Mars 在移动网络的探索和实践

garryyan(闫国跃)

自我介绍

2009.09-2012.10

- 武汉大学软件工程专业

2012.10-至今

- 跨平台基础组件
- 微信终端运维门户
- 高性能日志模块
- Mars 项目

内容大纲

Contents

01 **移动网络概述**

02 移动网络优化

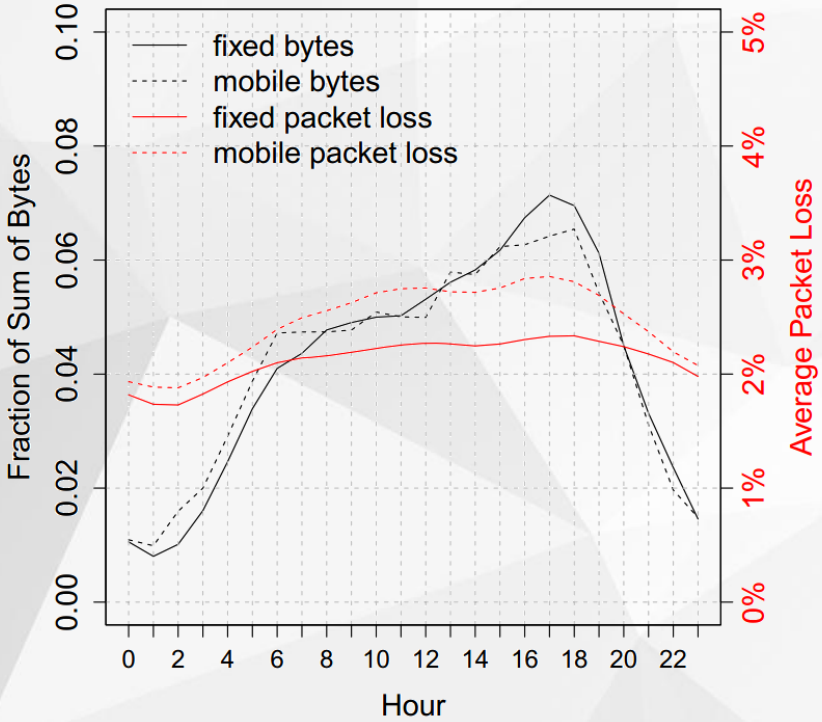
03 技术方案

04 Benchmark

05 如何快速接入

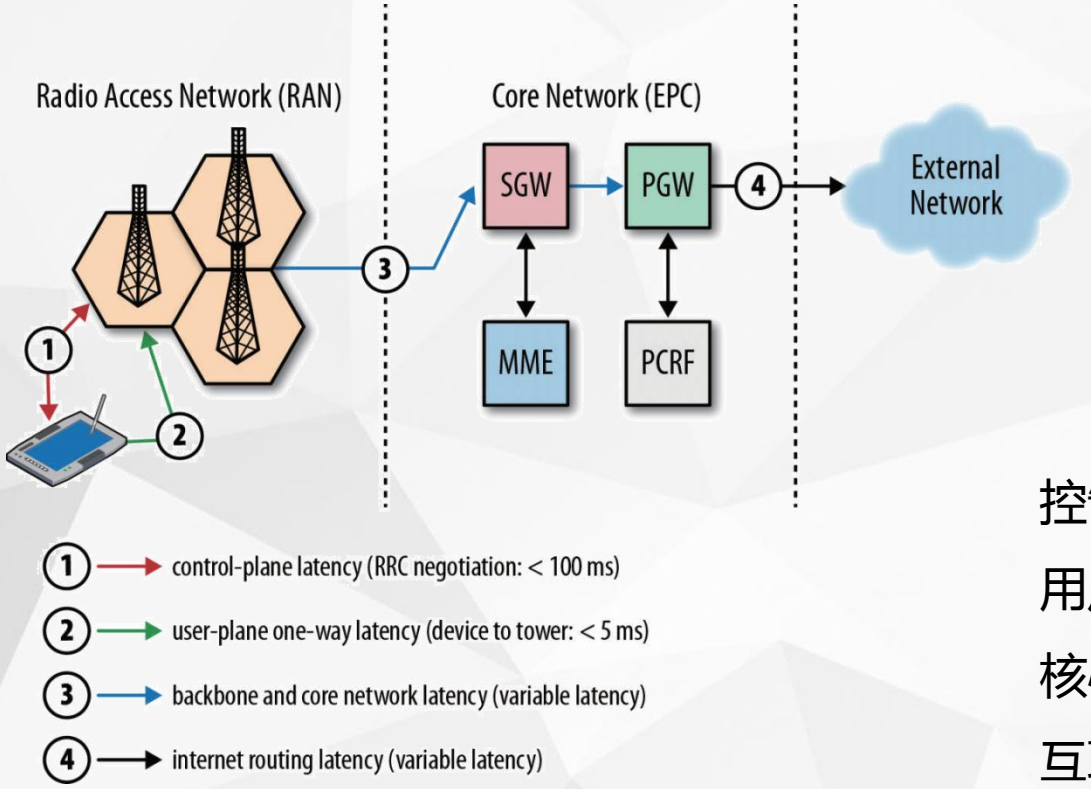
06 未来规划

丢包和误码



网络类型	BER
移动网络	10^{-4} 到10^{-6}
有线以太网	10^{-12}
光纤	10^{-15}

核心网络架构



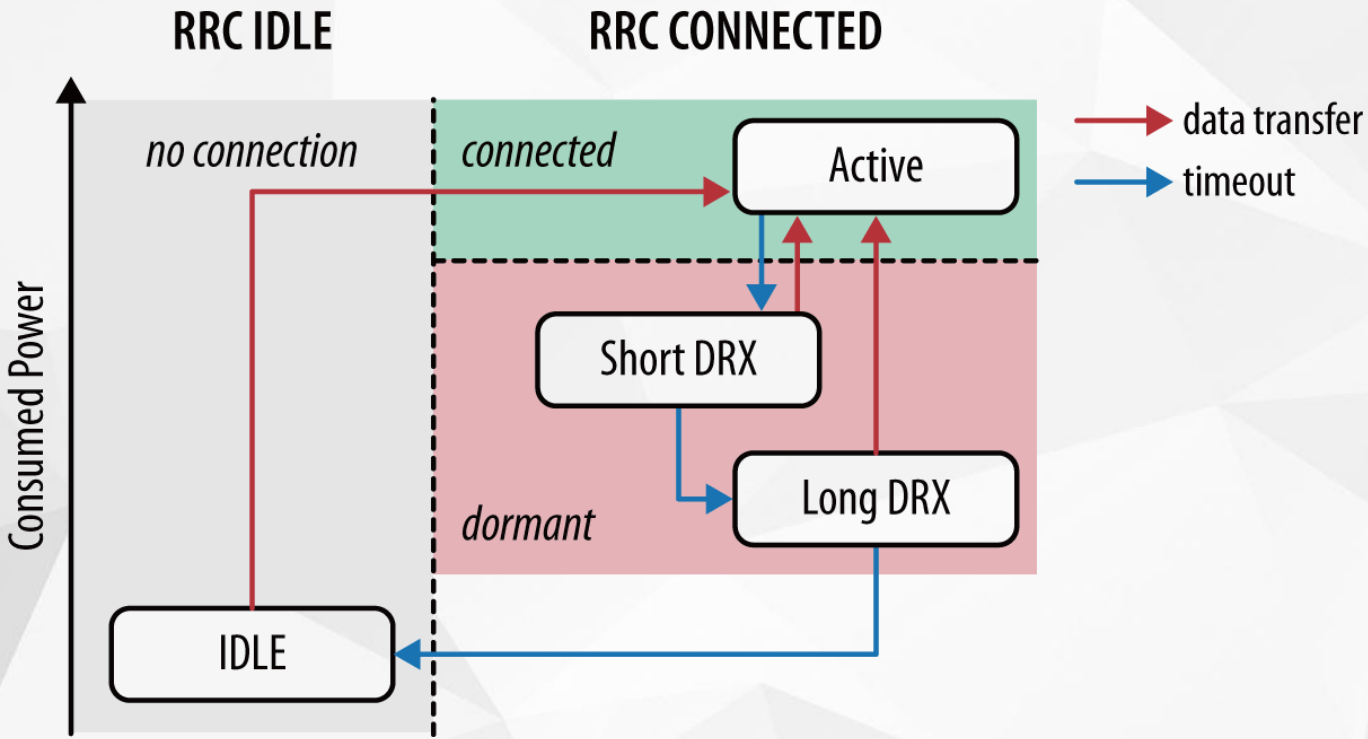
控制面延迟 < 100ms

用户面延迟 < 5ms

核心网路延迟 30-100ms

互联网路由延迟 不确定

RRC



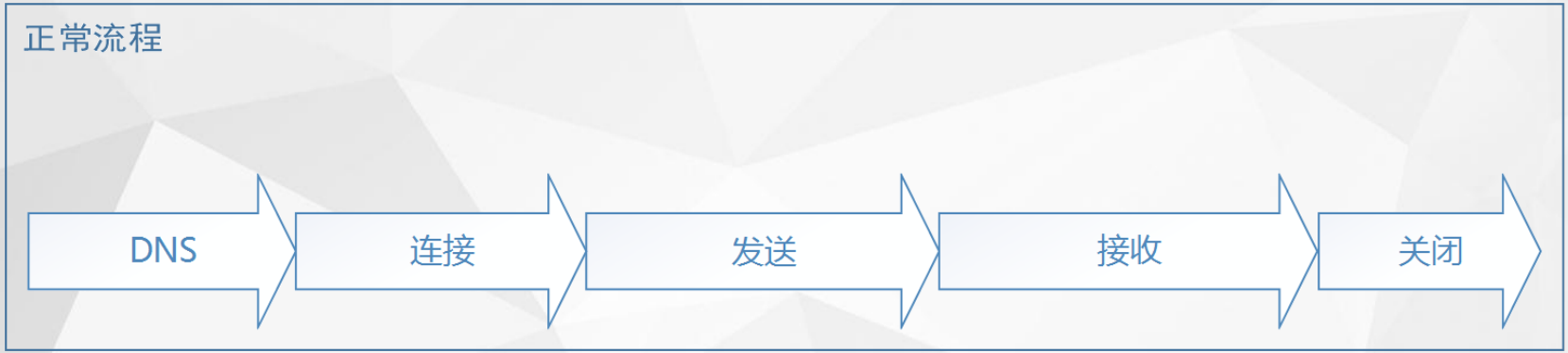
RRC空闲　只监听来自网络的控制信号，客户端没有无线电资源。
RRC连接　分配了专用的无线电资源，客户端可以收发数据。

移动网络特点

误码率高	丢包率高	不稳定的延迟	不稳定的带宽
• 环境电波 • 用户距离远	• 信号问题 • 用户过多 • 误码包 • 用户移动 • 基站切换	• 用户数量 • 信令分配 • 丢包 • 误码包	• 基站距离 • 用户数量 • 拥塞控制

行业水平

网络 / 来源	GRPS (2.5G)	EDGE (2.75G)	HSPA (3G)	HSPA+ (3.5G)	LTE (3.9G)	LTE+/WiMAX (4G)
AT&T	600-750 ms	600-750 ms	150-400 ms	100-200 ms	40-50 ms	-



100ms 100ms 50ms 50ms

RRC!

内容大纲

Contents

01 移动网络概述

02 **移动网络优化**

03 技术方案

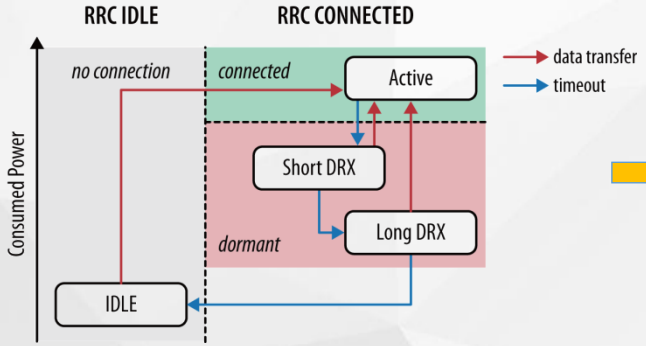
04 Benchmark

05 如何快速接入

06 未来规划

连接之前

RRC



SignallingKeeper

- UDP
- 不鼓励用
- 需要时再用

DNS

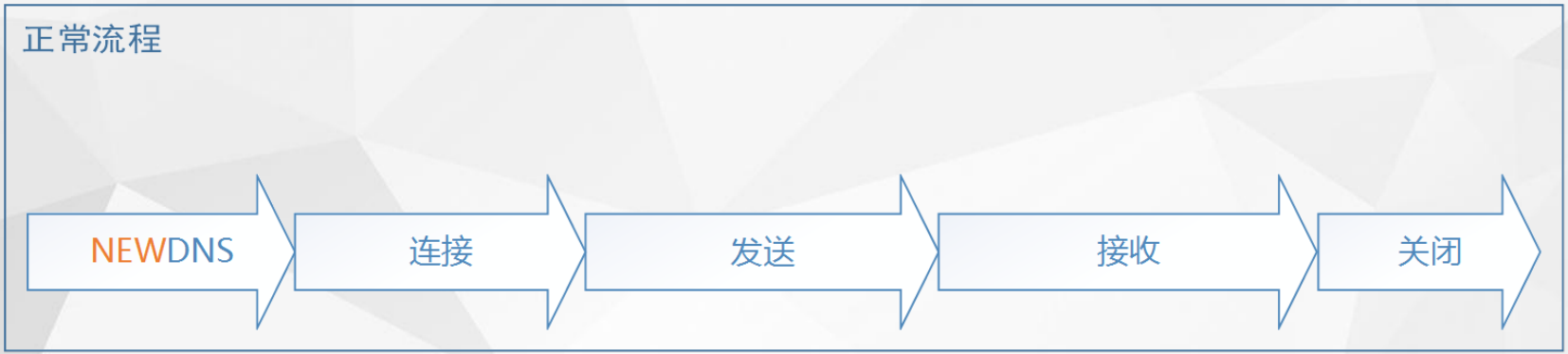
域名劫持

解析转发

更新缓慢

NewDNS

```
<domain name="your.domain1" timeout="1800">
    <ip>111.111.11.111</ip>
    <ip>111.111.11.112</ip>
</domain>
<domain name="your.domain2" timeout="1800">
    <ip>111.111.11.113</ip>
    <ip>111.111.11.114</ip>
</domain>
```


正常流程
NEWDNS
连接
发送
接收
关闭

复合连接

并发连接

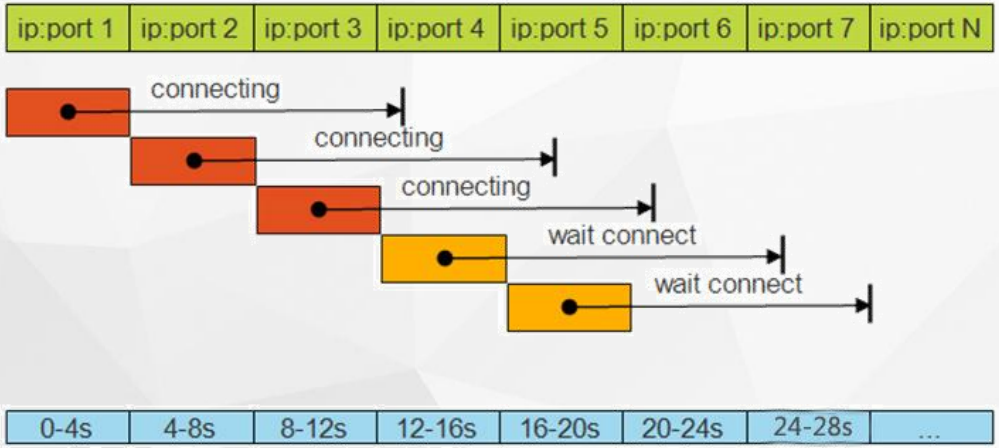
- 网络资源竞争
- 服务器负载
- 最快可用

串行连接

- 资源占用少
- 无服务器负载问题
- 超时选择困难
- 最慢可用

复合连接

复合连接



1. ip1+port1 0s连接，10s超时
2. ip2+port2 4s连接，14s超时
3. ip3+port3 8s连接，18s超时

…

任何一连接成功，其他连接关闭

✓ 连接成功率提升5%

✓ 更快找到可用链路和IP轮转

连接超时

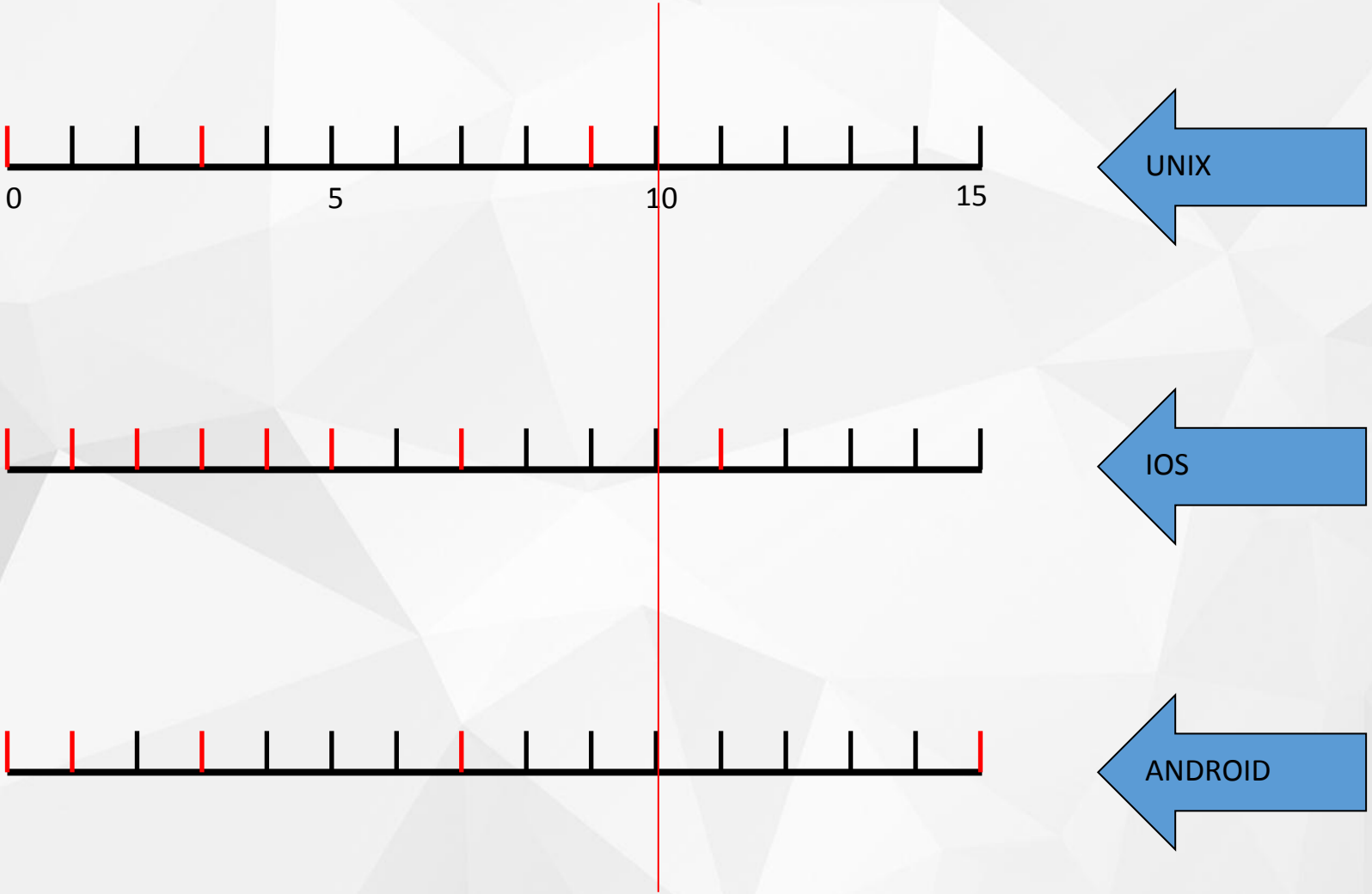
- Android超时机制（1, 2, 4, 8, 16）

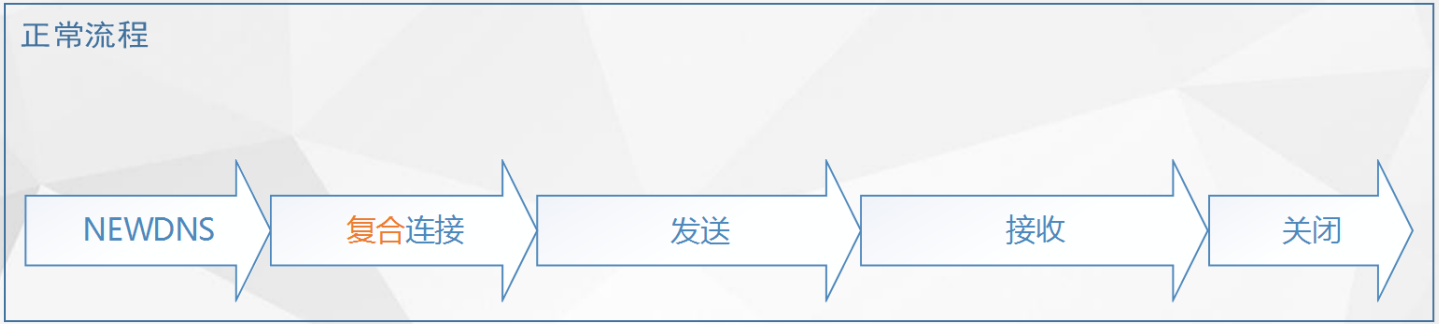
No.	Time	Source	Destination	Protocol	Length	Info
22	0.000000	192.168.1.223	121.42.149.193	TCP	76	33074 > http [SYN] Seq=0 Wi
26	0.991704	192.168.1.223	121.42.149.193	TCP	76	[TCP Retransmission] 33074
42	2.000020	192.168.1.223	121.42.149.193	TCP	76	[TCP Retransmission] 33074
54	4.010132	192.168.1.223	121.42.149.193	TCP	76	[TCP Retransmission] 33074
69	8.009847	192.168.1.223	121.42.149.193	TCP	76	[TCP Retransmission] 33074
105	16.040015	192.168.1.223	121.42.149.193	TCP	76	[TCP Retransmission] 33074

- iOS超时机制（1, 1, 1, 1, 1, 2, 4, 8, 16, 32）

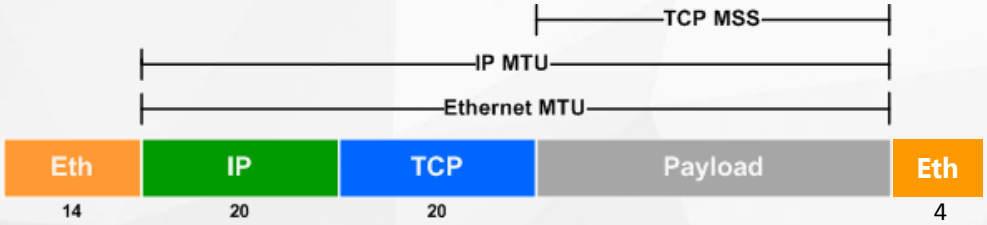
3	0.000000	172.20.10.5	121.42.149.193	TCP	78	50919 > http [SYN] Seq=0 Wi
6	1.003969	172.20.10.5	121.42.149.193	TCP	78	[TCP Retransmission] 50919
8	0.998387	172.20.10.5	121.42.149.193	TCP	78	[TCP Retransmission] 50919
9	1.001082	172.20.10.5	121.42.149.193	TCP	78	[TCP Retransmission] 50919
11	1.001123	172.20.10.5	121.42.149.193	TCP	78	[TCP Retransmission] 50919
13	1.000357	172.20.10.5	121.42.149.193	TCP	78	[TCP Retransmission] 50919
17	2.002216	172.20.10.5	121.42.149.193	TCP	78	[TCP Retransmission] 50919
19	4.000908	172.20.10.5	121.42.149.193	TCP	78	[TCP Retransmission] 50919
23	8.003164	172.20.10.5	121.42.149.193	TCP	78	[TCP Retransmission] 50919
52	16.011359	172.20.10.5	121.42.149.193	TCP	78	[TCP Retransmission] 50919
73	32.005287	172.20.10.5	121.42.149.193	TCP	62	[TCP Retransmission] 50919

连接超时




正常流程
NEWDNS
复合连接
发送
接收
关闭

Nagle



糊涂窗口综合症

Nagle

TCP_NODELAY

发送超时

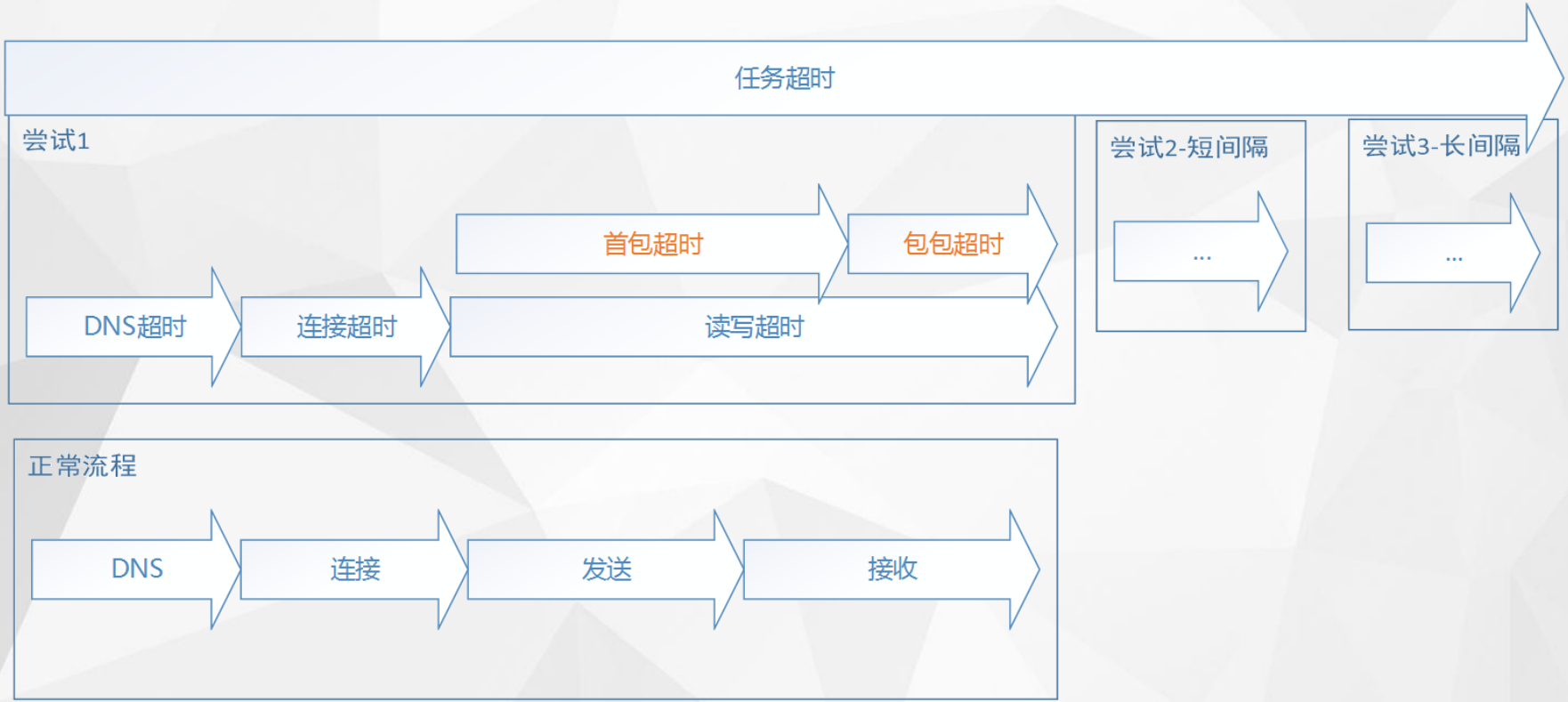
[1，2，4，8，16，32，64，64，64 ...]

```
1    0.0                bsdi.1029 > svr4.discard: S 1747921409:1747921409(0)
```

762	1.001761	172.20.10.3	121.42.149.193	TCP	63 [TCP Retransmission]
764	1.002003	172.20.10.3	121.42.149.193	TCP	63 [TCP Retransmission]
765	1.002253	172.20.10.3	121.42.149.193	TCP	63 [TCP Retransmission]
766	1.900673	172.20.10.3	121.42.149.193	TCP	63 [TCP Retransmission]
775	4.622541	172.20.10.3	121.42.149.193	TCP	63 [TCP Retransmission]
777	9.037153	172.20.10.3	121.42.149.193	TCP	63 [TCP Retransmission]
789	13.458430	172.20.10.3	121.42.149.193	TCP	63 [TCP Retransmission]
798	26.727231	172.20.10.3	121.42.149.193	TCP	63 [TCP Retransmission]
799	26.704635	172.20.10.3	121.42.149.193	TCP	63 [TCP Retransmission]
803	26.707128	172.20.10.3	121.42.149.193	TCP	63 [TCP Retransmission]
811	27.461498	172.20.10.3	121.42.149.193	TCP	63 [TCP Retransmission]
814	25.967649	172.20.10.3	121.42.149.193	TCP	63 [TCP Retransmission]
823	26.713700	172.20.10.3	121.42.149.193	TCP	63 [TCP Retransmission]
827	26.720471	172.20.10.3	121.42.149.193	TCP	40 61837 > http [RST, AC

```
13   182.488164 (64.0018)  bsdi.1029 > svr4.discard: P 15:23(8) ack 1 win 4096
14   246.489921 (64.0018)  bsdi.1029 > svr4.discard: P 15:23(8) ack 1 win 4096
15   310.491678 (64.0018)  bsdi.1029 > svr4.discard: P 15:23(8) ack 1 win 4096
16   374.493431 (64.0018)  bsdi.1029 > svr4.discard: P 15:23(8) ack 1 win 4096
17   438.495196 (64.0018)  bsdi.1029 > svr4.discard: P 15:23(8) ack 1 win 4096
18   502.486941 (63.9917)  bsdi.1029 > svr4.discard: P 15:23(8) ack 1 win 4096
19   566.488478 (64.0015)  bsdi.1029 > svr4.discard: R 23:23(0) ack 1 win 4096
```

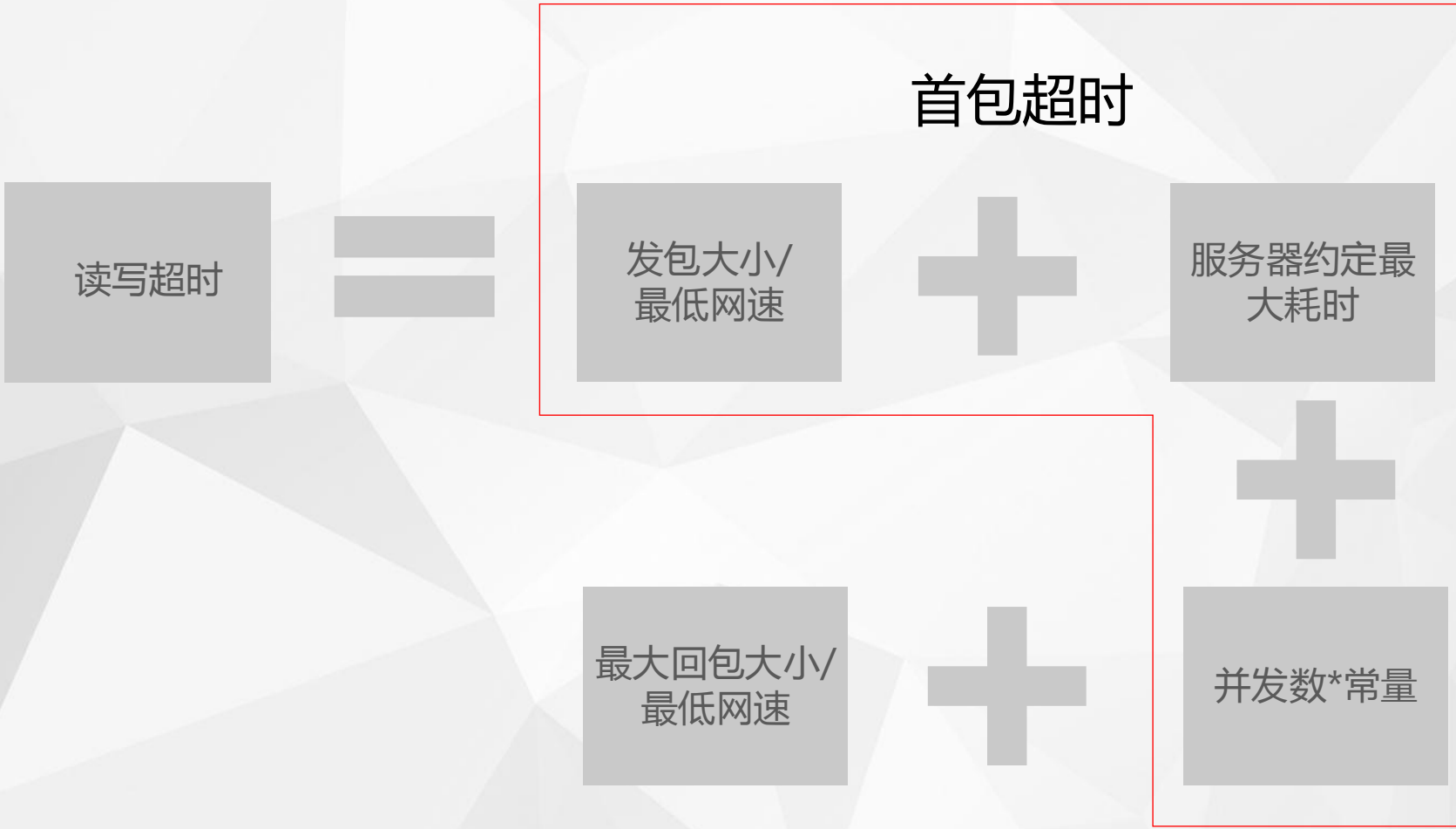
发送超时


任务超时
尝试1
首包超时
包包超时
DNS超时
连接超时
读写超时
尝试2-短间隔
...
尝试3-长间隔
...
正常流程
DNS
连接
发送
接收

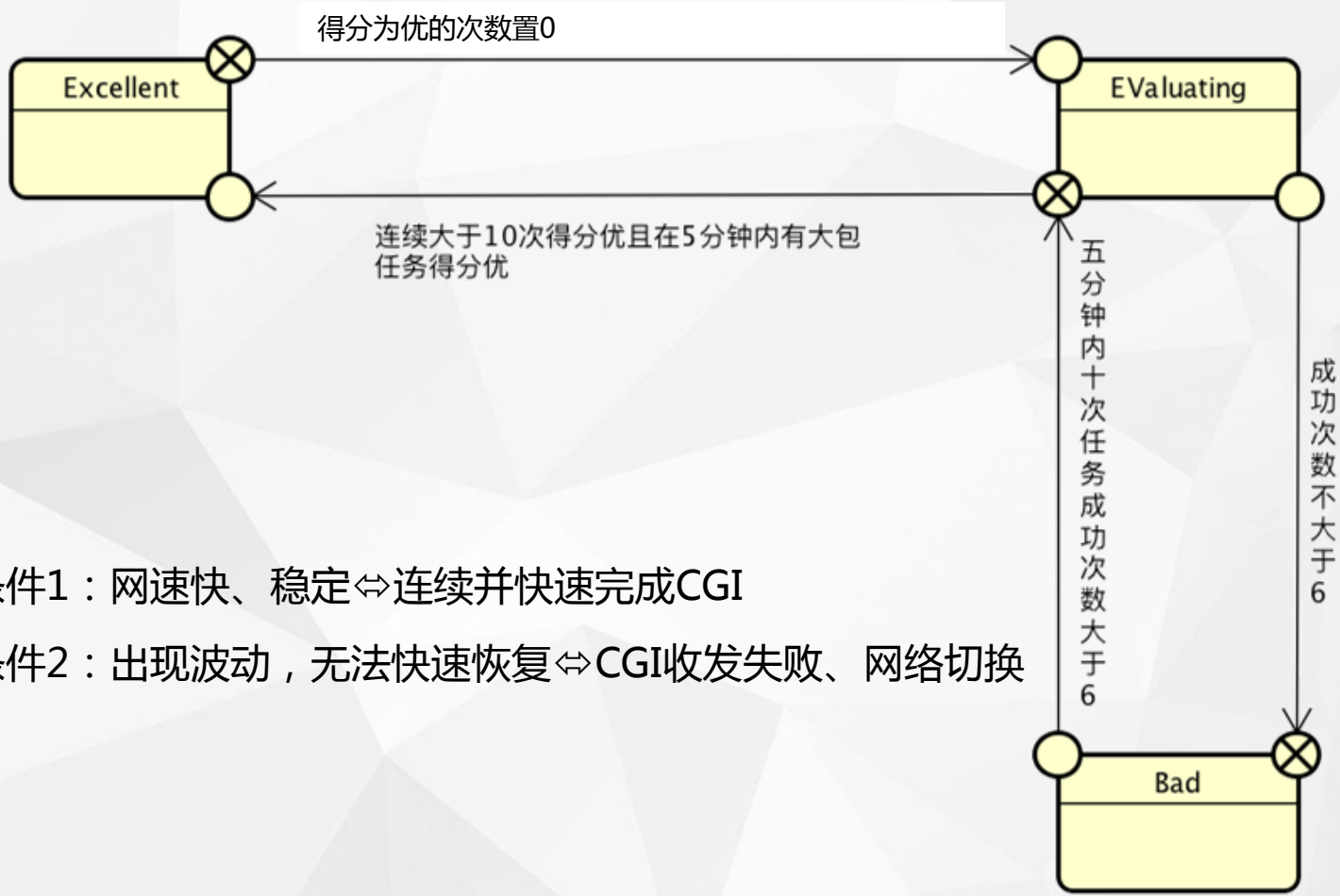
首包超时

首包超时 = 发包大小/最低网速 + 服务器约定最大耗时 + 并发数*常量

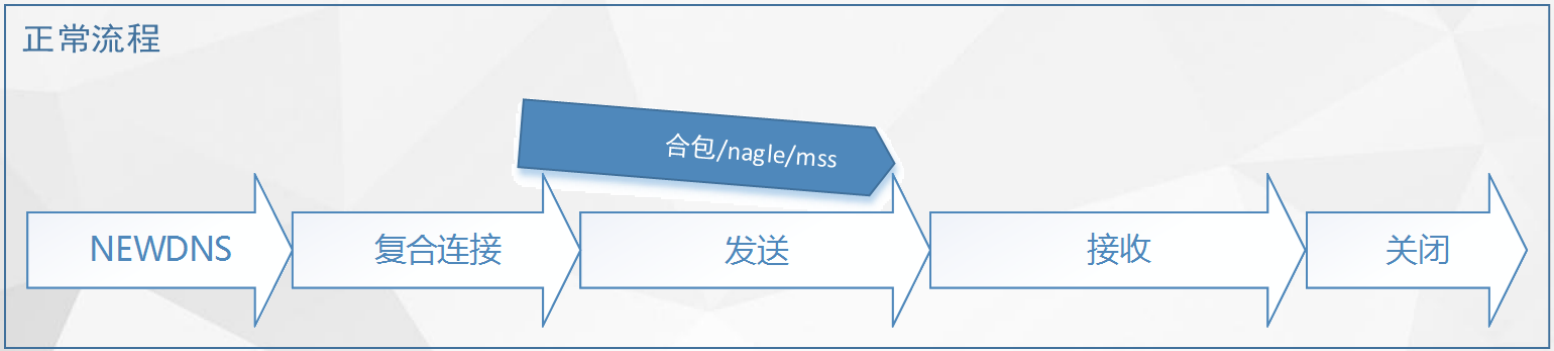
读写超时



动态超时



- 扭转条件1：网速快、稳定⇔连续并快速完成CGI
- 扭转条件2：出现波动，无法快速恢复⇔CGI收发失败、网络切换


正常流程
合包/nagle/mss
NEWDNS
复合连接
发送
接收
关闭

- 其实“接收”没有啥好说的
- 循环接收的buffer不要太小（减小系统调用次数）
- 业务处理线程和网络线程的分离

长连接

- 消息及时
- 省电省流量
- 提高发送速度

地区 / 网络	NAT 超时时间
中国移动 3G 和 2G	5 分钟
中国联通 2G	5 分钟
中国电信 3G	大于 28 分钟
美国 3G	大于 28 分钟
台湾 3G	大于 28 分钟

内容大纲

Contents

01 移动网络概述

02 移动网络优化

03 **技术方案**

04 Benchmark

05 如何快速接入

06 未来规划

应用特点

- 随时启动与中止

  用户退出或更改账户、手机休眠与唤醒……

- 并发少状态多

  主要功能收发、网络的有无、用户的活跃状态……

- 尽量少的资源尽量快的网络

  省电、省流量、网络要敏感……

- 基础库

  不重复造轮子

  防止被拖入适配深坑

  用多少编多少不用担心体积

- 线程模型

  多线程

  并行或并发提高速度

  死锁或析构时序异常

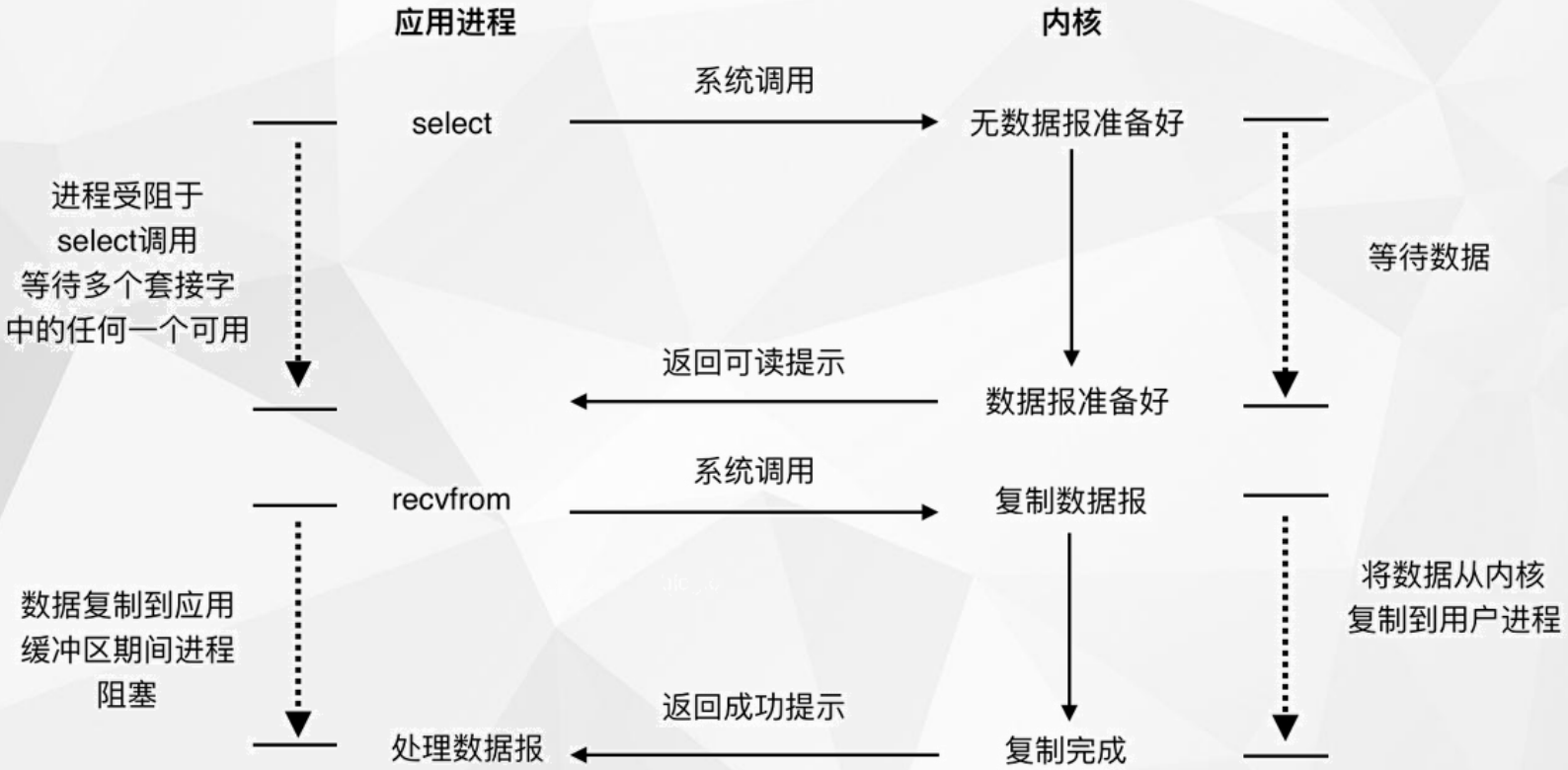
  VS

  消息队列

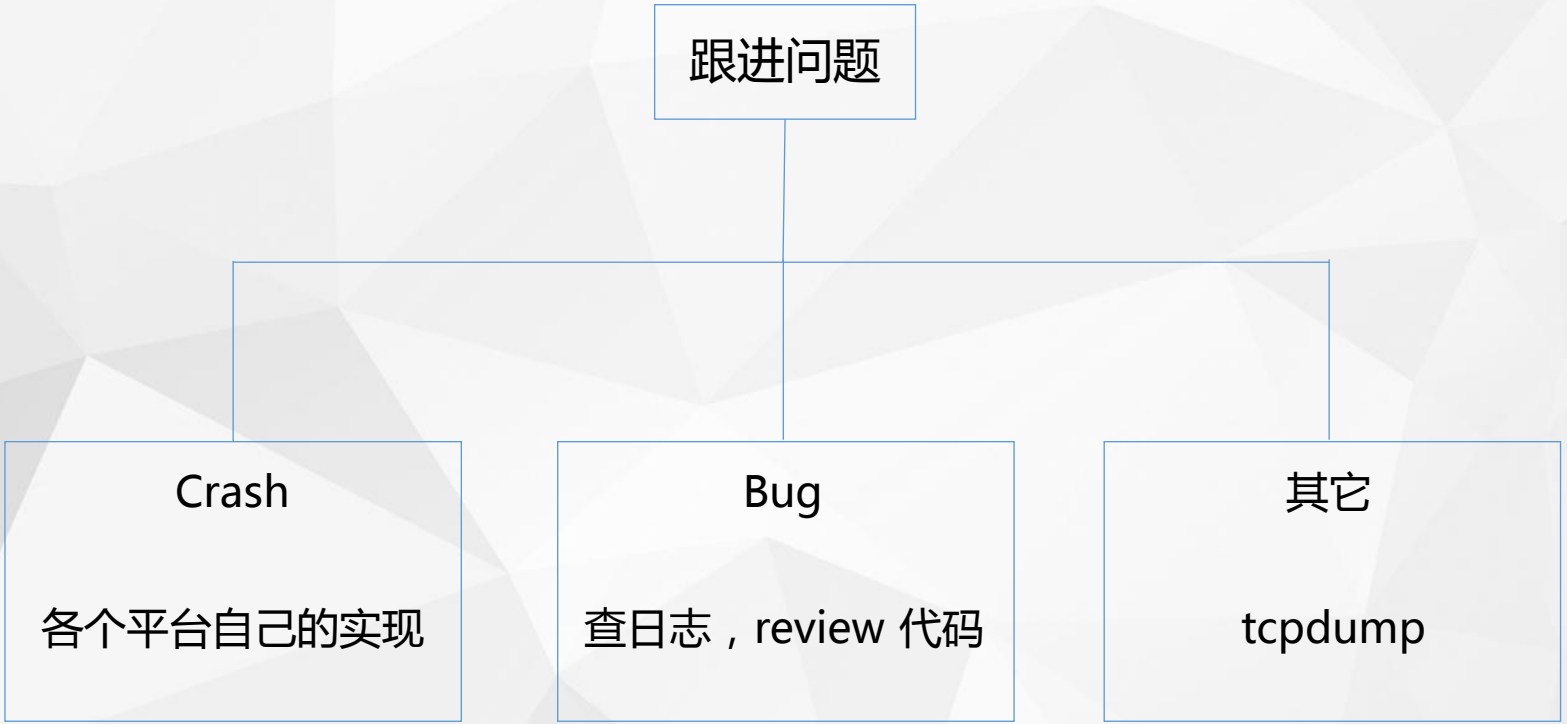
  时序清晰编程简单

  串行运行效率略低

I/O复用，事件驱动


应用进程
内核
系统调用
select
无数据报准备好
进程受阻于
select调用
等待多个套接字
中的任何一个可用
等待数据
返回可读提示
数据报准备好
系统调用
recvfrom
复制数据报
数据复制到应用
缓冲区期间进程
阻塞
将数据从内核
复制到用户进程
返回成功提示
处理数据报
复制完成

跟进问题



内容大纲

Contents

01 移动网络概述

02 移动网络优化

03 技术方案

04 **Benchmark**

05 如何快速接入

06 未来规划

Benchmark

- 测试方法：初始设置 100% Loss 网络参数，分别在不同的时间后，恢复为正常网络，观察 Mars 和 AFNetworking 的表现。
- 每种时间段进行 3 次测试，记录完成时间的差值，正数表明 Mars 先完成 CGI，负数表明竞品先完成。

Config	Mars - AFNetworking R1	Mars - AFNetworking R2	Mars - AFNetworking R3
5s 100% Loss	1	0.98	1.13
10s 100% Loss	0	0	0
15s 100% Loss	3.07	3.06	3.07
20s 100% Loss	13.9	14.1	14.1
25s 100% Loss	7.99	8.09	8.3
30s 100% Loss	3.9	3.96	3.99
35s 100% Loss	37.2 suc - 61.0 fail	37.4 suc - 60.3 fail	37.2 suc - 60.3 fail

Benchmark

- 测试方法：初始设置 100% Loss 网络参数，分别在不同的时间后，恢复为正常网络，观察 Mars 和 OKHttp 的表现。
- 每种时间段进行 3 次测试，记录完成时间的差值，正数表明 Mars 先完成 CGI，负数表明竞品先完成。
- OkHttp 的连接超时由默认的 10s 修改为 60s

Config	Mars VS OkHttp R1	Mars VS OkHttp R2	Mars VS OkHttp R3
5s 100% Loss	-0.036	-0.027	-0.022
10s 100% Loss	2.9	2.86	2.86
15s 100% Loss	13.9	13.85	14
20s 100% Loss	7.95	7.97	7.93
25s 100% Loss	3.1	3.11	3.1
30s 100% Loss	31.98 suc - 60.0 fail	31.91 suc - 60.0 fail	31.9 suc - 60.0 fail

Benchmark

测试：请求成功 500 次需要的尝试次数和总耗时

Config	Mars		AFNetworking	
	尝试次数	总耗时	尝试次数	总耗时
3G	500	226	500	229
	500	231	500	239
2G	500	883	500	874
	500	872	500	881
30% Loss	508	958	504	1229
	512	907	505	1119
40% Loss	547	2175	514	2599
	546	2310	518	2953
50% Loss	662	5095	568	7492
	659	4999	560	7273

iOS

Config	Mars		OkHttp	
	尝试次数	总耗时	尝试次数	总耗时
3G	500	406	500	403
	500	402	500	399
2G	500	1018	500	1013
	500	1019	501	1028
30% Loss	549	2576	666	2809
	538	2400	661	2688
40% Loss	669	5195	882	5651
	653	5037	884	5475
50% Loss	873	9915	1293	11023
	883	10077	1338	11678

Android

内容大纲

Contents

01 移动网络概述

02 移动网络优化

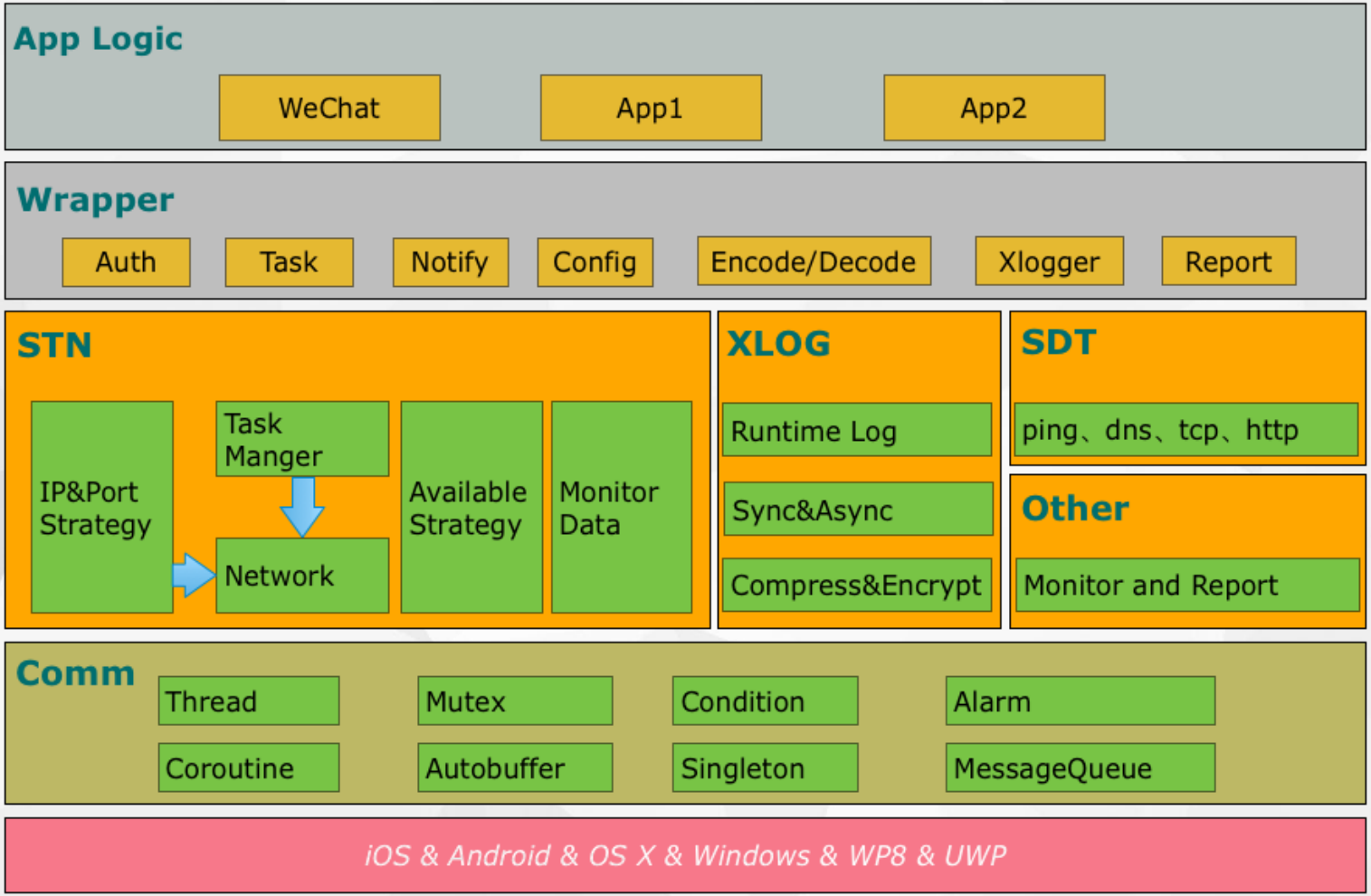
03 技术方案

04 Benchmark

05 **如何快速接入**

06 未来规划

mars 简介


App Logic
WeChat
App1
App2
Wrapper
Auth
Task
Notify
Config
Encode/Decode
Xlogger
Report
STN
IP&Port Strategy
Task Manger
Network
Available Strategy
Monitor Data
XLOG
Runtime Log
Sync&Async
Compress&Encrypt
SDT
ping、dns、tcp、http
Other
Monitor and Report
Comm
Thread
Mutex
Condition
Alarm
Coroutine
Autobuffer
Singleton
MessageQueue
iOS & Android & OS X & Windows & WP8 & UWP

Xlog

高可靠性高性能的运行期日志组件

STN

信令分发网络模块，也是 Mars 最主要的部分

- 跨平台，终端组件
- 针对小数据做优化
- 长连私有协议友好
- 短连支持部分Http协议

快速接入

内容大纲

Contents

01 移动网络概述

02 移动网络优化

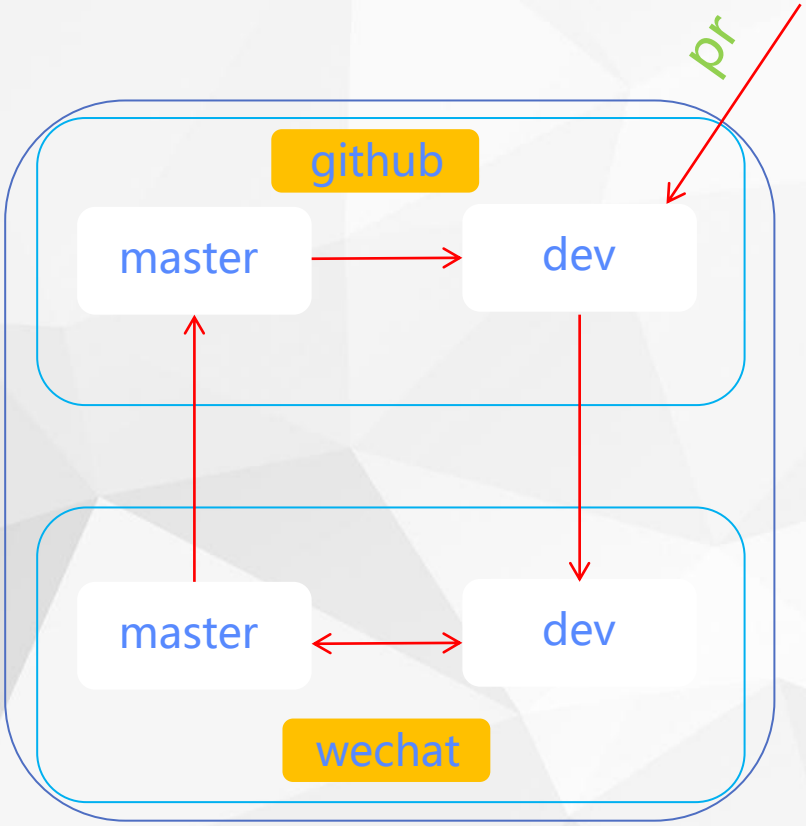
03 技术方案

04 Benchmark

05 如何快速接入

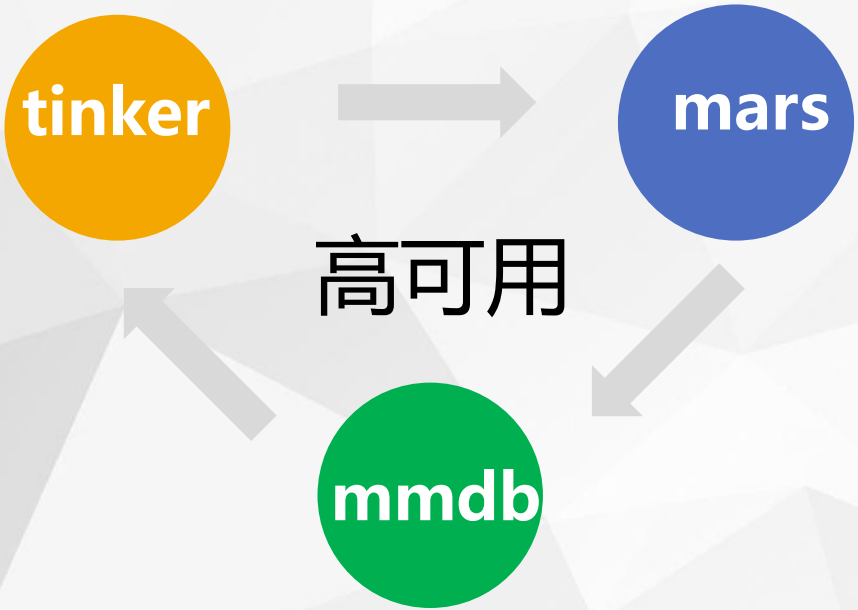
06 **未来规划**

未来规划



- 日志提供加密默认实现
- 长连接支持http、socks5代理
- Windows 平台支持
- ……

开源计划



微信终端开发公众号

Q&A
谢谢！

https://github.com/Tencent/mars