Operations

sina新浪

“秒开”体验背后的CDN优化之路

新浪内容加速平台 - 徐永健

2018年5月

Operations 新浪

目录

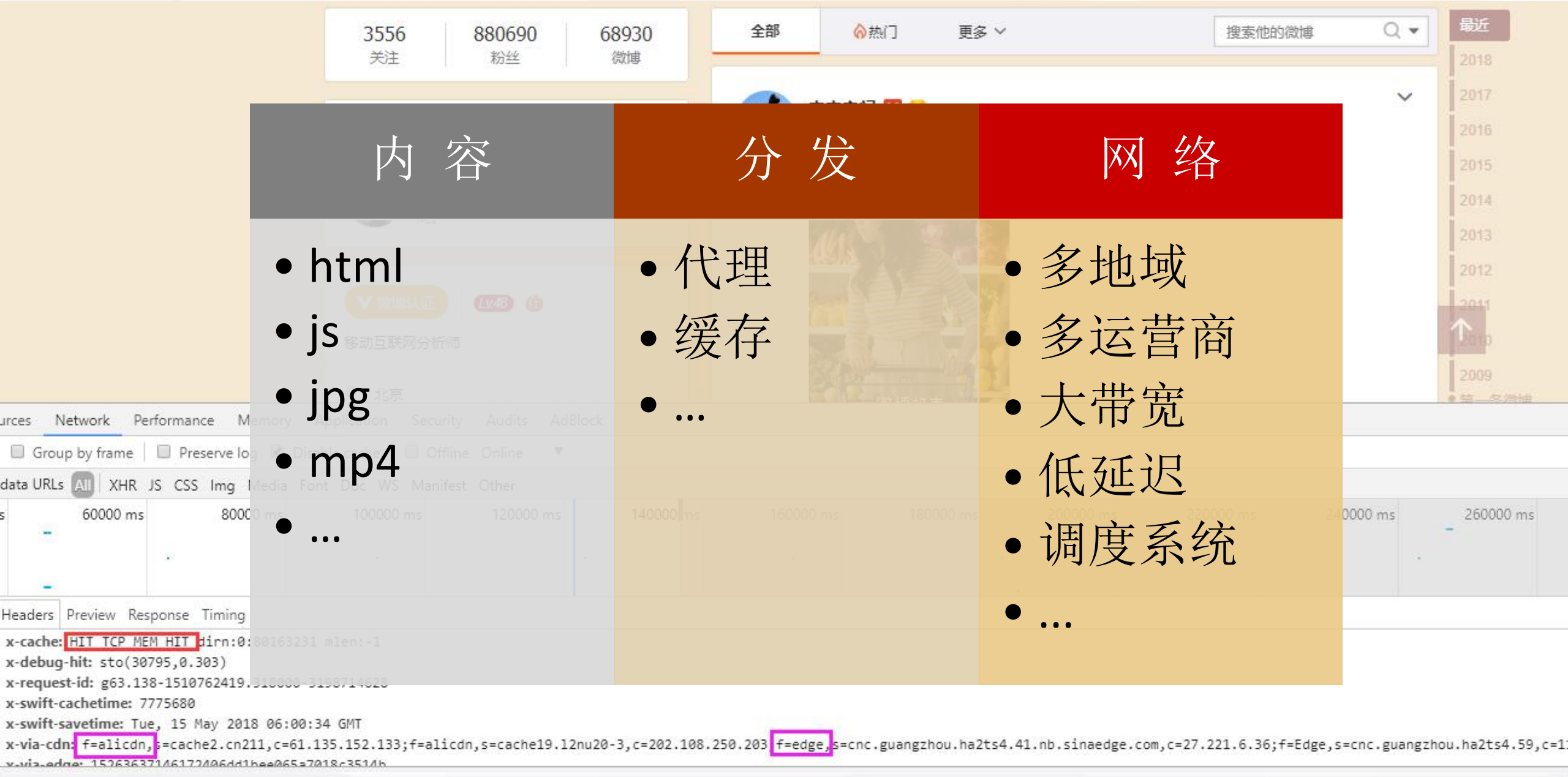
SinaEdge 架构介绍

01

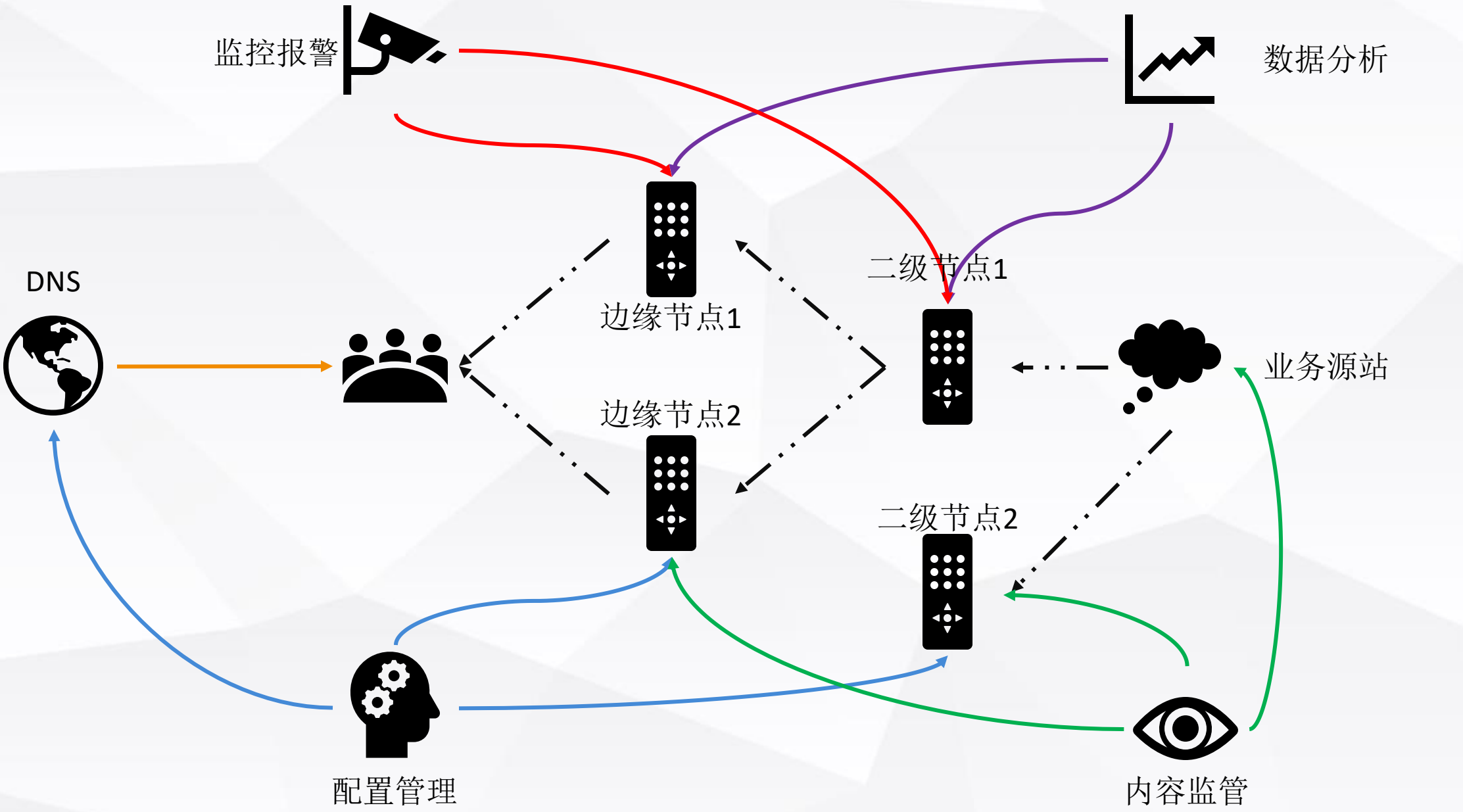
什么是CDN?

Operations 新浪

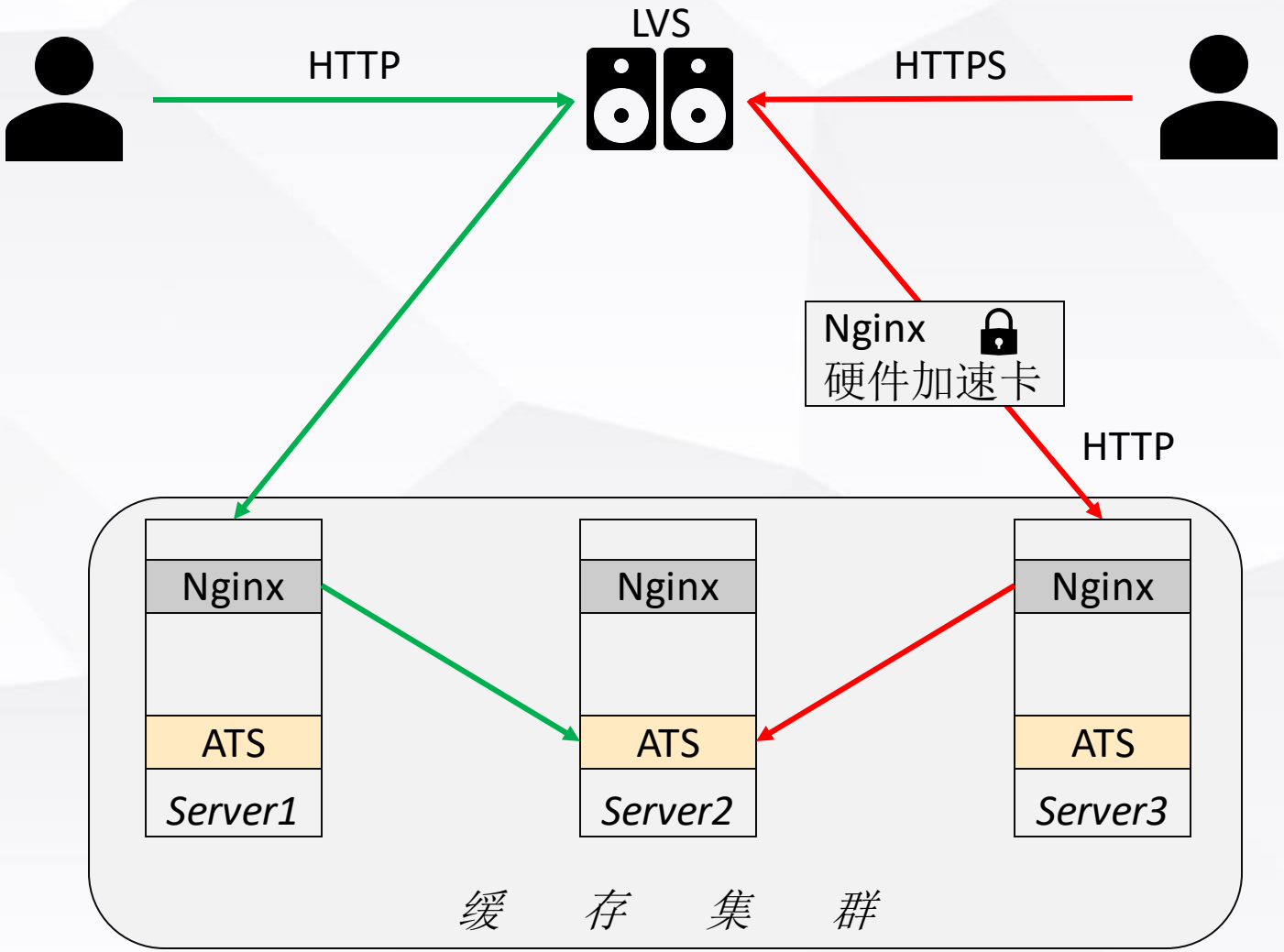
内 容	分 发	网 络
• html • js • jpg • mp4 • …	• 代理 • 缓存 • …	• 多地域 • 多运营商 • 大带宽 • 低延迟 • 调度系统 • …



系统架构



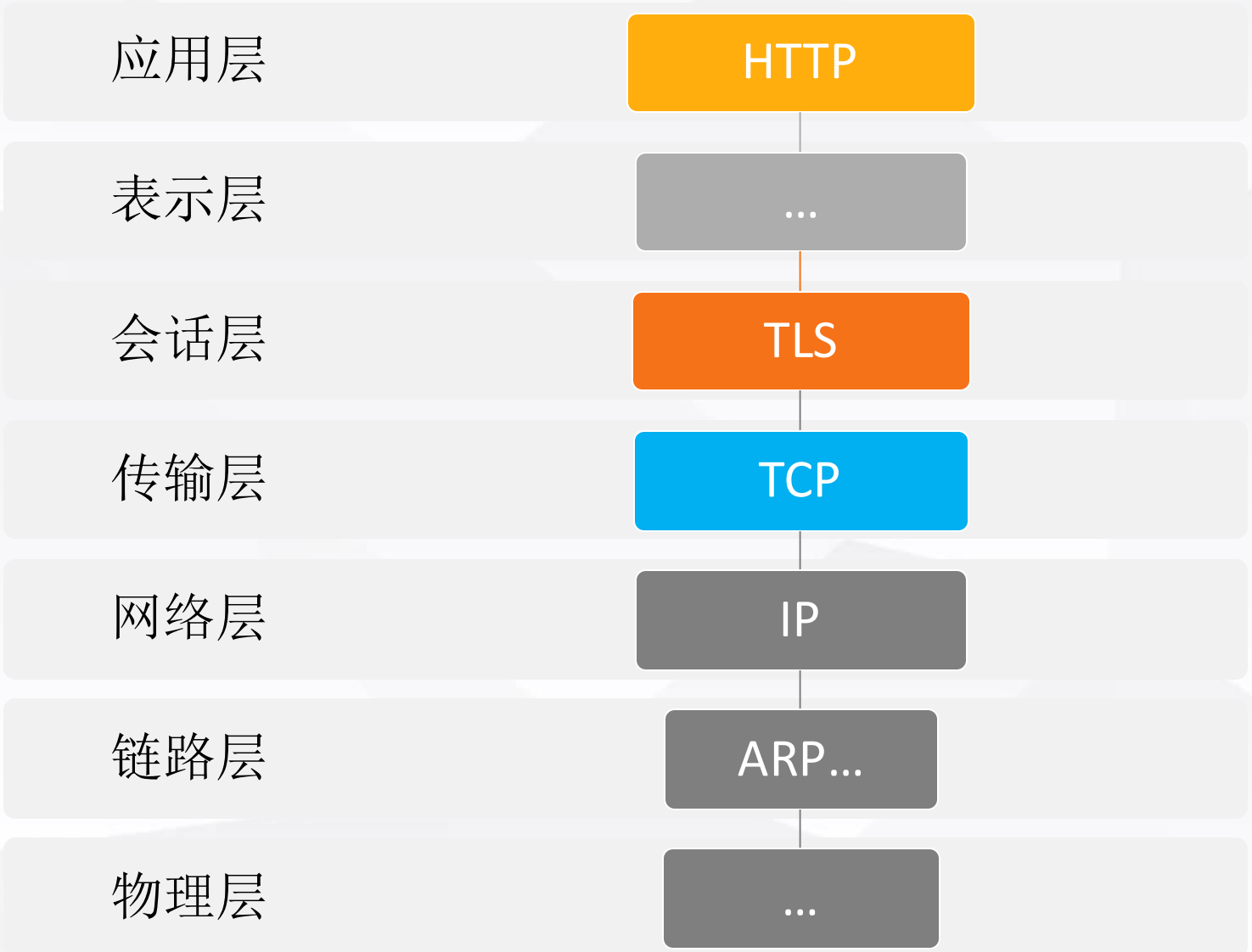
缓存节点架构


LVS
HTTP
HTTPS
Nginx
硬件加速卡
HTTP
Nginx
ATS
Server1
Nginx
ATS
Server2
Nginx
ATS
Server3
缓 存 集 群

CDN 服务优化要点

02

CDN 基础技术栈



基础服务

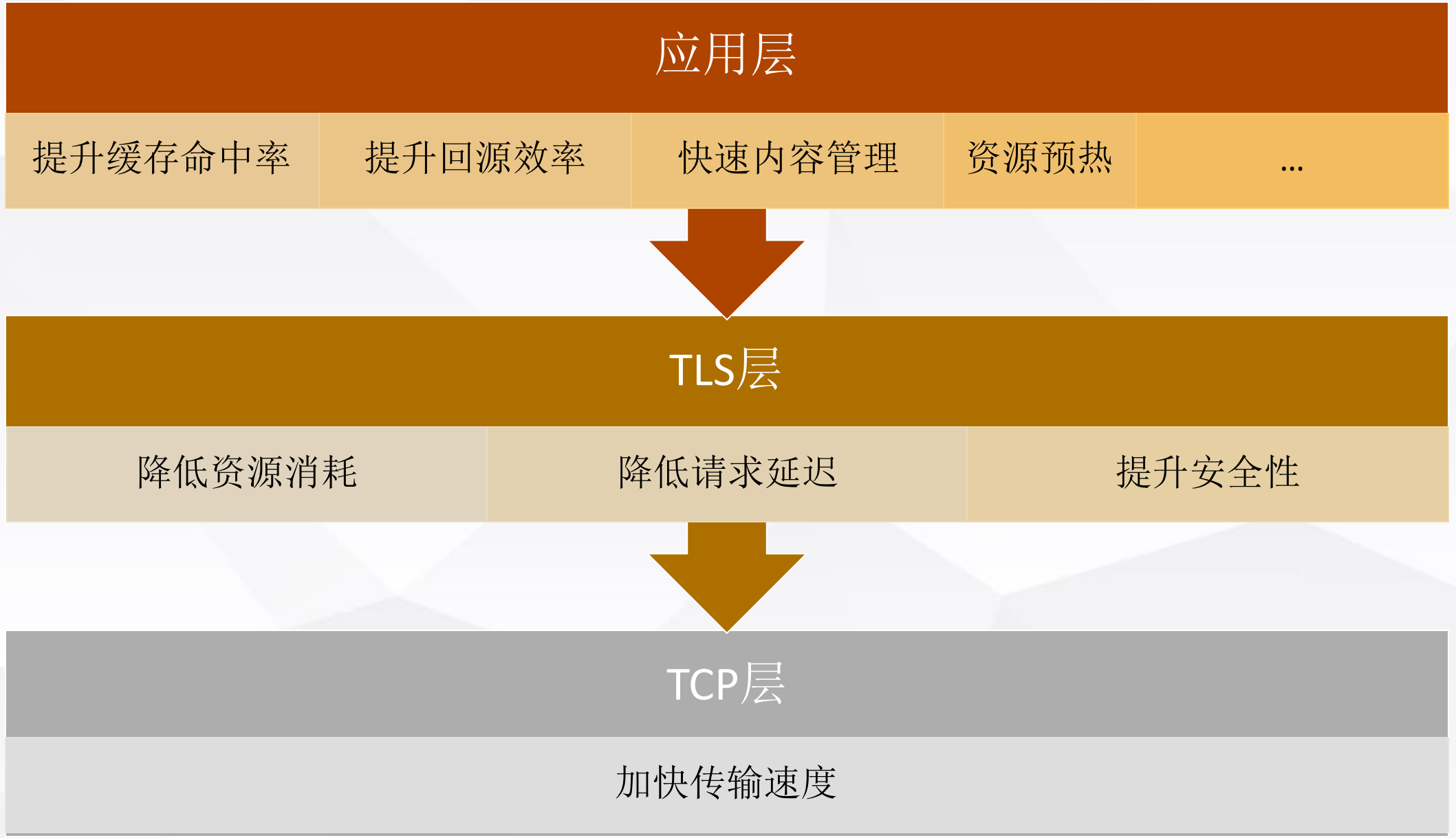
调度

四层负载均衡

七层负载均衡

缓存

CDN 服务分层优化要点



特殊服务场景与优化

03

用户A：xx视频特别火，你看了没？

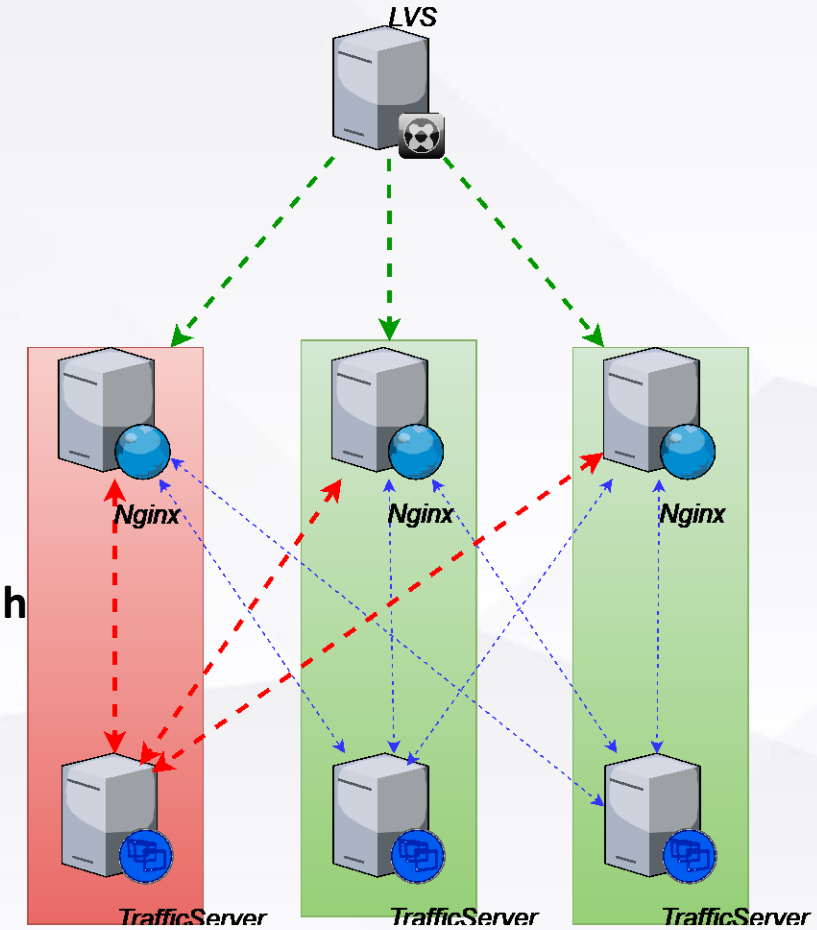
用户B：快发给我！

10秒钟后...视频还没有打开

热点文件问题与解决方案

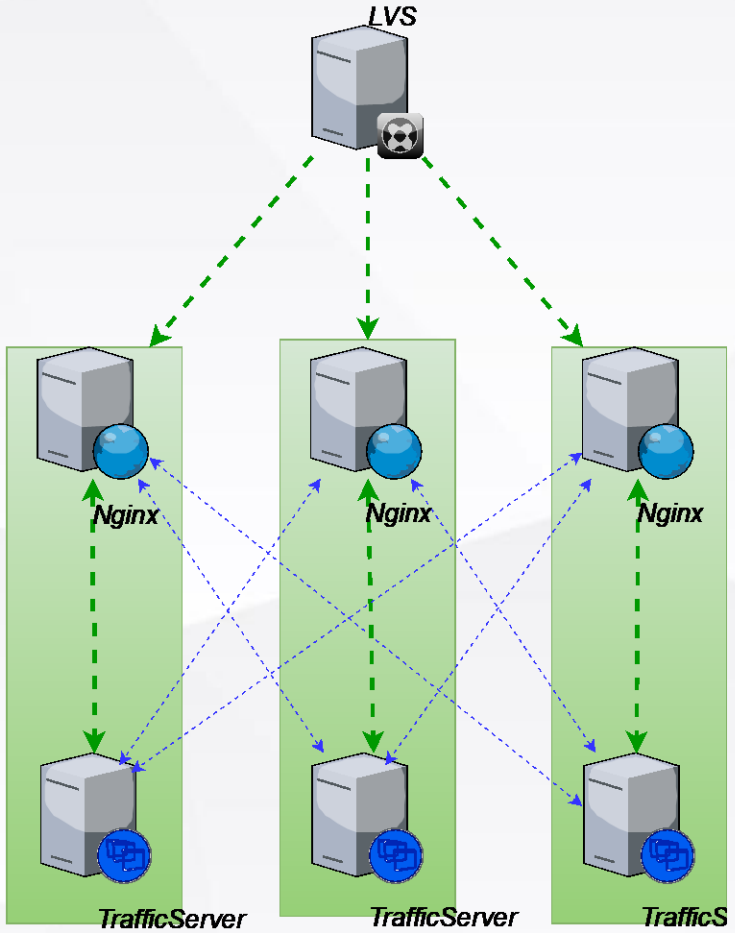
热点文件导致某缓存过载

一致性hash



发现热点文件后直接本地服务

Local only



小猿：老板，业务上线了！

老板：干的漂亮，晚餐加鸡腿！

1分钟后...服务器瘫痪

资源预热

Operations

新浪

应对情况	秒杀等活动页面资源
	定时更新或发布的资源
	准备新上业务
	...

预热等级	仿真预热（边缘节点）
	普通预热（二级节点）

网监：这张图片社会影响很不好！

小猿：稍等，我清理一下

1分钟后...服务器IP被封

URL快速封禁

Operations sina新浪

普通的缓存清除操作

PURGE源站 •耗时不确定，甚至无法清除

清理CDN二级节点 • < 30秒

清理CDN边缘节点 • < 30秒

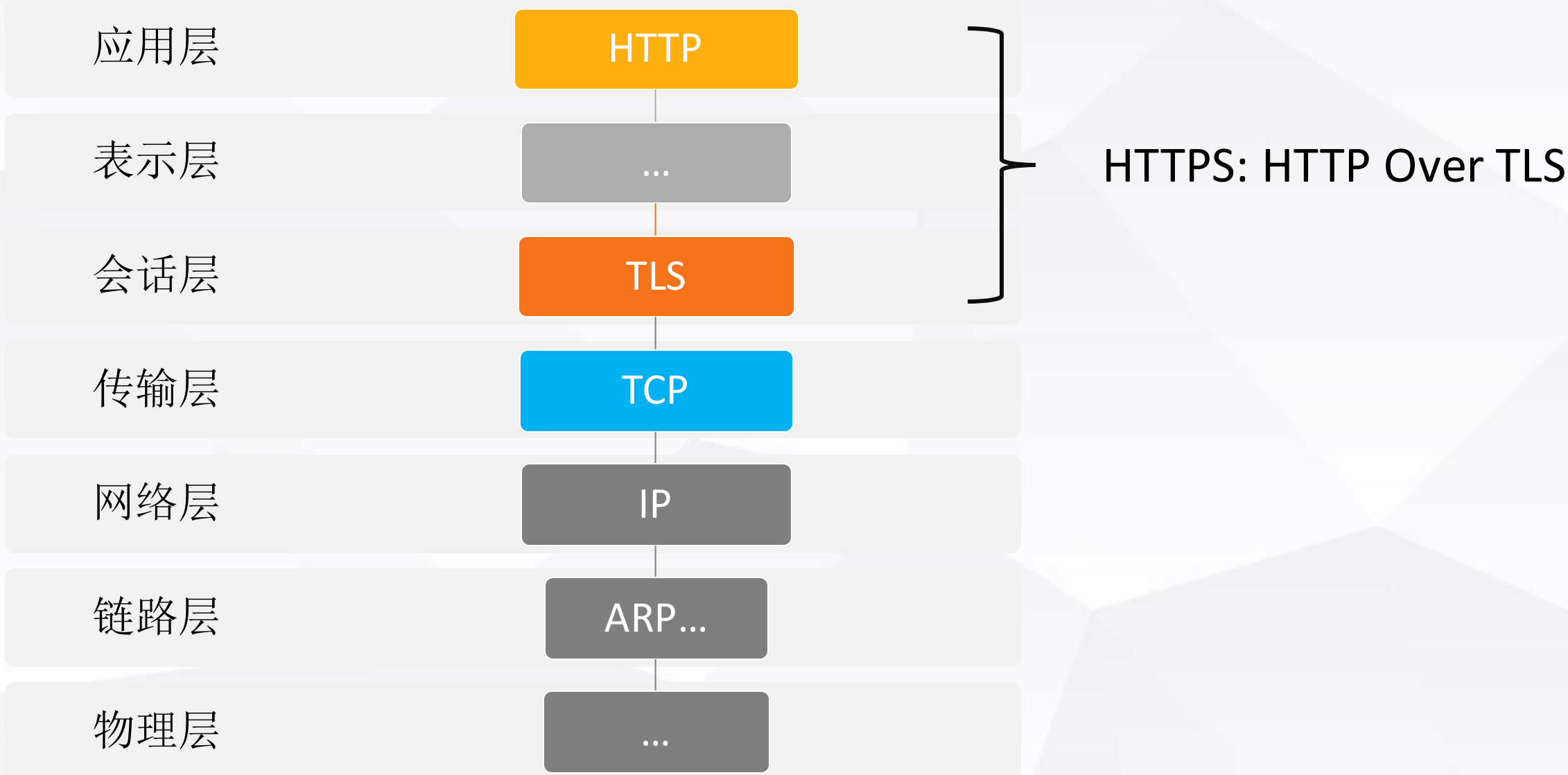
快速封禁

发送封禁请求至CDN边缘节点 • < 30秒

HTTPS 优化实践

04

HTTPS


应用层
HTTP
表示层
...
会话层
TLS
传输层
TCP
网络层
IP
链路层
ARP...
物理层
...
HTTPS: HTTP Over TLS

HTTPS 大势所趋

Operations

新浪

劫持

隐私泄露

钓鱼网站

...

HTTPS 带来新的问题

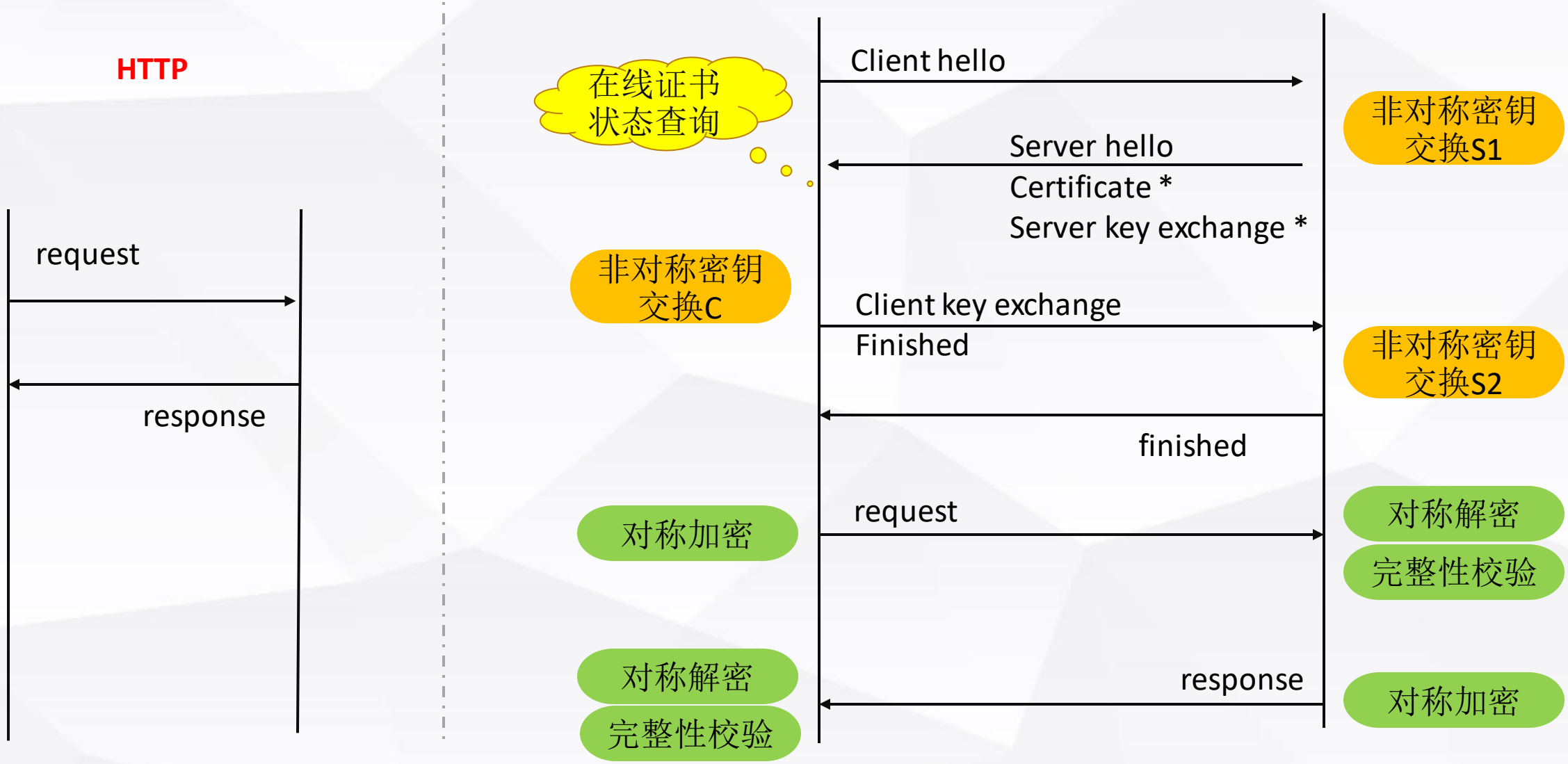
响应慢

服务器压力大

证书管理麻烦

...

HTTPS 慢且耗能


HTTP
request
response
HTTPS
在线证书
状态查询
Client hello
非对称密钥
交换S1
Server hello
Certificate *
Server key exchange *
非对称密钥
交换C
Client key exchange
Finished
非对称密钥
交换S2
finished
request
对称解密
完整性校验
对称加密
response
对称解密
完整性校验
对称加密

HTTPS 现有解决方案

在边缘节点中加入专门的HTTPS代理集群（带SSL硬件加速卡）

HTTPS 重点优化内容

Operations 新浪

降低消耗

硬件加速卡

远程异步计算

加密套件选择

...

减少延迟

Keepalive

HTTP/2

TLS session恢复

OCSP Stapling

...

安全增强

私钥安全

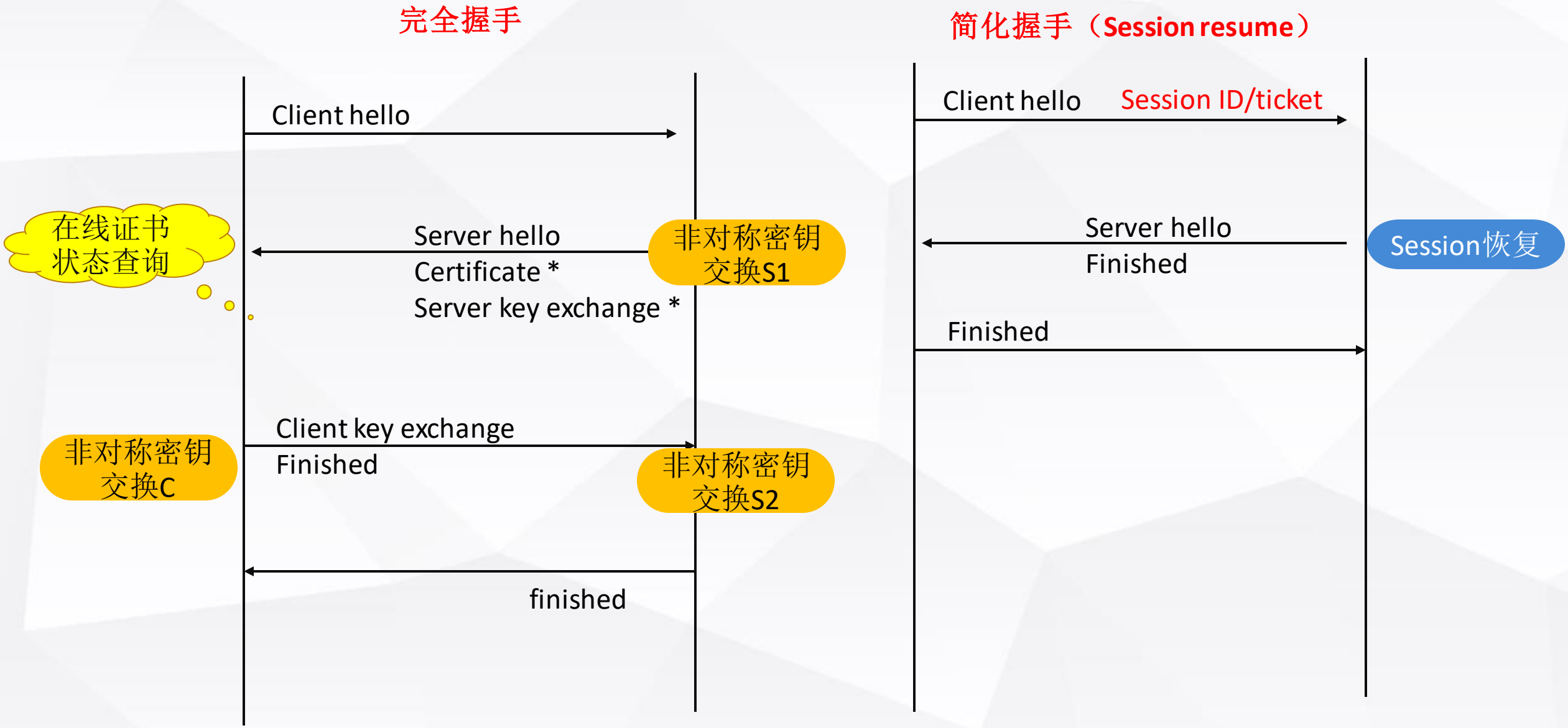
session ticket key

自动更换

HTTP强制跳转

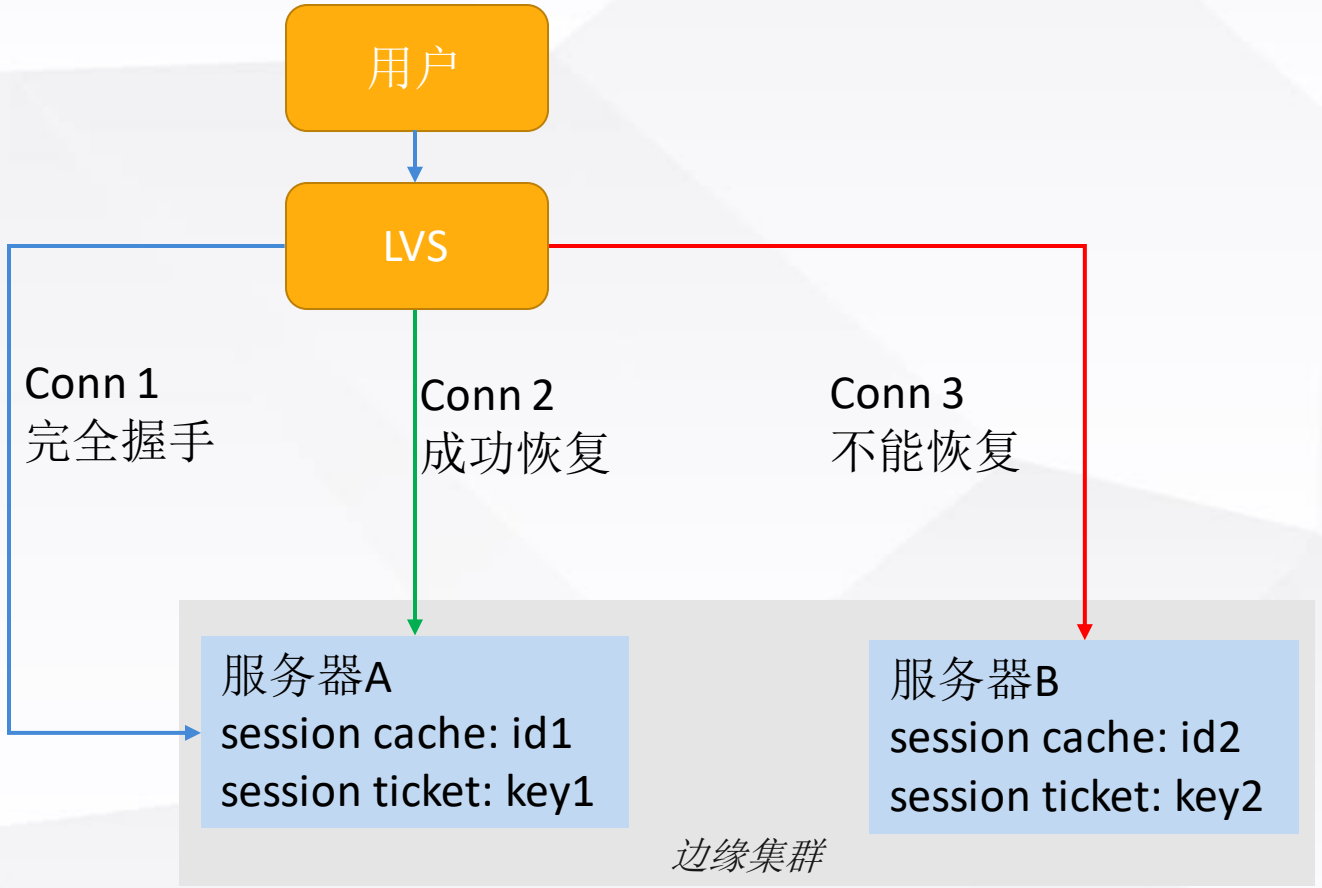
...

TLS session resume



TLS session resume的障碍

Nginx仅支持单机的session cache



我们还需要：

- session cache 多机共享
- session ticket key 多机同步

TLS session cache 多机共享

Operations

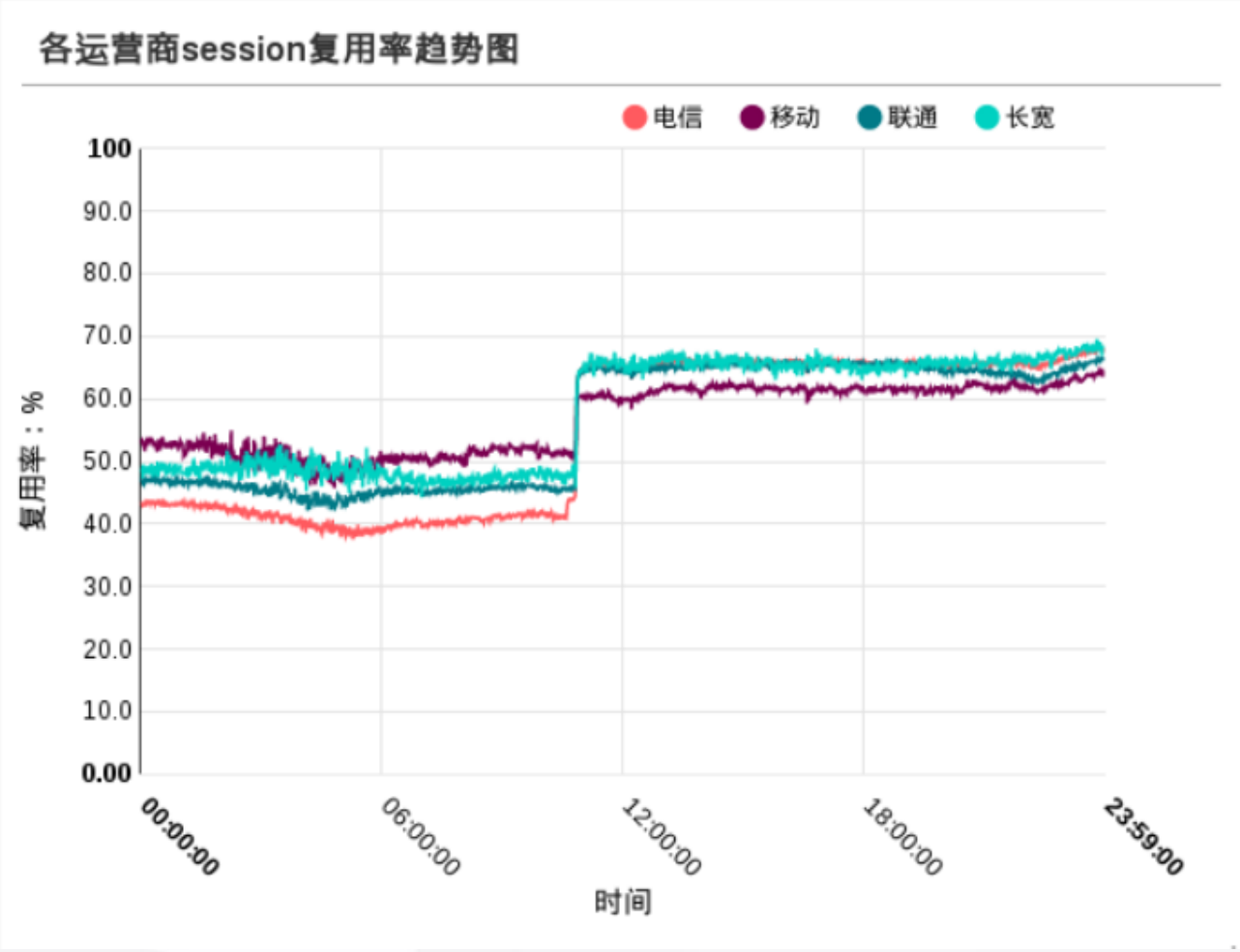
新浪

Nginx + lua

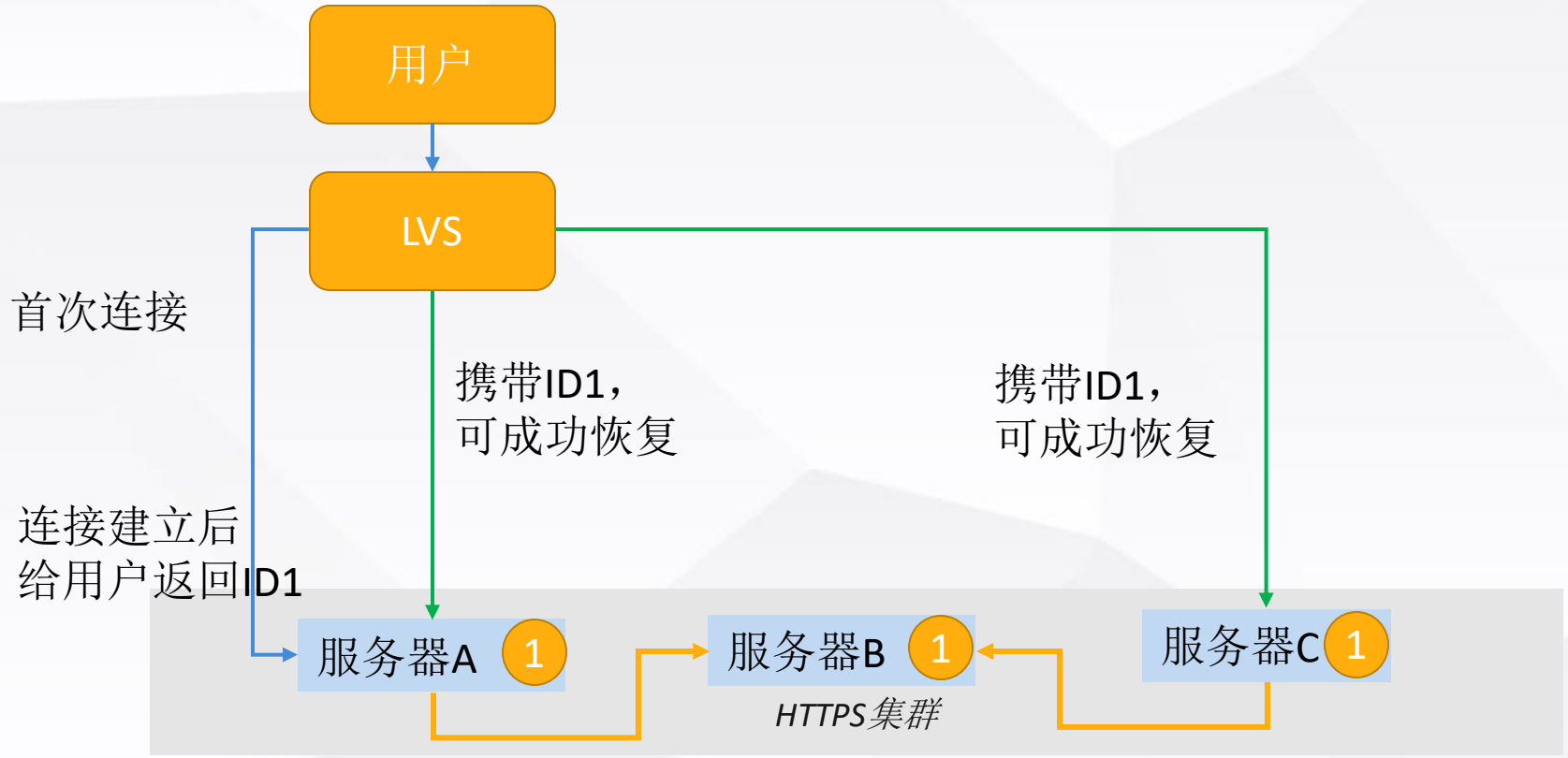
不依赖其他进程或中心节点

一致性 hash

健康检查



TLS session cache 多机共享



1 存储ID1到本机

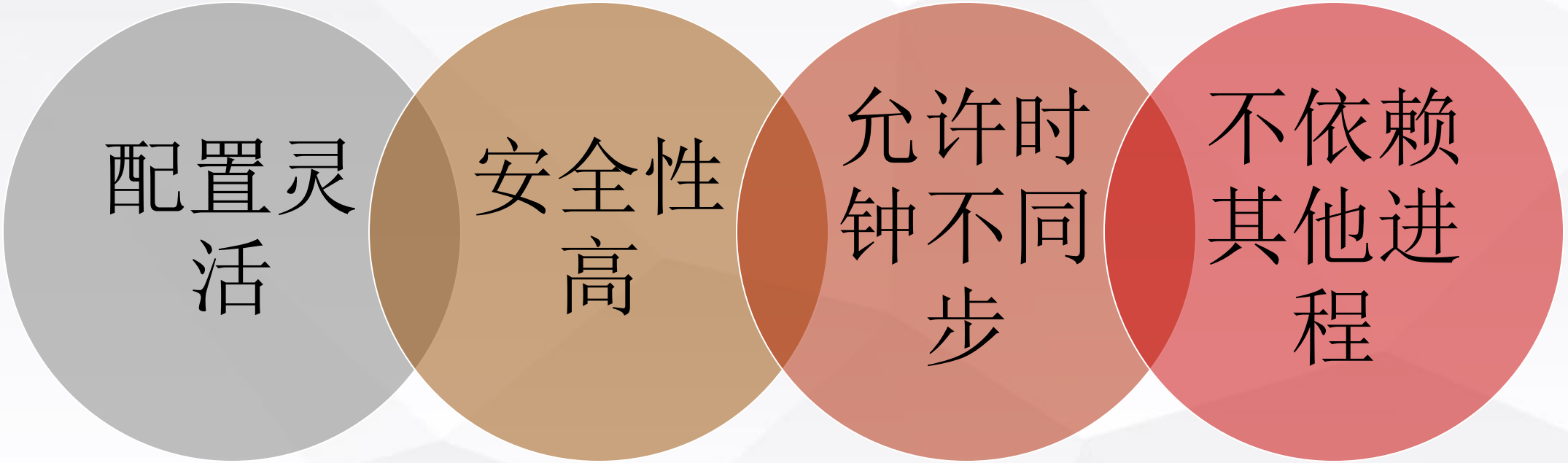
2 通过一致性hash运算出应该同时存储到服务器B

1 本机查找失败

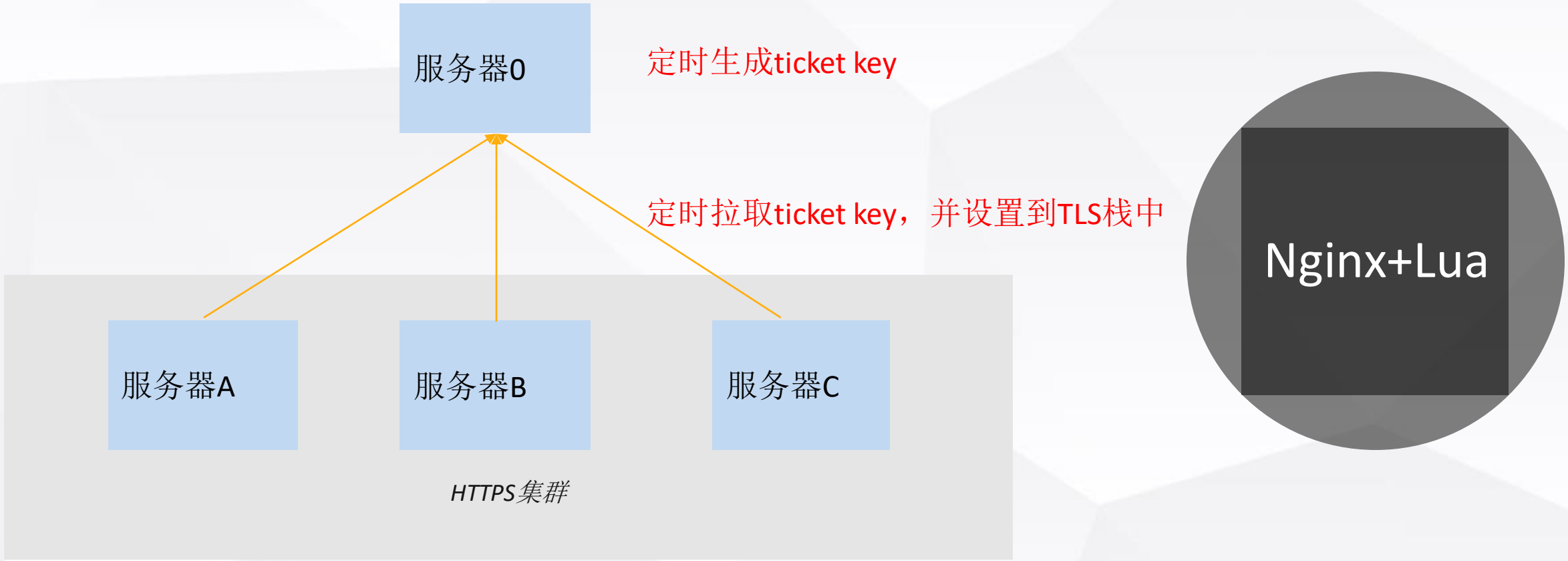
2 通过一致性hash运算出应该去服务器B查找

3 存储到本机

TLS session ticket key 多机同步与自动更新



TLS session ticket key 多机同步与自动更新



TLS session ticket key 兼容时钟不同步

Operations 新浪

仅链头Key用作生成Ticket

所有Key都可用来解密Ticket

预加载下一时刻Key到链尾
用于兼容时钟不同步

Server A（时钟快）	Server B
Key_20180531_15:00	**Key_20180531_14:00**
Key_20180531_14:00	Key_20180531_13:00
Key_20180531_13:00	Key_20180531_12:00
Key_20180531_12:00	Key_20180531_11:00
Key_20180531_11:00	Key_20180531_10:00
Key_20180531_10:00	Key_20180531_09:00
…	…
Key_20180531_16:00	Key_20180531_15:00

持续优化

Operations 新浪

HTTPS

- OCSP Stapling
- 远程异步计算（已有初版）
- Keyless（开发中）
- …

TCP

- fasttcp（自研TCP加速）
- BBR（已开始大规模部署）
- …

yongjian3@staff.sina.com.cn

徐永健 @ SINAEDGE

Operations | sina 新浪