



# 爱奇艺高可用高性能服务器编程架构实践

2017.11



悦球互享技术大品 质

## • 对于高可靠、高可用框架的要求

- 1、应用服务的不间断性：
  - 1.1、服务程序可在线热升级
  - 1.2、数据及配置可在线重新加载
- 2、应用服务可运维可监控：
  - 2.1、服务异常自动重启
  - 2.2、服务异常自动报警
  - 2.3、服务配置项一致性
  - 2.4、服务进程统一管理

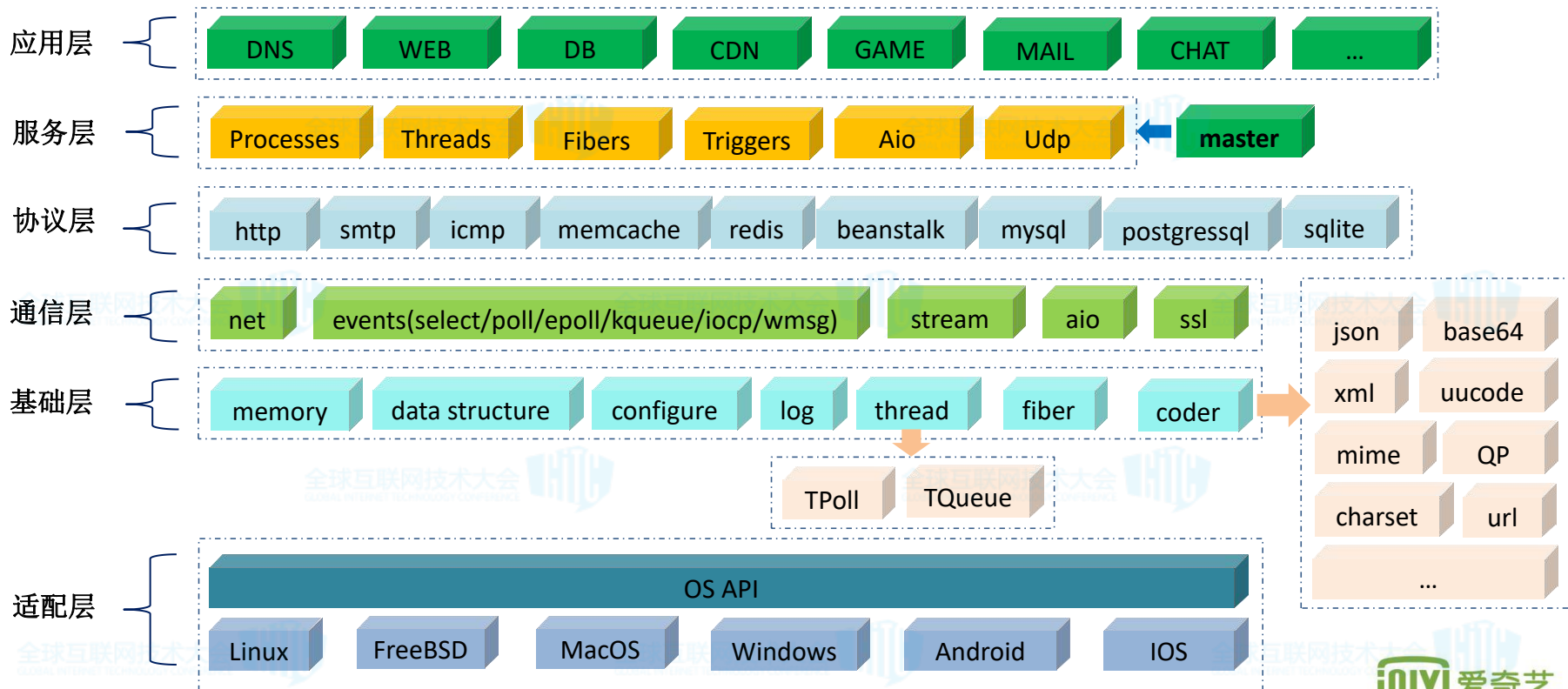
## • 为什么需要高并发、高性能框架？

- 1、互联网尤其是移动互联网的快速蓬勃发展对服务端应用的压力与日俱增
- 2、高性能、低延迟的业务需求对应用服务提出了更高的要求
- 3、内网中 RPC 通信、连接池的大量存在要求服务端支持大并发、高性能

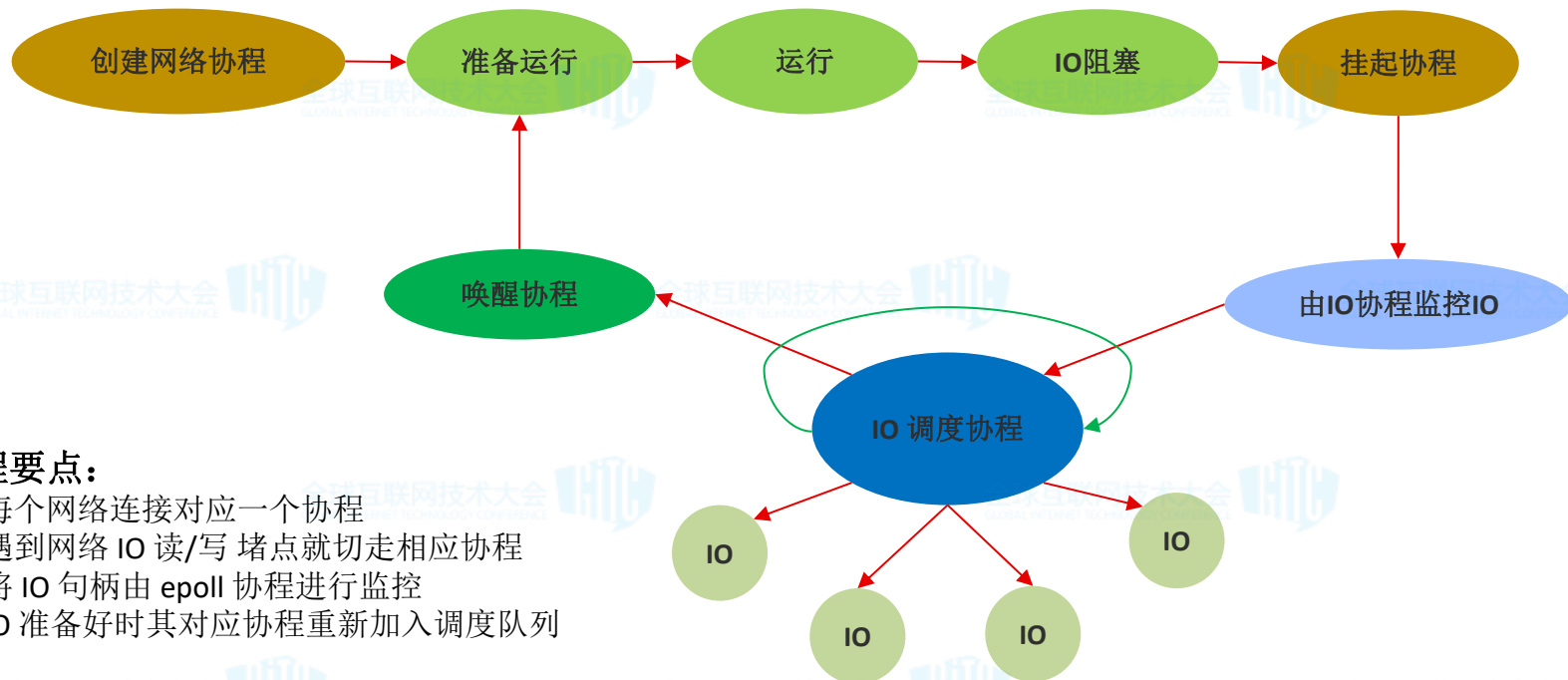
# Acl 服务器编程框架

内容分类	功能项
服务器模型	多进程服务模型、多线程服务模型、多线程服务模型、非阻塞服务模型、触发器服务模型、网络协程服务模型
安全运行机制	严格的用户权限限制，防止越权操作；IP 地址访问控制，防止非法地址访问
配置驱动方式	独立的服务配置文件，丰富的配置项：进程数、线程数、防问控制、监听地址、进程预启动等均由配置项决定
服务监听	支持绑定 TCP 套接口、UDP 套接口以及 UNIX 域套接口，支持同一进程同时绑定多个地址及端口
服务管理	支持应用服务在线热升级，升级过程不会中断服务；提供 WEB 管理接口及管理工具，方便远程管理
运维监控	提供应用服务进程管理模块，在线监控各个应用服务的存活状态，对异常崩溃的服务实时报警
运行方式	生产环境由 acl_master 服务管理进程控制管理；开发环境可以手工单独运行
开发调试	用服务器生成向导直接生成服务器程序框架；用 valgrind 检查服务程序的内存问题

# Acl 服务器框架模型



# 网络协程调度 --- 设计原理



## 编程要点:

- 1、每个网络连接对应一个协程
- 2、遇到网络 IO 读/写 堵点就切走相应协程
- 3、将 IO 句柄由 epoll 协程进行监控
- 4、IO 准备好时其对应协程重新加入调度队列

# 网络协程应用场景

## 一、问答式应用服务

- 基于 HTTP 协议的服务应用，诸如：网站
- 基于 SMTP/POP3/IMAP 协议的服务应用

## 二、生产者 - 消费者类应用服务

- 如消息队列类应用

## 三、reactor 和 proactor 两种模式的结合

- 统一的事件引擎监控所有的网络连接，有一个连接就绪时创建协程独立处理
- 此类应用如聊天服务、游戏服务等无状态的应用服务

## 四、大并发类应用服务

- 因为通过协程方式，将上层应用的堵塞式在底层转为非阻塞模式，所以非常容易以较低资源支持大并发类应用
- 如内网的大多数应用服务为提高效率都支持连接池模式，需要服务端支持非常大的并发

## 五、网络限流

- 在协程中可以直接 sleep，非常容易控制网络流量

# 爱奇艺高性能高可用服务的工程实践

- **背景：**

- 1、如何支撑超 5 亿月活用户的视频收看？
  - --- 海量用户、海量存储、海量带宽，每项都是巨大的挑战
- 2、如何达到热播剧的快速分发？
  - --- 新发视频快速“送达”用户最近观影点，即发即播
- 3、如何达到视频流的快速开播？
  - --- 不能让用户多等一秒
- 4、如何针对大文件进行有效回源？
  - --- 这是 Squid/Apache Traffic/Nginx 目前都未解决的问题



# CDN 回源系统开发历程

## • 遇到的坑：

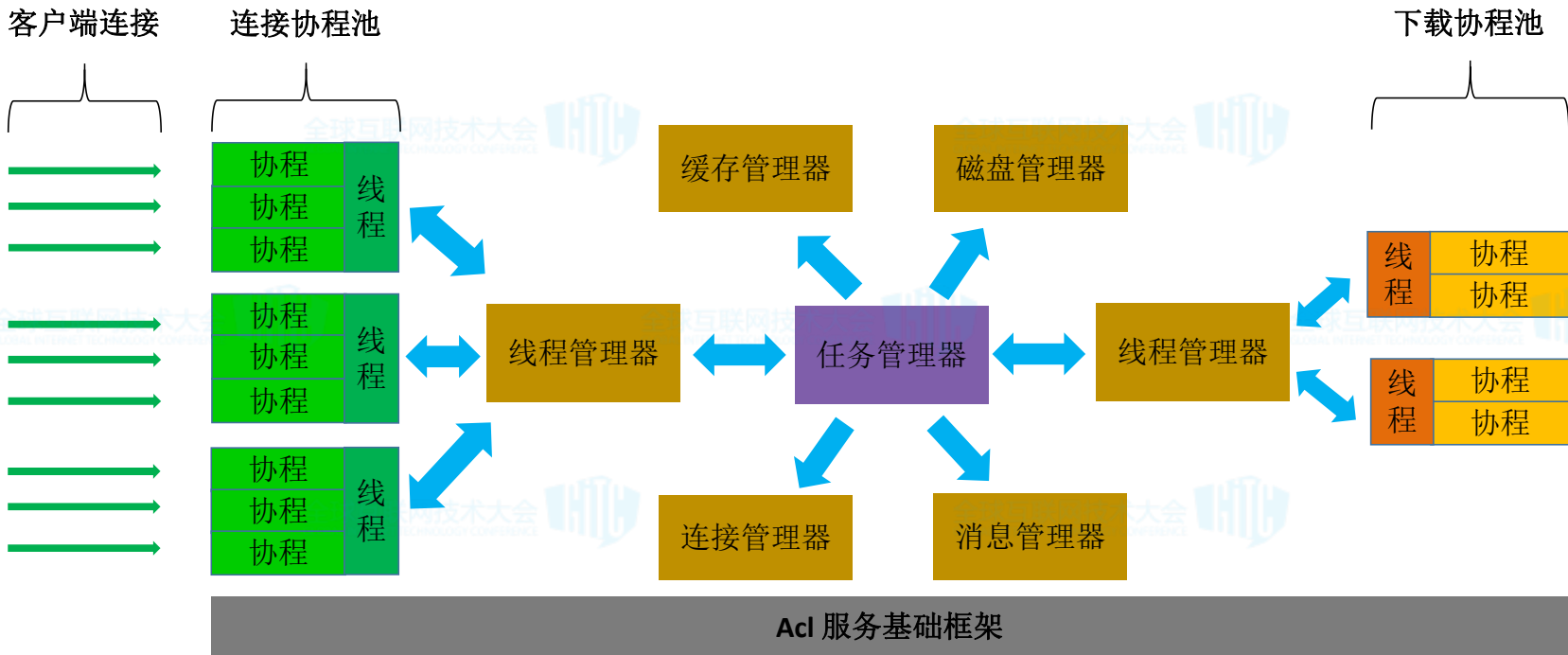
- 1、Squid , Apache Traffic , Nginx 无法针对大文件进行有效合并回源 ☹
- 2、多线程/多协程方式下的各种死锁 ☹
- 3、缓存对象提前释放导致的各种崩溃 ☹
- 4、内存“暴涨”导致的瞬间崩溃 ☹
- 5、用户跳跃式点播视频导致回源带宽的大量浪费 ☹
- . . .

## • 踩坑方案：

- 1、合并相同请求块数据，支持缓存部分数据 ☺
- 2、优化线程锁及协程锁设计，避免死锁 ☺
- 3、有效隔离缓存对象的内存队列 ☺
- 4、内部禁止大内存分配，同时采用 sendfile 等零拷贝技术，内存仅占几百兆 ☺
- 5、调整数据视频回源优先级，优先保证在线收看请求 ☺
- 6、支持流式数据传输，提升视频开播速度 ☺
- . . .



# 爱奇艺 CDN 回源软件架构

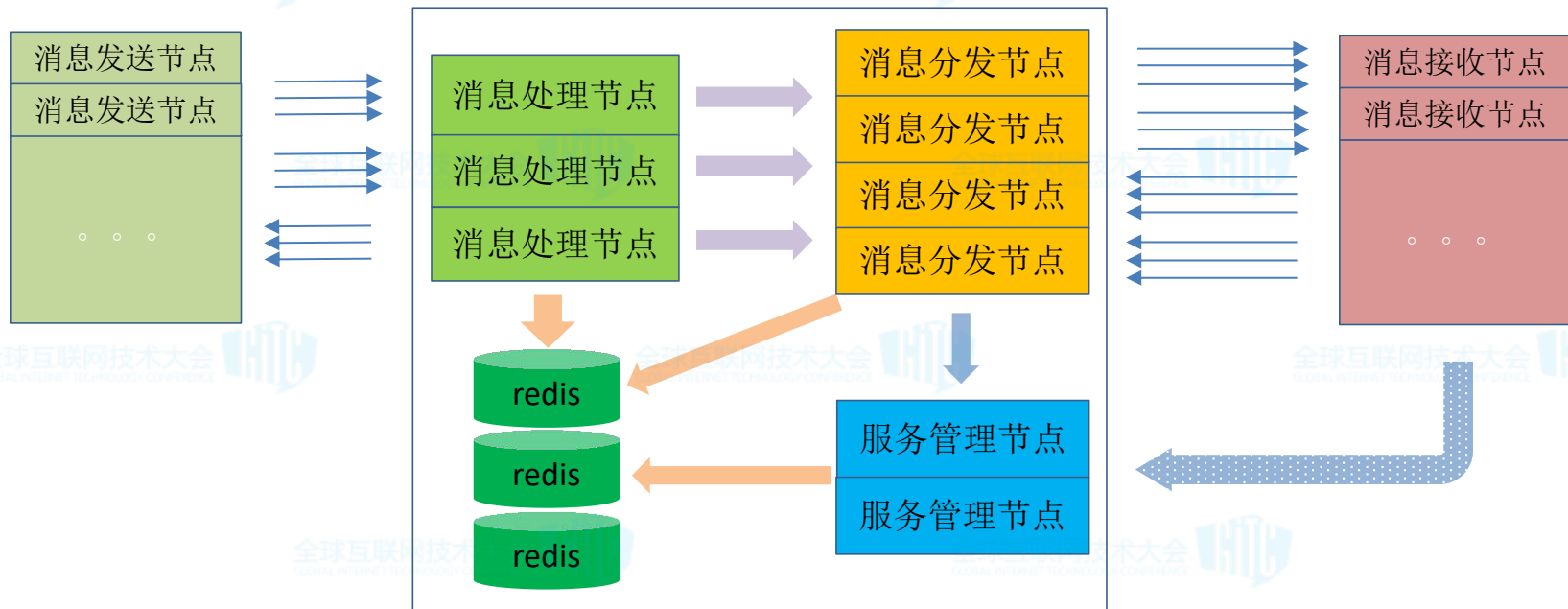


# 爱奇艺 CDN 回源方案解析

特性	说明
高并发	采用网络协程方式，支持高并发接入，同时简化程序设计
高性能	采用线程池 + 协程 + 连接池 + 内存池技术，提高业务处理性能
高吞吐	采用磁盘内存映射及零拷贝技术，提升磁盘及网络 IO 吞吐能力
低回源	合并相同请求，支持部分回源及部分缓存，大大降低回源带宽
开播快	采用流式数据读取方式，提升视频开播速度
可扩展	模块化分层设计，易于扩展新功能
易维护	采用统一服务器编程框架，易管理，好维护



# 爱奇艺 CDN 分发同步系统架构



## 方案特点：

- 1、消息处理模块、消息分发模块、服务管理模块均为多节点部署方式，消除故障单点；
- 2、采用集群版 redis 存储状态数据，降低消息丢失，同时提高消息处理能力；
- 3、服务管理模块检控所有服务的运行状态，可动态增加、删除服务节点；
- 4、消息接收模块通过服务管理模块自动获得消息分发模块接入点，可自动负载均衡及容错。

# 爱奇艺高性能 DNS

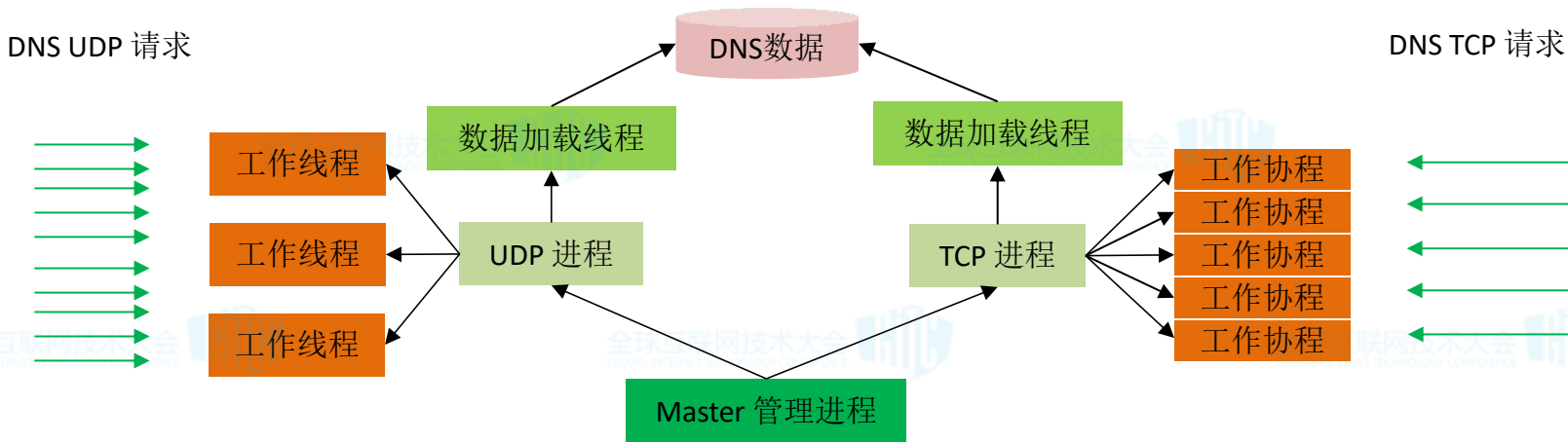
- 爱奇艺为什么要开发自己 DNS ?
  - 1、爱奇艺本身业务的需要，如 CDN 智能调度的需要
  - 2、高性能 QPS 以满足业务频繁访问需要，同时应对一定 DDos 攻击

- 开发 DNS 面临的挑战：

- 1、目标达到百万级 QPS
- 2、QPS 与 View 数量无关
- 3、Zone 频繁更新不影响正常服务
- 4、主从切换无感知



# 爱奇艺高性能 DNS 架构



## 一、UDP 服务进程:

- 1、采用 RUSEPORT 方式，提升工作线程并行收包能力
- 2、采用 recvmmsg/sendmmsg 多读多写方式，提升每次 IO 包收发能力
- 3、采用内存预分配策略，减少内存动态分配/释放时的“锁”冲突

## 二、TCP 服务进程：每个 TCP 请求分配一个协程进行业务处理

## 三、UDP 及 TCP 服务进程在更新数据时采用 RCU 方式

## 四、UDP 及 TCP 服务进程均由 master 管理进程统一管理

# 爱奇艺高性能 DNS 架构分析

全球互联网技术大会

优势	说明
高性能	启用 Linux 3.0 内核的 REUSEPORT 功能，提升多线程并行收发包的能力
	采用 Linux 3.0 内核的 recvmsg/sendmsg API，提升单次 IO 数据包收发能力
	采用内存预分配策略，减少内存动态分配/释放时的“锁”冲突
	针对 TCP 服务模式，采用网络协程框架，最大化 TCP 并发处理能力
高可用	采用 RCU（Read Copy Update）方式更新视图数据及配置项，无需停止服务，且不影响性能
	网卡 IP 地址变化自动感知（即可自动添加新 IP 或摘除老 IP 而不必停止服务）
	采用 Keepalived 保证服务高可用
易管理	由 Acl master 服务管理模块管理 DNS 进程，控制 DNS 进程的启动、停止、重读配置/数据、异常重启及异常报警等

全球互联网技术大会

全球互联网技术大会

全球互联网技术大会

# 爱奇艺生产环境数据真实展现

- 1、视频卡顿比  $< 2\%$
- 2、CDN 视频回源带宽  $< 1\%$
- 3、高性能 DNS 处理能力（非DPDK版本）：单机性能  $> 200$  万次/秒
- 4、单机数据分发能力：视频文件达**千万/日**，静态图片文件达**亿/日**



# 欢迎加入爱奇艺大家庭

- 公司总部：北京-爱奇艺创新大厦
- 地址：北京海淀区海淀北一街2号爱奇艺创新大厦11层
- 邮编：100080
- 电话：+86 10 6267.7171
- 传真：+86 10 6267 7000
- 联系方式：<http://zhaopin.iqiyi.com/>



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



悦 享 品 质