

# MySQL性能监控与诊断

苏普@Taobao DBA



追風堂



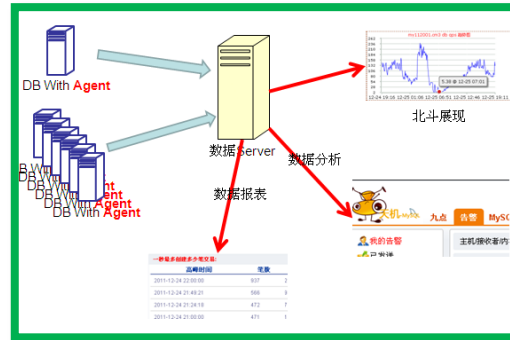
# 目录



- 监控系统天机：
  - 群组、Dashboard ...



- 监控架构



- 其他工具

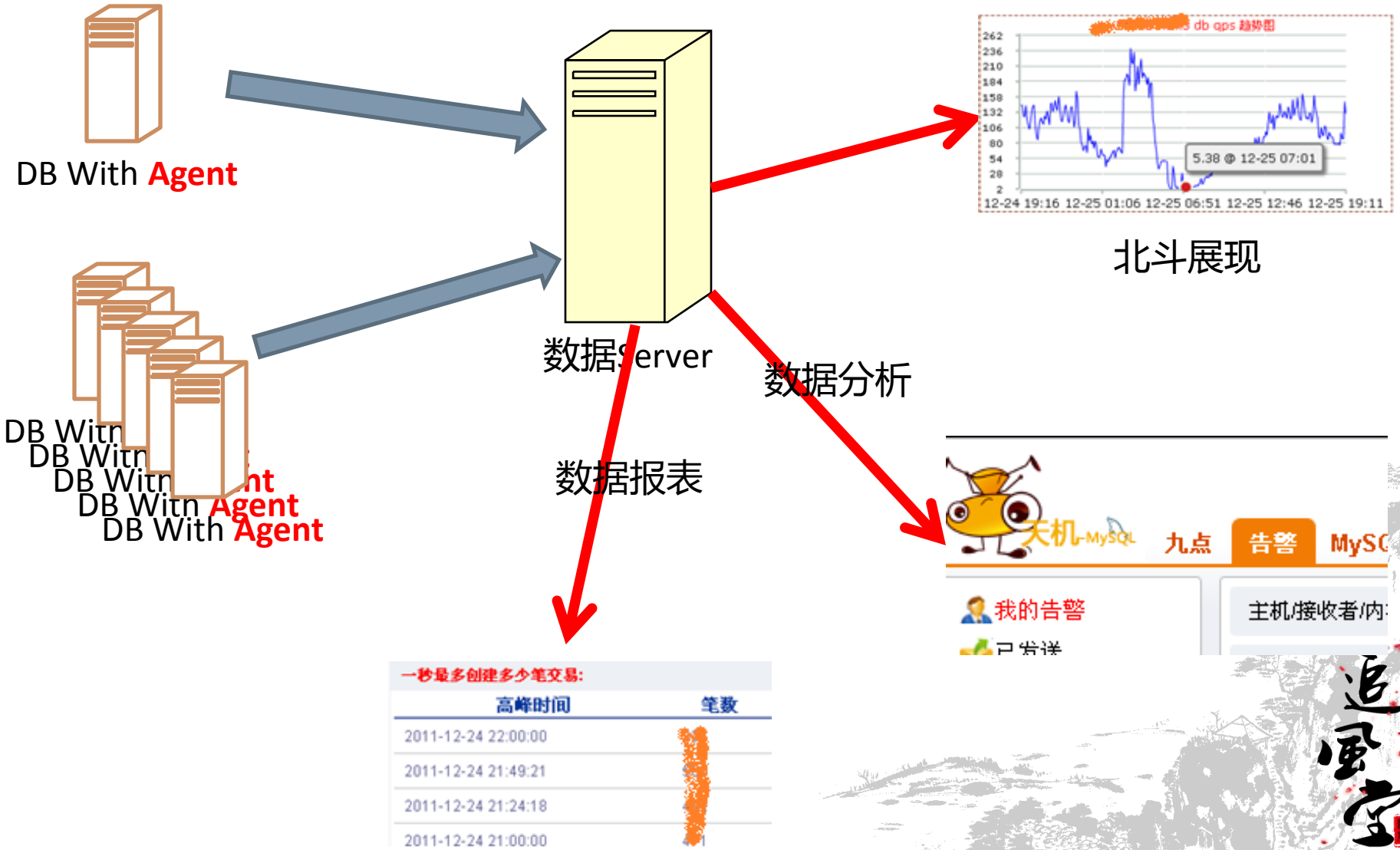


**TCPDUMP** & LIBPCAP

**PROFILE**

追風堂

# MySQL监控：告警架构





天机-MySQL

九点

告警

MySQL

Dashboard

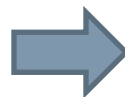
群组

- 找到你关心的主机：

告警 MySQL Dashboard 群组

主机/库名/IP  submit

MySQL群组余量预估



告警 MySQL Dashboard 群组 Re

主机名  submit

追風堂

# MySQL监控：Dashboard



天机 ==> Dashboard

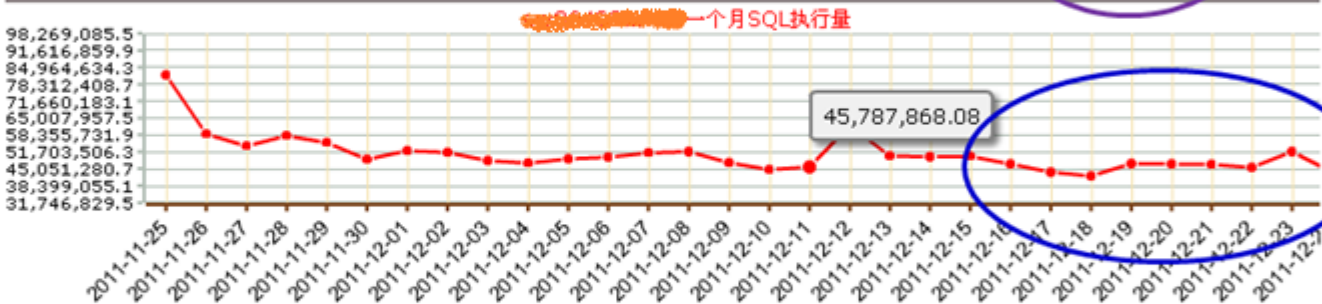


九点 告警 MySQL Dashboard 群组 Redis 垂直业务 JADE Oracle HBase 安装 Labs Wiki

- 88 概况
- 趋势图
- 当前连接
- 当前应用
- 空间状况
- 实时状态
- 备份信息
- 备份详情
- 告警信息
- 主机快照
- 群组状态
- 余量分析
- Top SQL
- SQL Detail
- 表空间
- 响应时间

主机名  submit

主机名	IP地址	角色	版本	u01.free	uptime	check time	Owner	库名
...	...	master	5.1.	1552.44/934.18	527 days	2011-12-25 18:48:52	...	...



KPI	Now	Last Day
QPS	651.08	627
TPS	57.35	68
Load	1.11	1.1
iowait	7.99	6

```
Running SQL refresh STOP!! 查看详细
SELECT id, ...
active status enabled status ...
FROM ...
WHERE id = ...
```

Advise



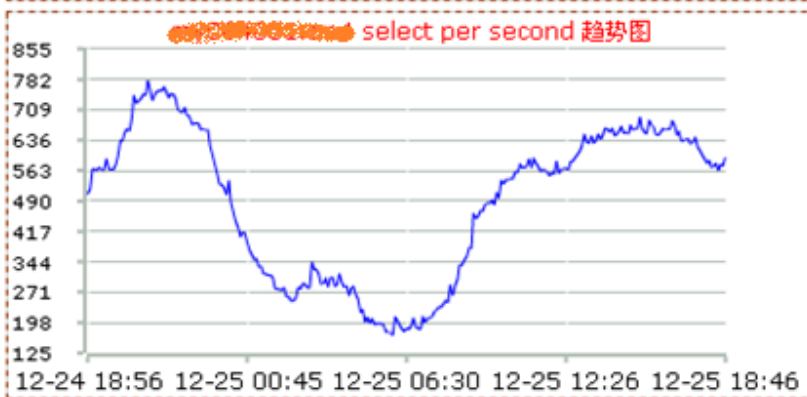
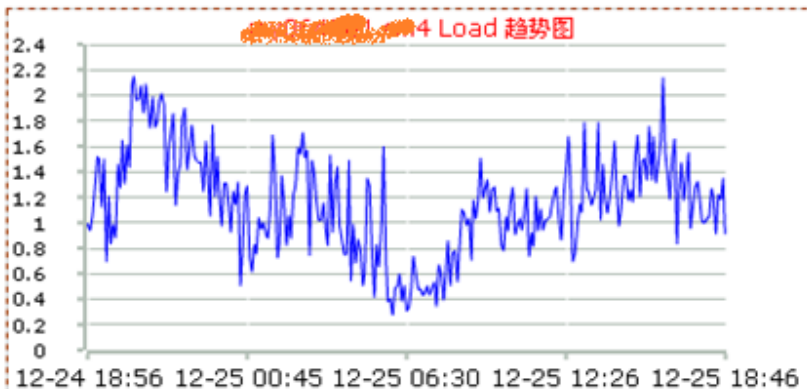
# MySQL监控：性能指标



天机 ==> Dashboard ==> 趋势图

- slow log
- network out
- Com commit
- user cpu
- ior
- Response Time
- sys cpu
- low
- SELECT / IUD
- cpu iowait
- phr
- memn useage
- phw
- network in
- slave lag

红色为基线 小图 最近3天 最近1天 最近12小时 最近3小时



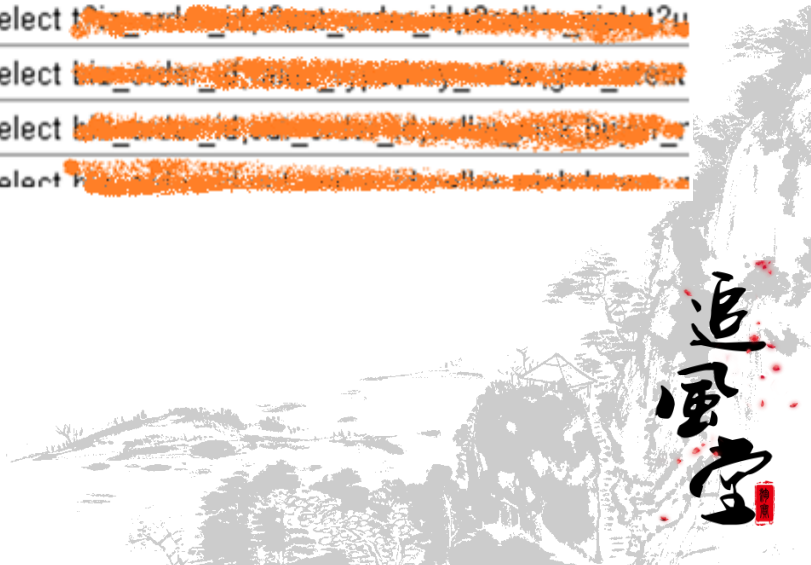
追風堂

# MySQL监控：SQL运行



天机 ==> Dashboard ==> Top SQL

queryid	calls	call_pct	rt_avg	rt_pct	fingerpr
B32BB699843C5D71	374929	17.52%	232.4419	11.08%	select [redacted]
F832C5B31D12081E	322926	15.09%	135.5688	5.57%	select [redacted]
30B30631148CD5DD	284756	13.31%	136.5046	4.94%	select [redacted]
MISC	257520	12.03%	255.1006	8.35%	
15ABFAE5C60384F7	179703	8.40%	579.0109	13.23%	select [redacted]
52C3BFFB0DE58A93	179245	8.38%	503.4439	11.47%	select [redacted]
E87B47C86174D6AA	80900	3.78%	269.5909	2.77%	select [redacted]
9621DD616015ED94	58451	2.73%	2235.9891	16.61%	select [redacted]
AF87F44BA6280EA7	44138	2.06%	174.6998	0.98%	select [redacted]
9DDC893C48584B6E	37335	1.74%	549.8567	2.61%	select [redacted]
F68C1A8BFRDAAEF141	23158	1.08%	905.8727	2.67%	select [redacted]



# MySQL监控：群组监控

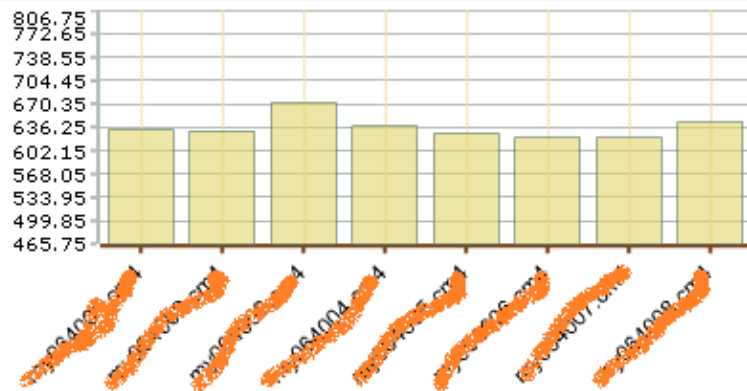


天机 ==> 群组 ==> KPI趋势

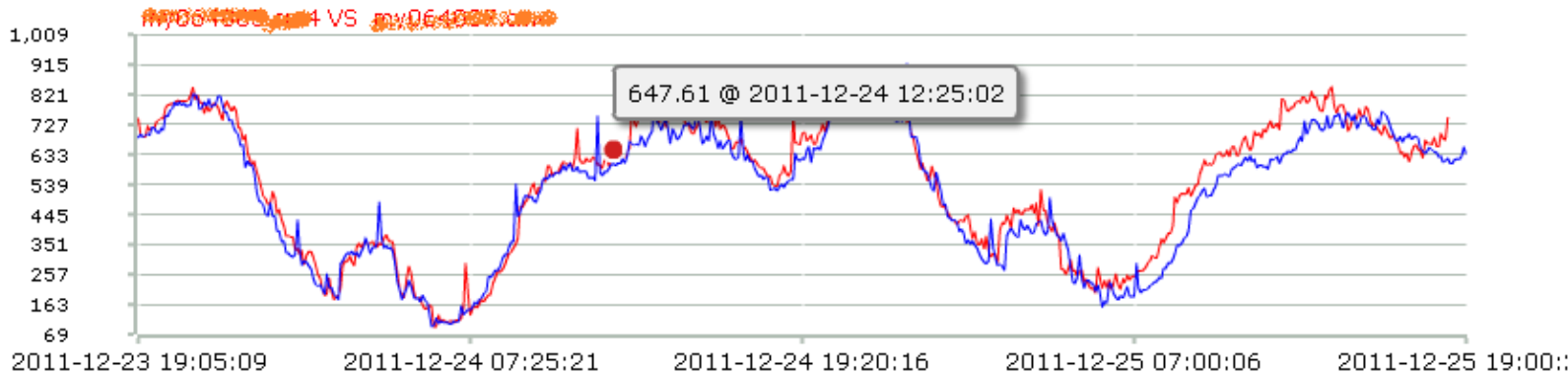
告警 MySQL Dashboard **群组** Redis 垂直业务 JADE Oracle HBase 安装 Labs Wiki 回北斗

主机组:  KPI:  时间: 2011-12-25 18:01 主备:

supu	最大	mysql-main	mysql-main	673.00
supu	最小	mysql-main	mysql-main	621.00
supu	平均	--ip	supu	636.12
supu	最大	supu	supu	0 ip
supu	最小	supu	supu	0 ip
supu	平均	--ip	supu	0.00



抽样



mysql-main 一个月SQL执行量

追風堂



# MySQL监控：实时监控



天机 ==> 群组 ==> 实时趋势



九点 告警 MySQL Dashboard 群组 Redis 垂直业务 JADE Oracle HBase

KPI趋势图

群组余量

余量细节

实时趋势

主机组:

主备:

```
=====
Welcome to use online real-stat tools   Group: mysql-main
Powered by Taobao DBA   Interval: 10s
===== Date : 2011-12-25 19:02:48 =====
```

-----MySQL Status Information-----

Time	insert	update	delete	select	commit	lor	low	phyr	phyw
19:03:35	109	420	35	4896	7	68125	8213	1445	540
19:03:24	111	404	35	5228	7	77960	7734	1509	552
19:03:13	131	415	40	5417	11	69907	7931	1498	622
19:03:01	105	391	34	4741	7	64307	7642	1352	559
baseline	125	356	39	4686	8	0	0	0	0

You can use script do more work, but sadly it make more jobs come to u

追風堂

# MySQL监控：告警



天机 ==> 告警 ==> 已发送



九点

告警

MySQL

Dashboard

群组

Redis

垂直业务

JADE

Oracle

+

我的告警

已发送

分析报告

告警过滤

待发送

告警历史

脚本错误

告警主机组

告警用户组

告警用户

主机阈值

MySQL阈值

Oracle阈值

MySQL归属

主机/接收者/内容  时间  到  类型:

已发送告警(总计3294条)

时间	告警组	类型	告警内
2011-12-25 19:00:34	[REDACTED]	旺旺	[REDACTED] swap
2011-12-25 19:00:34	[REDACTED]	旺旺	[REDACTED] 9071210C Reading from net NULL
2011-12-25 19:00:34	[REDACTED]	旺旺	[REDACTED] load mon
			[REDACTED] 11225 19:00:06 [M "db_... user: "user" host: "192.168.1.100" packets)11225 19:00:06 [Warning] Aborted connection 2817 "192.168.1.100" (Got an error reading comm

追風堂

# MySQL监控：告警阈值



天机 ==> 告警 ==> 主机阈值

1	pu	supu	10	30	70	90	10	30	idle_time	修改   删除
2	pu	supu	6	12	60	80	10	15	work_time	修改   删除
3	pu	supu	6	15	60	90	10	15	peak_time	修改   删除
4	pu	supu	20	30	70	90	45	50	idle_time	修改   删除
5	pu	supu	20	30	70	90	45	50	work_time	修改   删除
6	pu	supu	20	25	70	90	45	50	peak_time	修改   删除
7	pu	supu	20	25	70	90	5	10	idle_time	修改   删除
8	pu	supu	20	25	70	90	5	10	work_time	修改   删除
9	pu	supu	20	25	70	90	5	10	peak_time	修改   删除
10	u	supu	30	60	70	90	20	40	idle_time	修改   删除
11	u	supu	30	50	70	90	15	30	work_time	修改   删除
12	u	supu	30	40	70	90	15	30	peak_time	修改   删除

合并告警

多次确认发送



# 常见故障排查



- 慢日志

```
mk-query-digest \  
--type=slowlog \  
--since="2011-11-02 15:14:00" \  
--until="2011-11-02 15:16:25" \  
--filter '($event->{host} || $event->{ip} || "") =~ m/192.168.0.1/' \slow.log  
|more
```

```
# Profile  
# Rank Query ID           Response time      Calls R/Call      Item  
# ==== =====  
#      1 0xB1B47A21CCF9A417 12.3520 69.6%      8      1.5440 SELECT user_item_  
#      2 0xF44FE6ABE48FDFC7  2.9784 16.8%      2      1.4892 SELECT bask_item_  
#      3 0x594F0CBAEE137C67  2.4176 13.6%      2      1.2088 SELECT bask_item_  
  
# Query 1: 0.00 QPS, 0.00x concurrency, ID 0xB1B47A21CCF9A417 at byte 4519  
# This item is included in the report because it matches --limit.  
#          pct  total    min    max    avg    95%  stddev  median  
# Count          66      8  
# Exec time      69     12s     1s     2s     2s     2s    516ms     2s  
# Lock time      59    445us    45us   61us   55us   60us     5us    57us  
# Rows sent      10     328     14     87     41    84.10   29.52   43.20  
# Rows exam      59    9.26k    645   1.67k   1.16k   1.61k   393.97  918.49  
# Users          1    snsbask
```

# 常见故障排查(2)



- Orzdba --lazy

```
=====
| Welcome to use the orzdba tool !
| Yep...Chinese English~
|===== Date : 2011-12-25 =====

HOST: 192.168.1.100 IP: 192.168.1.100
DB : mysql
Var : innodb_buffer_pool_size[17G] innodb_flush_log_at_trx_commit[2] innodb_flush_method[O_DIRECT] i

-----load-avg-----cpu-usage---swap---
time | 1m 5m 15m |usr sys idl iow| si so|
19:40:54 | 0.78 0.66 0.57 | 1 0 97 2 | 0 0 |
19:40:55 | 0.78 0.66 0.57 | 0 0 97 2 | 0 0 |
19:40:56 | 0.78 0.66 0.57 | 0 0 98 1 | 0 0 |
19:40:57 | 0.78 0.66 0.57 | 2 1 94 3 | 0 0 |
19:40:58 | 0.72 0.64 0.57 | 0 0 96 3 | 0 0 |
19:40:59 | 0.72 0.64 0.57 | 0 0 98 2 | 0 0 |
19:41:00 | 0.72 0.64 0.57 | 0 0 97 3 | 0 0 |
19:41:01 | 0.72 0.64 0.57 | 1 0 95 3 | 0 0 |
19:41:02 | 0.72 0.64 0.57 | 4 0 93 2 | 0 0 |

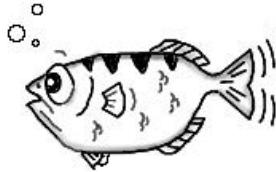
-QPS- -TPS- -Hit%-
ins upd del sel iud | lor hit|
0 0 0 0 0 | 0 100.00 |
10 6 0 125 16 | 7304 98.56 |
11 4 0 110 15 | 4987 98.78 |
10 5 0 148 15 | 5777 97.73 |
26 11 0 115 37 | 5571 97.77 |
18 8 0 88 26 | 2677 97.46 |
16 12 0 125 28 | 8110 98.16 |
15 10 0 147 25 | 5355 97.31 |
14 12 0 114 26 | 8814 99.02 |
```



# Internal Troubleshooting



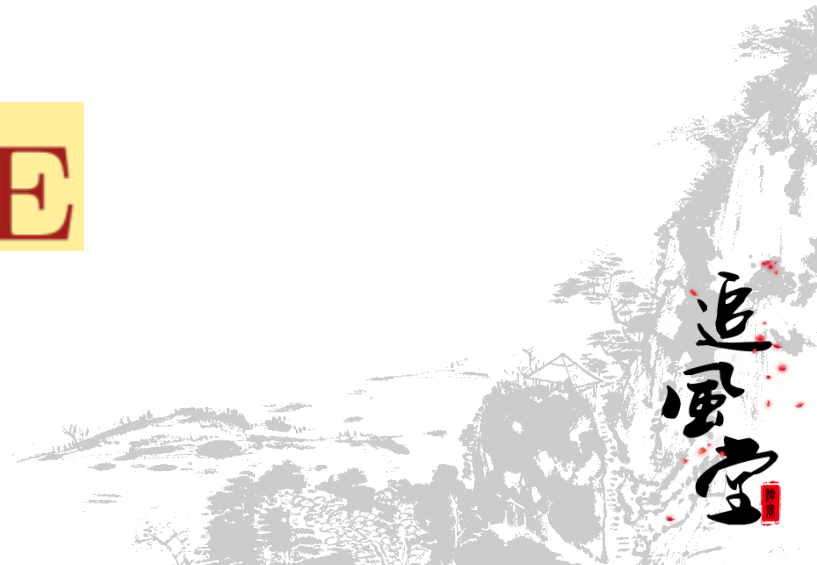
- **GDB**



- **tcpdump**



- **oprofiled**



- Quick And Simple

```
sudo sh -c 'gdb -ex "set pagination 0" -ex \  
"thread apply all bt" --batch \  
-p $(pidof mysqld) > bt.log'
```

```
sudo sh -c 'gdb -ex "set pagination 0" -ex \  
"thread apply all bt full" --batch \  
-p $(pidof mysqld) > bt.log.full'
```



- Sample:

```
#0  0x00002aaab3020481 in buf_LRU_search_and_free_block
#1  0x00002aaab302094e in buf_LRU_get_free_block
#2  0x00002aaab301643e in buf_page_init_for_read
#3  0x00002aaab30214d5 in buf_read_page_low
#4  0x00002aaab3021ec1 in buf_read_ahead_linear
#5  0x00002aaab3018d0e in buf_page_get_gen
#6  0x00002aaab300c612 in btr_pcur_move_to_next_page
#7  0x00002aaab30b8d6a in row_search_for_mysql
#8  0x00002aaab305133c in ha_innodb::general_fetch
#9  0x00000000006a7d97 in rr_sequential
#10 0x000000000061b756 in sub_select
#11 0x000000000062e0ad in do_select
#12 0x000000000063b0d4 in JOIN::exec
```







- **PMP: poor man's profiler**

```
for x in $(seq 1 $nsamples)
do
  gdb -ex "set pagination 0" -ex "thread apply all bt"
  -batch -p $pid
  sleep $sleeptime
done | \
awk '
BEGIN { s = ""; }
/Thread/ { print s; s = ""; }
/^\#/ { if (s != "" ) { s = s "," $4} else { s = $4 } }
END { print s }' | \
sort | uniq -c | sort -r -n -k 1,1
```



- tcpdump:

```
nohup tcpdump -n -nn -tttt -i bond0 \
-s 65535 'port 3306' -w tcpdump.ret -C 100 &
```

tcpdump -tttt -i bond0 -s 65535 -w tcpdump.ret -C 100 &

一个TCP包中包含多个MySQL协议包；一个协议包就是一条MySQL记录：

⊕ Transmission Control Protocol	⊖ MySQL Protocol
⊕ [2 Reassembled TCP Segments (0 - 1436 bytes captured)]	Packet Length: 448
⊕ MySQL Protocol	Packet Number: 72
⊕ MySQL Protocol	text: \220\216\201\201
⊕ MySQL Protocol	text: 49986964
⊕ MySQL Protocol	text: i3/T1s9FaxnxhxJD1upjX.jpg
⊕ MySQL Protocol	text: 50019938
⊕ MySQL Protocol	text: 760.00
⊕ MySQL Protocol	text: 760.00
⊕ MySQL Protocol	text: b





- tcpdump: Client端瓶颈

```
00:30:08.055266 192.168.1.104 3306 192.168.1.107 44711 234 [TCP segment of a reassembled PDU]
00:30:08.055270 192.168.1.104 3306 192.168.1.107 44711 2458 [TCP window Full] Response
00:30:08.055868 192.168.1.107 44711 192.168.1.104 3306 66 [TCP Zerowindow] 44711 > mysql [ACK] Seq=100000000
00:30:08.056468 192.168.1.107 44711 192.168.1.104 3306 66 [TCP window update] 44711 > mysql [ACK] Seq=100000000
00:30:08.056471 192.168.1.104 3306 192.168.1.107 44711 2626 [TCP window Full] Response
00:30:08.056832 192.168.1.107 44711 192.168.1.104 3306 66 [TCP Zerowindow] 44711 > mysql [ACK] Seq=100000000
00:30:08.057196 192.168.1.107 44711 192.168.1.104 3306 66 [TCP window update] 44711 > mysql [ACK] Seq=100000000
00:30:08.057201 192.168.1.104 3306 192.168.1.107 44711 2626 [TCP window Full] Response
00:30:08.057568 192.168.1.107 44711 192.168.1.104 3306 66 [TCP Zerowindow] 44711 > mysql [ACK] Seq=100000000
00:30:08.057931 192.168.1.107 44711 192.168.1.104 3306 66 [TCP window update] 44711 > mysql [ACK] Seq=100000000
```





- tcpdump: Server端瓶颈

Filter: ip.addr == 243.2.2.87 Expression... Clear Apply

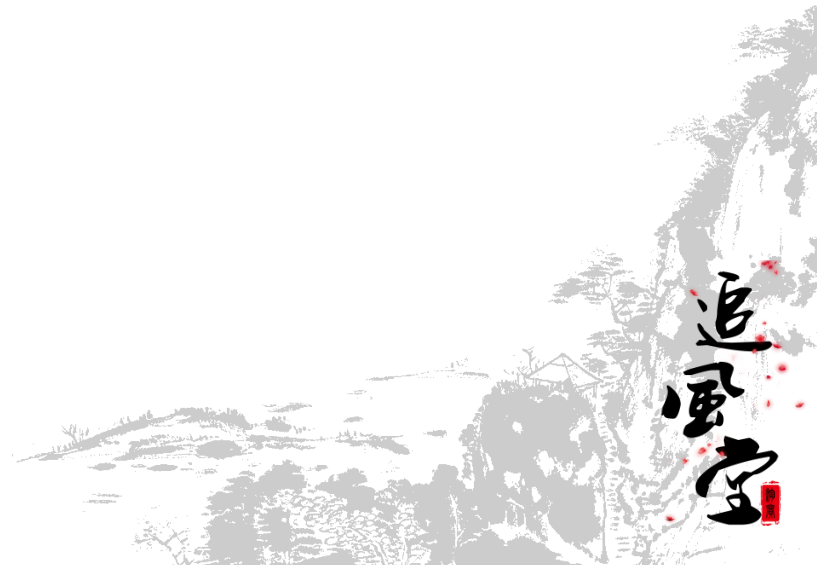
No.	Source	src port	Destination	dest port	Length	Info
00:44:16.110617	172.17.0.1	44711	172.17.0.16	3306	66	44711 > mysql [ACK] Seq
00:44:16.110626	172.17.0.16	3306	172.17.0.87	44711	1514	[TCP segment of a reass
00:44:16.110630	172.17.0.16	3306	172.17.0.87	44711	978	[TCP segment of a reass
00:44:16.110633	172.17.0.87	44711	172.17.0.16	3306	66	44711 > mysql [ACK] Seq
00:44:16.110635	172.17.0.87	44711	172.17.0.16	3306	66	44711 > mysql [ACK] Seq
00:44:16.110637	172.17.0.87	44711	172.17.0.16	3306	66	44711 > mysql [ACK] Seq
00:44:16.110639	172.17.0.87	44711	172.17.0.16	3306	66	44711 > mysql [ACK] Seq
00:44:16.110742	172.17.0.87	44711	172.17.0.16	3306	66	44711 > mysql [ACK] Seq
00:44:16.113845	172.17.0.87	44711	172.17.0.16	3306	66	44711 > mysql [ACK] Seq
00:44:20.986612	172.17.0.16	3306	172.17.0.87	44711	15994	[TCP segment of a reass
00:44:20.989779	172.17.0.87	44711	172.17.0.16	3306	66	44711 > mysql [ACK] Seq
00:44:20.989797	172.17.0.87	44711	172.17.0.16	3306	66	44711 > mysql [ACK] Seq
00:44:20.989876	172.17.0.87	44711	172.17.0.16	3306	66	44711 > mysql [ACK] Seq
00:44:20.989887	172.17.0.87	44711	172.17.0.16	3306	66	44711 > mysql [ACK] Seq
00:44:20.989891	172.17.0.87	44711	172.17.0.16	3306	66	44711 > mysql [ACK] Seq
00:44:20.989901	172.17.0.87	44711	172.17.0.16	3306	66	44711 > mysql [ACK] Seq
00:44:20.989905	172.17.0.16	3306	172.17.0.87	44711	5858	[TCP segment of a reass
00:44:20.989908	172.17.0.87	44711	172.17.0.16	3306	66	44711 > mysql [ACK] Seq
00:44:20.989911	172.17.0.87	44711	172.17.0.16	3306	66	44711 > mysql [ACK] Seq
00:44:20.989914	172.17.0.16	3306	172.17.0.87	44711	5858	[TCP segment of a reass



- tcpdump: Response Time Monitor

```
$/usr/bin/tcprstat -l [redacted] -t 1 -n 7
```

timestamp	count	max	min	avg	med	stddev	95_max	95_avg	95_std
1325743529	156	15787	22	364	129	1527	295	145	73
1325743530	143	9604	21	294	133	921	302	164	70
1325743531	194	4660	20	215	127	501	285	142	72
1325743532	186	9162	15	390	128	1228	305	138	78
1325743533	130	8661	24	381	133	1167	324	156	65
1325743534	141	5250	23	232	133	560	288	146	76
1325743535	66	382	94	177	138	72	284	168	65
1325743535	66	382	94	177	138	72	284	168	65





- oprofile:

```
sudo yum install binutils-devel
./configure --with-kernel-support && make
make install
sudo opcontrol --deinit
sudo modprobe oprofile timer=1
$ dmesg | grep oprofile | tail -n 1

sudo opcontrol --reset
sudo opcontrol --separate=lib --no-vmlinux \
--start --image=/opt/mysql/bin/mysqld
sudo opcontrol --dump
sudo opcontrol --shutdown
opreport -l /opt/mysql/bin/mysqld
```



- oprofile:

CPU: CPU with timer interrupt, speed 0 MHz (estimated)

Profiling through timer interrupt

samples	%	image name	symbol name
34043	72.3826	ha_innodb_plugin.so.0.0.0	buf_LRU_search_and_free_block
4837	10.2845	ha_innodb_plugin.so.0.0.0	ut_delay
2959	6.2915	ha_innodb_plugin.so.0.0.0	buf_LRU_free_block
1283	2.7279	libpthread-2.5.so	pthread_mutex_lock
851	1.8094	ha_innodb_plugin.so.0.0.0	buf_page_is_corrupted
546	1.1609	ha_innodb_plugin.so.0.0.0	mutex_spin_wait
315	0.6698	libc-2.5.so	memcpy
239	0.5082	ha_innodb_plugin.so.0.0.0	buf_calc_page_new_checksum
179	0.3806	ha_innodb_plugin.so.0.0.0	rec_get_offsets_func
162	0.3444	libpthread-2.5.so	pthread_mutex_unlock
107	0.2275	ha_innodb_plugin.so.0.0.0	row_search_for_mysql
98	0.2084	ha_innodb_plugin.so.0.0.0	row_sel_store_mysql_rec

# 数据查询:WebSQLPlus



WebSQLPlus ==> 分库分表

查询表, 模糊查询请带上 % .....

Welcome 表结构 SQL Plus 分表操作

搜索

分表逻辑名

- 分表逻辑名
- 分表逻辑名\_address
- 分表逻辑名\_account
- 分表逻辑名
- 分表逻辑名\_tra

分表字段 =

结果集

单行详情

1			24		pa
					**

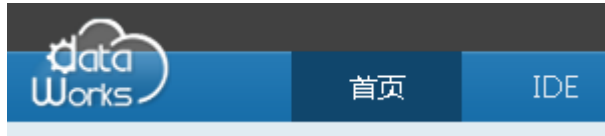
追風堂



# 数据查询:云梯



链接:



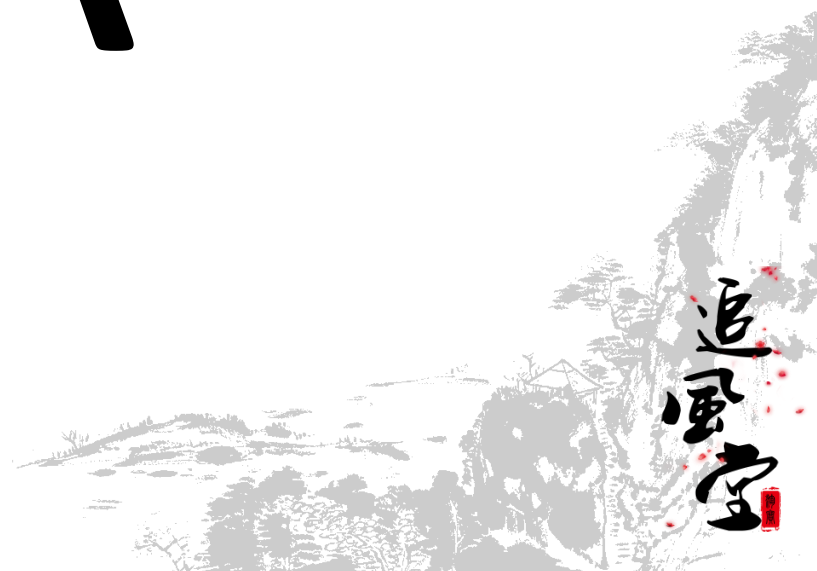
```
新建 保存 运行 运行并保存 停止 格式化 撤销  
AA-20...09:25* (x)  
1 select  
2     count(1)  
3 from  
4     [REDACTED]  
5 where  
6     pt='20111226000000'  
7     and days='50'  
8     and substr(gmt_create, 1, 10) = '2011-12-26'  
9     and [REDACTED];
```

```
hive>  
SELECT count(1)  
FROM s_xxx  
WHERE user_id=452751208 and pt=20111226000000;
```





# Q&A



追風堂