

管理MySQL

周振兴@**Taobao DBA**

<http://orczhou.com>

2012-04-06 DTCC2012

Powered by Taobao DBA

Agenda

- MySQL@Taobao
- DB Architecture@Taobao
- Monitor system
- Others:
 - version||Backup||debug

MySQL@Taobao

tem.taobao.com/item.htm?id=4523840615

豆瓣电台 豆瓣 Google 工作目录 每日常用 MySQL Linux 旅-游 淘宝购物 平时观看 TED

hi, orczhou! 退出

站内信

message

淘宝网首页 | 我要买 | 我的淘宝 | 卖家中心 | 服务中心

淘宝网

服饰 女装
生活 家居

母婴 食品 汽车 婚庆

箱包

数码 护肤

手机 手表

相机 饰品

家电 眼镜

配件 花鸟

笔记本 百货

想找什么

用户库

【四钻信誉】DOTA Tiny - Stone Giant 小小山岭巨人 魔兽手办

商品库



价 格：68.00 元

物流运费：江西南昌 | 至 浙江杭州 - 快递:10.00元 EMS:20.00元

30天售出：0件

评 价：★★★★★4.9分 | 1条评价

宝贝类型：全新 | 1757次浏览

购买数量：1 件 (库存11件)

立即购买

加入购物车

支付：信用卡分期 快捷支付

服务：

购物车

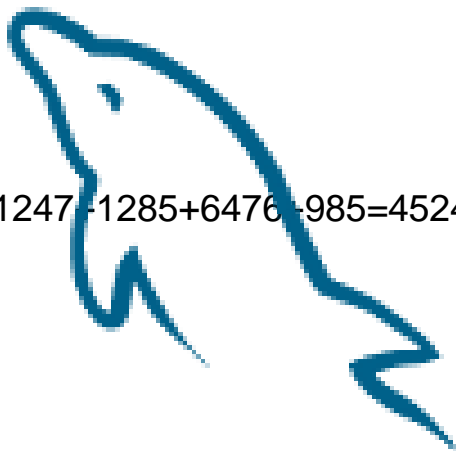
Queries Per second

197010+9519+284006+7700+13195+3161

= **51** 4591 (2011-11-11/12-12)

29069+6182+1247-1285+6476-985=45244

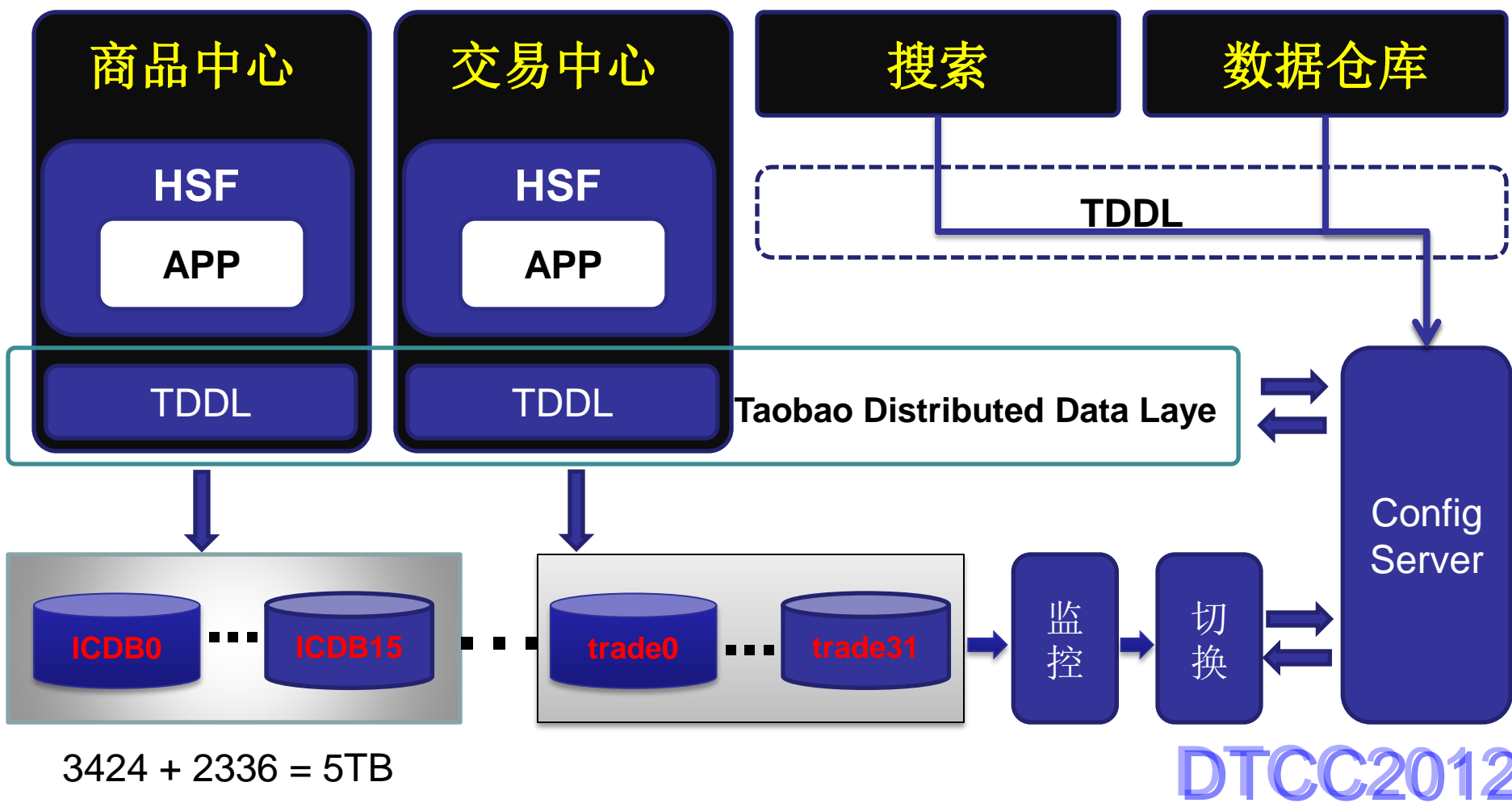
(rt ~ 410us)



DTCC2012

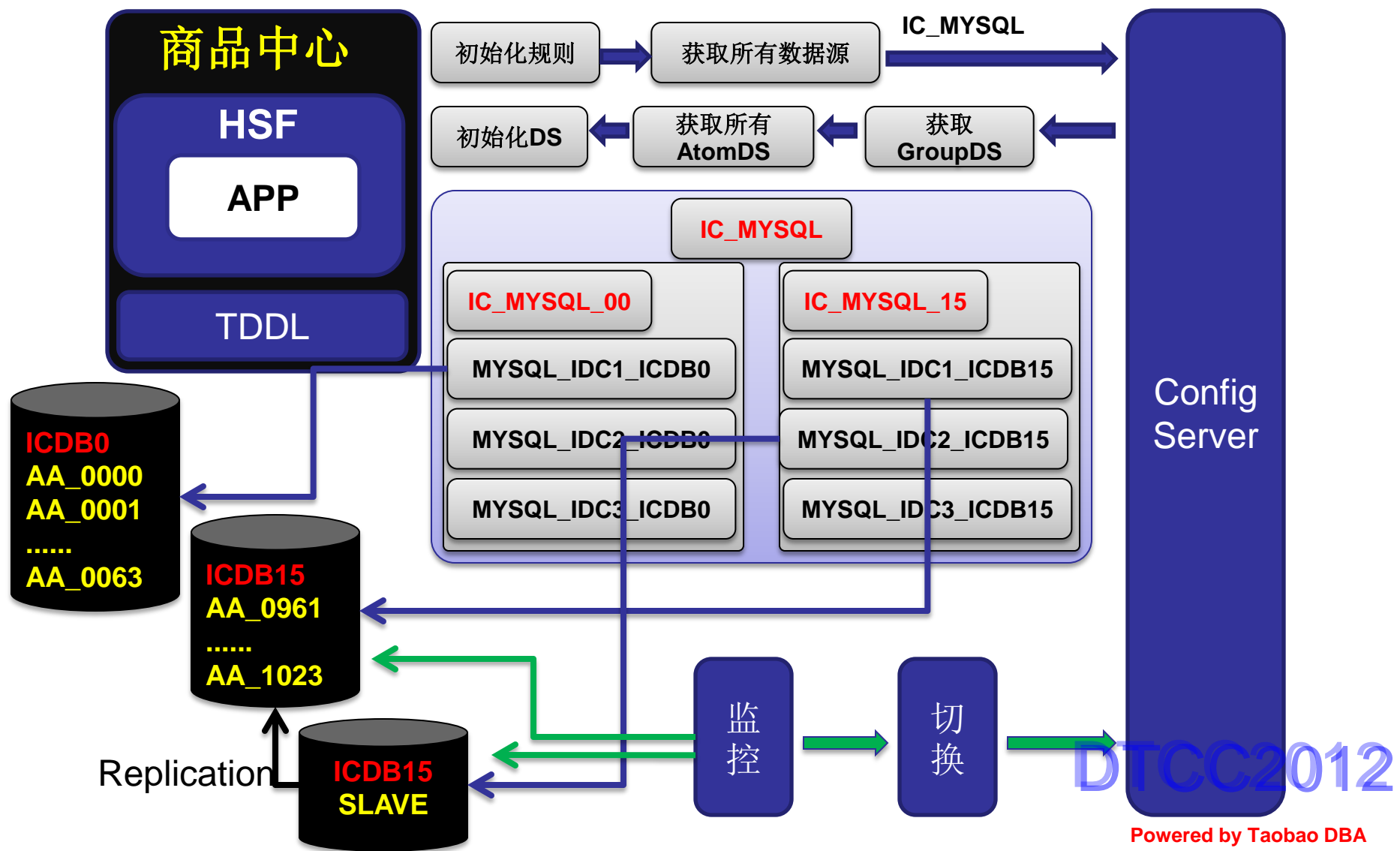
Powered by Taobao DBA

High Level: DB Arch

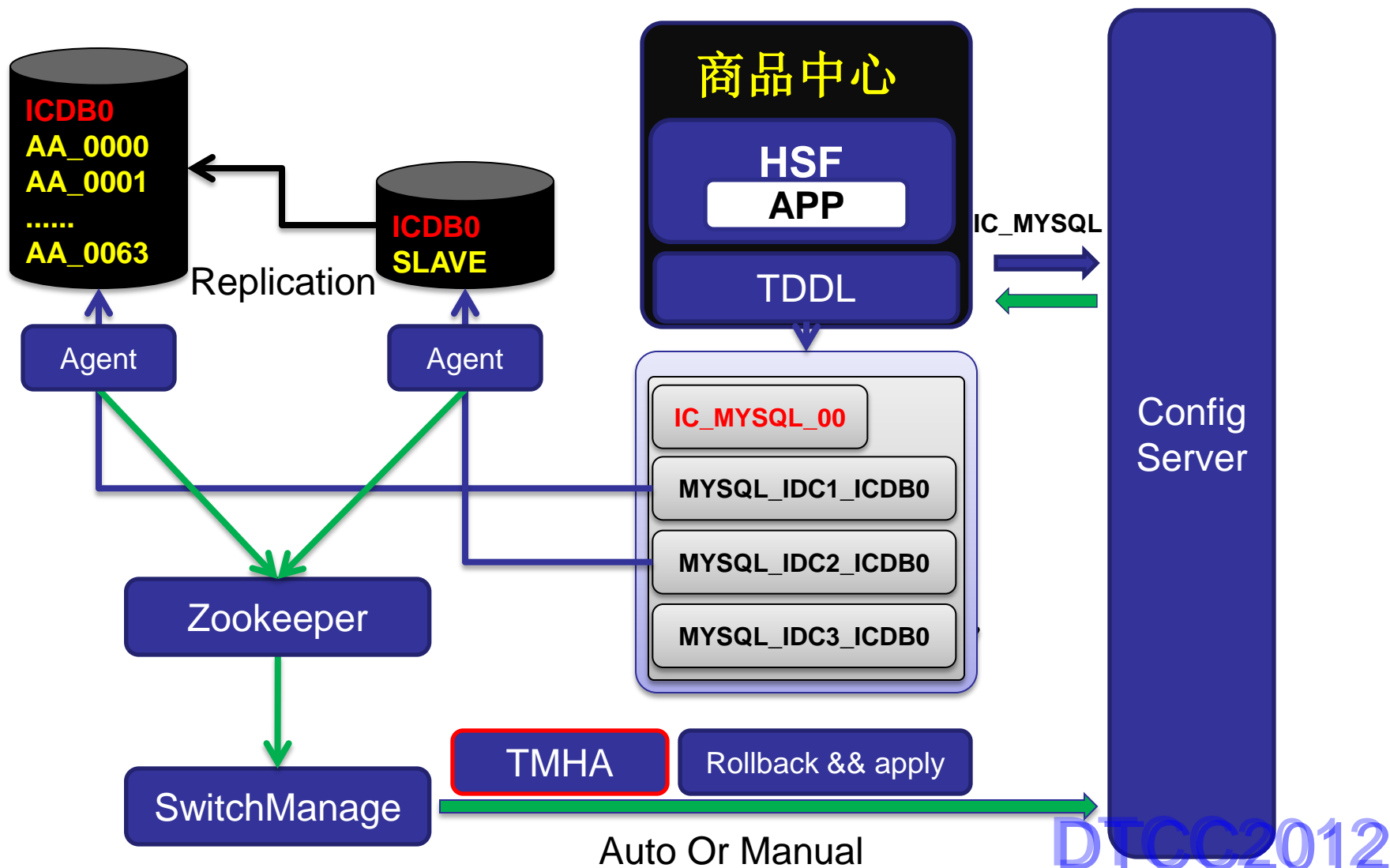


DTCC2012

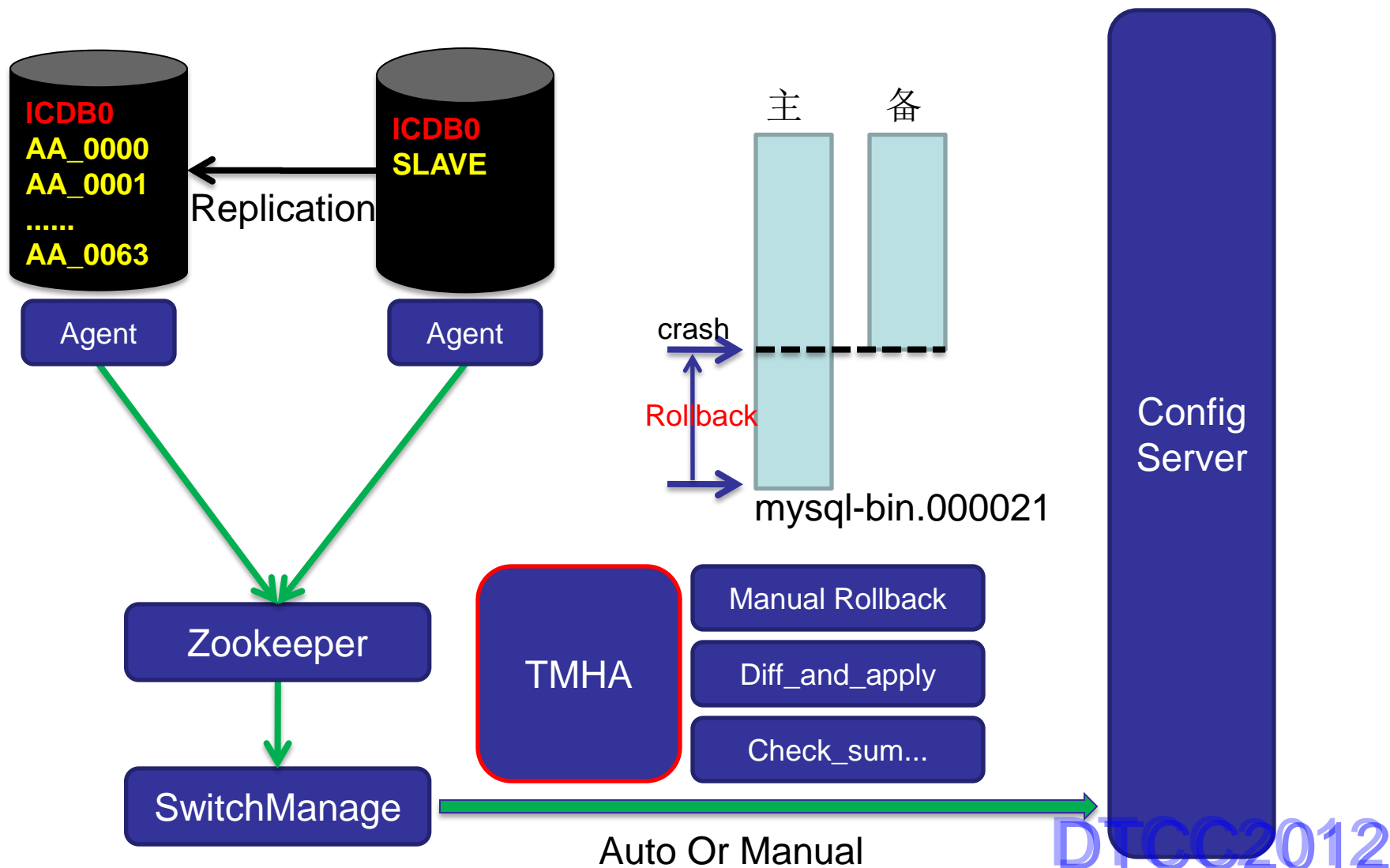
Low level: Groups



Low level: Switch

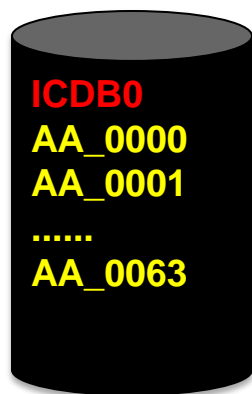


Low level: Consistent



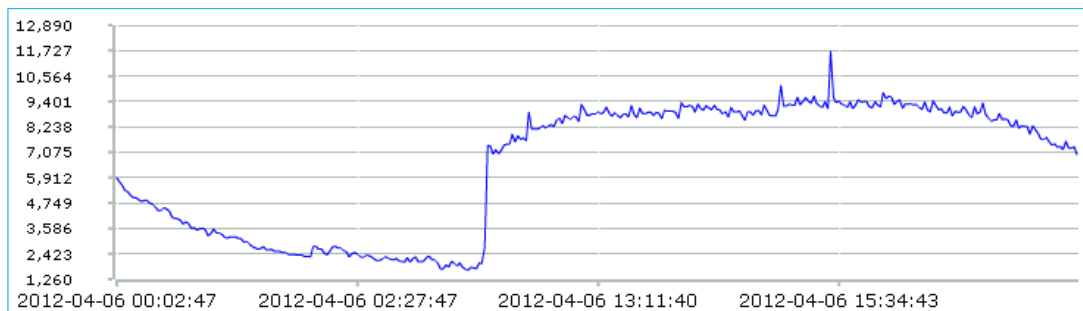
DTCC2012

Monitor system: Sample



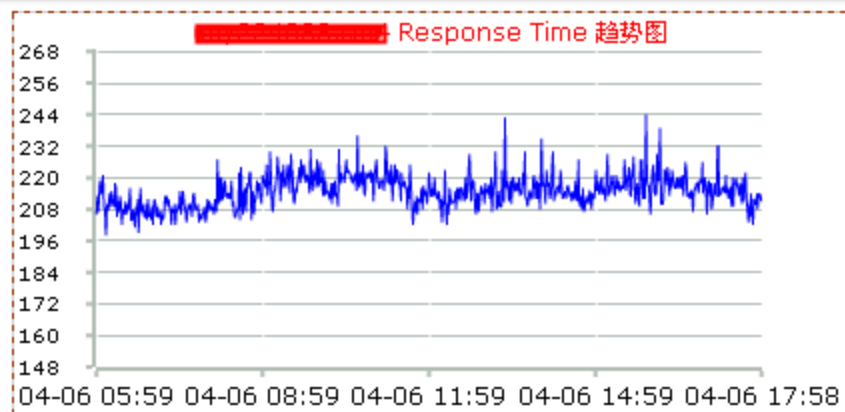
MAX QPS:

11721



Response Time:

230微妙



SQL 6BEE89278B

260887508/per day

MyAWR From 2012-04-05 00:00:00 To 2012-04-05 23:59:59

| Query ID | 总执行次数 | 执行比率/% | RT/微妙 | 总耗时比率/% | |
|------------------|-----------|--------|-------|---------|----------------|
| 6BEE89278B66C689 | 260887508 | 48.6 | 332 | 58.16 | select title,g |
| E699444EE1B56D5F | 144827348 | 26.98 | 191 | 18.57 | select id, us |
| 343033B849D860EA | 44090008 | 8.21 | 300 | 8.88 | select sku_i |
| 4005DF9BBC246C0A | 27780029 | 5.18 | 132 | 2.46 | select sku_i |

DTCC2012

Monitor system: Dashboard

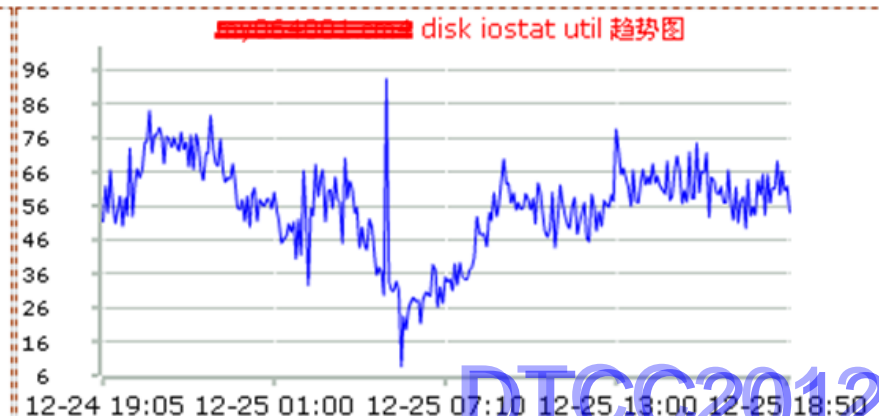
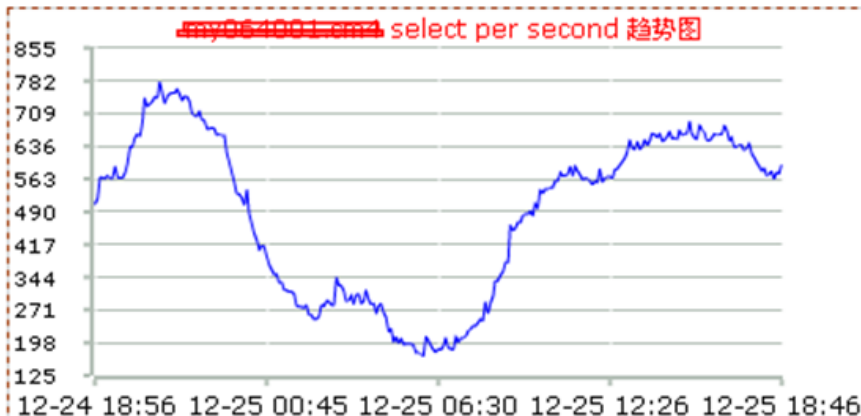
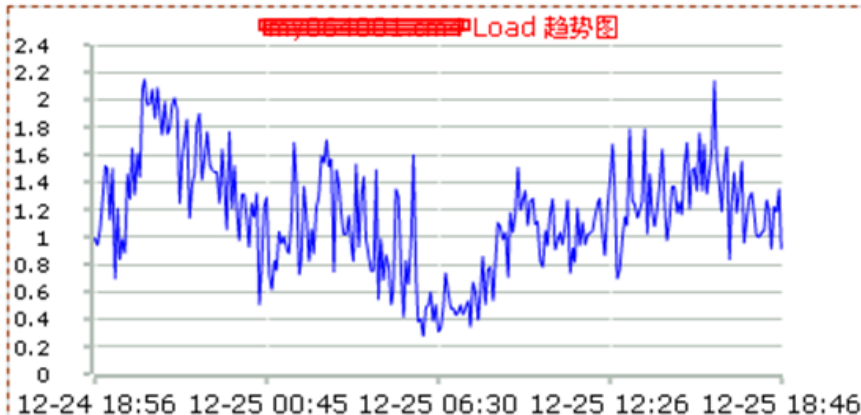


DTCC2012

Monitor system: Performance

☐ slow log ☐ user cpu ☐ sys cpu ☐ cpu iowait ☐ mem usage ☐ network in
☐ network out ☐ ior ☐ low ☐ phr ☐ phw ☐ slave lag
☐ Com commit ☐ Response Time ☐ SELECT / IUD

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



DTCC2012

Monitor system: AWR

MyAWR From 2012-04-10 00:00:00 To 2012-04-10 23:59:59

| Query ID | 总执行次数 | 执行比率/% | RT/微妙 | 总耗时比率/% | SQL Sample |
|------------------|-----------|--------|-------|---------|--|
| 6BEE89278B66C689 | 266476214 | 48.17 | 336 | 58.56 | select title,gmt_modified,starts,ends,pict_url,cate |
| E699444EE1B56D5F | 147789423 | 26.72 | 188 | 18.22 | select id, user_num_id, item_id, image_ |
| 343033B849D860EA | 45461574 | 8.22 | 306 | 9.1 | select sku_id,item_id,properties,quantity,price,sta |
| 4005DF9BBC246C0A | 27750872 | 5.02 | 134 | 2.43 | select sku_id,item_id,properties,quantity,price,sta |
| 3E7F4A61D084D3CA | 14610724 | 2.64 | 191 | 1.82 | select title,gmt_modified,starts,pict_url,category,r |
| 3AEAAD0E15D725B5 | 9347892 | 1.69 | 49 | 0.3 | SET autocommit=1 |
| 2E51F521E9BAE993 | 6477596 | 1.17 | 126 | 0.53 | select sku_id,item_id,properties,quantity,price,sta |

- tcpdump first: few packet will be dropped
 - -i eth0 port 3307 -s 65535 -x **-nn** -q -tttt **-w**
- pt-query-digest
 - --type=tcpdump **--watch-server** \$ip_addr:\$port

DTCC2012

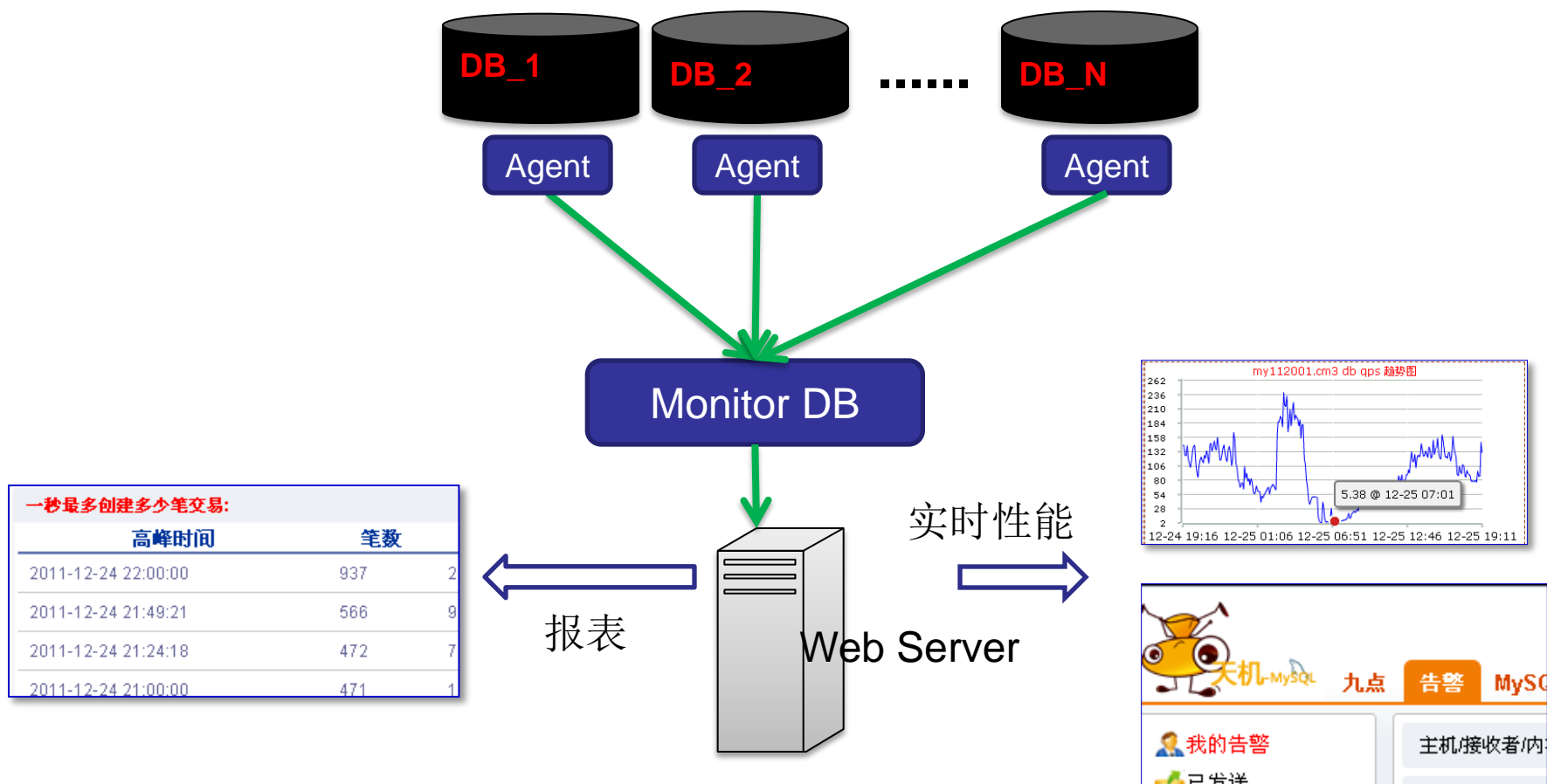
Monitor system:alert optimize

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

DTCC2012

Powered by Taobao DBA

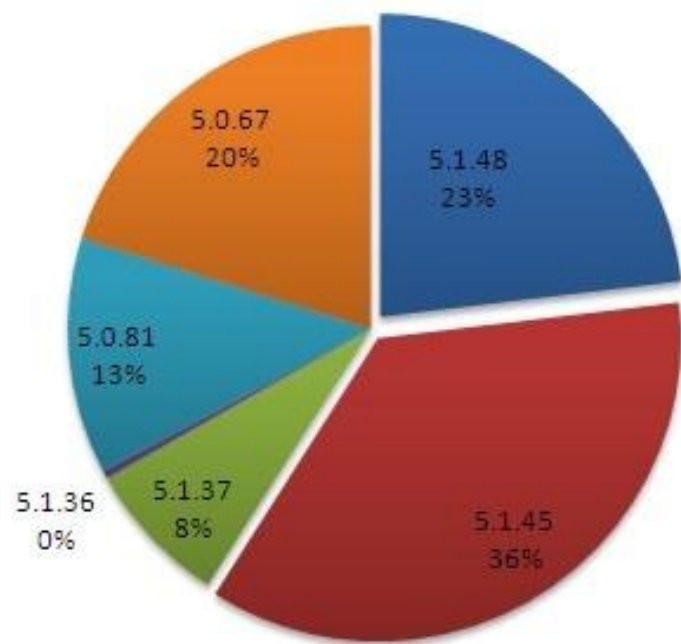
Monitor system: Architecture



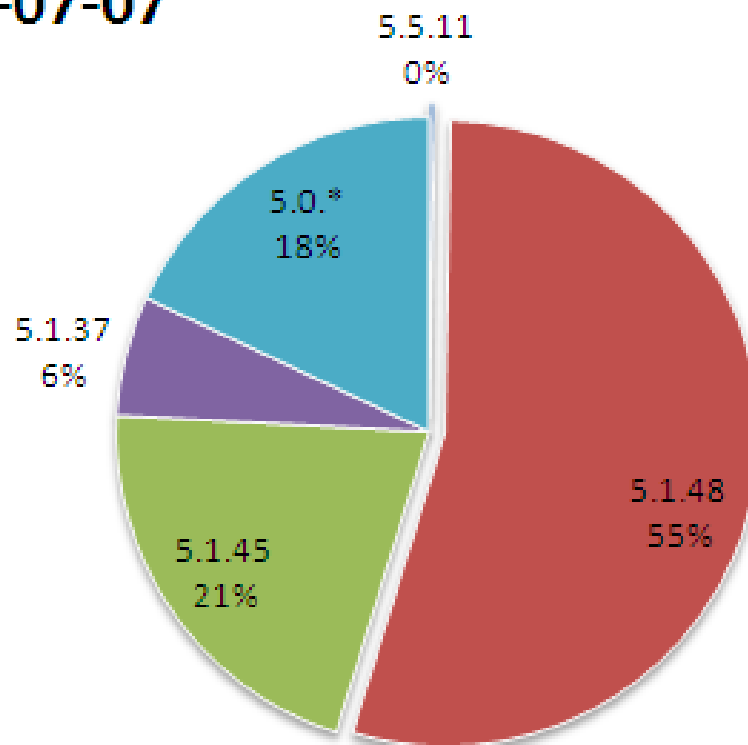
DTCC2012

How to choose MySQL version

2010年12月06日



2011-07-07



DTCC2012

Powered by Taobao DBA

How to choose MySQL version

No PSU; No Metalink; No SAR

Our principle:

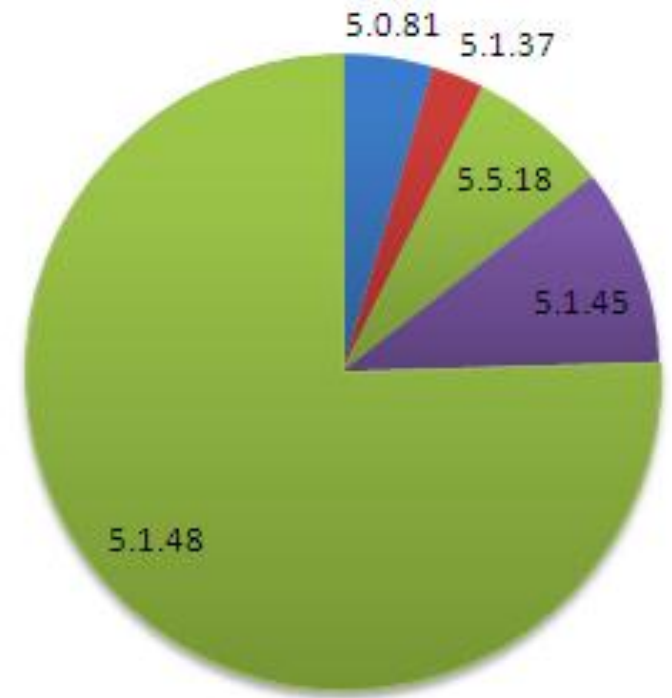
MySQL is basically **Stable**

The **Newer**, the better

Critical and **infamous Bugs**

Critical **new Feature**

Standard everything



2012年04月11日

DTCC2012

Powered by Taobao DBA

How to choose MySQL version

- **MySQL is basically Stable**
- Critical and infamous Bugs
 - DDL table during ddl(fixed in 5.1.61 5.5.19)
- Feature
 - Fast index create(InnoDB Plugin)
 - InnoDB Fast recovery(InnoDB Plugin1.0.9+)
 - Group Commit(Percona 5.5.19)
- Standard everything
 - Hardware Linux/service MySQL my.cnf

DTCC2012

Other:Debug

- Tools
 - MySQL Status
 - percona-toolkit + orzdba and so on.....
 - tcpdump
 - gdb
 - oprofile
 - perf

other-Debug:MySQL Status

- Mostly(50%)
 - Show full processlist

| Time | State | Info |
|------|----------------|---|
| 0.00 | Sorting result | SELECT id, user_id, user_nick, friend_url, url_size, rec_user_ids, gmt_create, gmt_modified FROM friend_url_0026 WHERE user_id = 803525978 AND DATE(gmt_create) = CURDATE() ORDER BY url_size, gmt_create, id LIMIT 0, 1 |
| 0.00 | Sending data | SELECT user_id, target_id, group_id, common_friend_count, common_friends FROM sns_universal_relations_0034 WHERE user_id = 789281698 ORDER BY gmt_modified DESC, common_friend_count DESC LIMIT 0, 200 |

other-Debug:pt(s)

- Slow log

```
/home/mysql/agent/bin/maatkit/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/172.24.27.52/' \  
\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
```

DTCC2012

other-Debug:pt(s)

- orzdba --lazy

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

HOST: 168.115.100.100 IP: 168.115.100.100
DB : snsbank|snsbank_00|snsbank_01|snsbank_02|snsbank_03
Var : innodb_buffer_pool_size[17G] innodb_flush_log_at_trx_commit[2] innodb_flush_method[O_DIRECT] innodb_file_per_instance[0]

-----load-avg-----cpu-usage---swap---
time | 1m | 5m | 15m | usr | sys | idl | iow | si | so | ins | upd | del | sel | iud | lor | hit |
19:40:54 | 0.78 | 0.66 | 0.57 | 1 | 0 | 97 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100.00 |
19:40:55 | 0.78 | 0.66 | 0.57 | 0 | 0 | 97 | 2 | 0 | 0 | 10 | 6 | 0 | 125 | 16 | 7304 | 98.56 |
19:40:56 | 0.78 | 0.66 | 0.57 | 0 | 0 | 98 | 1 | 0 | 0 | 11 | 4 | 0 | 110 | 15 | 4987 | 98.78 |
19:40:57 | 0.78 | 0.66 | 0.57 | 2 | 1 | 94 | 3 | 0 | 0 | 10 | 5 | 0 | 148 | 15 | 5777 | 97.73 |
19:40:58 | 0.72 | 0.64 | 0.57 | 0 | 0 | 96 | 3 | 0 | 0 | 26 | 11 | 0 | 115 | 37 | 5571 | 97.77 |
19:40:59 | 0.72 | 0.64 | 0.57 | 0 | 0 | 98 | 2 | 0 | 0 | 18 | 8 | 0 | 88 | 26 | 2677 | 97.46 |
19:41:00 | 0.72 | 0.64 | 0.57 | 0 | 0 | 97 | 3 | 0 | 0 | 16 | 12 | 0 | 125 | 28 | 8110 | 98.16 |
19:41:01 | 0.72 | 0.64 | 0.57 | 1 | 0 | 95 | 3 | 0 | 0 | 15 | 10 | 0 | 147 | 25 | 5355 | 97.31 |
19:41:02 | 0.72 | 0.64 | 0.57 | 4 | 0 | 93 | 2 | 0 | 0 | 14 | 12 | 0 | 114 | 26 | 8814 | 99.02 |
```

other-Debug:tcpdump

| Filter: ip.addr == <input type="text"/> Expression... Clear Apply | | | | | | |
|---|------------|----------|-------------|-----------|--------|-------------------------|
| No. | Source | src port | Destination | dest port | Length | Info |
| 00:44:16.110617 | 10.2.14.1 | 44711 | 172.17.0.2 | 3306 | 66 | 44711 > mysql [ACK] Seq |
| 00:44:16.110626 | 172.17.0.2 | 14 | 10.2.14.1 | 44711 | 1514 | [TCP segment of a reass |
| 00:44:16.110630 | 172.17.0.2 | 14 | 10.2.14.1 | 44711 | 978 | [TCP segment of a reass |
| 00:44:16.110633 | 10.2.14.1 | 44711 | 172.17.0.2 | 3306 | 66 | 44711 > mysql [ACK] Seq |
| 00:44:16.110635 | 10.2.14.1 | 44711 | 172.17.0.2 | 3306 | 66 | 44711 > mysql [ACK] Seq |
| 00:44:16.110637 | 10.2.14.1 | 44711 | 172.17.0.2 | 3306 | 66 | 44711 > mysql [ACK] Seq |
| 00:44:16.110639 | 10.2.14.1 | 44711 | 172.17.0.2 | 3306 | 66 | 44711 > mysql [ACK] Seq |
| 00:44:16.110742 | 10.2.14.1 | 44711 | 172.17.0.2 | 3306 | 66 | 44711 > mysql [ACK] Seq |
| 00:44:16.113845 | 10.2.14.1 | 44711 | 172.17.0.2 | 3306 | 66 | 44711 > mysql [ACK] Seq |
| 00:44:20.986612 | 172.17.0.2 | 14 | 10.2.14.1 | 44711 | 15994 | [TCP segment of a reass |
| 00:44:20.989779 | 10.2.14.1 | 44711 | 172.17.0.2 | 3306 | 66 | 44711 > mysql [ACK] Seq |
| 00:44:20.989797 | 10.2.14.1 | 44711 | 172.17.0.2 | 3306 | 66 | 44711 > mysql [ACK] Seq |
| 00:44:20.989876 | 10.2.14.1 | 44711 | 172.17.0.2 | 3306 | 66 | 44711 > mysql [ACK] Seq |
| 00:44:20.989887 | 10.2.14.1 | 44711 | 172.17.0.2 | 3306 | 66 | 44711 > mysql [ACK] Seq |
| 00:44:20.989891 | 10.2.14.1 | 44711 | 172.17.0.2 | 3306 | 66 | 44711 > mysql [ACK] Seq |
| 00:44:20.989901 | 10.2.14.1 | 44711 | 172.17.0.2 | 3306 | 66 | 44711 > mysql [ACK] Seq |
| 00:44:20.989905 | 172.17.0.2 | 14 | 10.2.14.1 | 44711 | 5858 | [TCP segment of a reass |
| 00:44:20.989908 | 10.2.14.1 | 44711 | 172.17.0.2 | 3306 | 66 | 44711 > mysql [ACK] Seq |
| 00:44:20.989911 | 10.2.14.1 | 44711 | 172.17.0.2 | 3306 | 66 | 44711 > mysql [ACK] Seq |
| 00:44:20.989914 | 172.17.0.2 | 14 | 10.2.14.1 | 44711 | 5858 | [TCP segment of a reass |

DTCC2012

other-Debug:gdb/pmp

#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 0x000000000006a7d97 in **rr_sequential**
#10 0x0000000000061b756 in sub_select
#11 0x0000000000062e0ad in do_select
#12 0x0000000000063b0d4 in JOIN::exec

other-Debug: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 |

other-Backup

- Xtrabackup
 - Try to update the last stable version
- mysqldump Tips(This really helps)
 - with --master-info
 - --default-character-set
 - restore it parallel
- mydumper(Excellent tools)
 - consistant

Finally

Q&A

DTCC2012

Powered by Taobao DBA