



# MySQL 5.5 Innodb/Xtradb 性能诊断与优化

季海东

Velocity China 2011

# 议程

- MyISAM pk InnoDB : 简要对比
- InnoDB Status
- InnoDB 的几个主要参数
- Percona Server 和其他产品开发流程介绍及参加开源软件的体会和建议
- 时间允许， Percona 工具简介  
很多基本知识，很多现场演示 ( live demo )

# 关于我

- 丰富的 IT 业界经验：从 Windows 到 Linux ，从微软 SQL Server 到开源 MySQL ，从行存储到列存储，从制造业到金融业到培训和咨询，都有广泛涉及和研究。
- 曾著有三本数据库书籍，并在时间允许的情况下，乐于分享和交流其经验和心得。
- [Http://www.haidongji.com](http://www.haidongji.com)

# MyISAM pk InnoDB

- MyISAM 不占地方 ( InnoDB 要更多硬盘空间)
  - MyISAM 支持全文搜索 (下版 InnoDB 可能会)
- 但是
- MyISAM 不支持 ACID 原则
  - MyISAM 很容易造成数据损害和丢失
  - MyISAM 的锁机制不适合并行处理
  - MyISAM 依靠文件系统缓存

# InnoDB Plugin 1.1

- MySQL 5.5 默认存储系统

SHOW ENGINES

SHOW PLUGINS

*Demo*

# 查看 InnoDB 状态

```
mysql> SHOW ENGINE INNODB STATUS\G
```

自 5.5 起，ENGINE 是必须的关键词

\G 便于阅读

# InnoDB 状态

- 数据采集应当超过 30 秒

```
=====
111126 17:41:40 INNODB MONITOR OUTPUT
=====
Per second averages calculated from the last 39 seconds
```

# InnoDB 状态组成部分

- Background Thread
- Semaphores
- Latest Foreign Key Error
- Latest Detect Deadlock
- File I/O
- Insert Buffer and Adaptive Hash Index
- Log
- Buffer Pool and Memory
- Row Operations
- Transactions

# ----- BACKGROUND THREAD -----

srv\_master\_thread loops: 11938931 1\_second,  
11935492 sleeps,

1193884 10\_second, 365 background, 365 flush

srv\_master\_thread log flush and writes:  
12087852

后台运行主线程

-----  
**SEMAPHORES**  
-----

OS WAIT ARRAY INFO: reservation count 32171153,  
signal count 31011552

Mutex spin waits 0, rounds 20027565614, OS  
waits 21875962

RW-shared spins 15207459, OS waits 1053752; RW-  
excl spins 95741267, OS

waits 2485654

# ----- SEMAPHORES -----

```
--Thread 8113 has waited at ibuf0ibuf.c line  
366 for 1.00 seconds the semaphore:
```

```
S-lock on RW-latch at 67e58d0 created in file  
dict0dict.c line 3706
```

```
number of readers 0, waiters flag 0
```

```
Last time read locked in file ibuf0ibuf.c line  
366
```

```
Last time write locked in file ibuf0ibuf.c line  
359
```

- 一般来讲，我们不想看到以上的 semaphore 等待，因为那意味着 OS Waits(耗资大)，而不是 spin-waits(耗资小)

-----  
**LATEST FOREIGN KEY ERROR**  
-----

- Demo

-----  
**LATEST DETECTED DEADLOCK**  
-----

- Demo

-----  
**FILE I/O (1 of 2)**  
-----

如果你有高端 I/O 系统，并且注意到有  
64\*innodb\_read\_io\_threads 数量的读等待，可  
以增加该参数值

[http://docs.oracle.com/cd/E17952\\_01/refman-5.5-e](http://docs.oracle.com/cd/E17952_01/refman-5.5-e)

- Demo

-----  
**FILE I/O (2 of 2)**  
-----

---

Pending normal aio reads: 0, aio writes: 0,  
ibuf aio reads: 0, log i/o's: 0, sync i/o's: 0  
Pending flushes (fsync) log: 0; buffer pool: 0  
151671 OS file reads, 94747 OS file writes,  
8750 OS fsyncs  
25.44 reads/s, 18494 avg bytes/read, 17.55  
writes/s, 2.33 fsyncs/s

# ----- INSERT BUFFER AND ADAPTIVE HASH INDEX -----

Ibuf: size 1, free list len 5, seg size 7,  
36923 inserts, 36924 merged recs, 27673 merges

Hash table size 3187567, node heap has 3949  
buffer(s)

2.00 hash searches/s, 17.08 non-hash searches/s

---  
**LOG**  
---

```
Log sequence number 1615338
Log flushed up to   1615338
Last checkpoint at  1615338
Max checkpoint age   7782360
Checkpoint age target 7539162
Modified age         0
Checkpoint age       0
0 pending log writes, 0 pending chkp writes
8 log i/o's done, 0.00 log i/o's/second
```

# LOG

- InnoDB 默认的重做日志文件 (redo log) , `ib_logfile0` 和 `ib_logfile1` 是 5MB , 绝大多数情况下这个太小
- $\text{Checkpoint age} = \text{最高 LSN} - \text{最后一次 checkpoint}$
- 当 `checkpoint age` 超过日志文件大小的 75% 时 , InnoDB 强制 `checkpoint` , SQL 工作必须停止
- 结果 : 系统运行缓慢

---  
**LOG**  
---

解决方案：

- <http://www.mysqlperformanceblog.com/2008/11/21/how>
- Demo

-----  
**BUFFER POOL AND MEMORY**  
-----

---

Demo

-----  
**ROW OPERATIONS**  
-----

---

Demo

# 几个重要参数

- innodb\_buffer\_pool\_size
  - 越多越好
- innodb\_log\_file\_size
  - 在重负荷时采样得出
- innodb\_io\_capacity
  - 根据硬盘 IOPS



**Haidong.Ji@percona.com**  
**Haidong.Ji@gmail.com**

**We're Hiring! [www.percona.com/about-us/careers/](http://www.percona.com/about-us/careers/)**