

服务器负载分析和问题排查

云计算工作室-池育龙



池育龙

介绍：

- 2009年毕业，12年7月加入博雅互动
- 熟悉C++、网络编程、系统架构、操作系统原理
- 行业心得：流程和工具，让工作更简单

授课经历：

- 《Linux地址空间分布及参数值传递》
- 《高效率工具分享》
- 《自动测试及工具》
- 《麻将架构及行牌流程解析》

博雅大学：

云计算开发部技术通道
中级讲师

猜一猜,谁说的?

在线要掉光了,服

这谁写的代码?

欧总,管理后台挂了啊!!!?!?!?

在策划陆哥清秋,帮我拉最近一

怎么不上正式服就这么

在我这是网络问题嘛!来看看

多问题啊!下缓存试试

你装的什么版本?

猜一猜,谁说的?

在线要掉光了,服

老是掉线。。。。老是掉线。。。。老是掉线。。。。老是掉线。。。。

欧总,这管理后台挂了啊!!!??!!

在策划陆哥清秋,帮我拉最近一

怎么弄上正式服就这么

坚决抵制博雅问题啊!有没有霍扶雍庆的人儿啊?加个

好友一起打牌呗~

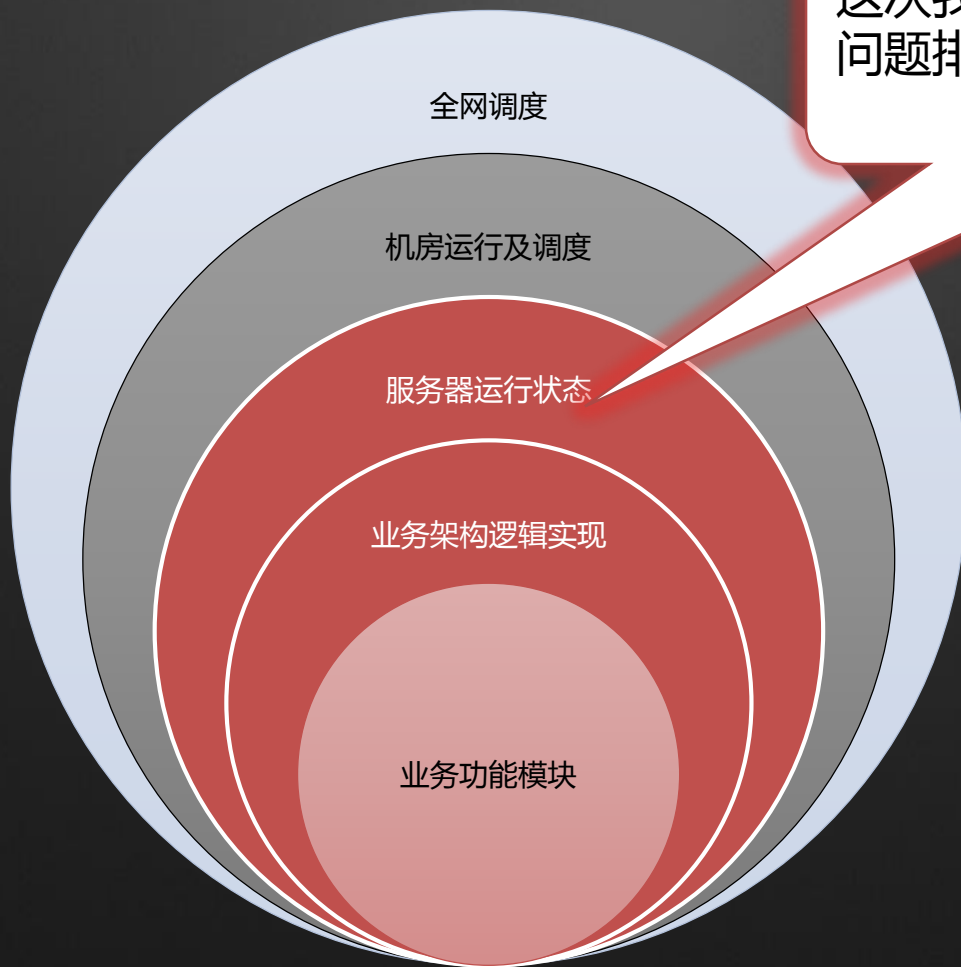
***谁弄的系统垃圾一天到晚卡啊

你装的什么版本?

课程目标

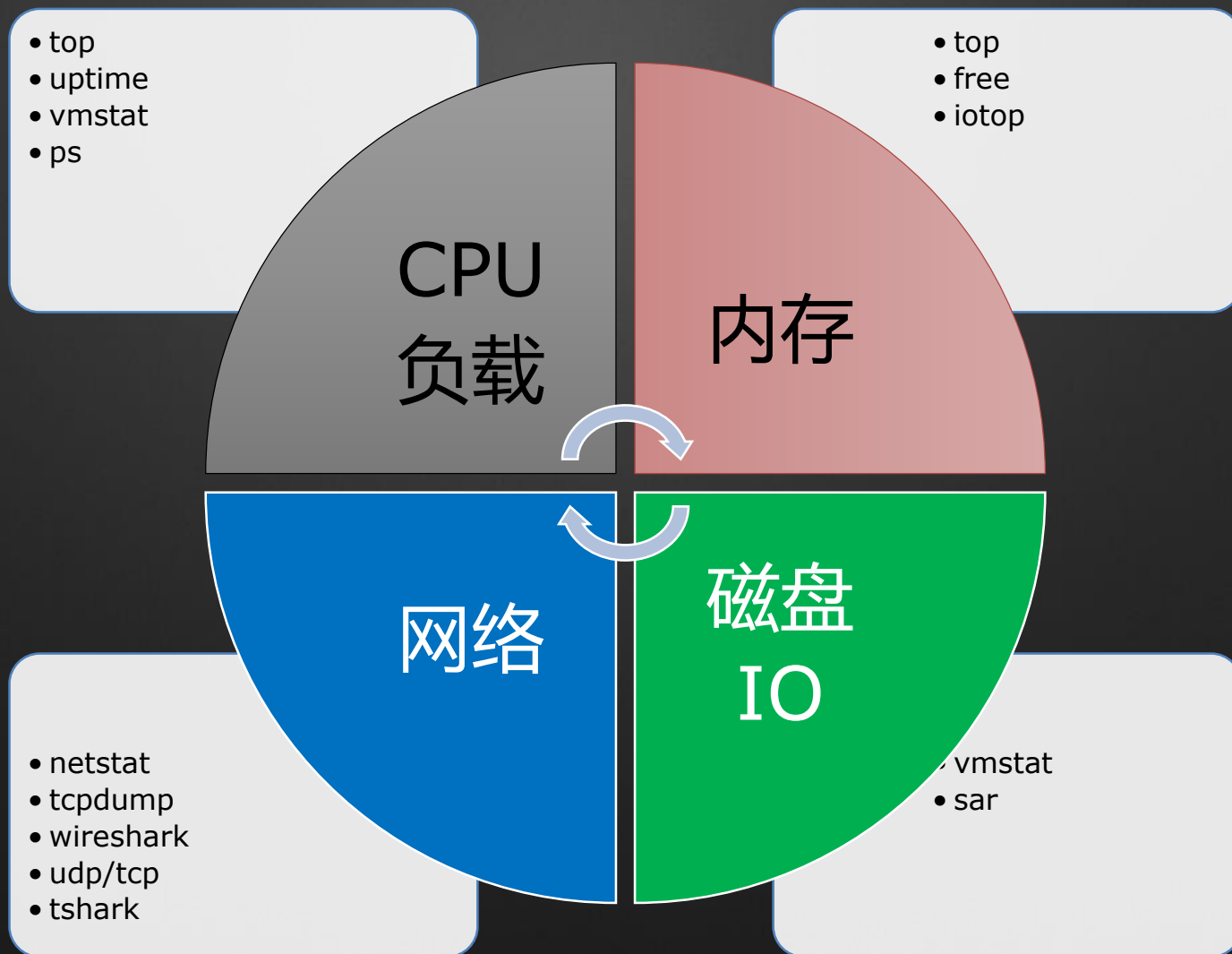
介绍从不同纬度分析服务器运行状态的方法；掌握对网络问题的常见排查流程和常用工具。

我所做的



这次我们主要关注服务器运行的
问题排查和故障分析

维度



网络

RST包、窗口为零等异常包数量;

DNS是影响性能的一个**关键**;

路由节点也会影响全网接入;

TCP: 主要关注重传率、**接收队列和发送队列长度**;

UDP:接收缓冲区和发送缓冲区是否已满

注意**流量**是否超标

磁盘IO

iowait% < 20%

redis持久化使用不当对机器性能影响较大，要重点关注

mysql分库分表、索引建立对磁盘IO的影响

控制日志写入数量，日志分级

CPU

$us \leq 70\%$, $sy \leq 35\%$, $us + sy \leq 70\%$;

每个处理器的可运行队列 ≤ 3 个线程;

上下文切换和CPU利用率关联, 如果CPU利用率良好, 那么大量的进程上下文切换也是可以接受的;

内存

swap in(**si**)==0, swap out(**so**)==0

应用程序可用内存/系统物理内存 $\leq 70\%$

Buff:物理内存用来缓存读写操作的buffer大小，linux内核希望尽可能产生缺页终端（从文件存储区读），并且尽可能避免主缺页终端（从磁盘读），随着次缺页终端的增多，文件缓存区也逐渐增大，直到物理内存减少才释放不用的页

Cache：缓存进程地址空间cache的大小

良好指标



CPU

- CPU利用率: $\text{user time} \leq 70\%$, $\text{system time} \leq 35\%$, $\text{user time} + \text{system time} \leq 70\%$
- 每个处理器的可运行队列 ≤ 3 个线程
- 上下文切换和CPU利用率关联, 如果CPU利用率良好, 那么大量的进程上下文切换也是可以接受的



内存

- $\text{swap in}(\text{si}) = 0$, $\text{swap out}(\text{so}) = 0$
- 应用程序可用内存/系统物理内存 $\leq 70\%$
- Buff: 物理内存用来缓存读写操作的buffer大小, linux内核希望尽可能产生缺页终端(从文件存储区读), 并且尽可能避免主缺页终端(从磁盘读), 随着次缺页终端的增多, 文件缓存区也逐渐增大, 直到物理内存减少才释放不用的页
- Cache: 缓存进程地址空间cache的大小



磁盘IO

- $\text{iowait}\% < 20\%$
- 关注日志写入数量、redis持久化、mysql对磁盘IO的影响
- 持久化使用不当对机器性能影响较大, 要重点关注



网络IO

- TCP: 主要关注重传率、接收队列和发送队列长度
- RST包、发送窗口、异常包数量
- UDP: 接收缓冲区和发送缓冲区是否
- DNS是影响性能的一个关键, 路由节点也会影响全网接入

网络:例一

在两台服务器上拉取支付中心的商品列表，不知道为何原因拉取配置速度差别很大，如下图：

```
[chiyl@bogon tmp]$ ifconfig | grep -e 'inet' -e 'eth'
```

```
eth0      Link encap:Ethernet  HWaddr 08:00:27:8E:81:9F
           inet addr:10.0.2.15  Bcast:10.0.2.255  Mask:255.255.255.0
           inet6 addr: fe80::a00:27ff:fe8e:819f/64 Scope:Link
```

```
eth1      Link encap:Ethernet  HWaddr 08:00:27:C3:C4:A9
           inet addr:192.168.56.102  Bcast:192.168.56.255  Mask:255.255.255.0
           inet6 addr: fe80::a00:27ff:fec3:c4a9/64 Scope:Link
           inet addr:127.0.0.1  Mask:255.0.0.0
           inet6 addr: ::1/128 Scope:Host
```

```
[chiyl@bogon tmp]$ time curl -o /dev/null pay.boyya.com
```

% Total	% Received	% Xferd	Average Speed	Time	Time	Time	Current
			Dload	Upload	Total	Spent	Left
103	2894	0	2894	0	0	186	0
--:--:--	0:00:15	--:--:--	9187				

```
real    0m15.490s
user    0m0.005s
sys     0m0.006s
```

```
[chiyl@bogon tmp]$ ifconfig | grep -e 'inet' -e 'eth'
```

```
eth0      Link encap:Ethernet  HWaddr 08:00:27:41:BC:28
           inet addr:10.0.2.15  Bcast:10.0.2.255  Mask:255.255.255.0
           inet6 addr: fe80::a00:27ff:fe41:bc28/64 Scope:Link
```

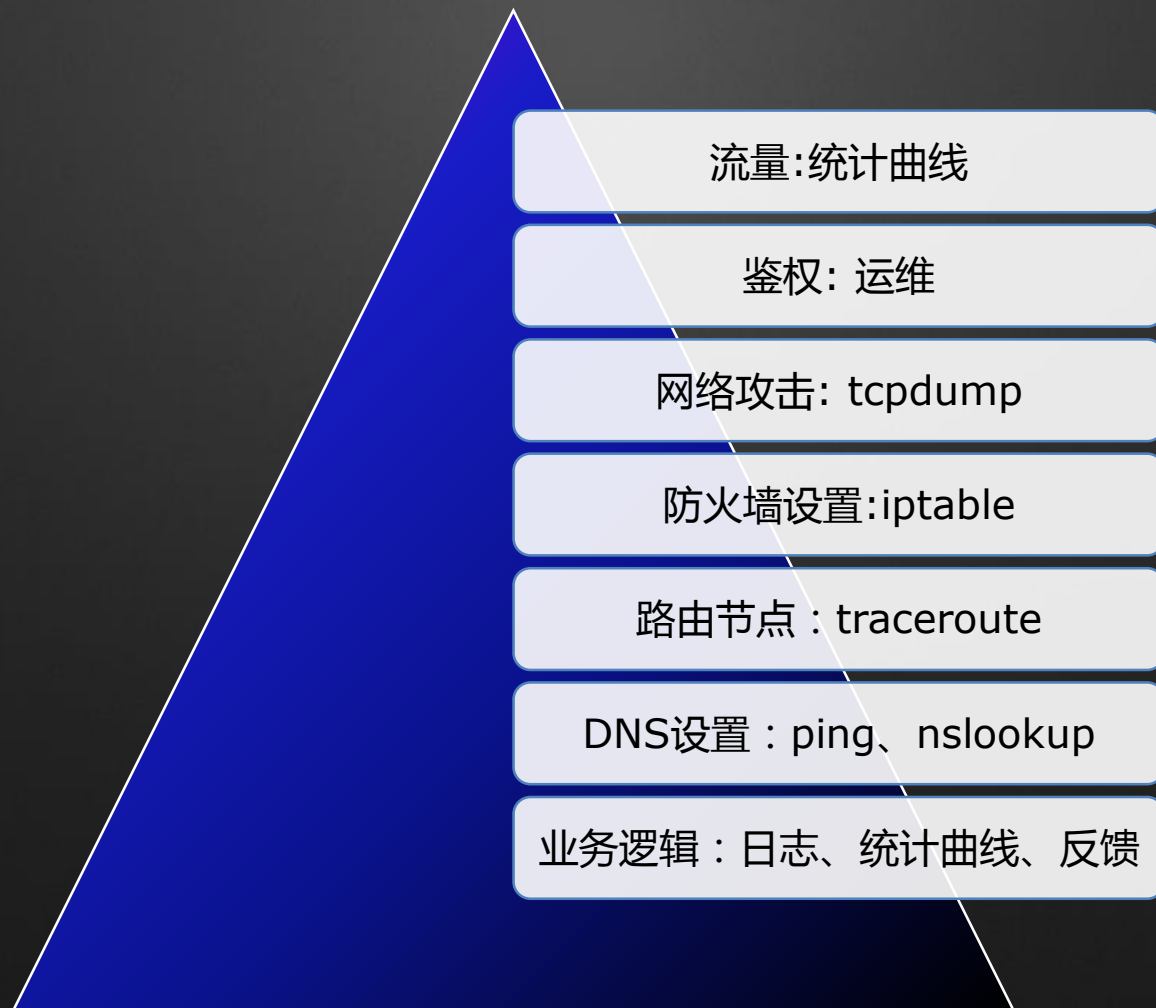
```
eth1      Link encap:Ethernet  HWaddr 08:00:27:11:1C:74
           inet addr:192.168.56.101  Bcast:192.168.56.255  Mask:255.255.255.0
           inet6 addr: fe80::a00:27ff:fe11:1c74/64 Scope:Link
           inet addr:127.0.0.1  Mask:255.0.0.0
           inet6 addr: ::1/128 Scope:Host
```

```
[chiyl@bogon tmp]$ time curl -o /dev/null pay.boyya.com
```

% Total	% Received	% Xferd	Average Speed	Time	Time	Time	Current
			Dload	Upload	Total	Spent	Left
103	2894	0	2894	0	0	2924	0
--:--:--	0:00:00	--:--:--	3499				

```
real    0m1.001s
user    0m0.003s
sys     0m0.014s
```

紧急性网络问题排查优先级



网络优化示例

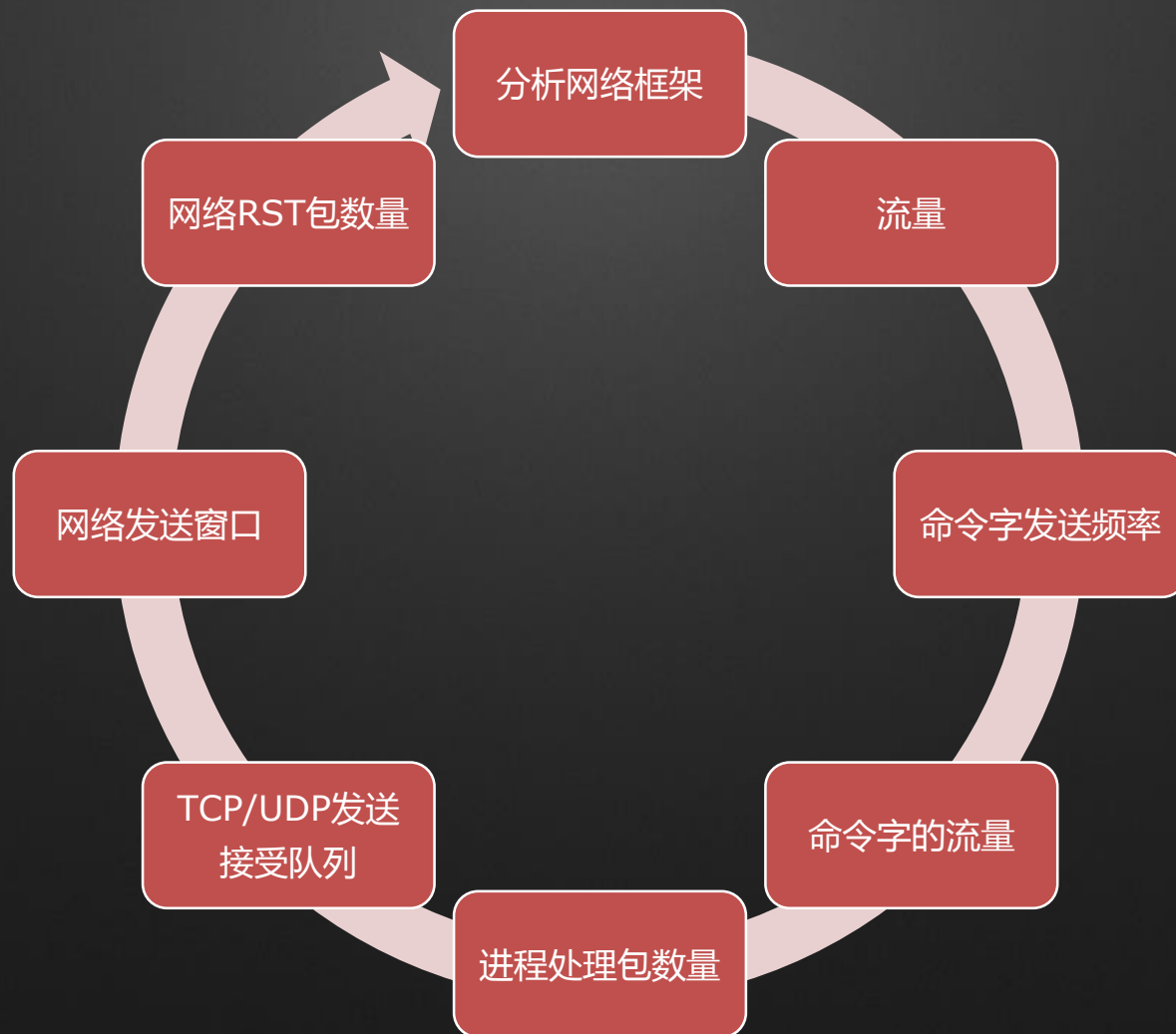
背景数据：

- 棋牌游戏，同时在线约2万，游戏数据互通，不区分分服
- 主要的服务器处于是转发层和逻辑层
- 单台机器中监听的端口有607个：
 - HallServer：136个，
 - AllocServer：84个，
 - UserServer:31个，
 - CacheUser: 13个，
 - MoneyServer: 16个，
 - redis: 4 个。

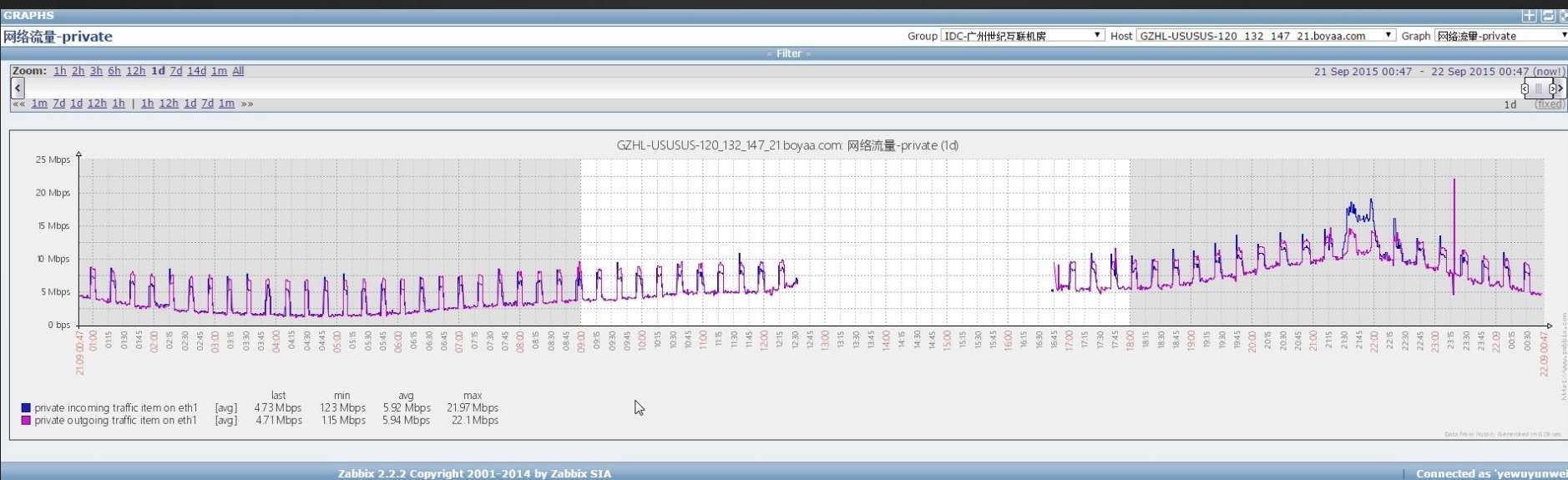
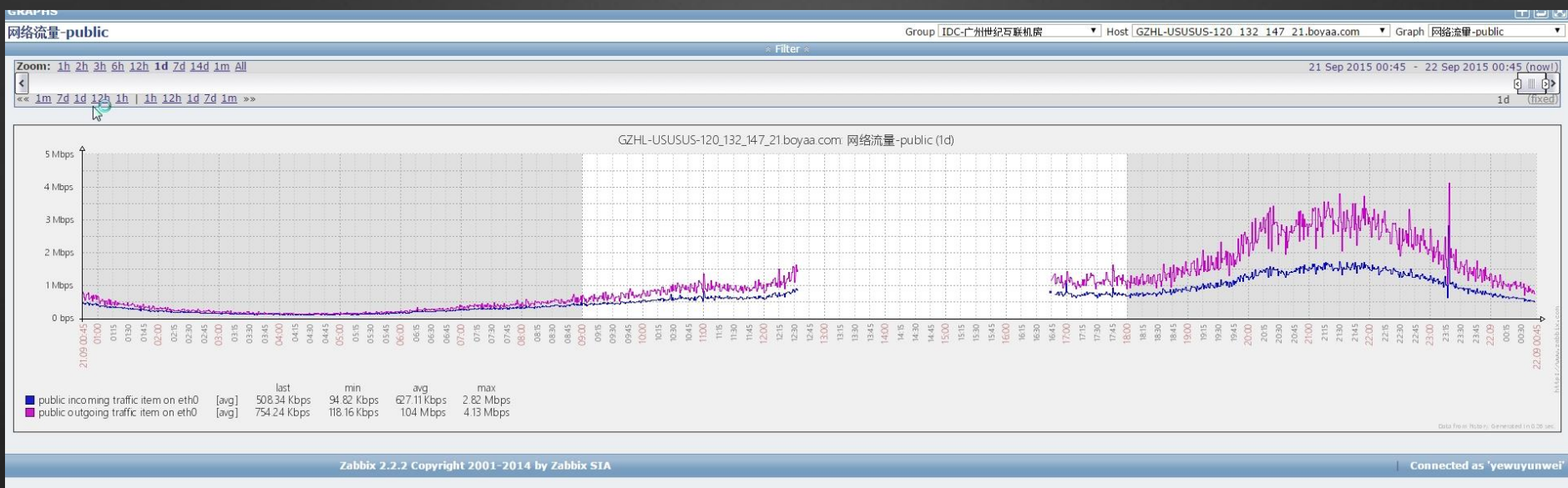
目标：

1. 减少玩家玩游戏的流量
2. 提高网络请求的响应速度
3. 找出现在系统中的风险点

优化型网络问题分析优先级

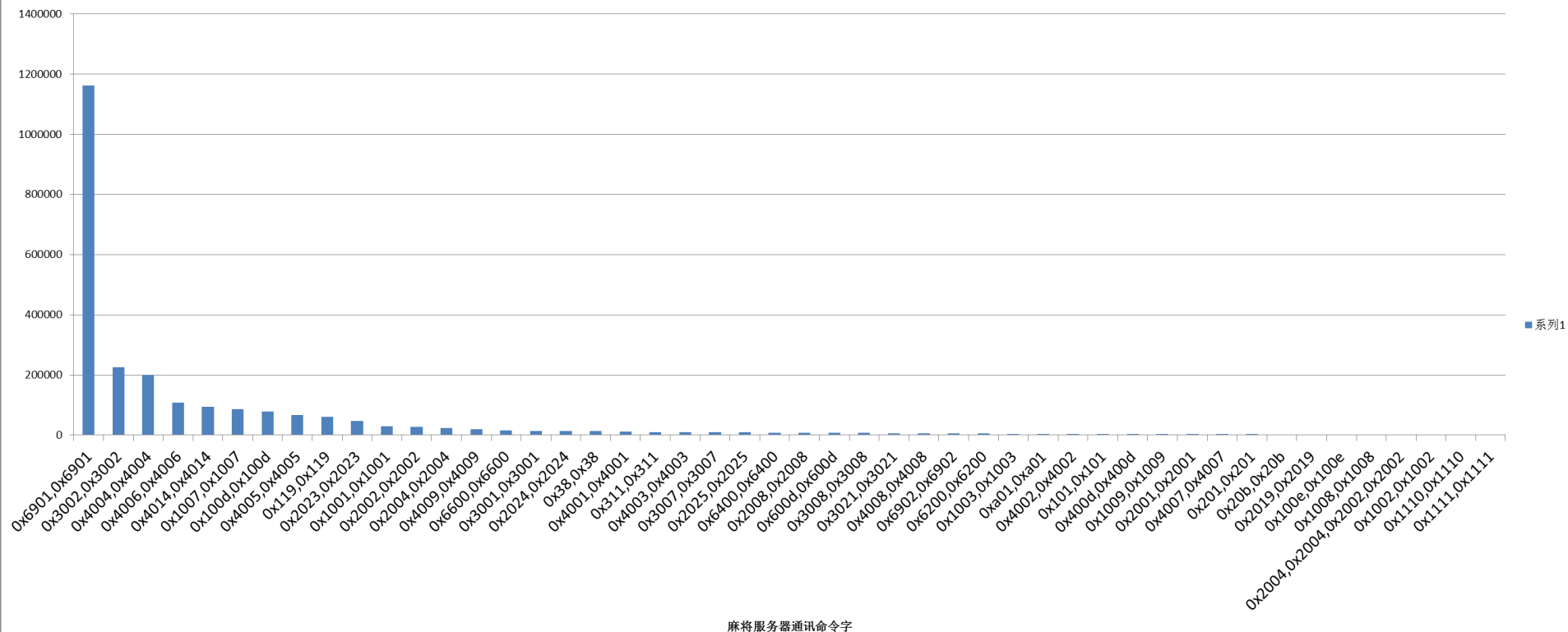


网络问题分析（例二）



外网命令字流量统计

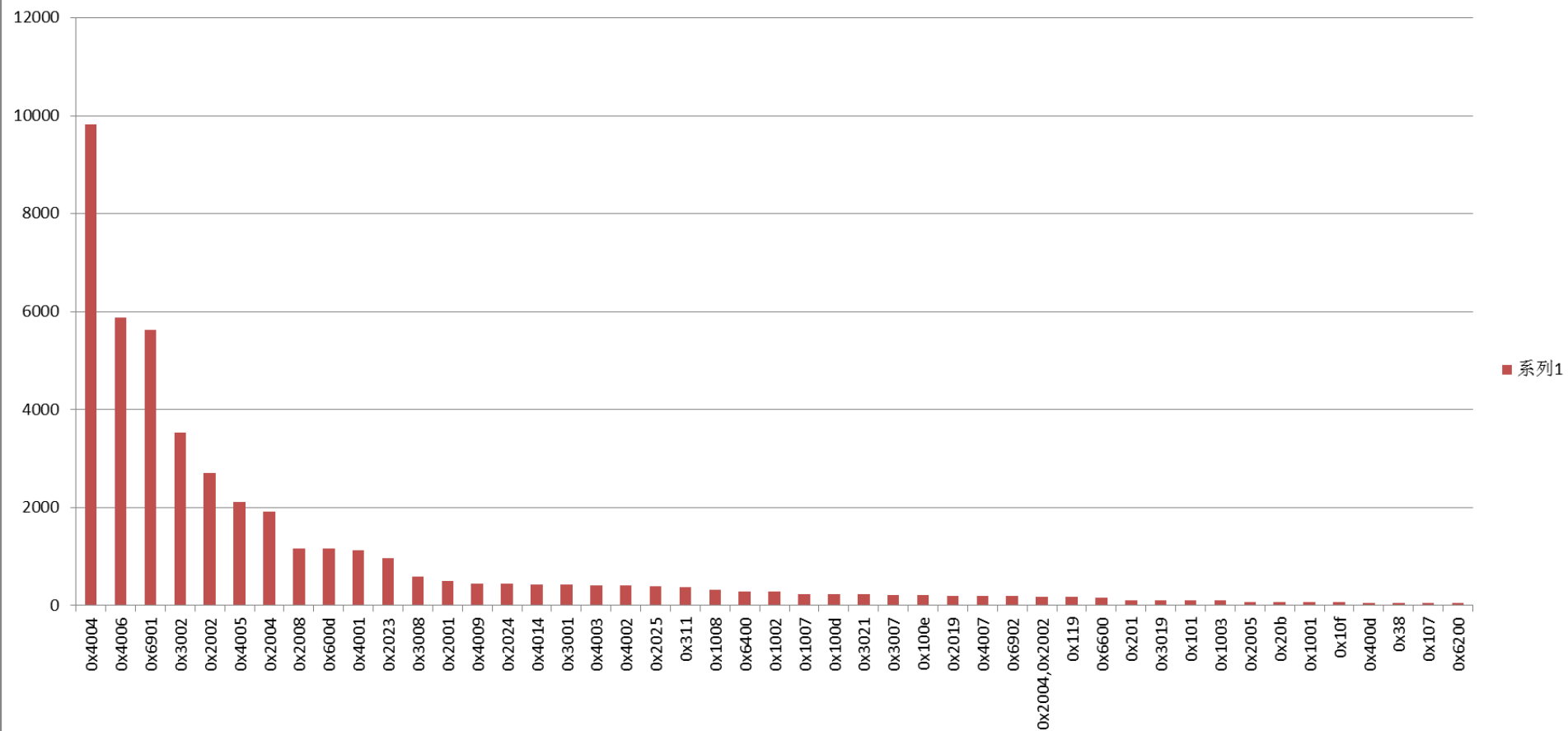
麻将外网各命令字包流量



麻将服务器通讯命令字

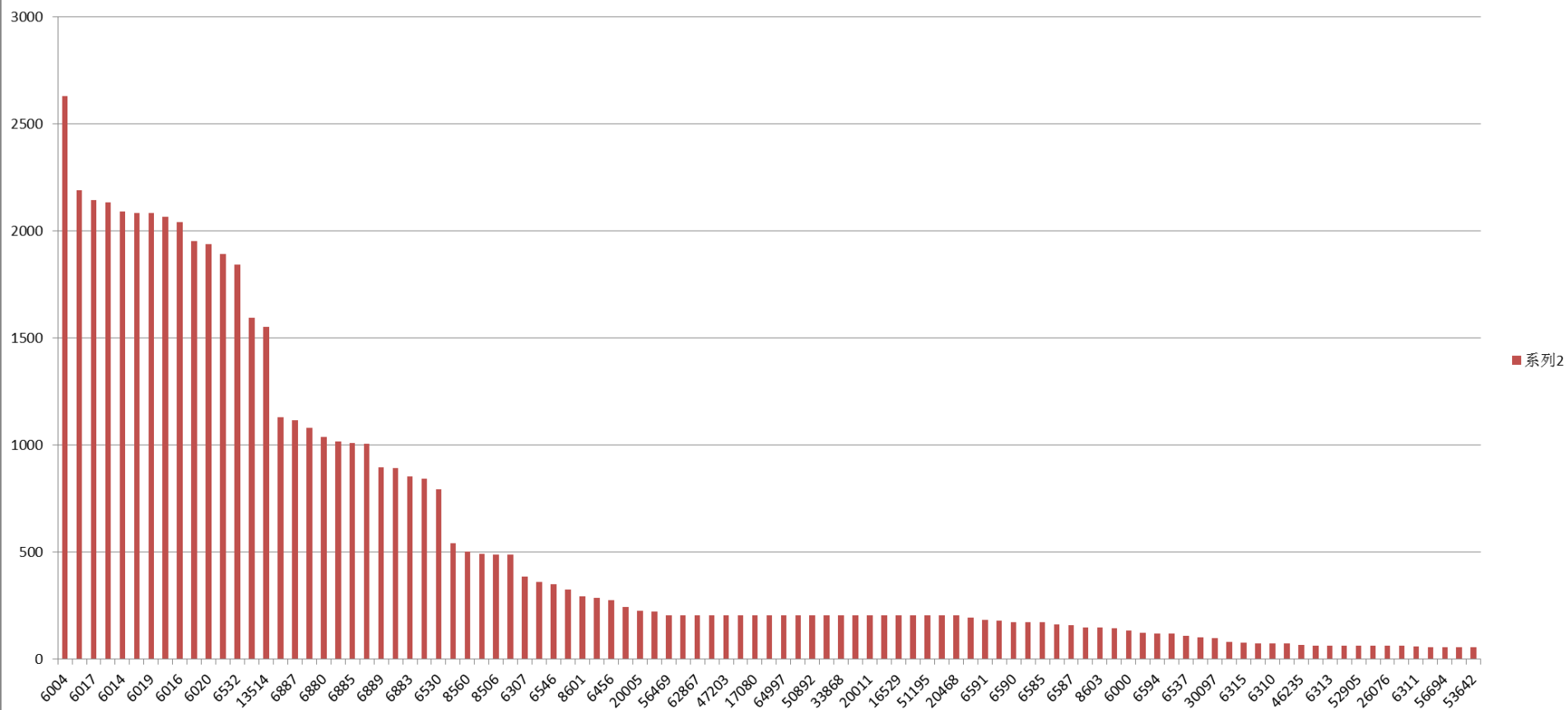
外网命令字包数量

麻将外网各命令字包数量



每个端口处理包数量统计

每个端口接收的包的数量（内网）



网络分析结论

流量优化结论：

1. 0x6091命令字的流量明显偏高，应尽早优化
2. 0x4004命令字发送频率偏高，应尽早优化
3. 外网流量有突发性的尖峰，应该继续跟进排查是否为统计方法问题，如果不是，找出原因
4. 内网流量尖峰更为明显，找出原因，如果为异常，也应该优化

故障点结论：

- AllocServer 处理包的数量不符合均匀分布的原则，0号Alloc负载明显偏高，可能路由算法有问题
- MemServer 负载偏高，而且其实服务器的业务单点，有可能成为性能瓶颈
- 外网RST包数量为71个，正常
- 内网RST包数量为 1557个，肯定有问题
- AutoAI、GoldLogServer进程的TCP接受队列不为空
- 大部分HallServer进程的发送队列不为空
- 少量UserSever的发送队列不为空
- MailBoxServer的TCP接受窗口为0

负载分析

```
top - 15:43:18 up 6 days, 6:41, 16 users, load average: 2.82, 2.10, 1.85
Tasks: 726 total, 2 running, 724 sleeping, 0 stopped, 0 zombie
Cpu0 : 20.7%us, 40.5%sy, 0.0%ni, 38.5%id, 0.0%wa, 0.0%hi, 0.3%si, 0.0%st
Cpu1 : 16.1%us, 28.4%sy, 0.0%ni, 55.2%id, 0.0%wa, 0.0%hi, 0.3%si, 0.0%st
Mem: 4043720k total, 3710744k used, 332976k free, 68324k buffers
Swap: 2096472k total, 403160k used, 1693312k free, 441100k cached
```

PID	USER	PR	NI	VIRT	RES	SHR	S	%CPU	%MEM	TIME+	SWAP	TIME	CODE	DATA	nFLT	nDRT	WCHAN	COMMAND
5012	mysql	15	0	1020m	499m	11m	S	0.0	12.6	14:32.41	521m	14:32	7704	976m	105	0	?	mysqld
27296	root	15	0	171m	113m	1788	S	0.0	2.9	0:26.47	58m	0:26	1268	113m	9	0	?	BackServer.er.v
3972	JunlyLu	15	0	148m	84m	2720	S	0.0	2.2	0:21.92	63m	0:21	1580	84m	2	0	?	Alloc.sc
9273	JunlyLu	15	0	133m	71m	1428	S	0.0	1.8	1:51.54	61m	1:51	644	76m	6	0	?	HTTPServer.sc
26517	JunlyLu	15	0	67620	45m	1052	S	0.0	1.2	0:00.08	20m	0:00	680	46m	8	0	?	DWFServer.sc
27079	root	15	0	67620	45m	1052	S	0.0	1.2	0:00.06	20m	0:00	680	46m	8	0	?	DWFServer.er
4165	root	15	0	391m	32m	3604	S	0.0	0.8	2:31.83	358m	2:31	4	151m	10	0	?	salt-minion
5275	www	15	0	77188	27m	2472	S	0.0	0.7	0:11.99	47m	0:11	476	26m	4	0	?	nginx
5272	www	15	0	76800	27m	2472	S	0.0	0.7	0:11.86	47m	0:11	476	25m	2	0	?	nginx
5274	www	15	0	76872	27m	2464	S	0.0	0.7	0:11.17	47m	0:11	476	26m	4	0	?	nginx
5271	www	15	0	76648	26m	2464	S	0.0	0.7	0:10.88	47m	0:10	476	25m	5	0	?	nginx
2702	root	10	-10	28704	22m	1744	S	0.0	0.6	0:00.00	6152	0:00	172	20m	3	0	?	iscsiuio
6051	JunlyLu	15	0	126m	21m	560	S	0.0	0.6	0:24.36	105m	0:24	76	112m	1	0	?	memcached
6725	JunlyLu	15	0	124m	20m	560	S	0.0	0.5	0:01.55	103m	0:01	76	110m	0	0	?	memcached
975	www	16	0	229m	19m	13m	S	0.0	0.5	0:03.44	210m	0:03	6692	6256	9	0	?	php-cgi
24995	www	15	0	227m	19m	13m	S	0.0	0.5	0:03.18	208m	0:03	6692	6100	27	0	?	php-cgi
24996	www	16	0	228m	19m	12m	S	0.0	0.5	0:03.24	209m	0:03	6692	6784	2	0	?	php-cgi
31780	www	16	0	229m	18m	12m	S	0.0	0.5	0:02.93	211m	0:02	6692	6252	8	0	?	php-cgi
4955	www	16	0	227m	18m	13m	S	0.0	0.5	0:03.55	208m	0:03	6692	5612	2	0	?	php-cgi
12463	www	16	0	227m	18m	12m	S	0.0	0.5	0:03.27	209m	0:03	6692	5976	10	0	?	php-cgi
12475	www	16	0	227m	18m	12m	S	0.0	0.5	0:03.13	208m	0:03	6692	5656	14	0	?	php-cgi
26281	www	16	0	229m	18m	12m	S	0.0	0.5	0:02.09	211m	0:02	6692	5972	0	0	?	php-cgi
31784	www	16	0	227m	18m	12m	S	0.0	0.5	0:02.79	209m	0:02	6692	5596	4	0	?	php-cgi
5347	JunlyLu	15	0	82940	18m	2712	S	0.0	0.5	0:01.12	62m	0:01	1580	16m	0	0	?	Alloc.sc
4463	JunlyLu	15	0	82964	18m	2712	S	0.0	0.5	0:01.11	62m	0:01	1580	16m	1	0	?	Alloc.sc
26836	www	16	0	229m	18m	12m	S	0.0	0.5	0:03.40	211m	0:03	6692	5596	10	0	?	php-cgi
5137	JunlyLu	15	0	82884	18m	2696	S	0.0	0.5	0:02.10	62m	0:02	1580	16m	0	0	?	Alloc.sc
24997	www	16	0	229m	18m	11m	S	0.0	0.5	0:03.30	211m	0:03	6692	6268	0	0	?	php-cgi
4246	JunlyLu	15	0	82876	18m	2712	S	0.0	0.5	0:01.02	62m	0:01	1580	16m	0	0	?	Alloc.sc
4536	JunlyLu	15	0	82876	18m	2712	S	0.0	0.5	0:01.15	62m	0:01	1580	16m	0	0	?	Alloc.sc

CPU

- 双核CPU，负载无压力，有压榨空间
- 系统CPU时间占用偏高，可能有过多的系统调用
- 用户CPU时间正常，而且还有提升的空间

内存

- Swap已经使用了
- 物理内存还剩三百多M，但是swap还有剩余，内存上还有压榨空间
- 看起来有进程内存泄露了

磁盘/网络

- 从系统调用频繁来看，应该是网络io或者磁盘io比较频繁

负载分析

```
[LawrenceChi@shijihulian-120-132-145-34 ~]$ vmstat 1
```

procs		memory				swap		io		system				cpu			
r	b	swpd	free	buff	cache	si	so	bi	bo	in	cs	us	sy	id	wa	st	
4	0	8106964	3376432	470136	6432288	0	0	38	297	0	0	12	9	77	2	0	
4	0	8106964	3369612	470136	6432436	0	0	144	1676	25138	49746	8	6	86	0	0	
3	0	8106964	3354236	470136	6432524	0	0	36	1388	25029	53902	8	6	85	0	0	
0	5	8106964	3377992	470136	6432620	0	0	104	23812	21665	40989	5	4	82	8	0	
4	0	8106964	3361236	470136	6432688	0	0	144	1588	24457	50662	8	6	76	9	0	
5	1	8106964	3378364	470140	6432888	0	0	160	2352	23303	50346	8	6	86	0	0	
2	0	8106964	3359128	470140	6432960	0	0	112	1748	23701	53998	8	6	86	0	0	
5	0	8106964	3338088	470148	6433164	0	0	212	1936	26986	58255	10	7	82	0	0	
6	0	8106964	3317372	470148	6433216	0	0	228	1844	25136	57444	12	8	79	1	0	
2	0	8106964	3356712	470148	6433324	0	0	36	3496	26710	65186	16	8	75	0	0	
9	0	8106964	3353304	470160	6433460	0	0	52	2264	21158	53898	16	15	69	0	0	
4	0	8106964	3351708	470160	6433572	0	0	80	2304	24425	67264	14	14	72	0	0	
5	0	8106964	3351696	470168	6433700	0	0	104	4676	23731	62224	19	19	61	0	0	
4	1	8106964	3363316	470168	6433872	0	0	80	15184	22163	48956	10	9	80	2	0	
8	0	8106964	3355496	470168	6433960	0	0	148	10932	27499	60012	9	6	76	9	0	
0	0	8106964	3377368	470168	6434032	0	0	164	1940	26010	58359	9	7	83	1	0	
2	0	8106964	3356024	470168	6434220	0	0	240	1488	21975	46604	7	5	87	0	0	
7	0	8106964	3362348	470168	6434396	0	0	184	1816	25185	55369	8	6	85	0	0	
1	0	8106964	3380748	470168	6434500	0	0	256	2064	28367	67780	13	8	79	1	0	
3	0	8106964	3374416	470168	6434648	0	0	64	1136	22594	49115	9	6	85	0	0	
5	4	8106964	3369960	470168	6434768	0	0	244	7908	25655	52376	8	6	78	8	0	
4	1	8106964	3361172	470168	6434864	0	0	392	1472	25582	55164	9	6	76	9	0	
5	0	8106964	3372092	470168	6434976	0	0	352	2060	27149	59225	12	7	80	1	0	
8	0	8106964	3360968	470168	6435228	0	0	104	2068	25674	59622	11	7	82	1	0	
14	6	8106964	3366620	470168	6435304	0	0	140	13388	28685	63797	11	7	81	1	0	
9	6	8106964	3354468	470168	6435484	0	0	112	11768	22907	47025	11	6	67	16	0	
5	0	8106964	3374696	470168	6435596	0	0	320	2796	24805	54396	12	7	76	5	0	
5	0	8106964	3362040	470168	6435776	0	0	128	1544	27043	61876	10	6	84	0	0	
0	0	8106964	3362288	470168	6435860	0	0	152	2232	24070	51133	9	6	85	1	0	

负载分析

四川麻将负载优化项目立项及项目状态周知 ☆

池育龙

hi All:

现在川麻中卡房的现象很严重，需要立项推动这个服务器优化的工作的进行。

这次川麻负载优化项目由 @池育龙 负责推动，大家协调进行，目标是**优化服务器的资源配置，减少卡房现象，提高服务器能够承载的最高在线的玩家数量。**

项目时间（两周）：2013-11-22 ~ 2013-12-07

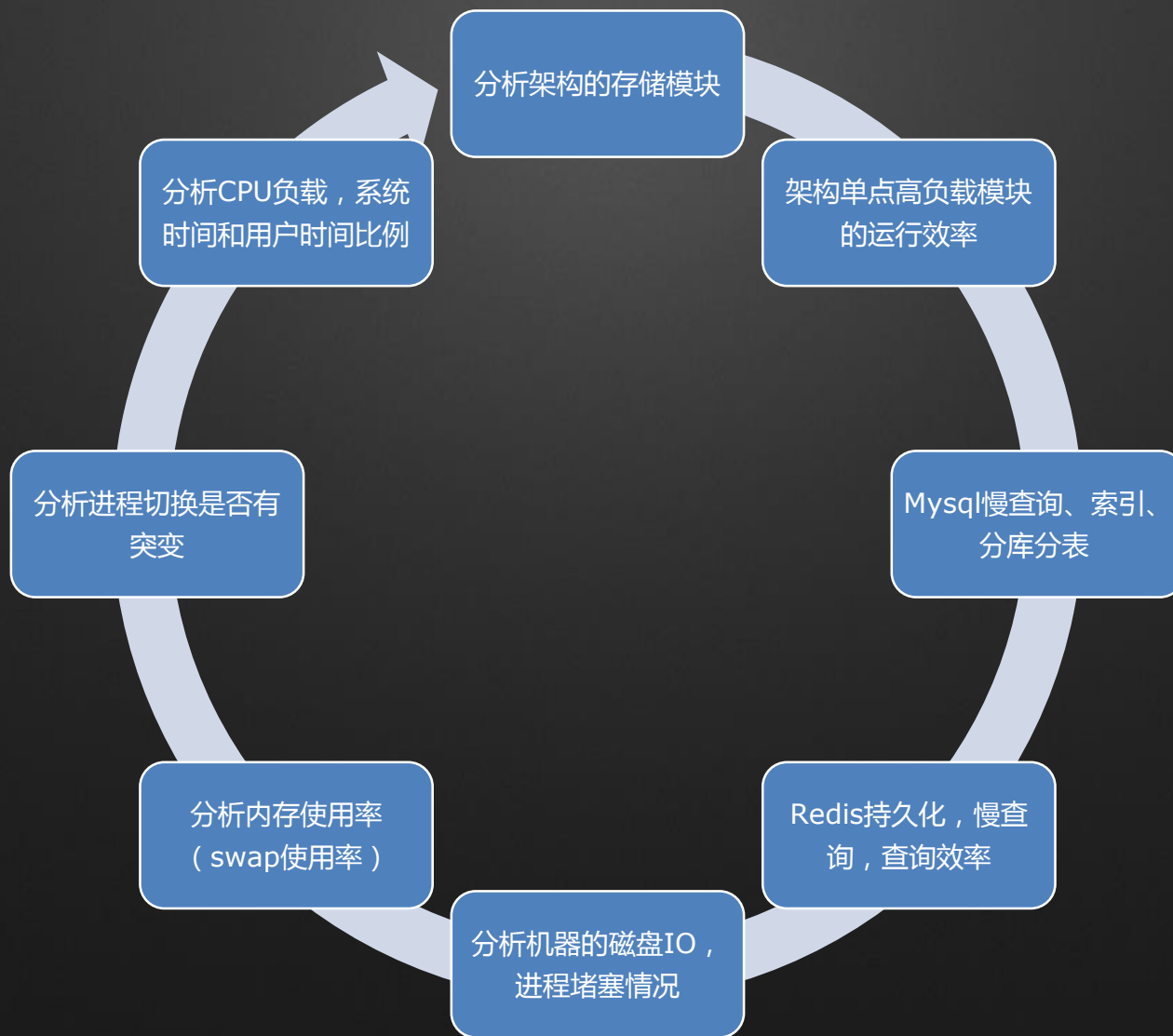
1. 四川的卡房现象已经严重影响了玩家的游戏体验：

①根据搜集玩家的反馈的结果，有非常多的玩家反馈现在玩游戏很卡：川麻搜集反馈活动中搜集到的反馈中，有**30%的玩家反馈房间卡**；

②高峰期的最高在线的曲线走势显示，当最高在线在1.3w左右的时候，**在线增长异常平缓。**



负载问题分析优先级



复盘八折

③moneyserver所处的DB机在高峰期压力较大，有时候会有高达36个进程被堵塞，截图时间就是晚上8点左右。

4	0	4750708	17670243	76664	2818032	0	0	176	1192	3568	36813	8	12	78	1	0
2	0	4750708	17664876	76688	2818444	0	0	88	1832	3963	32407	7	11	81	1	0
11	0	4750708	17664496	76700	2821472	0	0	32	1210	3469	29762	7	9	83	0	0
6	1	4750708	17671992	76748	2819092	0	0	100	30336	10826	36213	13	11	74	2	0
56	1	4750708	17609476	77088	2818652	0	0	68	8288	25078	67239	31	25	42	2	0
3	29	4750708	20225180	1564	337956	0	0	5088	3292	18489	46712	25	27	17	31	0
11	23	4750708	20217680	3320	341436	0	0	5904	240	9957	38355	9	21	21	49	0
9	32	4750708	20166764	4962	361300	0	0	17984	6268	11869	49281	16	21	27	37	0
6	36	4750708	20091156	6380	365736	0	0	5024	3112	10804	47748	23	29	16	32	0
6	4	4750708	20196600	7536	370888	0	0	4864	1916	3177	43052	30	25	13	32	0
4	2	4750708	20301288	8884	374088	0	0	2626	2126	2567	40682	8	17	63	13	0

④DB机器的慢查询一直居高不下。

2.数据搜集

为了便于直观地比较卡房情况，需要搜集以下数据：

- ①玩牌掉线率（统计不同时段中，断开链接的数目，总链接数目）
- ②强制托管的次数（按照托管次数/牌局）
- ③moneyserver 的同步查询耗时。
- ④一局牌中被丢弃的玩家操作的次数（即玩家的操作窗口过了，他的请求才到，导致请求被丢弃）

这些数据都纳入到监控系统中进行监控，并且每天输出。

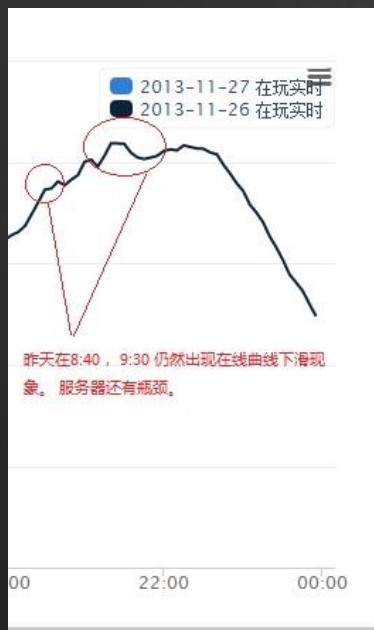
3.应对措施

- ①增开HallServer端口的数量（已经开到47个），检查程序中是否存在bug。
- ②减少DB的慢查询的次数（①GameServer对Membertable表的查询已经彻底去掉 ②php优化DB的使用）
- ③将moneyserver从DB机迁移到Server机（下周一）
- ④调整各个场次的出牌时间、操作时间（等候产品方案）。
- ⑤redis持久化策略的调整。

4.检查点。

- 检查点一：周一(2013-11-25)下班前定好迁移moneyserver的配置、php更新脚本。预计周二凌晨 6:00 开始将DB机中的Mserver、TJServer全部迁移到 Server机。
- 检查点二：周二(2013-11-26)下午，上线监控系统，和四川GameServer、HallServer带监控模块版本，用于搜集上面提到的四项数据。
- 检查点三：周四(2013-11-28)上午，讨论程序中可能出现问题点地方，并针对解决。
- 检查点四：周五(2013-11-29)输出简报。

负载分析



```
4 5 4692544 17755364 56952 3009660 0 0 18536 7868 8170 33707 19 32 42 8 0
4 5 4692544 17737180 57168 3033644 0 0 20940 6104 8202 32857 14 19 46 21 0
7 26 4692544 20095448 1568 664144 0 0 11348 7524 7944 44596 18 13 41 29 0
1 36 4692544 20413440 2580 358524 0 0 11284 876 7284 19446 5 4 33 58 0
4 26 4692544 20397172 3952 372512 0 0 13016 1336 8707 31135 14 21 17 48 0
2 24 4692544 20372168 4504 399568 0 0 24500 5176 7658 29014 13 9 26 52 0
6 21 4692544 20404232 5316 428112 0 0 27188 4104 8424 32702 12 10 25 53 0
4 18 4692544 20367720 5756 459368 0 0 25868 4776 7446 28366 11 9 32 48 0
2 18 4692544 20345932 6112 477212 0 0 17676 2252 8227 26065 9 9 38 44 0
3 12 4692544 20343148 6528 495232 0 0 17840 1976 7580 24850 10 13 46 31 0
8 15 4692544 20313144 6808 527000 0 0 29932 1388 10661 33818 13 12 39 36 0
9 8 4692544 20293356 7152 557652 0 0 29340 3088 7890 41379 14 14 42 31 0
0 18 4692544 20278068 7392 583220 0 0 23868 3012 8183 23223 11 6 56 28 0
5 12 4692544 20251388 7728 603744 0 0 21388 424 10464 32516 13 18 36 34 0
```

已连接 120.132.145.34:3600。

My Screenshot: 11-27-13 00:00:24
SSM2-sterm-131x58-58.1 1会话 CAP NUM

张李桂 博雅 22:01:10

```
5 1 4052300 19042424 34124 1542044 0 0 32 1552 10201 34755 9 12 79 0 0
7 0 4692588 19641804 54148 1342820 0 0 32 1328 8972 35338 10 13 77 0 0
4 0 4692588 19641192 54160 1343128 0 0 24 1488 15824 48007 13 13 73 0 0
36 2 4692588 19554080 54536 1342888 0 0 144 7960 29747 70665 35 25 39 1 0
3 24 4692588 20627284 1812 338804 0 0 5048 4508 16237 41719 18 15 18 49 0
5 14 4692588 20607968 3492 343048 0 0 6312 760 12395 42364 15 30 16 38 0
16 22 4692588 20566372 5756 349360 0 0 8536 1844 16609 52278 20 31 16 33 0
8 11 4692588 20573312 7292 357352 0 0 4384 6448 13315 37303 19 35 20 26 0
```

开始时间: 2013-11-26 00:00:00

结束时间: 2013-11-27 00:00:00

开始时间: 2013-11-25 00:00:00

结束时间: 2013-11-26 00:00:00

慢查询排名前10的DB:

序号	机器IP	慢查询条数	业务	类别
1	218.213.244.136	3142.00	香港大老二日志库	棋牌
2	216.12.196.24	2721.00	繁体poker-haoyou-db	poker
3	111.223.38.6	1838.00	泰语桌球-泰语斗地主-代理博雅游记-db	棋牌
4	216.12.194.282	1788.00	印尼poker-db	poker
5	216.12.208.204	1667.00	印尼虫虫主DB	SG
6	120.132.145.34	1489.00	新浪麻将db	棋牌
7	208.43.186.73	1468.00	反馈中心	棋牌
8	111.223.38.4	1250.00	泰语king-zlave-db	棋牌
9	111.223.32.220	1212.00	泰语虫虫	棋牌
10	216.12.194.247	837.00	支付中心新加坡db-dr	公共

慢查询排名前10的DB:

序号	机器IP	慢查询条数	业务	类别
1	111.223.33.1	5583.00	泰语poker-db	poker
2	218.213.244.136	3903.00	香港大老二日志库	棋牌
3	216.12.196.24	2215.00	繁体poker-haoyou-db	poker
4	120.132.145.34	1990.00	新浪麻将db	棋牌
5	216.12.208.204	1867.00	印尼虫虫主DB	SG
6	216.12.194.282	1581.00	印尼poker-db	poker
7	111.223.38.6	1368.00	泰语桌球-泰语斗地主-代理博雅游记-db	棋牌
8	111.223.38.4	1355.00	泰语king-zlave-db	棋牌
9	208.43.186.73	1241.00	反馈中心	棋牌
10	111.223.32.220	1089.00	泰语虫虫	棋牌

【检查点2】回复: 四川麻将负载优化项目立项及项目状态周知 ☆

池育龙

邮件回复时间: 2013/12/3 16:33:19。

从观测的结果分析得到如下结论：

- ①昨天晚上的最高在线，还是出现了波峰平缓的现象，判断Server还有瓶颈。
- ②DB机的慢查询操作已经有从1990条降低到1489条，排名下降2位。（membertable表的慢查询已经完全没有了）。
- ③DB机的写操作已经大大减少，玩牌高峰期的 iops持续在 170左右（原来持续在230左右）
- ④DB机还是会有偶然性的进程切换突高现象，突高时候会堵塞大量进程；
- ⑤DB机在凌晨 00:00,会有持续约40s左右的进程堵塞；
- ⑥DB机的iowait一直维持在1~3，极少时候是0.

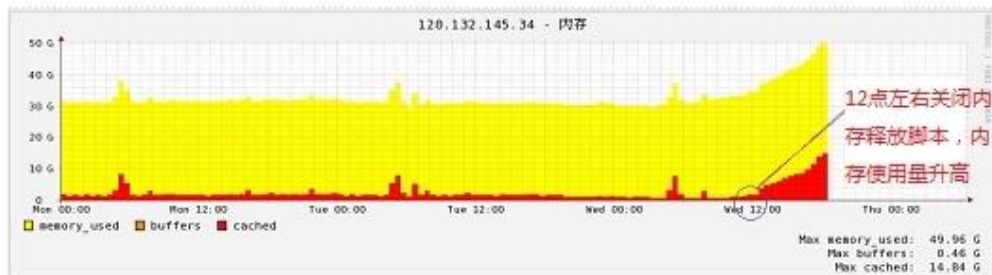
经过各位同学今天的努力，已经对这些问题找到以下原因：

DB机偶然性的进程切换突高，导致进程大量堵塞

这个是由于server的内存优化策略导致的，运维脚本每隔20分钟，就会释放一次内存，每次释放内存的时候，就会导致进程堵塞，进程切换突高。现在这个策略已经关闭。关闭之后，不会再出现进程切换突高现象了，但是服务器的内存占用升高，影响如下图：

memory

开始时间: 2013-11-25 00:00 结束时间: 2013-11-28 06:25 查询



负载分析

【检查点2】回复: 四川麻将负载优化项目立项及项目状态周知 ☆

池育龙

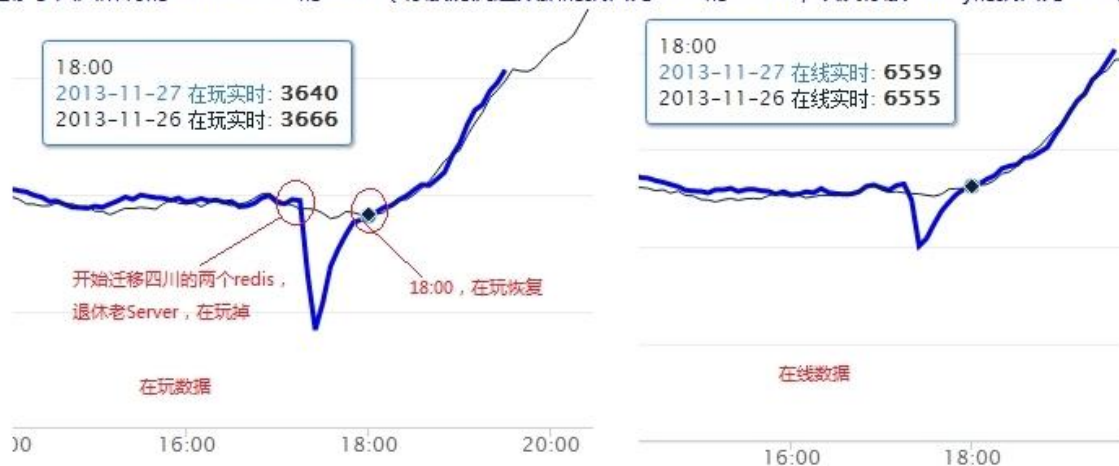
邮件回复时间: 2013/12/3 16:33:19。

DB机在凌晨 00:00会有持续40s左右的进程堵塞

- ①nginx的日志在00:00通过gzip进行压缩, 占了很高的cpu。现在已经改成不再gzip压缩, 直接重命名, 等过期之后再删除。
- ②PHP会在00:00进行一些日志统计, 排行计算的操作。现在已经调整算法, 优化对DB的访问。
- ③DB中对表的切换(调用rename来进行切表, 速度很快, 影响不大)

玩牌逻辑已经和牌局日志分析、入库隔离开来, DB机的负载再高也不应该影响到Server机

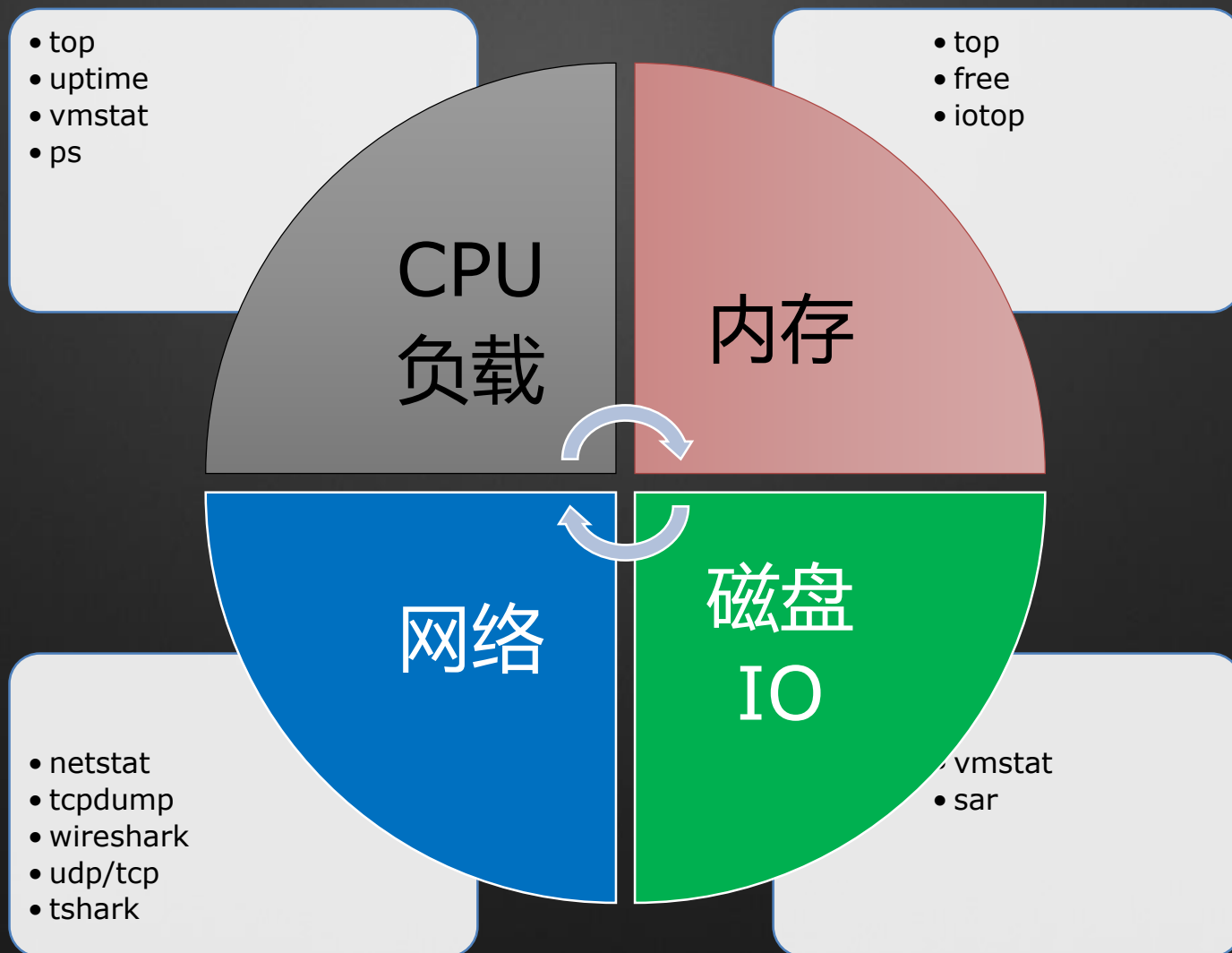
经过讨论, 认为DB机的进程堵塞有可能影响到Redis, 导致GameServer访问Redis的时间过长(因为GameServer对Redis的访问是同步读写)从而影响到玩家玩牌, 为了避免这中情况, 在下午已经把GameServer所连接的两个Redis从DB机迁移到了Server机。目前仅迁移了四川麻将的GameServer的Redis(存放防沉迷数据的端口为4530的Redis, 以及存放mtkey的端口为4532的Redis)。



接下来的工作

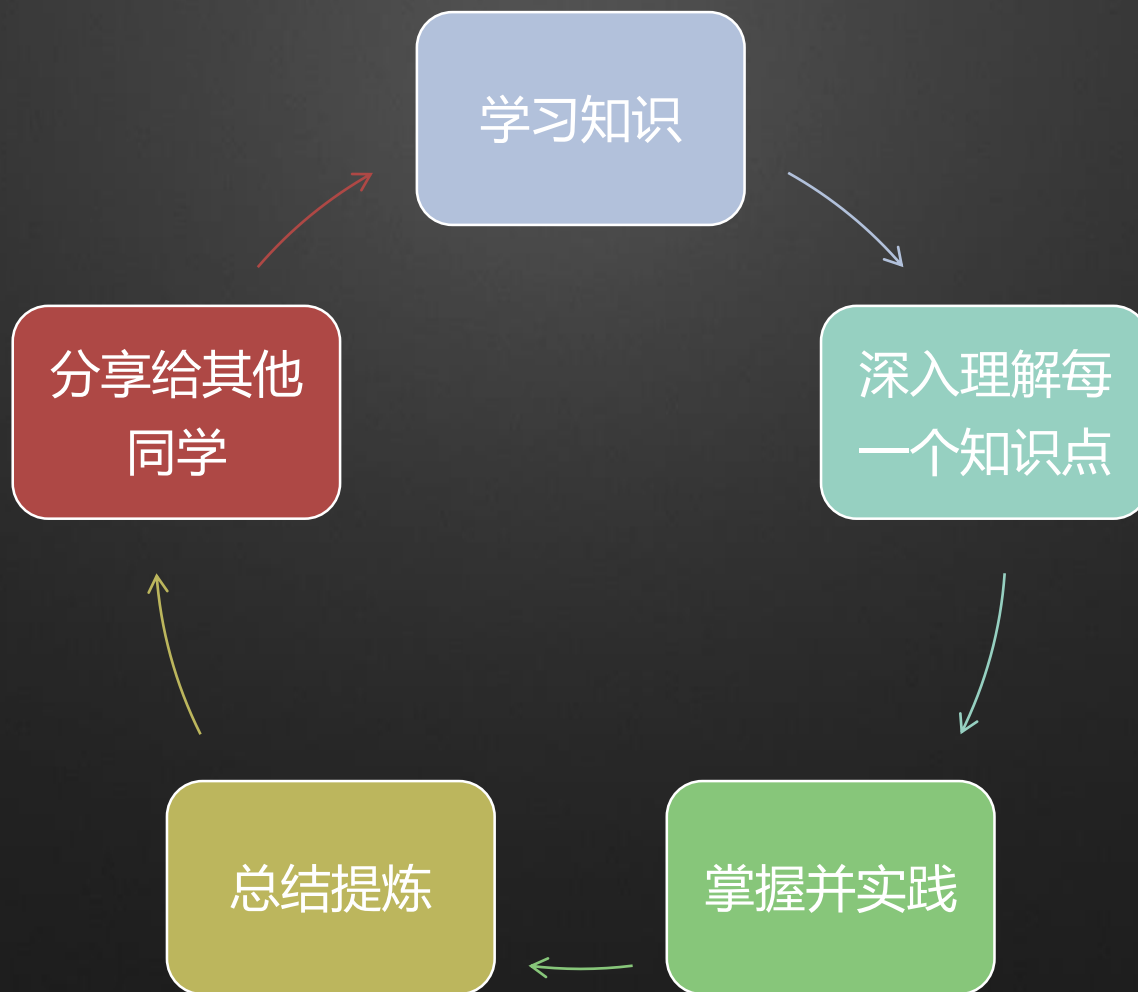
- ①Server机的写磁盘次数过高, 写日志比较凶。GameServer需要对日志进行分级, 控制写日志的量。
- ②DB机内存占用很高, 有可能影响到Server的安全运行, 需要持续关注。
- ③对DB的访问优化

总结



next...

知识管理



Thank you!