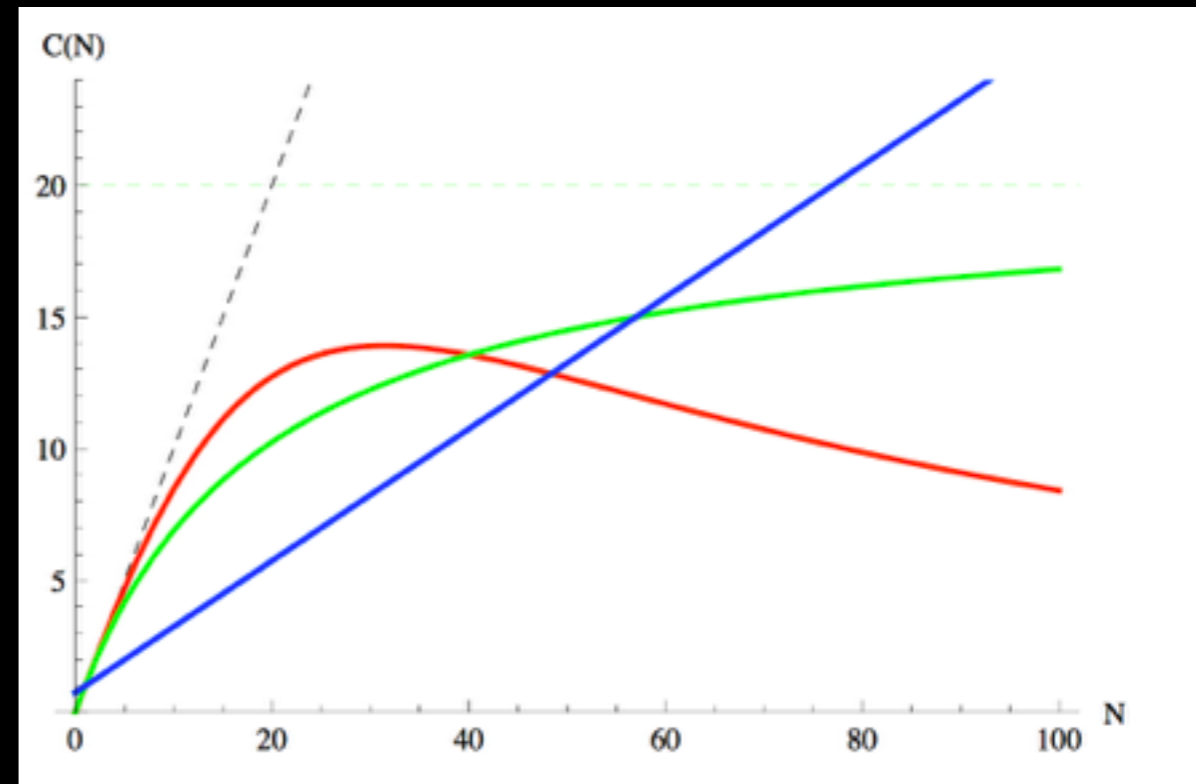# MySQL性能诊断与实践

洪 斌

大纲

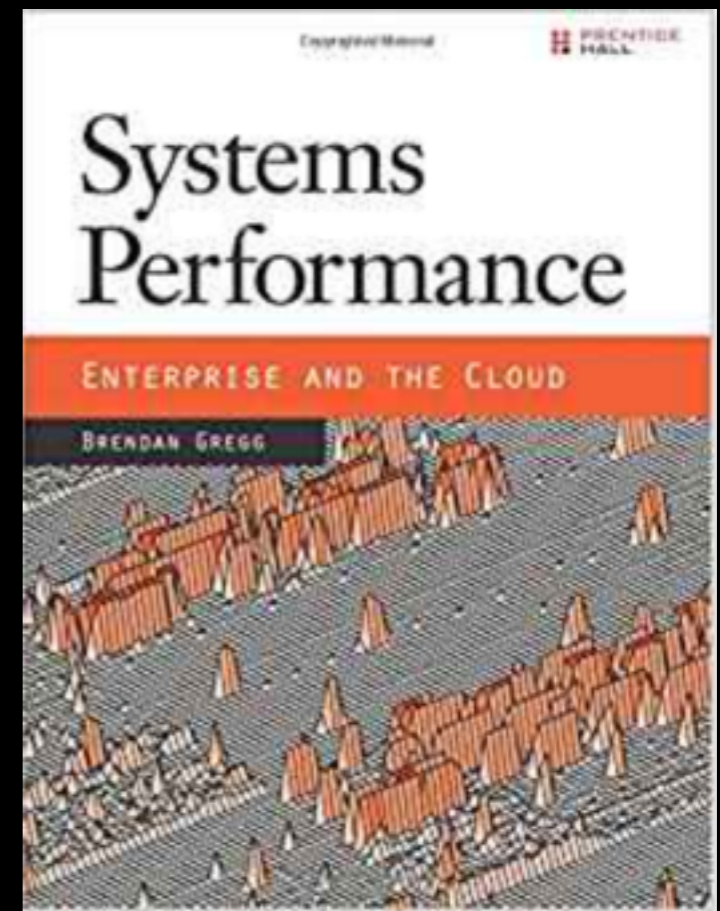- 了解有关性能诊断的方法

- 介绍一些观测工具用法

- 分享两个案例

几个定律

- Little's Law (queueing theory)

- Amdahl's Law (1967)

- Universal Scalability Law (1993)
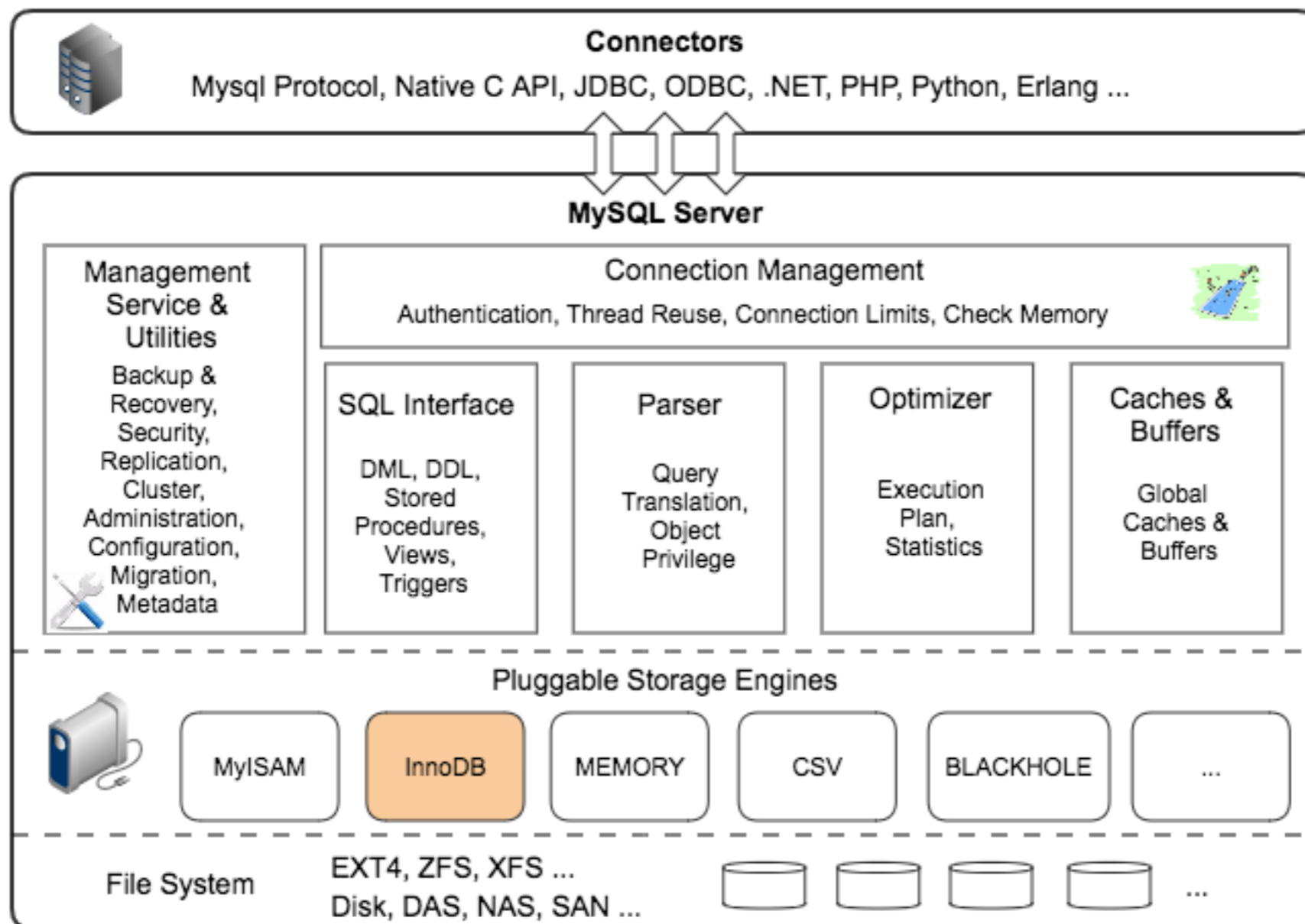
通用方法

- USE(Utilization Saturation and Errors)

- 火焰图on-cpu & off-cpu

- 观测工具很重要

- 基线对比

MySQL体系结构

快速诊断

- top 判断主机负载情况

- dmesg | tail 是否存在oom-killer 或tcp drop等错误信息

- vmstat 1 检查r、free、si、so、us, sy, id, wa, st列

- mpstat -P ALL 1 检查CPU使用率是否均衡

- pidstat 1 检查进程的CPU使用率，多核利用情况

- iostat -xz 1 检查r/s, w/s, rkB/s, wkB/s, await, avgqu-sz, %util

- free -m 检查内存使用情况

- sar -n DEV 1 检查网络吞吐量

- sar -n TCP,ETCP 1 检查tcp连接情况active/s, passive/s, retrans/s

MySQL诊断工具

- error log & slow log & general log
- MySQL SHOW [SESSION|GLOBAL] STATUS
- SHOW PROCESSLIST
- InnoDB 存储引擎状态 SHOW ENGINE INNODB STATUS
- Explain 查看执行计划
- performance schema

诊断步骤

1. 检查系统全局资源负载

2. 检查MySQL错误日志

3. 检查MySQL在做什么

4. 检查InnoDB事务情况

5. 检查MySQL复制状态

InnoDB

- InnoDB表必须有主键或唯一索引

SELECT t.table_schema, t.table_name FROM information_schema.tables t
LEFT JOIN information_schema.table_constraints c
ON (t.table_schema = c.table_schema AND t.table_name = c.table_name AND c.constraint_type IN ('PRIMARY KEY','UNIQUE'))
WHERE t.table_schema NOT IN ('mysql','information_schema', 'performance_schema') AND t.engine = 'InnoDB' AND
c.table_name IS NULL;

- 主键应使用较小数据类型且有序

- 避免大事务(运行时间长或变更记录多)
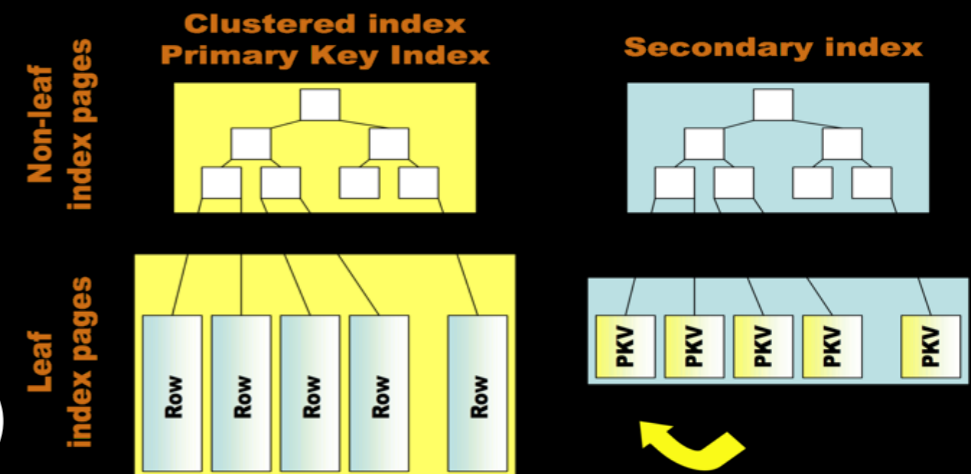
SELECT a.requesting_trx_id '被阻塞事务ID' ,b.trx_mysql_thread_id '被阻塞线程ID', TIMESTAMPDIFF(SECOND,b.trx_wait_started,NOW())
'被阻塞秒数', b.trx_query '被阻塞的语句' , a.blocking_trx_id '阻塞事务ID' ,c.trx_mysql_thread_id '阻塞线程ID',d.INFO '阻塞事务信息' FROM
information_schema.INNODB_LOCK_WAITS a
 INNER JOIN information_schema.INNODB_TRX b ON a.requesting_trx_id=b.trx_id
 INNER JOIN information_schema.INNODB_TRX c ON a.blocking_trx_id=c.trx_id
 INNER JOIN information_schema.PROCESSLIST d ON c.trx_mysql_thread_id=d.ID ;

重要参数

- max_connection

- innodb_buffer_pool_size

- Innodb_flush_neighbors

- Innodb_io_capacity

- Innodb_log_file_size
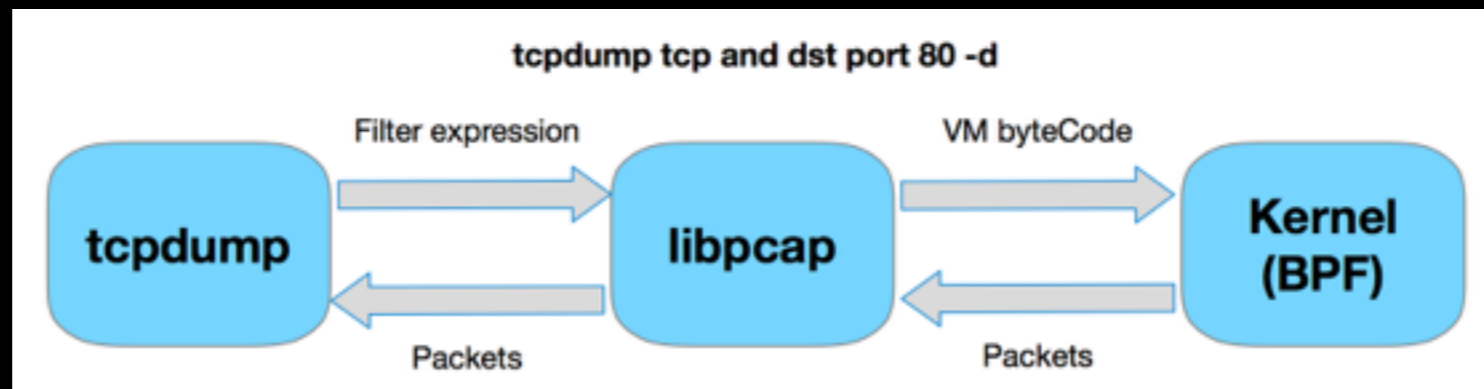
- innodb_thread_concurrency

SQL优化
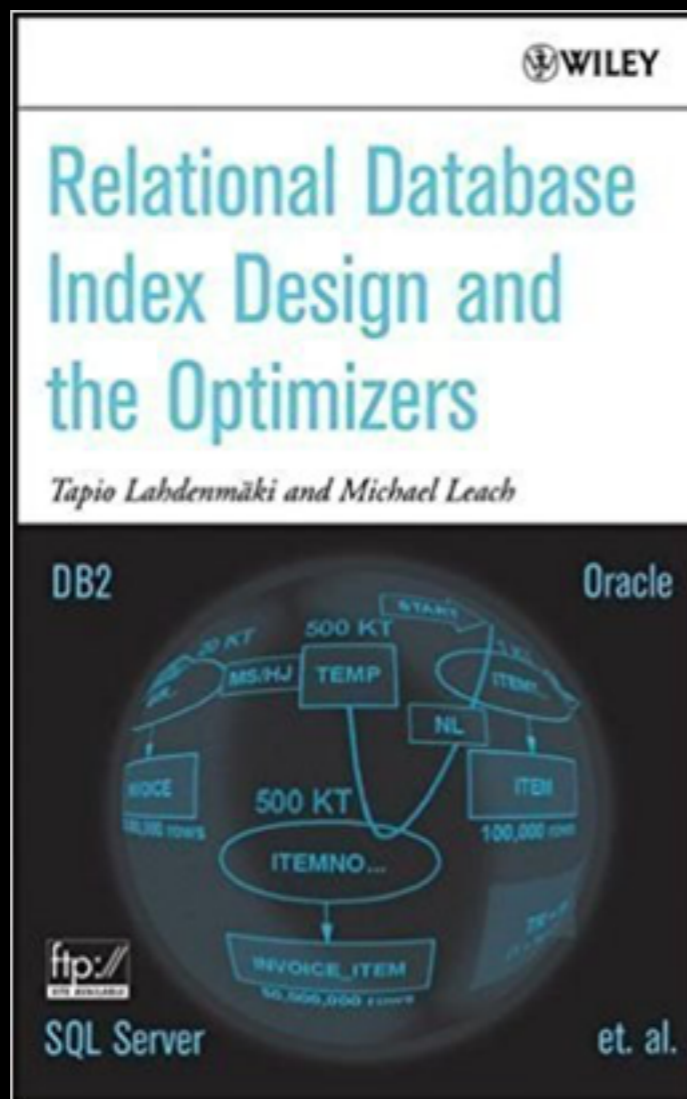
Note

- 优化的核心就是"少做事"

- 切勿盲目追求最优配置模板

- 避免过早优化

BPF是什么

- BPF = Berkeley Packet Filter

- The Berkeley Packet Filter (BPF) provides a raw interface to data link layers, permitting raw link-layer packets to be sent and received.


- Since version 3.18, the Linux kernel includes an extended BPF virtual machine, termed extended BPF (eBPF). It can be used for non-networking purposes

http://www.tcpdump.org/papers/bpf-usenix93.pdf

前提条件

- Linux kernel 4.4+ (推荐 4.9+)

- 安装Bcc https://github.com/iovisor/bcc/blob/master/INSTALL.md

- MySQL 编译 -DENABLE_DTRACE=1 & 安装 systemtap-sdt-devel

Bcc 工具箱

1. execsnoop
2. opensnoop
3. ext4slower
4. biolatency
5. biosnoop
6. cachestat
7. tcpconnect
8. tcpaccept
9. tcpretrans
10. gethostlatency
11. runlat
12. profile

Linux bcc/BPF Tracing Tools
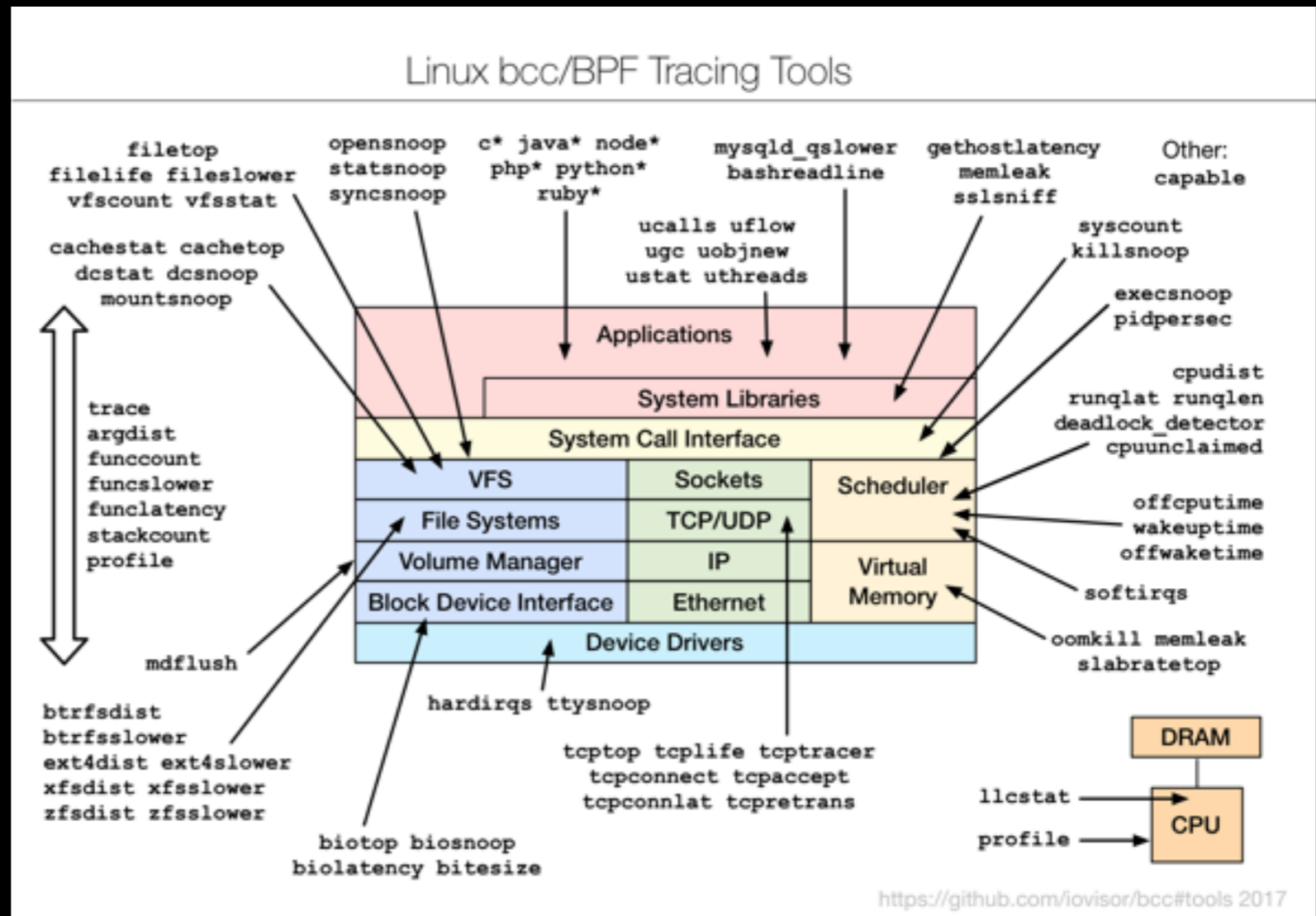
Query延迟分布


```
//Only select

root@R820-08:/usr/share/bcc/tools# ./dbstat -p `pidof mysqld` -u -- mysql
Tracing database queries for pids 4754 slower than 0 ms...
^C[11:20:53]
     query latency (us)  : count      distribution
         0 -> 1          : 0          |                                        |
         2 -> 3          : 0          |                                        |
         4 -> 7          : 0          |                                        |
         8 -> 15         : 0          |                                        |
        16 -> 31         : 0          |                                        |
        32 -> 63         : 0          |                                        |
        64 -> 127        : 400308     |****************************************|
       128 -> 255        : 148021     |***************                         |
       256 -> 511        : 261        |                                        |
       512 -> 1023       : 3          |                                        |
      1024 -> 2047       : 0          |                                        |
      2048 -> 4095       : 1          |                                        |
      4096 -> 8191       : 3          |                                        |
      8192 -> 16383      : 9          |                                        |
```

```
// Select and update
root@R820-08:/usr/share/bcc/tools# ./dbstat -p `pidof mysqld` -u -- mysql
Tracing database queries for pids 4754 slower than 0 ms...
^C[11:20:33]
     query latency (us)  : count      distribution
         0 -> 1          : 0          |                                        |
         2 -> 3          : 0          |                                        |
         4 -> 7          : 0          |                                        |
         8 -> 15         : 0          |                                        |
        16 -> 31         : 0          |                                        |
        32 -> 63         : 0          |                                        |
        64 -> 127        : 9198       |**************                          |
       128 -> 255        : 25826      |****************************************|
       256 -> 511        : 17629      |***************************             |
       512 -> 1023       : 14568      |**********************                  |
      1024 -> 2047       : 12533      |*******************                     |
      2048 -> 4095       : 9840       |***************                         |
      4096 -> 8191       : 4031       |******                                  |
      8192 -> 16383      : 463        |                                        |
     16384 -> 32767      : 33         |                                        |
     32768 -> 65535      : 20         |                                        |
     65536 -> 131071     : 20         |                                        |
```

慢Query抓取

```
// Select and update
root@R820-08:/usr/share/bcc/tools# ./dbslower -p `pidof mysqld` -m 5 -- mysql
Tracing database queries for pids 4754 slower than 5 ms...
TIME(s)              PID            MS QUERY
0.956044             4754        5.358 UPDATE sbtest1 SET k=k+1 WHERE id=514
0.956199             4754        5.837 UPDATE sbtest1 SET k=k+1 WHERE id=505
0.956876             4754        5.257 UPDATE sbtest1 SET k=k+1 WHERE id=503
0.955977             4754        6.656 UPDATE sbtest1 SET k=k+1 WHERE id=503
0.956287             4754        6.801 UPDATE sbtest1 SET k=k+1 WHERE id=503
0.955870             4754        7.554 UPDATE sbtest1 SET k=k+1 WHERE id=498
0.956329             4754        7.121 UPDATE sbtest1 SET k=k+1 WHERE id=497
...
```

VFS 延迟分析

Ext4 延迟分析

```
//Insert data

root@R820-08:/usr/share/bcc/tools# ./ext4slower 1
Tracing ext4 operations slower than 1 ms
TIME       COMM          PID     T BYTES  OFF_KB    LAT(ms) FILENAME
21:59:40 mysqld        4754    S 0      0          3.56 ib_logfile1
21:59:40 mysqld        4754    S 0      0          8.42 sbtest1.ibd
21:59:41 mysqld        4754    S 0      0          3.83 ib_logfile1
21:59:41 mysqld        4754    S 0      0          8.35 sbtest1.ibd
21:59:42 mysqld        4754    S 0      0          8.50 sbtest1.ibd
21:59:42 mysqld        4754    S 0      0          3.53 ib_logfile1
21:59:42 mysqld        4754    S 0      0          8.34 sbtest1.ibd
21:59:43 mysqld        4754    S 0      0          2.69 ib_logfile1
21:59:43 mysqld        4754    S 0      0          8.41 sbtest1.ibd
21:59:44 mysqld        4754    S 0      0          8.37 sbtest1.ibd
21:59:44 mysqld        4754    S 0      0          4.13 ib_logfile1
21:59:44 mysqld        4754    S 0      0          8.38 sbtest1.ibd
21:59:45 mysqld        4754    S 0      0          8.52 sbtest1.ibd
```

```
root@R820-08:/usr/share/bcc/tools# ./ext4slower 10
Tracing ext4 operations slower than 10 ms
TIME       COMM          PID     T BYTES      OFF_KB    LAT(ms) FILENAME
22:03:14 dd           42639   W 1073741824 0        873.20 test1.img
22:03:15 mysqld        4754    W 1048576 1024     16.48 ibdata1
22:03:15 mysqld        4754    W 507904  2048     13.98 ibdata1
22:03:15 mysqld        4754    W 1048576 1302528  15.10 sbtest1.ibd
22:03:15 mysqld        4754    S 0       0        110.94 ibdata1
22:03:16 mysqld        4754    W 1048576 1306624  22.35 sbtest1.ibd
```

块设备延迟分析

```
//Select and update

root@R730-117:/usr/share/bcc/tools# ./biolatency -D 2
Tracing block device I/O... Hit Ctrl-C to end.


disk = 'sdb'
        usecs                : count     distribution
          0 -> 1             : 0         |                    |
          2 -> 3             : 0         |                    |
          4 -> 7             : 0         |                    |
          8 -> 15            : 0         |                    |
         16 -> 31            : 0         |                    |
         32 -> 63            : 4694      |********************|
         64 -> 127           : 3399      |**************      |
        128 -> 255           : 2211      |*********           |
        256 -> 511           : 2250      |*********           |
        512 -> 1023          : 642       |**                  |
       1024 -> 2047          : 0         |                    |
       2048 -> 4095          : 0         |                    |
```

```
root@R730-117:/usr/share/bcc/tools# ./biolatency -D 2
Tracing block device I/O... Hit Ctrl-C to end.


disk = 'sdb'
        usecs                : count     distribution
          0 -> 1             : 0         |                                        |
          2 -> 3             : 0         |                                        |
          4 -> 7             : 0         |                                        |
          8 -> 15            : 0         |                                        |
         16 -> 31            : 0         |                                        |
         32 -> 63            : 0         |                                        |
         64 -> 127           : 0         |                                        |
        128 -> 255           : 0         |                                        |
        256 -> 511           : 2         |**************************              |
        512 -> 1023          : 0         |                                        |
       1024 -> 2047          : 0         |                                        |
       2048 -> 4095          : 3         |****************************************|
```

MySQL文件IO压力分析

```
root@R820-08:/usr/share/bcc/tools# ./filetop -p `pidof mysqld` -C 5
Tracing... Output every 5 secs. Hit Ctrl-C to end

22:26:30 loadavg: 7.50 5.28 4.87 18/1925 44235


TID    COMM        READS  WRITES  R_Kb    W_Kb    T FILE
39956  mysqld      0      115     0       462     R ib_logfile1
40075  mysqld      0      107     0       424     R ib_logfile1
39900  mysqld      0      1220    0       137     R R820-08.log
38046  mysqld      0      1263    0       142     R R820-08.log
39085  mysqld      0      101     0       332     R ib_logfile1
38957  mysqld      0      114     0       425     R ib_logfile1
39959  mysqld      0      1       0       2       R ibmPAQIO
4780   mysqld      0      4       0       28      R ib_logfile1
40266  mysqld      0      107     0       361     R ib_logfile1
39984  mysqld      0      111     0       414     R ib_logfile1
39991  mysqld      0      1211    0       136     R R820-08.log
37224  mysqld      0      104     0       449     R ib_logfile1
40259  mysqld      0      109     0       340     R ib_logfile1
39958  mysqld      0      107     0       342     R ib_logfile1
39969  mysqld      0      1214    0       137     R R820-08.log
39966  mysqld      0      1275    0       144     R R820-08.log
39937  mysqld      0      1227    0       138     R R820-08.log
```

临时表文件生命周期观测

```
root@R820-08:/usr/share/bcc/tools# ./filelife
TIME      PID    COMM            AGE(s)  FILE
22:17:01  43687  cron            0.00    tmpfgHF5vY
22:22:21  39170  mysqld          5.30    #sql1292_59a1f_0.frm
```

短连接分析

案例1

开发：MySQL数据库怎么时快时慢，什么原因?

- 什么类型的请求慢，查询? 写入? 个别? 全部?

- 利用USE方法检查系统资源

- 检查MySQL 线程状态和存储引擎状态

案例2

xtrabackup不是热备么，怎么还会堵住业务

```
mysql> show processlist;
+----+----------+-----------------+------+---------+------+-----------------------------+-----------------------------------+
| Id | User     | Host            | db   | Command | Time | State                       | Info                              |
+----+----------+-----------------+------+---------+------+-----------------------------+-----------------------------------+
| 31 | root     | localhost       | NULL | Query   |    0 | starting                    | show processlist                  |
| 32 | msandbox | localhost:62019 | test | Query   |   19 | Waiting for global read lock | UPDATE sbtest1 SET k=k+1 WHERE id=676  |
| 33 | msandbox | localhost:62020 | test | Query   |   10 | Waiting for global read lock | UPDATE sbtest1 SET k=k+1 WHERE id=2177 |
| 34 | msandbox | localhost:62026 | test | Query   |   10 | Waiting for global read lock | UPDATE sbtest1 SET k=k+1 WHERE id=1074 |
| 35 | msandbox | localhost:62024 | test | Query   |   10 | Waiting for global read lock | UPDATE sbtest1 SET k=k+1 WHERE id=1716 |
| 36 | msandbox | localhost:62027 | test | Query   |   10 | Waiting for global read lock | UPDATE sbtest1 SET k=k+1 WHERE id=2129 |
| 37 | msandbox | localhost:62023 | test | Query   |   10 | Waiting for global read lock | UPDATE sbtest1 SET k=k+1 WHERE id=2971 |
| 38 | msandbox | localhost:62022 | test | Query   |   10 | Waiting for global read lock | UPDATE sbtest1 SET k=k+1 WHERE id=2370 |
| 39 | msandbox | localhost:62028 | test | Query   |   10 | Waiting for global read lock | UPDATE sbtest1 SET k=k+1 WHERE id=3830 |
| 40 | msandbox | localhost:62025 | test | Query   |   10 | Waiting for global read lock | UPDATE sbtest1 SET k=k+1 WHERE id=3648 |
| 41 | msandbox | localhost:62021 | test | Query   |   10 | Waiting for global read lock | UPDATE sbtest1 SET k=k+1 WHERE id=308  |
| 42 | msandbox | localhost:62029 | test | Query   |   10 | Waiting for global read lock | UPDATE sbtest1 SET k=k+1 WHERE id=524  |
+----+----------+-----------------+------+---------+------+-----------------------------+-----------------------------------+
12 rows in set (0.00 sec)
```

```
thread_list=$(gdb -p $1 -q -batch -ex 'info threads'| awk '/mysqld/{print $1}'|grep -v '*'|sort -nk1)
for i in $thread_list; do
  echo ">>>>>> thread $i <<<<<"
  grl=`gdb -p $1 -q -batch -ex "thread $i" -ex 'p do_command::thd->thread_id' -ex 'p do_command::thd->global_read_lock'|grep -B3 GRL_ACQUIRED_AND_BLOCKS_COMMIT`
  if [[ $grl =~ 'GRL_ACQUIRED_AND_BLOCKS_COMMIT' ]]; then
    echo "$grl"  ; break
  fi
done
```

PHPCON 官网：www.phpconchina.com
全部 PPT 下载：https://github.com/ThinkDevelopers/PHPConChina

官方QQ群：34449228、135615537
官方公众号：ThinkInLAMP