

拥抱开源：阿里集团的实践与经验

章文嵩（正明）

Velocity China

2012.12.4



- 一、阿里集团的简介
- 二、淘宝软件基础设施
- 三、事例：图片存储、CDN与DB
- 四、阿里开源成果
- 五、阿里开源流程
- 六、小结

淘宝网
Taobao.com

C2C

淘宝商城
mall.taobao.com

B2C

一淘
淘宝旗下网站

Search & pricing
comparison

Alibaba.com®
Global trade starts here.™

Global e-commerce platform
for small businesses

阿里巴巴
1688.com

E-commerce platform for
Chinese small businesses

聚划算
-juhuasuan.com-
品质团购每一天

Group Shopping

阿里云计算
Alibaba Cloud Computing

Cloud OS +
Service

YAHOO!
中国雅虎

China Yahoo!

支付宝™
Alipay.com

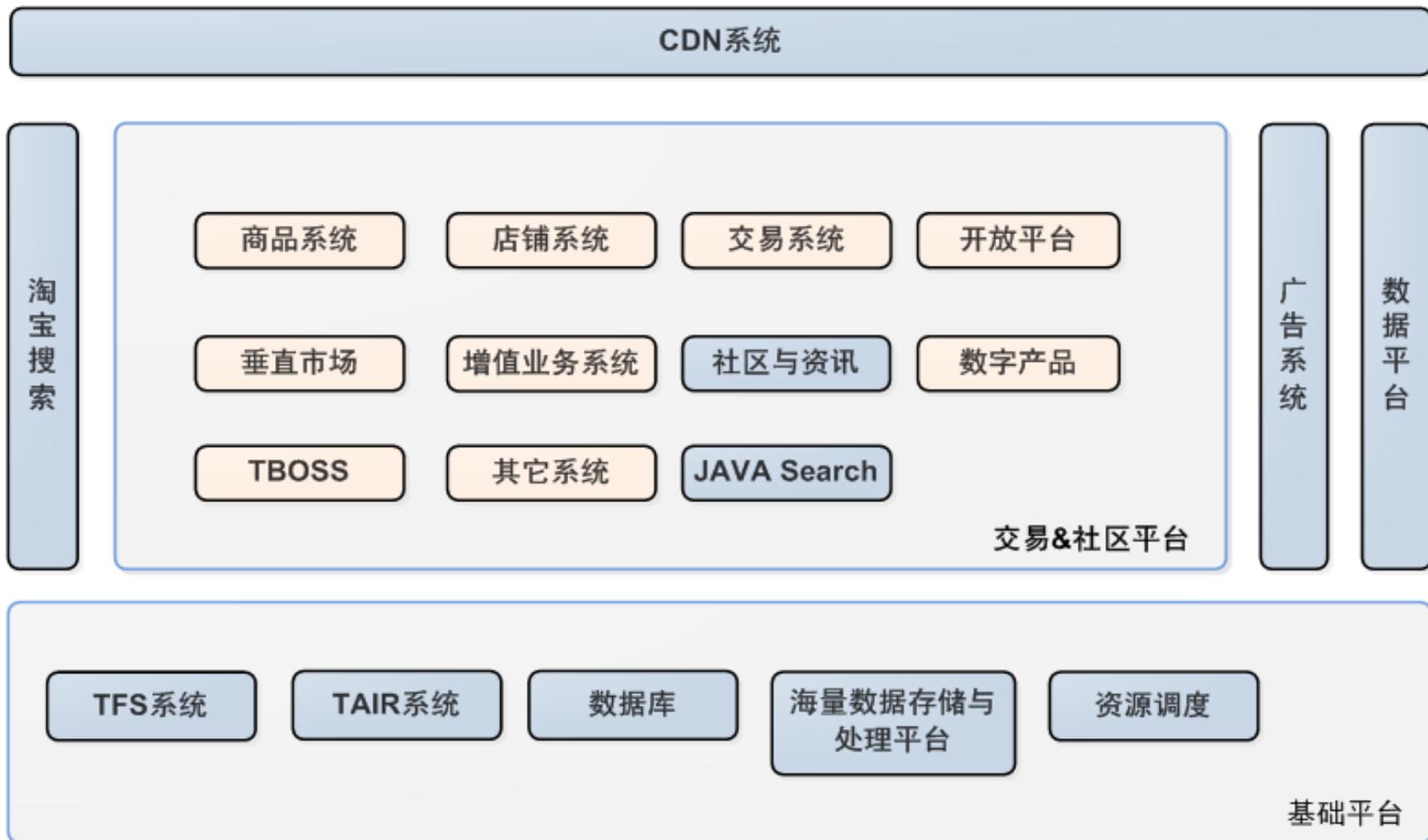
payment

- 2011年网购交易额约7849亿元，淘宝网占8成；
淘宝创造的直接就业机会是270.8万
- 网络流量排名（Alexa统计）
 - 国际：11（10~18）
 - 国内：3
- 今年双11大促活动的一些数据
 - 双11购物狂欢节支付宝总销售额191亿
 - 第1分钟超过1000万人涌入，1分钟成交19.2万笔
 - 全天有2.13亿独立访客，占中国网民总数4成
 - CDN最高峰值流量到达2100Gbps

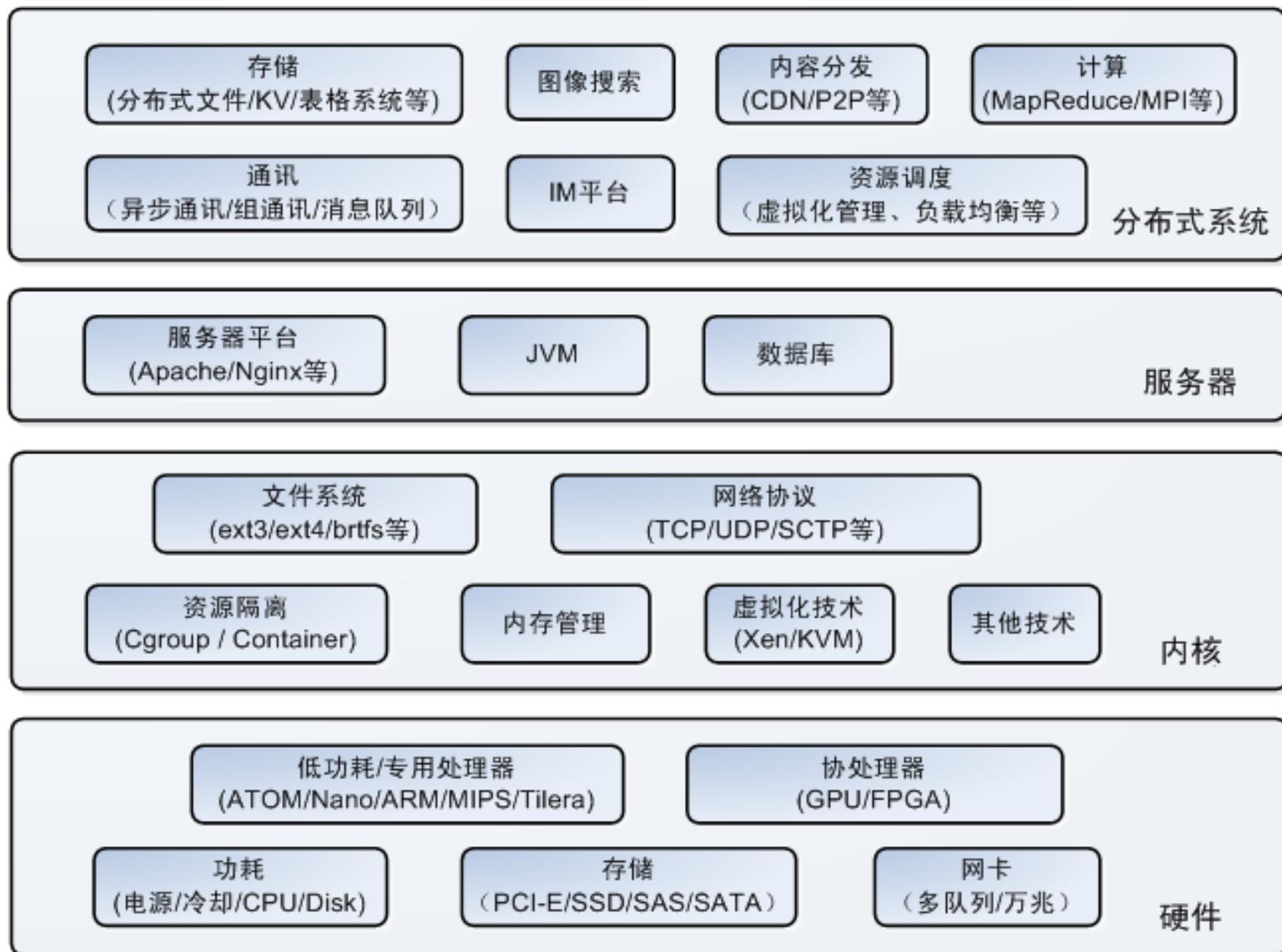


- 一、阿里集团的简介
- 二、淘宝软件基础设施
- 三、事例：图片存储、CDN与DB
- 四、阿里开源成果
- 五、阿里开源流程
- 六、小结

淘宝系统框架示意图



基础平台的规划



- CDN：世界上流量最大的、面向图片的CDN系统
 - 基于开源软件LVS+Haproxy+Squid/TS+Bind上开发的CDN系统
 - 现有140个节点，能承载2400Gbps流量的能力
- TFS：自主开发的分布式对象存储系统
 - 可存储容量13.36P，实际使用10.28P容量
 - 图片空间每GB每年存储与运维成本从7.2元降到3.4元
- TAIR：淘宝的分布式缓存和K/V存储
 - 集成了开源的Redis和LevelDB存储引擎
 - 提供跨机房容灾的解决方案
- OceanBase：淘宝的分布式数据库系统
 - 支持千亿条记录级别的数据库、支持事务

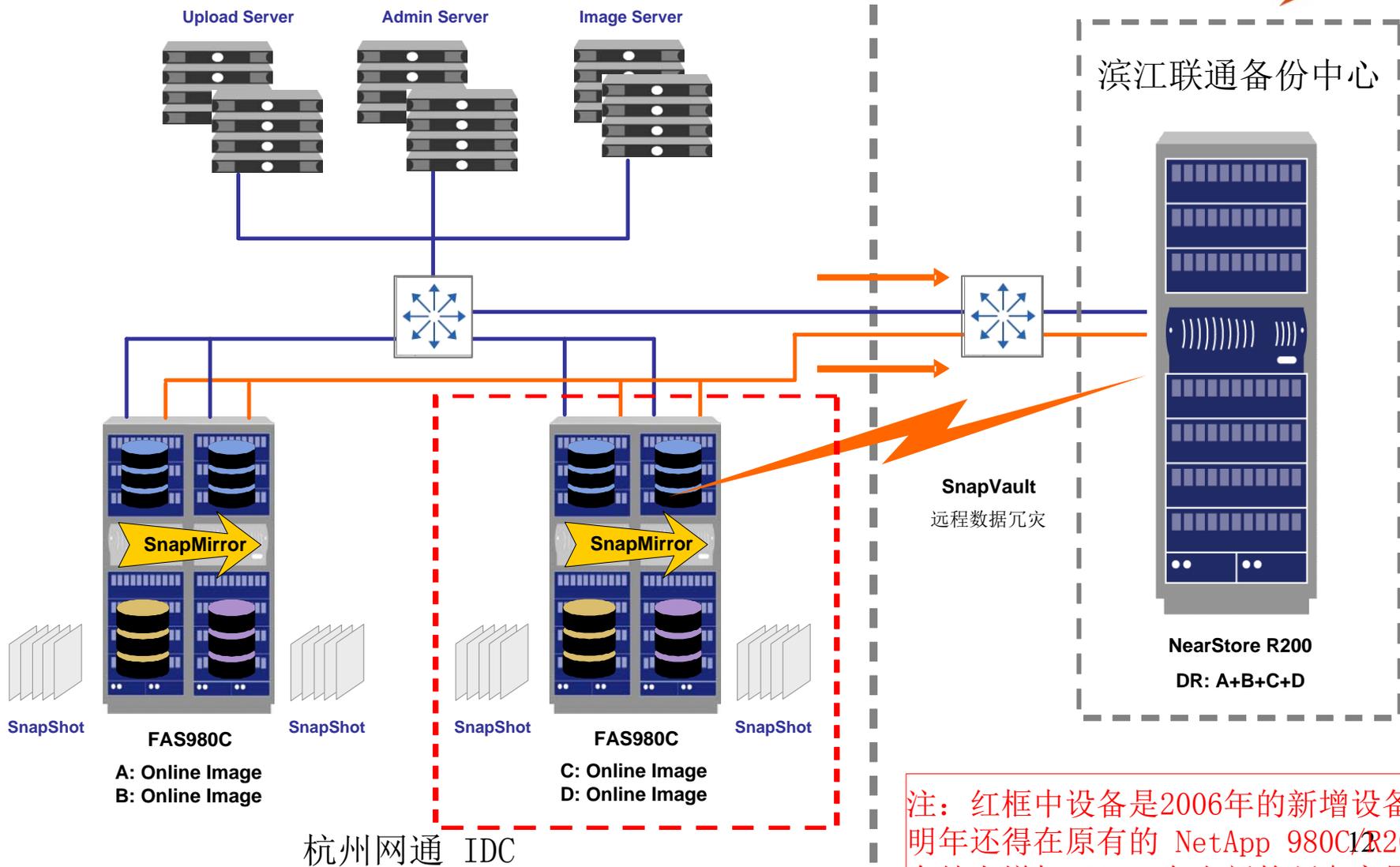
- 海量数据：采用开源的Hadoop平台
 - 现在单一集群到3200台服务器的规模
 - 系统可存储容量为76PB，已使用52.5PB，历史数据为压缩存储
 - 现在每天存储净增量约为259TB
 - 每天运行的作业数约为15~18万
- 核心数据库：采用开源的MySQL，加上高速的非易失存储，以及多层级的系统优化
- 旺旺平台：自主开发；最高在线1231万，全年可用99.99%，机器数约1800台，服务了淘宝和B2B
- 服务器平台：Nginx 部署130个应用，约2000台机器，占有率20%；完成TMD等重要基础模块，Tengine 项目开源

- 底层的支撑软件：
 - 在OpenJDK基础上开发和维护Taobao JVM
 - 在Red Hat基础上维护自己的Linux内核
 - 基于cgroup的轻量级弹性计算平台T4
 - 在Sheepdog上实现了KVM的虚拟化弹性计算平台
 - 在LVS基础上实现负载均衡解决方案
 - 用开源软件实现了高流量的网络镜像项目
- 可以说淘宝网平台建立在开源软件和自主开发的基础上。



- 一、阿里集团的简介
- 二、淘宝软件基础设施
- 三、事例：图片存储、CDN与DB
- 四、阿里开源成果
- 五、阿里开源流程
- 六、小结

2007年之前的图片存储系统



注：红框中设备是2006年的新增设备，明年还得在原有的 NetApp 980C/R200 存储上增加 20TB 左右新的硬盘容量。

- 2007年6月

淘宝自主开发的分布式的文件系统

TFS (Taobao File System) 1.0上线运行

主要解决海量小文件的分布式存储

集群规模: 200台PC Server(146G*6 SAS 15K Raid5)

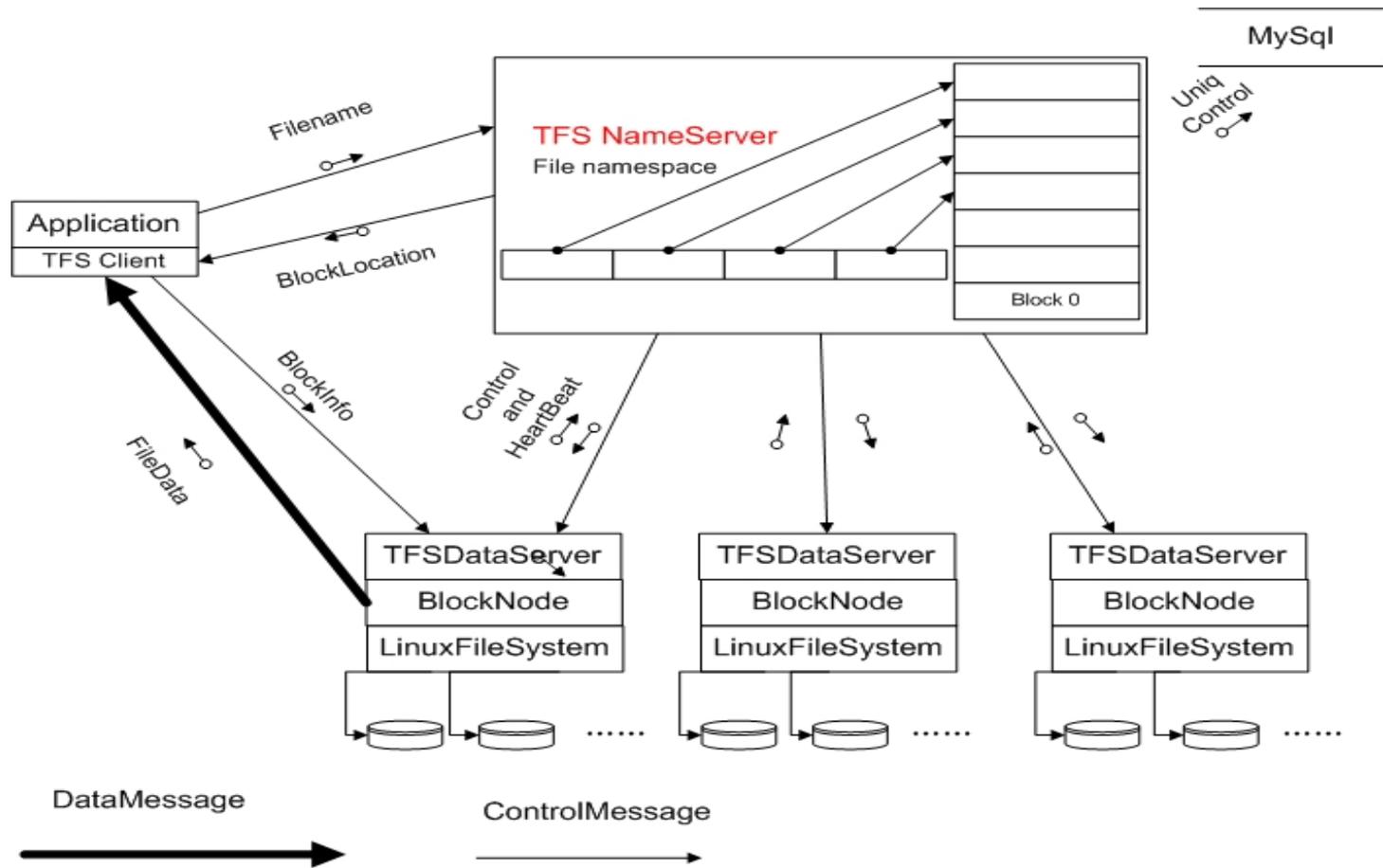
文件数量: 亿级别

系统部署存储容量: 140 TB

实际使用存储容量: 50 TB

单台支持随机IOPS 200+, 流量3MBps

TFS 1.0的逻辑结构



- 文件名内置元数据信息，用户自己保存TFS文件名与实际文件的对照关系 - 使得元数据量特别小
 - 如T2auNFXXBaXXXXXXXXXX_!!140680281.jpg，名字中含有逻辑的block_no和object_no等
- 集群由主备Name Server和多台Data Server构成
- Data Server运行在挂很多硬盘的Linux主机上
- 以block文件的形式存放数据文件(一般64M一个block)
- block存多份保证数据安全
- 利用ext3文件系统存放数据文件
- 磁盘raid5做数据冗余

- 2009年6月
TFS (Taobao File System) 1.3上线运行
- 集群规模 (2010.8.22)
 - 440台PC Server (300G*12 SAS 15K RPM) + 30台PC Server (600G*12 SAS 15K RPM)
 - 文件数量: 百亿级别
 - 系统部署存储容量: 1800 TB
 - 当前实际存储容量: 995TB
 - 单台Data Server支持随机IOPS 900+, 流量15MB+
 - 目前Name Server运行的物理内存是217MB (服务器使用千兆网卡)

- TFS1.3提供了一些重要的功能特性
 - 所有的元数据全部都内存化
 - 清理磁盘空洞
 - 容量和负载的均衡策略
 - 平滑的扩容
 - 数据安全性的冗余保证
 - 几秒内完成Name Server故障自动切换
 - 容灾策略
 - 性能大幅提升

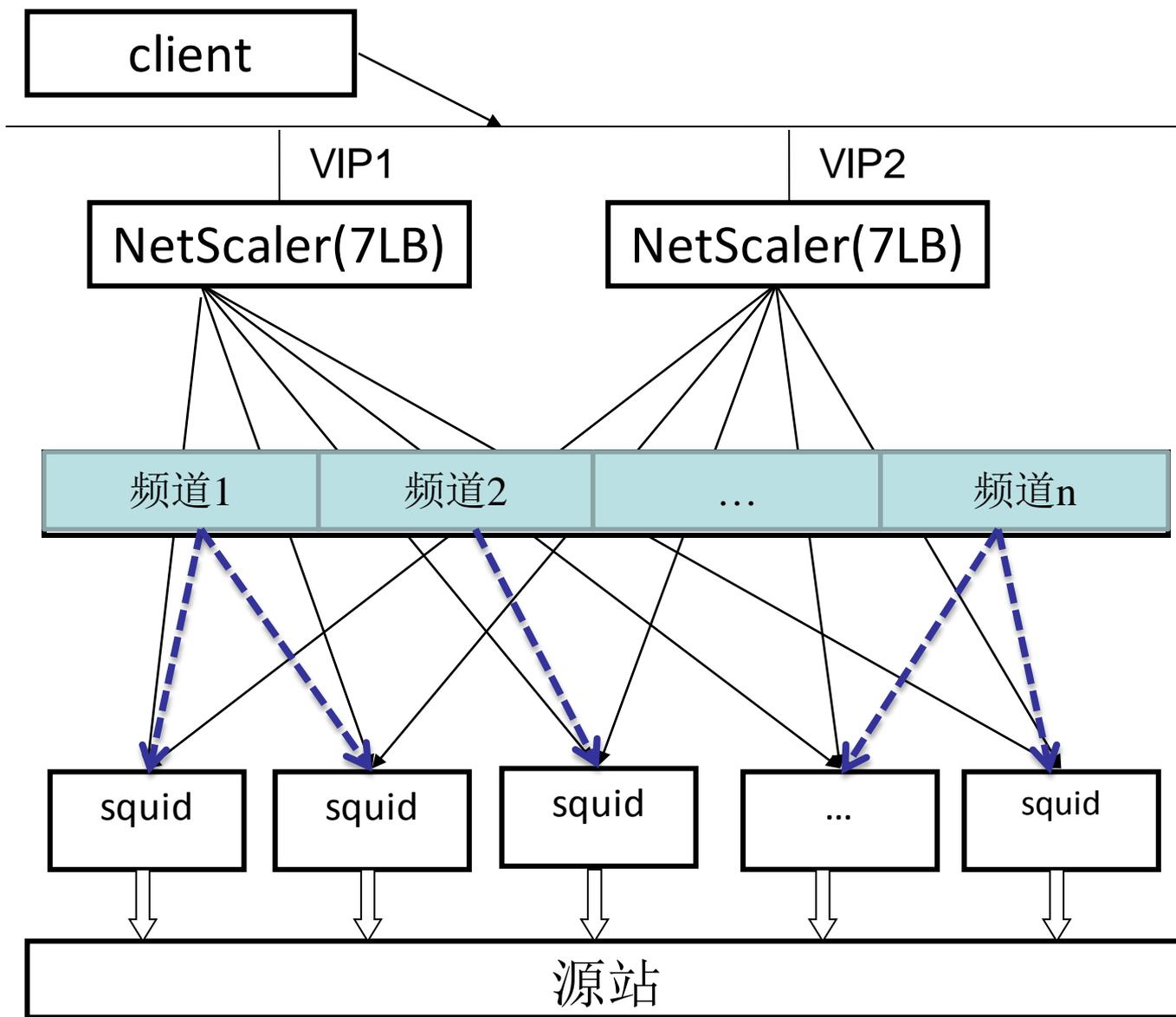
- TFS在2010年9月开源，希望更多人来使用和改进
- TFS 2.0已经在生产系统中使用
 - 支持大文件存储
 - 在外围通过MySQL集群向应用提供目录支持
 - 加入资源中心，控制集群级别的权限
- 后续开发
 - 优化性能，提高扩展性，降低存储成本
 - RAID (Erasure Coding)
 - 分级存储机制 (SSD/SATA)，动态文件迁移等



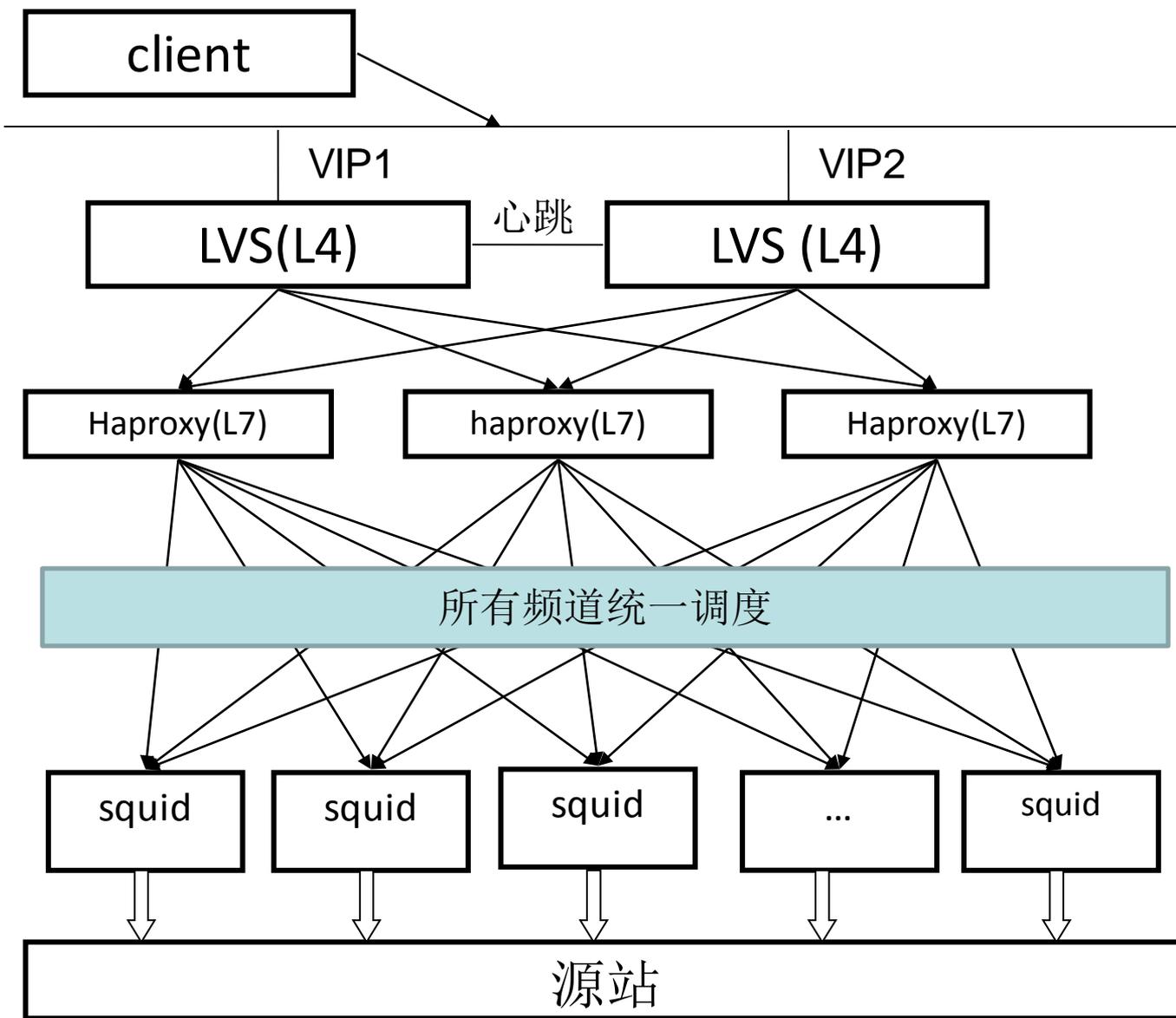
- 一、阿里集团的简介
- 二、淘宝软件基础设施
- 三、事例：图片存储、CDN与DB
- 四、阿里开源成果
- 五、阿里开源流程
- 六、小结

- 主要解决现有的问题
 - 商用产品的性能瓶颈、功能欠缺，以及不稳定性
 - 整个系统的规模、性能、可用性和可管理性
- 开发完全自主的CDN系统
 - CDN节点的新架构和优化
 - CDN监控平台
 - 全局流量调度系统支持基于节点负载状态调度和基于链路状态调度
 - CDN实时图片删除
 - CDN访问日志过滤系统
 - 配置管理平台

CDN节点的架构对比-老架构



CDN节点的架构对比-新架构



CDN节点的架构对比

对比项 \ 节点	新架构	老架构
流量分布均匀性	☆☆☆☆☆	☆☆☆
可维护性	☆☆☆	☆☆☆
抗攻击能力	☆☆☆☆	☆☆☆☆
自主控制能力	☆☆☆☆☆	☆☆☆
价格	☆☆☆☆☆	☆☆☆
扩展能力	☆☆☆☆☆	☆☆
灵活性	☆☆☆☆☆	☆☆

- 流量分布均匀性：所有的频道统一调度到128台squid，而不是将squid按频道分组，可提高命中率2%以上
- 扩展能力：在一个VIP上新架构可以扩展到近100G的流量（当然要用万兆网卡）
- 灵活性：一致性Hash调度方法使得增加和删除服务器非常方便，只有 $1/(n+1)$ 的对象需要迁移

- 2010: LVS + Haproxy + Squid + GTM
 - 利用DELL 2950和混合存储 (1SSD + 4*SAS + 1SATA)
 - Squid优化, ext2文件系统
 - 32个节点, 320Gbps以上的能力
- 2011: LVS + Haproxy + Squid + GTM
 - 低功耗服务器和混合存储 (1SSD + 3*SATA)
 - Squid优化, ext4+nojournal
 - 103个节点, 1000Gbps以上的能力
- 2012: LVS + Haproxy + TS + Pharos为主
 - 40Gbps节点, Xeon L处理器+万兆网卡+6*SSD
 - 104个节点, 320Gbps以上的能力
 - 开发了轻量级高性能的Cache软件Swift

- CDN系统的研发与运维
 - 持续提高节点性能（应用软件、操作系统等）
 - 精细化和自动化全局调度系统
 - 优化视频支持（P2P结合）、移动网络支持
 - 持续提高CDN系统可运维性，服务质量监测
 - 面向音视频通讯的中转支持
- CDN系统的建设
 - 系统进一步整合，优化不同规模节点的硬件配置，建设中型和大型节点
 - 定制化和快速部署

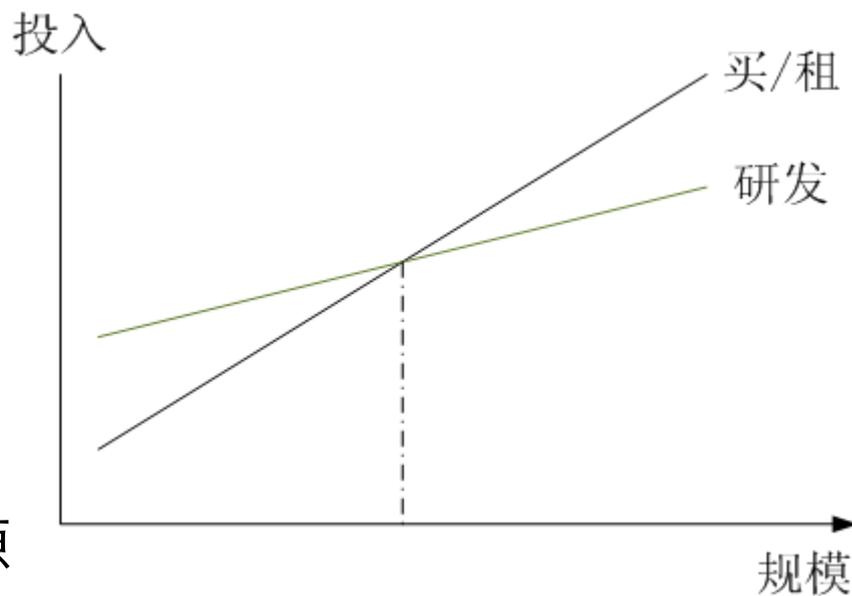


- 一、阿里集团的简介
- 二、淘宝软件基础设施
- 三、事例：图片存储、CDN与DB
- 四、阿里开源成果
- 五、阿里开源流程
- 六、小结

- IOE = IBM + Oracle + EMC
- 优点：
 - 稳定，功能非常强大
 - 支持完善，方便运维
- 缺点：
 - License贵，软硬件成本高
 - 集中式架构，不利扩展
 - 软件黑盒子

- 2008年开始周边的数据库使用MySQL
- 2010年 组建数据库开发团队
 - 结合高速的非易失存储设备，多层优化MySQL
 - 实现分库分表的TDDL中间件成熟
 - 核心业务开始迁移到MySQL
- 2011年 核心数据库迁移到基于MySQL解决方案
 - 商品库：MySQL 16*2
 - 交易库：MySQL 16*2，IOE成本是2千万，现在成本是近400万，TPS从9000提升到12.8万

- 商用软件不能满足大规模系统的需求
- 采用开源软件与自主开发相结合，有更好的可控性，更高的可扩展性
- 规模效应，研发投入都是值得的
- 在软件和硬件多个层次优化，优化是长期持续的过程
- 先平台后业务 vs 先业务后平台，后者更顺
 - 发挥边际效应，提高资源利用率





- 一、阿里集团的简介
- 二、淘宝软件基础设施
- 三、事例：图片存储、CDN与DB
- 四、阿里开源成果
- 五、阿里开源流程
- 六、小结

- 阿里自开源以来，共开源自主开发软件100余个
- 涵盖前端、后端、数据库、文件系统、硬件等多方面
- 对阿里使用的若干项目贡献了代码



- <http://kernel.taobao.org>
- Ext4文件系统核心开发团队之一
- 被Linux官方接受150+个patch
- 全球最活跃Linux开发团队排115名
 - 自2006年起统计
- 开源虚拟化项目sheepdog
 - 主要代码贡献者和维护者

⊕ No. 112	MathEmbedded Consulting	160(0.06%)
⊕ No. 113	General Electric	158(0.06%)
⊕ No. 114	Real-Time Remedies	156(0.06%)
⊕ No. 115	Tao Bao	140(0.05%)
⊕ No. 116	Open Nandra	137(0.05%)
⊕ No. 117	Tilera	135(0.05%)
⊕ No. 118	US National Security Agency	134(0.05%)
⊕ No. 119	Barco	129(0.05%)
⊕ No. 120	CSR	128(0.05%)
⊕ No. 121	secunet Security Networks AG	124(0.05%)
⊕ No. 122	MSC Vertriebs GmbH	123(0.05%)
⊕ No. 123	RisingTide Systems	118(0.04%)
⊕ No. 124	VIA Technologies, Inc.	115(0.04%)
⊕ No. 125	OKI SEMICONDUCTOR	114(0.04%)
⊕ No. 126	Realtek	113(0.04%)
⊕ No. 127	Ksplice	109(0.04%)
⊕ No. 128	Eukrea Electromatique	108(0.04%)
⊕ No. 129	Solid Boot	106(0.04%)
⊕ No. 130	Nuvoton Technology	104(0.04%)
⊕ No. 131	Fusion-io	103(0.04%)
⊕ No. 132	Adaptec	101(0.04%)
⊕ No. 133	AudioScience	98(0.04%)
⊕ No. 134	Selenic Consulting	94(0.03%)
⊕ No. 135	CISCO	89(0.03%)
⊕ No. 135	MEV Limited	89(0.03%)
⊕ No. 135	Gaisler Research	89(0.03%)
⊕ No. 135	Holoscopio Tech.	89(0.03%)
⊕ No. 139	Candela Tech.	88(0.03%)
⊕ No. 140	GNU	80(0.03%)
⊕ No. 141	Telargo	79(0.03%)
⊕ No. 141	Tower Technologies	79(0.03%)
⊕ No. 141	IDT	79(0.03%)
⊕ No. 144	OMICRON electronics	77(0.03%)
⊕ No. 145	Apple	73(0.03%)
⊕ No. 145	Digi International	73(0.03%)
⊕ No. 147	Tensilica	70(0.03%)
⊕ No. 147	SMSC	70(0.03%)
⊕ No. 149	Wacom	67(0.02%)
⊕ No. 149	Infor	67(0.02%)
⊕ No. 149	University of Queensland	67(0.02%)
⊕ No. 152	KFKI Research Institute	66(0.02%)
⊕ No. 152	emlix GmbH	66(0.02%)
⊕ No. 154	Coraid	65(0.02%)
⊕ No. 154	Microgate	65(0.02%)

Source: http://www.remword.com/kps_result/all_whole.html

- 基于Nginx的Web服务器
- <http://tengine.taobao.org>
- 特点
 - 动态模块加载支持
 - 输入过滤器机制支持
 - 组合多个CSS、JavaScript的访问请求变成一个请求
 - 可以对后端的服务器进行主动健康检查
 - 支持管道（pipe）和syslog形式的日志以及日志抽样
 - 监控系统的负载和资源占用从而对系统进行保护
 - 很多运维和易用性增强

- 基于OpenJDK VM
 - 性能提升
 - 针对淘宝需求的定制化改进
- 2011年部署率14%，计划2012年三淘全面部署
- 向Oracle JVM贡献16个通用性patch
- 向Oracle JVM团队输出了人才：)

- 维护阿里MySQL分支
- <http://mysql.taobao.org>
- 2010开始规模使用，替换商品、交易、用户等原基于IOE方案的核心数据库，目前已部署千台规模
- 和Oracle, Percona, Mariadb等上游厂商有良好合作，共向上游提交20多个patch

- Taobao File System
- <http://tfs.taobao.org>
- 分布式文件系统
 - 高可靠性
 - 高可用性
 - 为海量小文件特殊优化
 - 大数据量高并发
 - 低成本
 - 线性扩容
 - 支持自定义文件名

- Taobao Pairs
- <http://tair.taobao.org>
- K/V 储存特点
 - 分布式
 - 自动容灾
 - 支持持久化
 - 支持多种存储引擎
 - Mdb
 - Redis
 - LevelDB
 - ...

- 分布式数据库系统
- <http://oceanbase.taobao.org>
- 特点
 - 支持千亿条记录级别的数据库
 - 兼容MySQL协议，融合SQL与NOSQL技术
 - 支持事务
 - 支持线性扩展
 - 在淘宝大数据量得到验证

- HDFS, MapReduce, Hive, HBase等都维护自己的版本，同时兼容社区版本，支持Hadoop生态系统中Pig, Mahout等软件。
- Hadoop官方中文文档翻译
- Hive
 - 为社区贡献patch超过20个，功能包括mutli-distinct aggregation, JDBC接口和认证体系等
- HBase
 - 改进HBase稳定性，缩短HBase的宕机恢复时间，让HBase的在线服务能力大大提高。
 - 向社区提交47个patch，占总数的3%

- TimeTunnel（简称TT）是一个基于thrift通讯框架搭建的实时数据传输平台
- <http://timetunnel.taobao.org>
- 特点：
 - 具有高性能、实时性、顺序性、高可靠性、高可用性、可扩展性

- DataX是一个在异构的数据库/文件系统之间高速交换数据的工具
- <http://code.taobao.org/p/datax/info/>
- 特点：
 - 在异构的数据库/文件系统之间高速交换数据
 - 采用Framework + plugin架构构建，Framework处理了缓冲，流控，并发，上下文加载等高速数据交换的大部分技术问题，提供了简单的接口与插件交互，插件仅需实现对数据处理系统的访问
 - 运行模式：stand-alone
 - 数据传输过程在单进程内完成，全内存操作，不读写磁盘，也没有IPC
 - 开放式的框架

- MetaQ是类似于kafka的Java消息中间件
- <http://metaq.taobao.org/>
- 主要特性：
 - 高性能，单机2K消息，TPS在4.5W/S以上，网卡可跑满
 - 支持顺序消息
 - 支持服务端消息过滤
 - 支持Consumer端事务消费消息
 - 支持完全的分布式部署，Producer、Consumer、Server都可以分布
 - 支持广播消费、集群消费两种消费方式

- TDDL(Taobao Distributed Data Layer) 是一套分布式数据访问引擎
- https://github.com/taobao/tb_tddl
- 核心的目标是为了解决以下三个问题：
 - 数据访问路由。将针对数据的读写请求发送到最合适的地方
 - 数据的多向非对称复制。一次写入，多点读取
 - 数据存储的自由扩展。不再受限制于单台机器的容量瓶颈与速度瓶颈，平滑迁移

- Diamond是一个持久配置管理中心，提供稳定的配置访问
- <http://code.taobao.org/p/diamond/wiki/index/>
- 特点
 - 简单
 - HTTP接口获得配置简单灵活
 - 可靠
 - 多级保护和容灾支持。在任何异常情况下不影响应用的启动
 - 配置变更通知可靠，保证客户端能拿到最新数据

- 一套分布式服务框架
- <http://code.alibabatech.com/wiki/display/dubbo>
- 很多阿里外部的公司在使用
- 特点
 - 远程服务调用
 - 服务动态发现
 - 集群软负载均衡
 - 集群失败容错

- TBSchedule的目的让一种批量任务或者不断变化的任务，能够被动态的分配到多个主机的JVM中，不同的线程组中并行执行。所有的任务能够被不重复，不遗漏的快速处理
- <http://code.taobao.org/p/tbschedule/wiki/index/>
- 特点
 - 任务可以动态创建和停止，并分布在多个机器上执行
 - 可以指定任务执行的时间周期
 - 有统一集中的控制台，进行各种参数的动态调整
 - 使用简单，只需要实现一个简单接口
 - 目前在淘宝的很多系统，已经外部公司的系统使用

- 一个轻量级的脚本引擎，作为一个嵌入式规则引擎在业务系统中使用
- <http://code.taobao.org/p/QLExpress/wiki/index/>
- 特点：
 - 支持标准的JAVA语法，
 - 支持自定义操作符号、操作符号重载、函数定义、宏定义、数据延迟加载等
 - 可以自定义语法。
 - 类编译执行，执行效率比较高
 - 目前在淘宝的很多系统，已经外部公司的系统使用

- 一个非常好用的JDBC数据库连接池
- <https://github.com/AlibabaTech/druid/wiki>
- 特点
 - SQL监控/Web-Spring-SQL关联监控
 - 提供Filter-Chain模式的扩展机制
 - 支持监控统计、日志
 - 防SQL注入攻击
 - 解决Oracle PSCache内存占用问题

- 一个非常高效的JSON库
- <http://code.alibabatech.com/wiki/display/FastJSON/Home>
- 在第三方评测的数据显示，FastJSON比Jackson、Gson、JSON-Lib和Hessian都快
- 易于使用，方便扩展

- 一套强大的Java Web框架
- <http://www.openwebx.org/>
- 特点
 - 基于Java Servlet API
 - 扩展性良好
 - 在大流量网站上的稳定性得到了验证

- 一个强大的JavaScript框架
- <http://docs.kissyui.com>
- 特点
 - 模块清晰
 - 扩展方便
 - 功能全面
 - 组件丰富

- 一个开放、简单、易用的前端基础类库
- <http://aralejs.org>
- 能从底层让代码的书写规范、组件的构建方式、模块的调用方式等保持整体一致性

- 一个适用于 Web 端的模块加载器
- <http://seajs.org>
- 特点
 - SeaJS 追求简单、自然的代码书写和组织方式
 - 提升代码的可维护性。如果一个网站的 JS 文件超过 3 个，就适合用 SeaJS 来组织和维护代码

- 轻量级服务器及应用信息监控和数据收集软件
- <http://tsar.taobao.org>
- 特点
 - 部署简单，稳定可靠，系统开销小
 - 模块化设计，方便添加自定义的采集模块
 - 支持实时终端显示和集中式数据库存储及查询
 - 支持与Nagios联动，发送报警信息

- Sqlautoreview是一个SQL自动化审核程序
- <http://code.taobao.org/p/sqlautoreview/src/>
- 主要有三方面的功能:
 - 从SQLMap中解析出SQL
 - 为每台SQL生成建index的脚本
 - 将同一个表的建index的脚本进行自动聚合,减少索引变动

- 可以在生产环境长期使用的Java性能分析工具
- <http://github.com/taobao/tprofiler>
- 特点
 - 同时支持剖析和采样两种方式
 - 记录方法执行的时间和次数，生成方法热点、对象创建热点 线程状态分析等数据
 - 为查找系统性能瓶颈提供数据支持.

- ZooKeeper的监控与报名系统
- <http://code.taobao.org/p/taokeeper/src/>
- 特点
 - 能够统计ZK集群连接数，Watcher数目，节点数等系列信息
 - 能够通过设置一些阈值来达到监控报警的效果
 - 能够按天按周进行ZK运行状况的报表展现
 - ZK实时运行状态的统计和机器存活性的检查功能

- 淘宝自动化测试组自主研发的一套界面自动化测试框架
- <http://code.taobao.org/p/AutoMan/wiki/index/>
- 特点
 - 框架的核心是基于界面模型的设计，将“元素查找”和“控件操作”分开。
 - 元素查找的方式定义在PageModel的Web服务器上，在脚本中只说明使用控件的名称和对该控件的操作方式
 - 上手快、易维护

- <http://toast.taobao.org>
- 自动化测试任务调度平台。目标是在复杂环境下(多种业务逻辑, 多种应用环境)的通用调度平台, 主要提供任务的定时运行、手动运行以及基于持续集成等方式的触发执行; 同时提供了测试机的简单监控和管理功能, 支持多测试机并行执行任务以及任务的分阶段执行。

定制的低功耗服务器



- 每个热插拔模块2个nodes
- 每个NODES 3块硬盘
- 支持 24 x 2.5" SATA/SSD
- 选择2U 8 nodes 的原因：
 - SuperServer: 2U 8 nodes 支持热插拔设计。
 - 降低功耗 (2U 8nodes 共享4个系统FAN)
 - 成本更低 (8nodes共享1个机箱, 1u 4nodes 共享1个机箱)
 - 2U TWIN机箱和所使用的主板是成熟产品
- 单服务器配置：
 - Intel® Atom™ D525 with 2 cores
 - Intel® ICH9R Chipset
 - 4GB memory DDR23 800MHZ SO-DIMM w/o ECC
 - LAN: Intel 82574L 2*1 GB
 - HDD:
 - 1* SSD 80G ,
 - 2* 2.5" SATA 500GB

- 开源网站 <http://www.greencompute.org/>

开源绿色计算

English Version

首页

项目介绍

设计规范

合作赞助

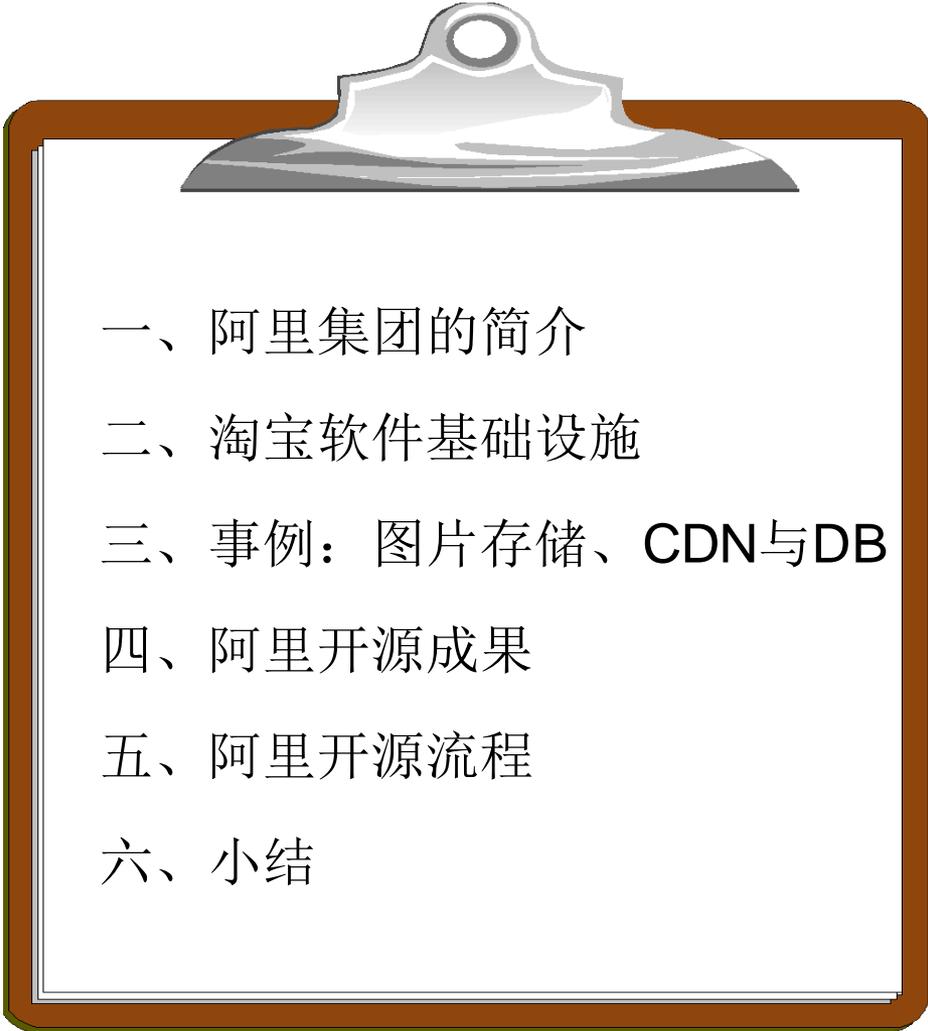
论坛讨论

联系方式

新闻公告



设计规范名称	采用CPU型号	版本	发布时间	下载地址	面向应用
主板设计规范	Intel Atom D525	V1.0	2011-9-27	中文版 英文版	CDN的缓存服务应用
机箱和电源设计规范	Intel Atom D525	V1.0	2011-9-27	中文版 英文版	CDN的缓存服务应用
服务器测试规范	Intel Atom D525	V1.0	2011-9-27	中文版 英文版	CDN的缓存服务应用

- 
- 一、阿里集团的简介
 - 二、淘宝软件基础设施
 - 三、事例：图片存储、CDN与DB
 - 四、阿里开源成果
 - 五、阿里开源流程
 - 六、小结



- 目前有10人
- 人员构成，
 - 工程师为主
 - 法务
 - 安全
- 协助推动开源的过程
 - 流程
 - 辅导

- 由项目的开发工程师发起
- 获得团队主管和部门主管的同意
 - 确认是否具有开源意义
- 通知开源委员会备案
- 开源小组协助开源协议问题
- 代码和文档通过安全review
- 代码和网站host到淘蝌蚪上
 - 二级域名与网站
- 可选择同步到github上

对开源文化的鼓励

- 开源项目的成就感是对工程师的最大鼓励
- 内部的开源项目经验交流
- 优秀开源项目奖励

- GPL vs. BSD vs. MIT vs. Apache
 - GPL对项目的后续开源保护最好
 - 其他协议比较宽松
- 阿里的选择
 - 大部分项目使用GPL。
 - 一些开源的库用BSD或Apache等
 - 非阿里初始项目遵守原项目的license
- 阿里集团是版权持有人
 - (C) 2007-2012 Alibaba Group Holding Limited

- code.taobao.org

- 目的

- 开源平台本身开源
- 国内访问方便

- 现状

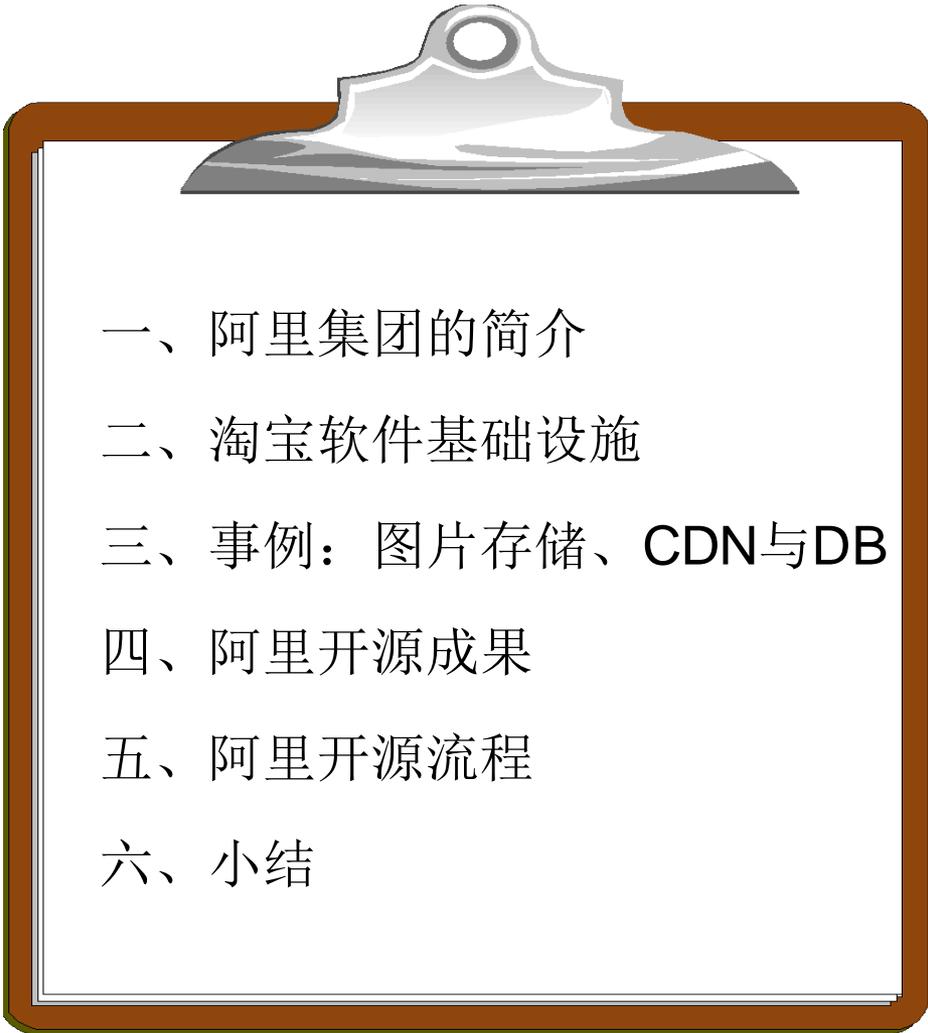
- 现有项目373个
- 成熟项目以淘宝的为主
- 非淘宝项目逐渐开始活跃，并出现了一些不错的项目



- code.alibabatech.com



The screenshot shows the Open Sesame website dashboard. The top navigation bar includes "Dashboard", "Browse", "Log In", and a search box. The main content area is divided into two columns. The left column features the "OPEN SESAME" logo and a welcome message. Below this, there are links for "Feed Builder" and "People Directory". A "Spaces" section lists several projects: Bazar, Cobar, CobarClient, Druid, and Dubbo. Each project entry includes a brief description and links to documentation or code. The right column displays a list of recent blog posts, each featuring a user profile picture, the author's name, the post title, and the date. The posts are: 1. "Dubbo 2.4.4 released on 2012-08-16" by 李鼎, dated Aug 16, 2012. 2. "Dubbo 2.5.0 released on 2012-08-13" by 李鼎, dated Aug 13, 2012. 3. "Dubbo 2.4.3 released on 2012-08-03" by 梁飞, dated Aug 03, 2012. 4. "Dubbo 2.4.2 and 2.3.6 released on 2012-07-26" by 李鼎, dated Jul 25, 2012.

- 
- 一、阿里集团的简介
 - 二、淘宝软件基础设施
 - 三、事例：图片存储、CDN与DB
 - 四、阿里开源成果
 - 五、阿里开源流程
 - 六、小结



- 通过开源，可以得到很多开发人员和高手的帮助，不断提高自己的能力
- 通过开源，可以得到更多的用户
- 通过开源，可以让自己的代码有更长的生命周期
- 用户的用法，会超出工程师的想象

- 促进了开源软件本身的质量提高
 - 用户的意见反馈
 - 用户的bug报告
 - 用户的patch
- 阿里的技术能力和开放精神获得业界的肯定
 - 阿里技术工程师强烈的认同感
 - 吸引更多有能力的工程师加入

- 阿里是开源系统的受益者，并积极参与开源生态系统的建设，促进开源生态系统的发展，积累更好的口碑，凝聚人才，迎接未来更大的技术挑战
- 阿里在开源的过程中有付出，更有收获
- 阿里希望以更开放的方式与业界一起进行技术创新

- Q&A
- 谢谢!