

百度开放云

--多媒体数据流服务

<http://bce.baidu.com/solution/media.html>

DTCC

2015中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2015

大数据技术探索和价值发现



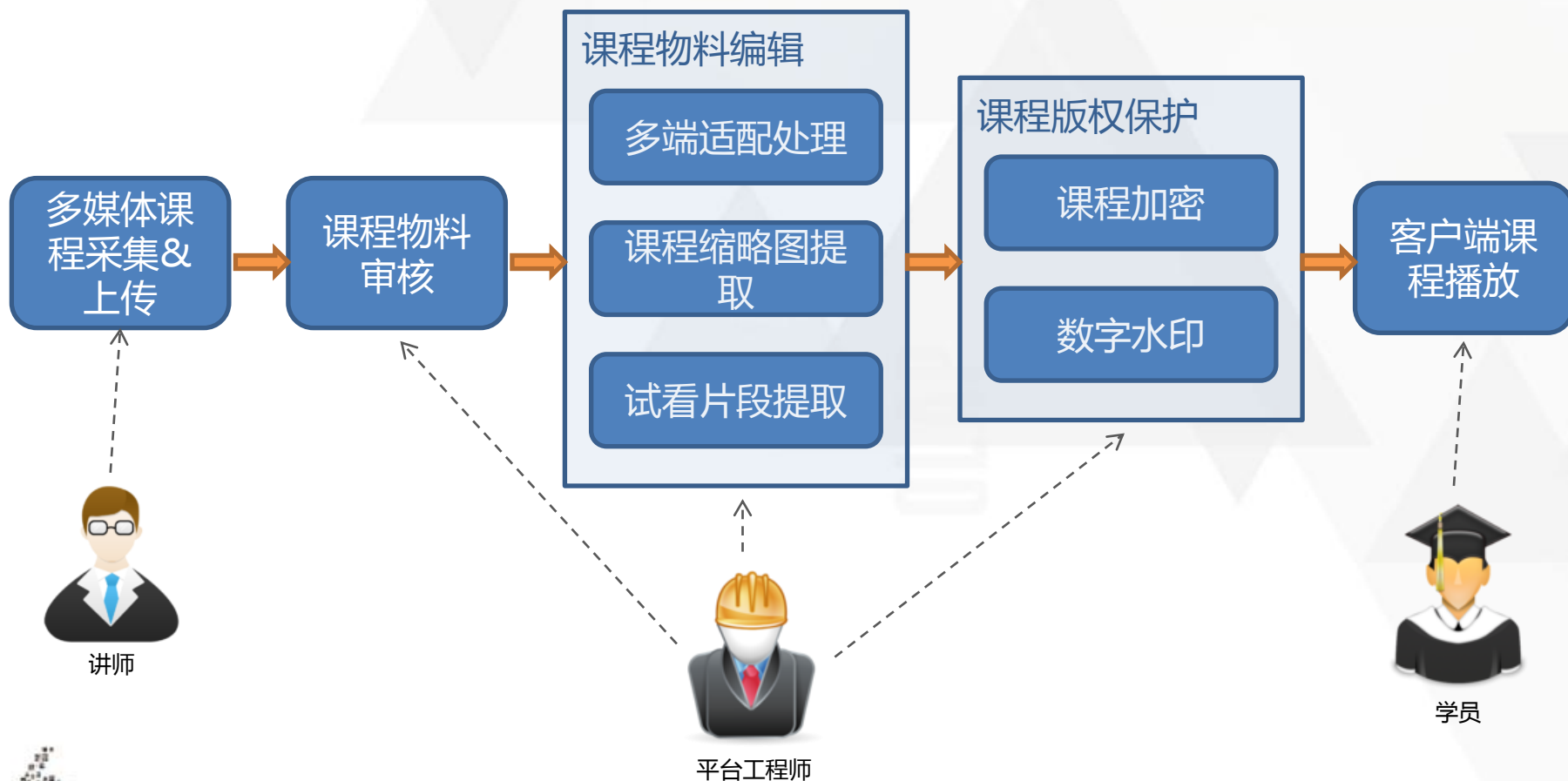
主要内容

- 典型业务场景
- 开放云解决方案
- 多媒体存储 – 开放云BOS系统
- 多媒体计算 – 开放云BMC系统
- 多媒体分发 – 开放云CDN系统
- 总结 & 展望

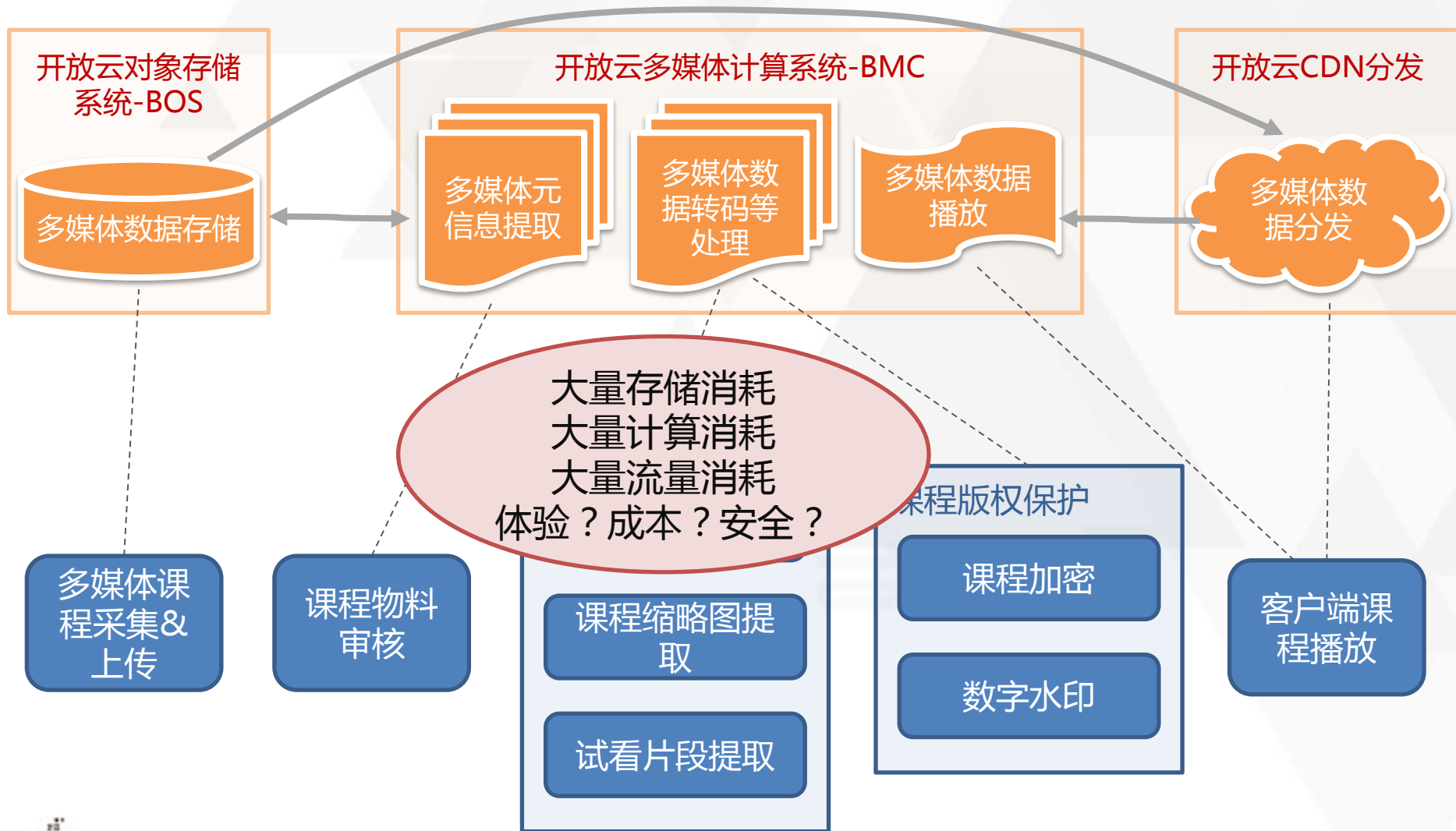


典型业务场景

在线教育平台



百度开放云解决方案 – 端到端解决能力



开放云对象存储系统

– BOS (Baidu Object Storage)

- ✓ 系统概念&架构
- ✓ 面临挑战
- ✓ 低成本存储
- ✓ 数据安全
- ✓ 数据可靠性



全国首批获得可信云服务认证



BOS-系统概念

提供HTTP服务，支持多语言SDK、Web Console等交互方式，用户可以通过它随时在Web上的任何位置存储和检索任意大小的数据。为开发者提供高扩展、可靠、安全、高效廉价的存储服务。

Object (对象)

✓ 用户操作的基本数据单元是Object，包含Key、Meta和Data。可以是文本、音乐、视频、二进制程序等任何类型的数据。

Bucket (容器)

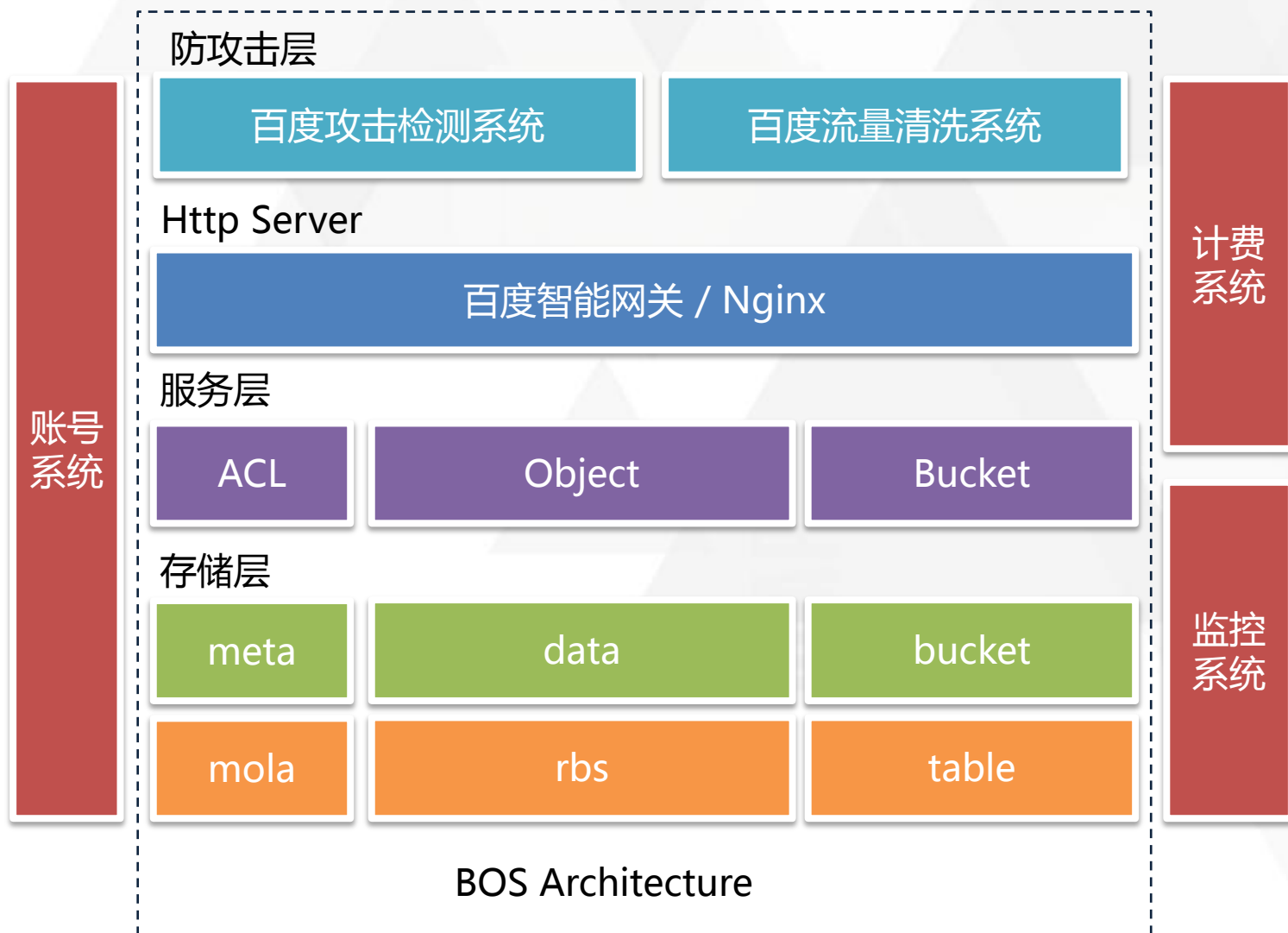
✓ Bucket可以理解为命名空间。Bucket名称在一个Region中具有全局唯一性，每个Object必须包含在某个Bucket中。

ACL (访问控制列表)

✓ BOS系统的鉴权服务。



BOS-系统架构



BOS-面临挑战

1. 如何低成本存储海量的多媒体数据？
2. 如何保证数据的存储安全？
3. 如何保证数据的可靠性？



BOS-系统特性：低成本海量存储

硬件

- ✓ 遍布国内的自研数据中心，内蒙、阳泉20万台机器超大规模数据中心
- ✓ 数据中心能耗指标PUE 1.36，国内第一
- ✓ 专属硬件：超大存储容量服务器、低功耗服务器（自研arm服务器）



软件

- ✓ 存储容量无限扩展，多线BGP网络确保全国各地访问流畅
- ✓ 并行/分块上传，有效存储多媒体视频等超大文件，支持最大5TB单文件存储
- ✓ 通过持续优化的EC编码，在保证数据可靠的前提下，可大幅降低存储空间占用和成本

支撑业务

- ✓ 以低成本下支撑百度网盘、百度地图、百度文库等大规模产品



BOS-系统特性：多维度数据安全保证

用户签名认证

- ✓ 采用Access Key与请求签名认证
- ✓ 多对Access Key ID/Secret Access Key



ACL权限管理&鉴权

- ✓ 通过ACL实现对Bucket的权限管理和校验
- ✓ 同时支持基于用户角色和资源的鉴权



防盗链等内容保护

- ✓ 支持基于referer/IP的防盗链
- ✓ 支持资源限时访问，用户可以提供一个自定义时间内有效的URL



BOS-系统特性：数据可靠性保障

数据可靠性：99.999999999%

服务可用性：99.99%



- ✓ 多副本数据备份：3副本
- ✓ 系统升级透明
- ✓ 完善的监控运维，7*24小时服务支持
- ✓ 支持跨集群/地域冗余机制



开放云多媒体计算系统

– BMC (Baidu Multimedia Computing)

- ✓ 面临挑战
- ✓ 分布式转码
- ✓ 版权保护
- ✓ 内容审核、多端播放器



BMC-面临挑战

1. 如何快速完成海量媒体文件的处理？
2. 如何保证音视频的质量？
3. 如何保护媒体内容的版权？



BMC-系统特性：高性能音视频转码

什么是音视频转码

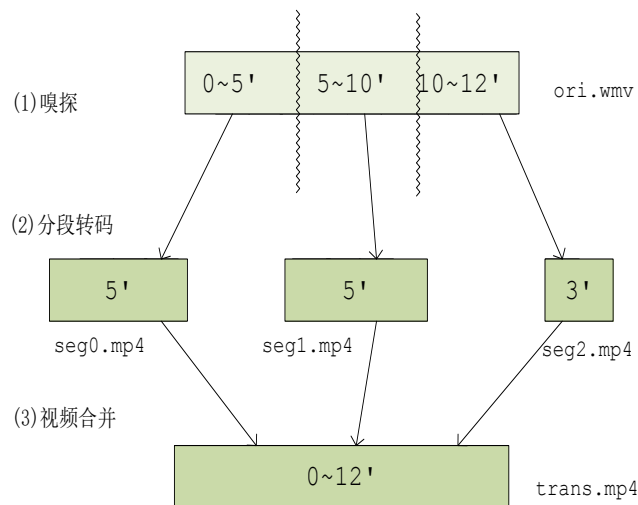
- ✓ 将已压缩编码的音视频码流转换成另一个码流，是一个先解码，再编码的过程。转换前后的码流可能遵循相同的编码标准，也可能不遵循相同的编码标准。
- ✓ 由于转码包含了一次编解码的全过程，会耗费大量计算。

为什么要做转码

- ✓ 适应不同网络带宽
- ✓ 适应不同终端的处理能力
- ✓ 适应不同的用户需求

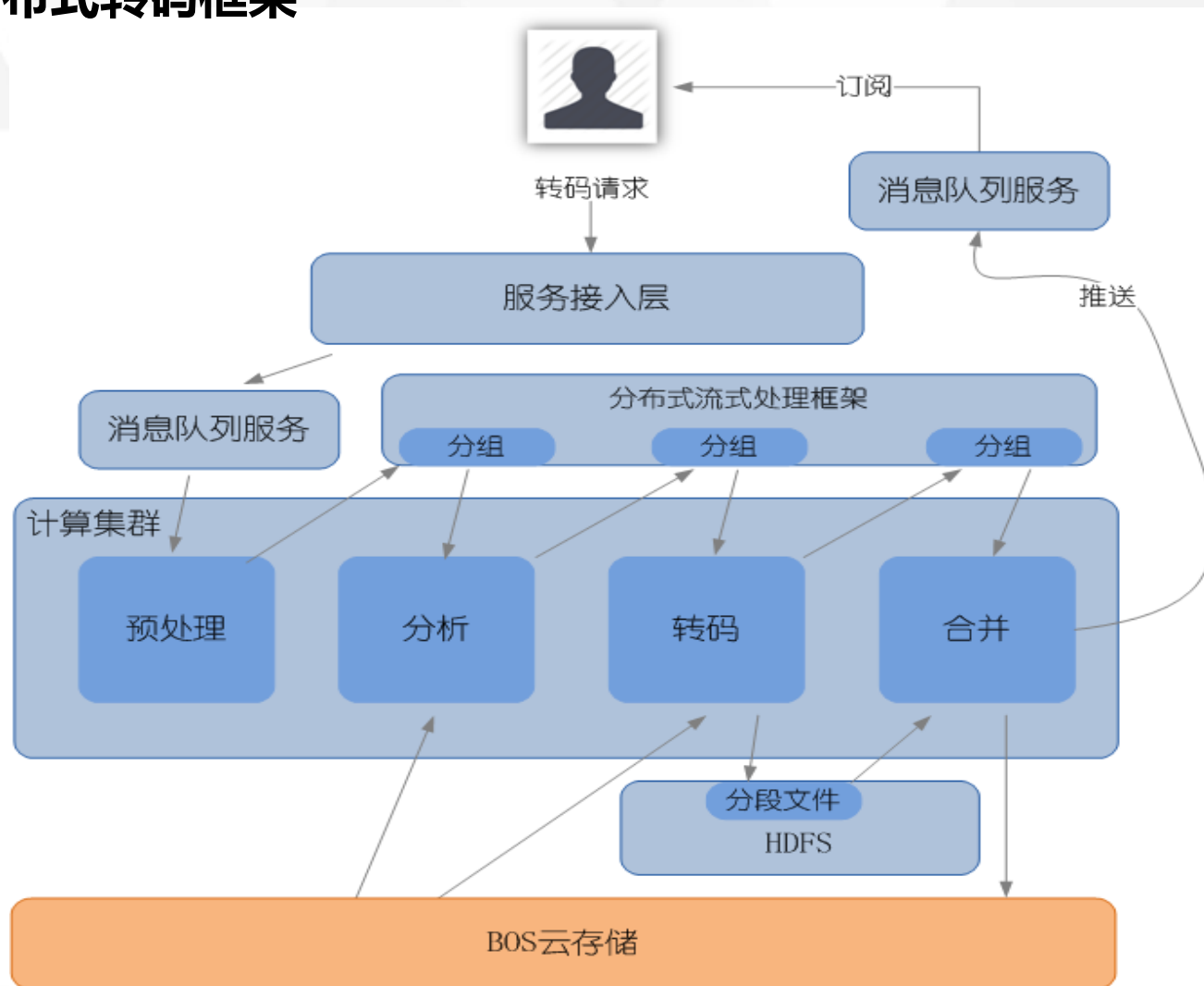
分布式转码

- ✓ 基于MapReduce的分布式计算思想，将整段视频分割成多段，Map阶段将每个分段并行完成转码，Reduce阶段完成所有分段的合并



BMC-系统特性：高性能音视频转码

BMC分布式转码框架



BMC-系统特性：高性能音视频转码

转码全面性

- ✓ 支持对主流视频格式进行转码
- ✓ 同步输出11种码流，5种音视频格式
- ✓ 自适应码率输出（ Adaptive HLS ）
- ✓ 试看视频提取
- ✓ 音轨提取转码

转码效率

- ✓ 分布式12倍速高效编解码（ 360fps ）
- ✓ 自动检测transmux，达到秒级格式转换
- ✓ 内网分布式上下传，10倍速度提升

转码质量

- ✓ 千级别计算集群使得视频能够高保真转码
- ✓ 同等带宽下视频质量优于竞品

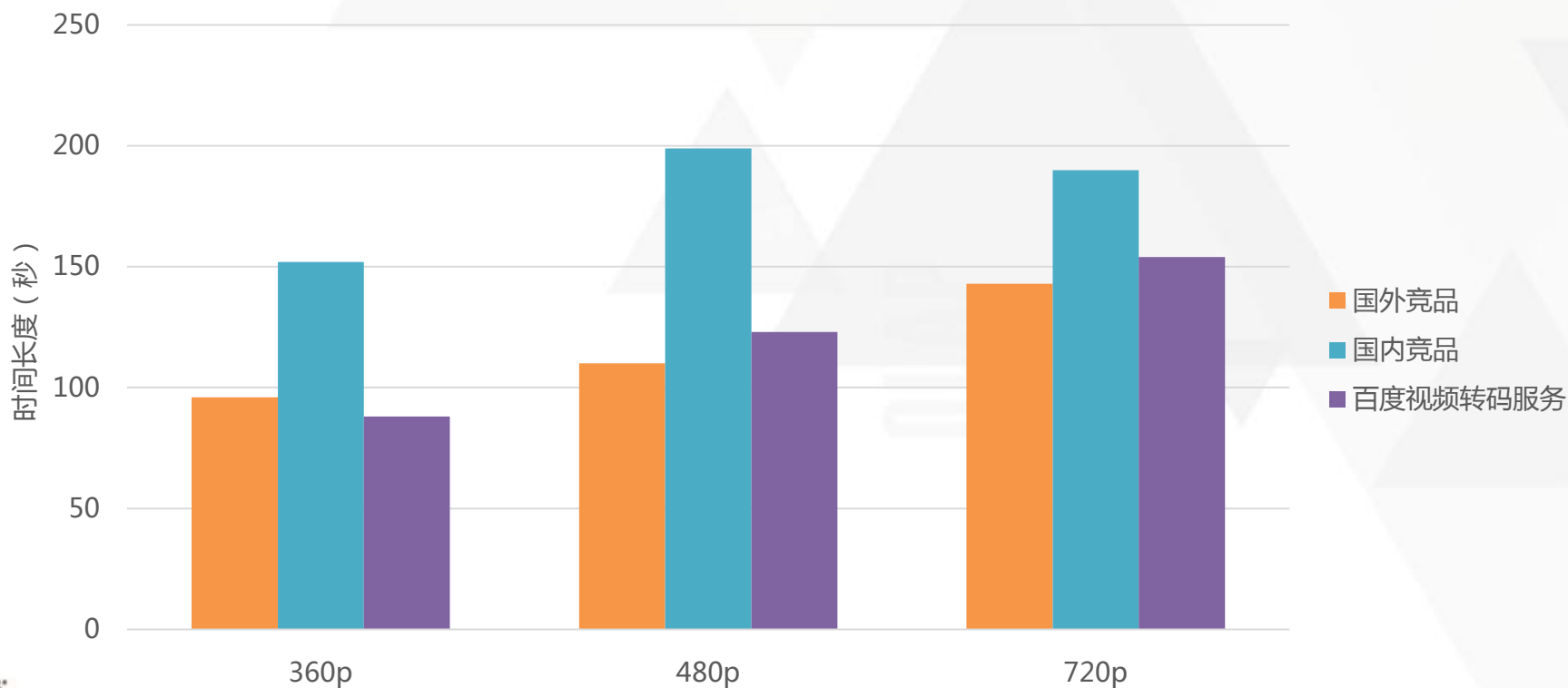


BMC-系统特性：高性能音视频转码

转码效率

- ✓ 衡量指标：转码任务所花时间
- ✓ 实验数据：1080p视频转码成720p, 480p和360p

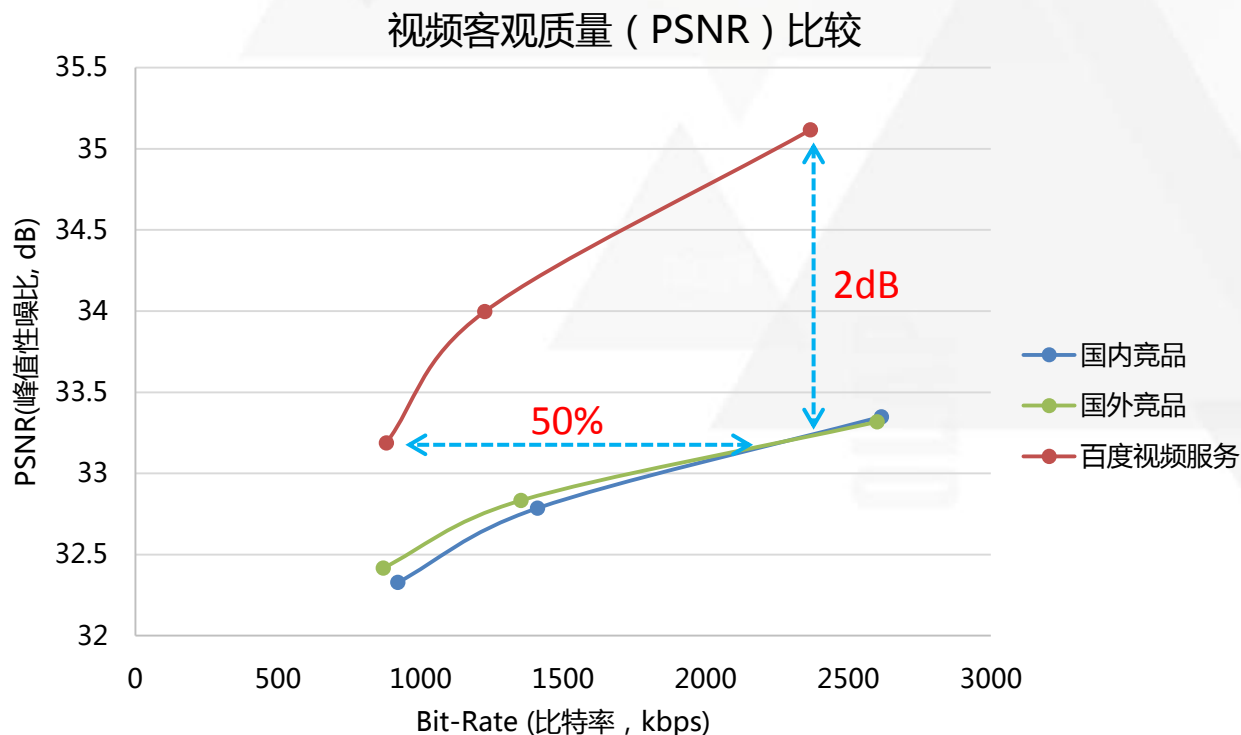
转码所耗时间对比



BMC-系统特性：高性能音视频转码

转码质量

- ✓ 同等带宽损耗下，百度转码PSNR（峰值信噪比）比同类产品高出1.5~2dB
- ✓ 同等视频质量下，百度转码比同类产品节省一半流量



BMC-系统特性：高性能音视频转码



竞品

(720P, 30fps, 2833kbps)



百度

(720P, 30fps, 2691kbps)



BMC-系统特性：媒体版权保护



BMC-系统特性：内容审核、多端播放器

内容审核

- ✓ 元信息提取
- ✓ 缩略图提取

多终端安全播放器

- ✓ Android Player SDK
- ✓ iOS Player SDK
- ✓ Web Player SDK (Flash/Html5)



多媒体分发服务

– **CDN (Content Delivery Network)**

- ✓ 面临挑战
- ✓ 硬件支撑
- ✓ 软件优化



CDN-面临挑战

1. 如何提升多媒体分发性能和播放体验？
2. 如何控制多媒体分发的成本？
3. 如何保证数据安全？



CDN-系统特性：大规模节点资源

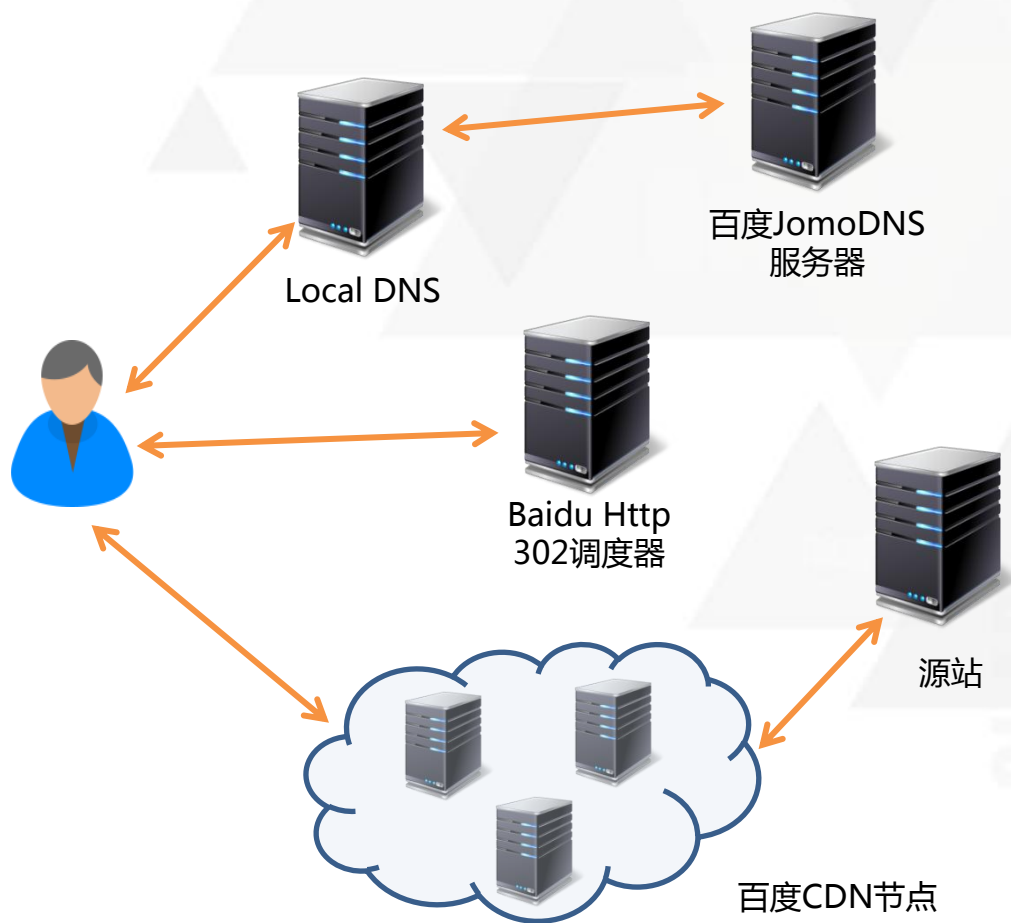
节点规模

- ✓ 60+节点（电信、联通、移动、铁通、教育网、鹏博士等主流运营线及BGP接入）
- ✓ 1万+缓存服务器
- ✓ 2Tbps+带宽
- ✓ 承载百度70%的流量，覆盖百度75%的产品线，包括百度搜索、百度网盘、百度地图等

百度自建CDN节点分布图



CDN-系统特性：优质的CDN节点



节点质量

- ✓ PB级存储，视频分发百G带宽，SSD读写
- ✓ 高命中率：百度个人网盘百PB级的业务，保证90%命中率
- ✓ 毫秒级响应，通过第三方评测机构，下载速度领先同行水平20%
- ✓ 节点故障自动切换



CDN-系统特性：多媒体分发优化

灵活的计费策略，有效控制成本

- ✓ 按流量计费
- ✓ 按峰值带宽计费

分发安全防护

- ✓ 基于referer/IP/回源校验的防盗链
- ✓ 自带Web应用防护系统-WAF防护

多媒体分发优化

- ✓ MP4/FLV伪流技术支持
- ✓ 超大文件分发
- ✓ 多媒体文件上传加速



总结 & 展望



总结 - 解决方案服务架构

开发者 / 合作伙伴

Console / REST API / SDK

开放云多媒体计算-BMC

分布式转码

版权保护

内容审核

数字水印

监控
报警



Web Player SDK



Android Player SDK



iOS Player SDK

开放云CDN

安全防护

视频加速

缓存刷新

监控报警

开放云对象存储-BOS

ACL服务

Bucket服务

Object服务

监控报警

...



总结 - 全流程的解决方案

一站式服务，便捷接入

- ✓ 上传->处理->分发->播放
- ✓ REST API, SDK, Console

高稳定性/性能+低成本

- ✓ 高服务SLA保证：99.99%
- ✓ 转码质量&效率优异
- ✓ 免费计算集群，低成本存储+分发

行业案例：CSDN教育频道、传课网构建基于百度开放云的在线教育方案。

CSDN
全球最大中文 IT 社区

Chuanke.com
学知识 上传课

全流程数据安全保护

- ✓ 存储访问多层控制
- ✓ 媒体版权保护
- ✓ 防盗链
- ✓ HTTPS (即将支持)

全流程加速

- ✓ 上传加速：分块/多线程上传
- ✓ 转码加速：分布式转码，高性能计算集群
- ✓ 分发加速：优质节点，冗余带宽



更多期待

- 视频直播服务
- 图片转码服务
- 硬件加速
- ...



IT168

ChinaUnix

ITPUB

IT168

THANKS