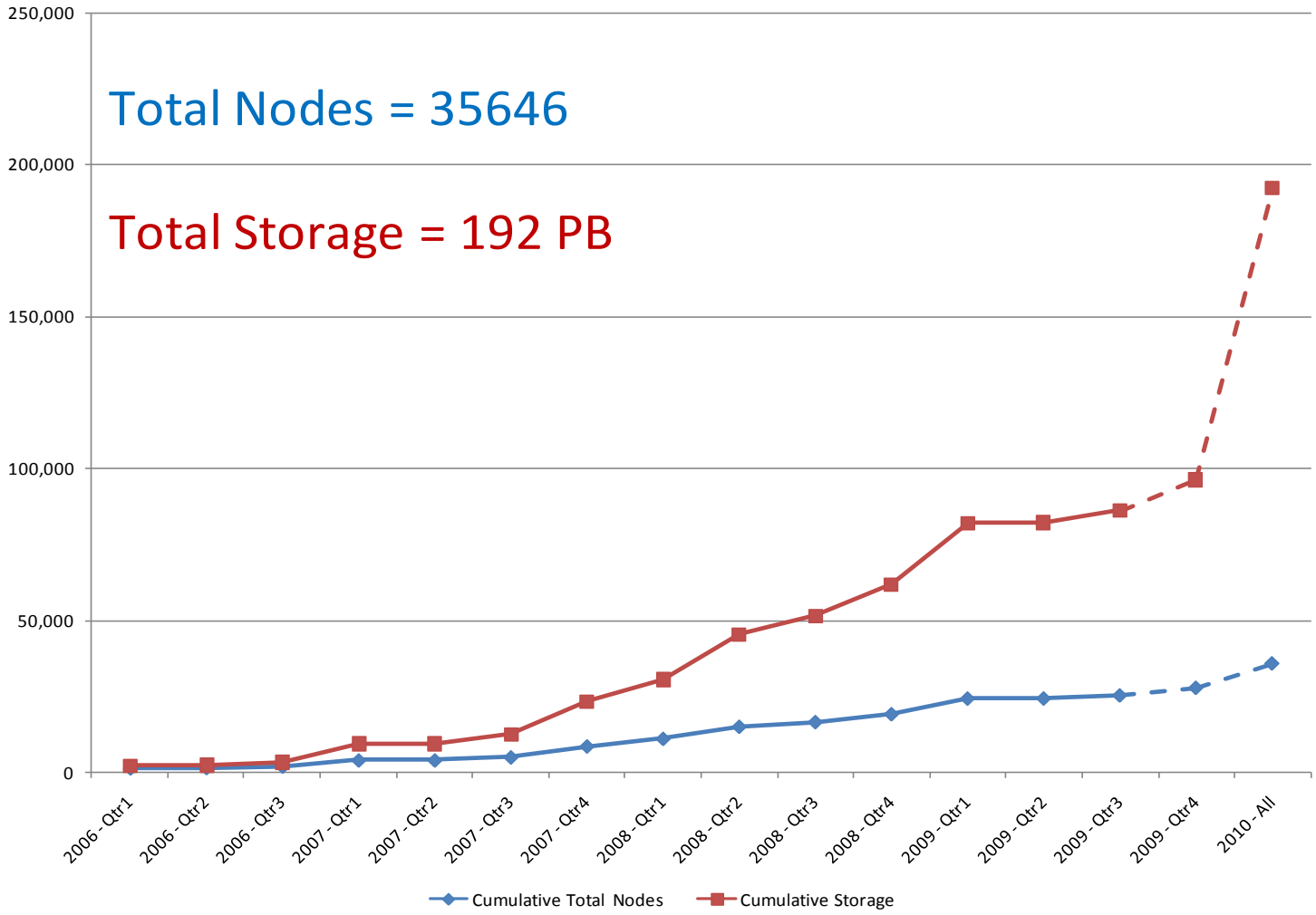
缘何 Yahoo! 选择 Hadoop

- 数据规模
  - 每月超过6亿的独立用户
  - 每天产生数十亿的transaction
  - PB级别的数据
- 分析和处理数据非常关键
  - 需要及时处理所有数据
  - 大量调研以寻找更好的模式，分析数据以生成各种报表
- 更低的成本需求
  - 使用低成本的通用硬件
  - 多项目之间共享资源
  - 在大规模集群上快速完成新的实验
  - 每天需要处理许多硬件故障
- Hadoop这一基础架构可以满足这些需求

Yahoo!使用Hadoop时间表

2004	2006	2007	2008	2009	2010
• HDFS 和 MapReduce 初版	• 启动 Apache Hadoop 项目 • 研究集群规模达到600个节点	• 两个超过1000节点的研究集群	• 每天载入数据量达到10 TB	• 12个集群，总规模超过25,000个节点 • 公布 *Yahoo! Hadoop*公开版	• 公布 *Yahoo! Hadoop Security*公开版

Yahoo!使用Hadoop趋势图



Yahoo! Front Page – 案例分析

Yahoo! Front Page – 案例分析

Yahoo! Front Page – 案例分析

Search Assist™ – 案例分析

- 使用Hadoop构建数据库
- 3年的日志数据
- 20步的map-reduce**计算**

	Hadoop 前	Hadoop 后
耗时	26 天	20 分钟
语言	C++	Python
开始时间	2-3 星期	2-3 天

OpenCirrus™ 筋斗云

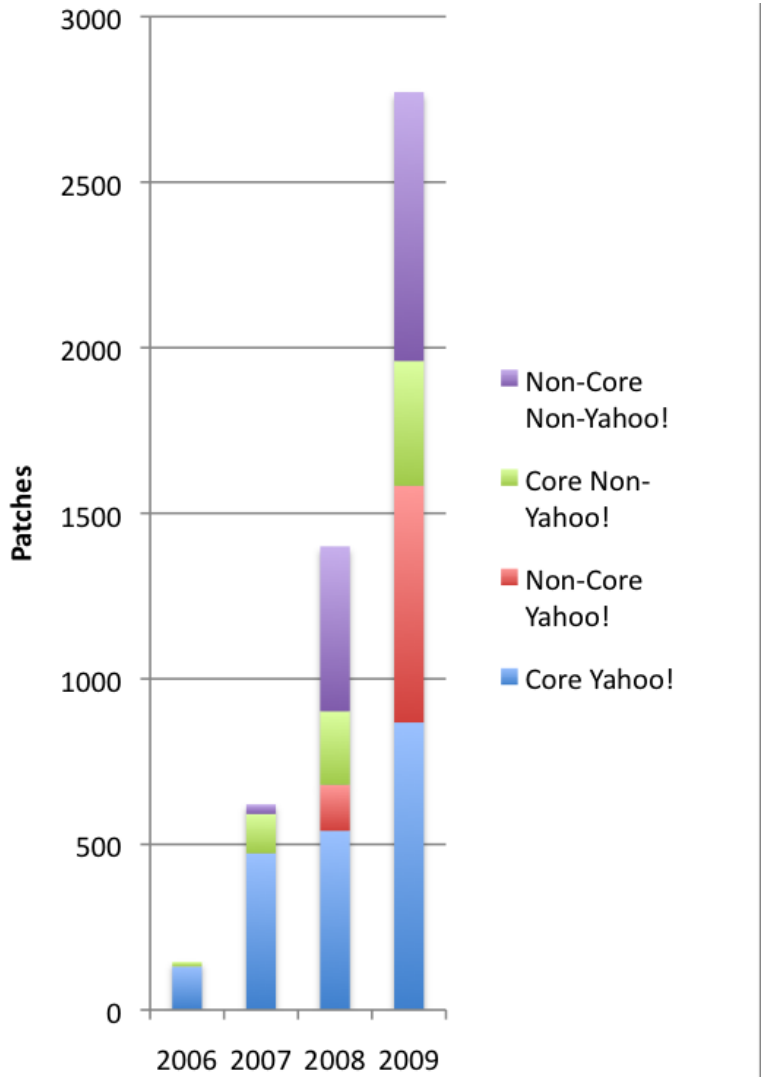
- OpenCirrus™筋斗云**互助式云计算试验平台** 成立于2009六月
- Yahoo! , HP,和Intel 联合发起
- 模拟一个现实的、全球化的互联网规模的环境
- 消除财务和后勤上的障碍
- 专注于数据密集型互联网计算的研究
- 给予研究人员前所未有的能力，来测试应用，测量基础设施以及构建于其上的服务的性能
- 促进产业界、学术界和政府部门的开放合作

问题?

郑皓
首席架构师
雅虎全球软件研究和开发中心 北京

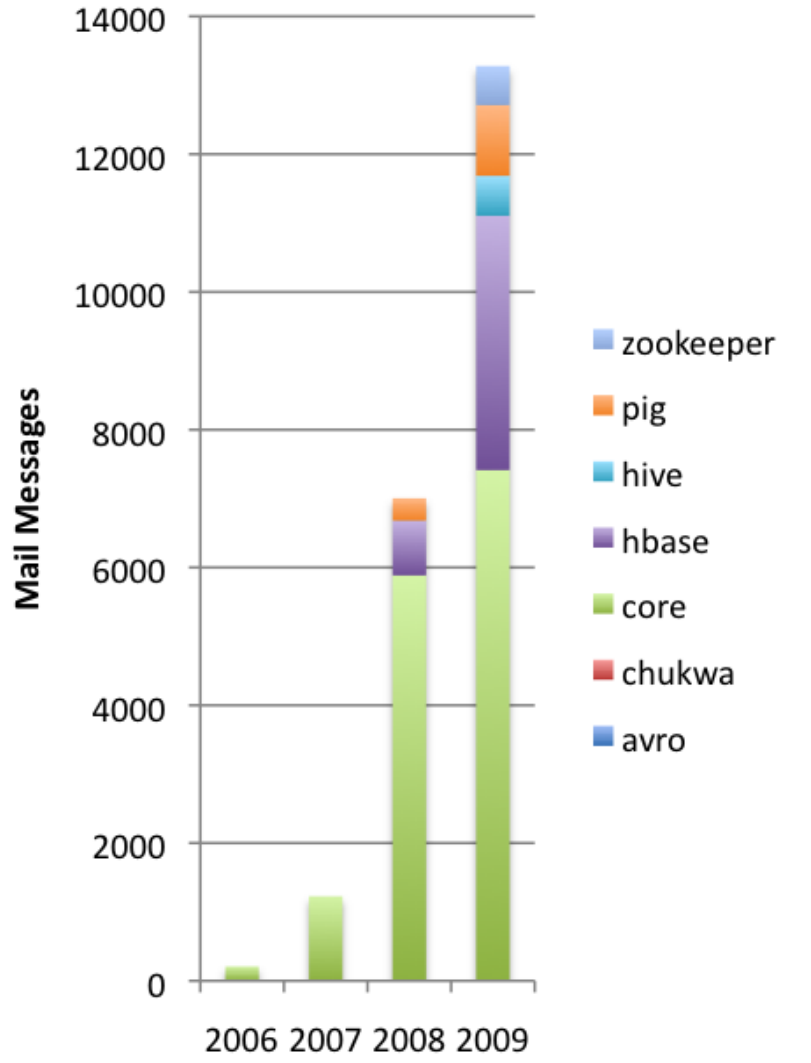
信息咨询:
http://hadoop.apache.org/
http://hadoop.yahoo.com/
http://beijing.yahoo.com/

Patch贡献年表



- 分为两个Sub-project
  - 核心 (HDFS & Map-Reduce)
  - 其他
- 所有分类都在增长
- 在2009 非核心>核心!
- 核心贡献者
  - 185 人 (30% 来自雅虎)
  - 72%的patch来自雅虎

子项目的用户列表



- 用户邮件量
  - 使用好的指标
  - 社团直接的衡量
  - 6个月旧的数据…
- 社团快速发展!
  - 更多的用户
  - 更多的子项目!
- 子项目exploding
  - 给母项目带来价值
  - 发展生态系统

大量应用

	2008	2009
Webmap	~70 小时用时 ~300 TB shuffling ~200 TB 输出	~73 小时用时 ~490 TB shuffling ~280 TB 输出 +55% 硬件
Sort benchmarks (Jim Gray contest)	1 Terabyte 排序 • 209 秒 • 900节点	1 Terabyte排序 • 62 秒, 1500 节点 **1 Petabyte** sorted • 16.25 小时, 3700节点
Largest cluster	2000 节点 • 6PB 硬盘 • 16TB 内存 • 16K 核	4000 节点 • 16PB 硬盘 • 64TB 内存 • 32K 核 • (快 40%!)