

MySQL管理之道

性能调优、高可用与监控

贺春旸 著

Management of MySQL

Performance optimization, High availability and Surveillance

- 51CTO技术博客之星、资深MySQL技术专家多年工作经验的结晶，指导MySQL DBA进阶修炼的最佳实践之作
- 包含大量来自实际生产环境的经典案例，为各种疑难问题提供了解决方法，从性能优化、高可用与监控角度帮助读者把握实践技巧



机械工业出版社
China Machine Press



CD

MySQL管理之道

性能调优、高可用与监控

Management of MySQL
Performance optimization, High availability and Surveillance

贺春阳 著



机械工业出版社
China Machine Press

图书在版编目 (CIP) 数据

MySQL 管理之道: 性能调优、高可用与监控 / 贺春旸著. —北京: 机械工业出版社, 2014.1
(数据库技术丛书)

ISBN 978-7-111-44906-5

I. M… II. 贺… III. 关系数据库系统 IV. TP311.138

中国版本图书馆 CIP 数据核字 (2013) 第 282228 号

版权所有·侵权必究

封底无防伪标均为盗版

本书法律顾问 北京市展达律师事务所

本书由资深 MySQL 专家撰写, 以最新的 MySQL 版本为基础, 以构建高性能 MySQL 服务器为核心, 从故障诊断、表设计、SQL 优化、性能参数调优、mysqldump 逻辑、xtrabackup 热备份与恢复、MySQL 高可用集群搭建与管理、MySQL 服务器性能和服务监控等方面多角度深入讲解了如何去管理与维护 MySQL 服务器。

书中内容以实战为导向, 所有内容均来自于笔者多年实践经验的总结和对新知识的拓展, 同时也针对运维人员、DBA 等相关工作者会遇到的有代表性的疑难问题给出了实用的情景模拟, 并给出了解决方案。不论你目前有没有遇到过此类问题, 相信你以后处理相关问题都会有所借鉴。本书适合所有希望构建和管理高性能、高可用性的 MySQL 数据库系统的开发者和 DBA 阅读。



机械工业出版社 (北京市西城区百万庄大街 22 号 邮政编码 100037)

责任编辑: 朱秀英

北京市荣盛彩色印刷有限公司印刷

2014 年 1 月第 1 版第 1 次印刷

186mm × 240mm · 20 印张

标准书号: ISBN 978-7-111-44906-5

ISBN 978-7-89405-202-5 (光盘)

定 价: 69.00 元 (附光盘)

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

客服热线: (010) 88378991 88361066

投稿热线: (010) 88379604

购书热线: (010) 68326294 88379649 68995259

读者信箱: hzjsj@hzbook.com

前 言

为什么要写这本书

很久以前，我就梦想着自己有一天能写一本书出来，但始终未能成行，直到今天，我终于鼓起勇气着手写作，把多年工作的积累毫无保留地写出来，展现给大家，这不仅对他人的技术成长会有些许帮助，同时对自己来说，也是一个重新学习的过程，何乐而不为？

2010年我有幸加入了梦寐以求的飞信公司，并以此正式开始了我的 MySQL 职业生涯。工作中能实际接触到海量数据、大并发是我在该领域发展的很重要的因素，工作的驱使和个人的兴趣成为我每天坚持学习 MySQL 的源动力。对于 MySQL，目前业内用得最多的是 MySQL5.1.X 版本，事实上，MySQL5.5 已经推出两年多了，其性能方面有了显著的提升，但市面上介绍 MySQL 该版本的图书还寥寥无几，大家都是通过阅读英文手册去获取新的知识，因此我决定基于这个版本来写作，把自己学到的新东西做一个系统性的总结。

读者对象

本书是一本面向应用的数据库类书籍，是日常工作中的积累。本书主要面向使用 MySQL InnoDB 存储引擎作为数据库后端的运维人员和 DBA。书中的大部分例子都是生产环境的实战，相信会对读者今后的工作有所帮助。

要想更好地学习本书的内容，要求具备以下条件：

- 有一定的 SQL 基础。
- 掌握基本的 MySQL 操作和 Linux 操作，以及数据库基本原理。
- 接触过 Shell 脚本语言。

如何阅读本书

本书从知识结构上分为五大部分：

第一部分（第 1 ~ 2 章）详细介绍了 MySQL5.5 的新特性、注意事项、安装和升级方法，以及半同步复制的日常管理维护，并针对 MySQL5.5 与 MySQL5.1 进行了性能测试和功能测试。

第二部分（第 3 ~ 6 章）为故障诊断与优化，内容为生产环境下 MySQL 故障处理以及性能调优，包括表设计阶段范式的理解、字段类型的选取、采用表锁还是行锁、MySQL 默认的隔离级别与传统 SQL Server 以及 Oracle 数据库默认的隔离级别的区别、SQL 语句的优化以及合理利用索引等，然后讲解了 my.cnf 配置文件的优化、如何定期整理数据库碎片，最后介绍了硬件的优化、内存对 InnoDB 引擎所起的决定性作用，以及 MySQL5.6 的新特性等内容。

第三部分（第 7 ~ 8 章）为高可用软件的搭建与维护，包括企业级高可用集群 RHCS、Heartbeat+DRBD 和 Keepalived+MySQL Replication 一主一从模式的搭建与维护，故障切换脚本、MySQL+MMM 读写分离架构搭建与维护，以及批量管理上百台 MySQL 服务器等内容。

第四部分（第 9 ~ 10 章）为主流监控软件的搭建与维护，包括监控软件 Cacti 和 Nagios 的安装与使用，以及线上 MySQL 服务器的监控脚本等内容。

第五部分（第 11 章）为项目案例讲解，内容包括生产环境数据库碎片整理方案、用户信息表水平拆表方案、阿里巴巴中间件 Cobar 水平拆表方案。

每个部分可以单独作为一本迷你书阅读，如果你未接触过 MySQL5.5，建议从第一部分开始阅读。

勘误和支持

由于作者的水平有限，编写的时间也很仓促，书中难免会出现一些错误或者不准确的地方，恳请读者批评指正。你可以将书中的错误发到我的邮箱 chunyang_he@sohu.com，我将尽量提供最满意的解答。如果你有更多的宝贵意见，也欢迎你发送邮件至我的邮箱，很期待能够听到你们的真挚反馈。

致谢

在飞信 3 年多的时间里，我得到了很多帮助。在这里感谢任志修经理对我的信任，让我始终挑战运维一线，使我得到了成长；感谢我的同事周海平（www.MySQLpub.com）鼓励我，让我把知识展现出去；还要感谢我的同事吴炳锡把我介绍到了飞信，而且在工作中给了我很多鼓励。没有你们的帮助，这本书是完成不了的。

同时感谢我的朋友邱治军参与监控部分的编写；感谢 51CTO 博客总编曹亚莉、博客管理员张丽妍多次邀请我参加 51CTO 举办的数据库大会，拓展了我的知识视野；感谢 51CTO 博客管理员高

阳帮我处理图片；感谢前同事李才荣提供 Linux 批量管理的工具 (Python)；感谢前同事李冉 (乐蜂网高级 DBA) 提供事件调度器故障切换后的注意事项；另外，还要感谢给我提建议的一些朋友，谢谢你们。

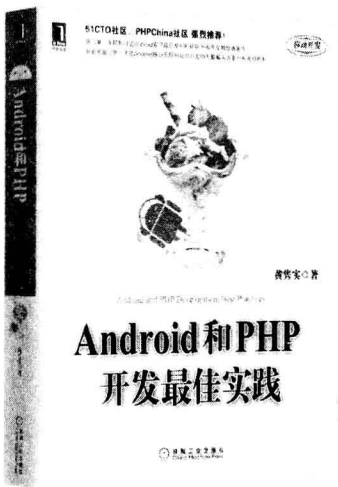
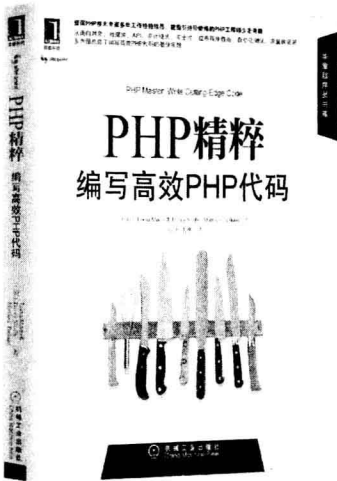
感谢机械工业出版社华章公司的编辑杨绣国，感谢你的魄力和远见，在这一年多的时间中始终支持我的写作，你的鼓励和帮助引导我顺利完成全部书稿。

谨以此书献给我最亲爱的家人，以及众多热爱 MySQL 的朋友们。

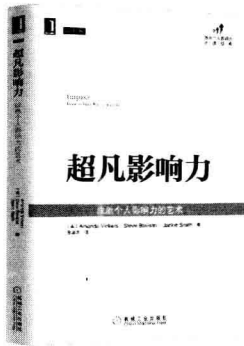
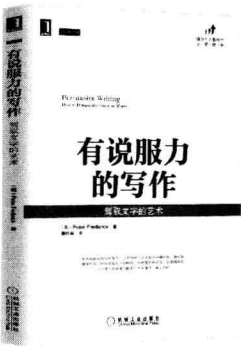
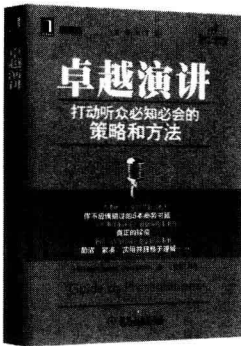
贺春暘

2013 年 8 月于北京

推荐阅读



程序员提升个人影响力和软实力必读经典



目 录

前言

第一部分 MySQL5.5 新特性篇

第 1 章 MySQL5.5 介绍 2

1.1 性能上的显著改变 2

1.1.1 MySQL5.5 默认存储引擎
的调整 2

1.1.2 充分利用 CPU 多核的
处理能力 7

1.1.3 提高刷新脏页数量和合
并插入数量, 改善磁盘
I/O 处理能力 8

1.1.4 增加自适应刷新脏页功能 9

1.1.5 让 InnoDB_Buffer_Pool 缓冲
池中的热数据存活更久 9

1.1.6 InnoDB 的数据恢复时间
加快 11

1.1.7 InnoDB 同时支持多个
BufferPool 实例 15

1.1.8 可关闭自适应哈希索引 17

1.1.9 在 InnoDB 中可选择
使用内存分配程序 18

1.1.10 提高默认 InnoDB 线程

并发数 21

1.1.11 预读算法的变化 22

1.1.12 首次在 Linux 上实现了
异步 I/O 23

1.1.13 恢复组提交 24

1.1.14 InnoDB 使用多个回滚
段提升性能 26

1.1.15 改善清除程序进度 26

1.1.16 添加删除缓冲和清除
缓冲 27

1.1.17 控制自旋锁 Spin Lock
轮训间隔 28

1.1.18 快速创建、删除、更改
索引 29

1.1.19 InnoDB 支持创建压缩
数据页 30

1.1.20 可动态关闭 InnoDB 更新
元数据的统计功能 37

1.2 安全性、稳定性的显著改变 38

1.2.1 复制功能加强 38

1.2.2 中继日志 relay-log 可
自我修复 39

1.2.3 开启 InnoDB 严格检查
模式 39

1.3	动态更改系统配置参数	39
1.3.1	支持动态更改独立表空间	39
1.3.2	支持动态更改 InnoDB 锁超时时间	40
1.4	InnoDB 新参数汇总	40
1.5	同步复制新参数汇总	48
1.6	SQL 语句写法的改变	53
1.6.1	delete 表连接语法改变	53
1.6.2	MySQL5.5 存储过程支持 limit 变量	54
1.7	MySQL5.1 升级为 MySQL5.5	55
1.7.1	采用 MySQL_upgrade 升级授权表方式升级	55
1.7.2	直接安装 MySQL5.5, 采用数据导出 / 导入方式升级	59
1.8	性能测试: MySQL5.5 与 MySQL5.1	60

第 2 章 半同步复制 62

2.1	半同步复制简介	62
2.2	半同步复制安装配置	63
2.3	参数说明	63
2.4	功能测试	64
2.4.1	如何验证半同步复制是否正常工作	64
2.4.2	半同步复制与异步复制的切换	65
2.5	性能测试	68
2.6	小结	70

第二部分 故障诊断与性能优化篇

第 3 章 故障诊断 72

3.1	影响 MySQL 性能的因素	72
3.2	系统性能评估标准	73
3.2.1	影响 Linux 服务器性能的因素	73
3.2.2	系统性能评估指标	74
3.2.3	开源监控和评估工具介绍	76
3.3	故障与处理	79
3.3.1	连接数过多导致程序连接报错的原因	79
3.3.2	记录子查询引起的宕机	84
3.3.3	诊断事务量突高的原因	87
3.3.4	谨慎设置 binlog_format=MIXED	90
3.3.5	未设置 swap 分区导致内存耗尽, 主机死机	94
3.3.6	MySQL 故障切换之事件调度器注意事项	95
3.3.7	人工误删除 InnoDB ibdata 数据文件, 如何恢复	97
3.3.8	update 忘加 where 条件 误操作恢复 (模拟 Oracle 闪回功能)	99
3.3.9	delete 忘加 where 条件 误操作恢复 (模拟 Oracle 闪回功能)	108

第 4 章 同步复制报错故障处理 112

4.1	最常见的 3 种故障	112
-----	------------	-----

4.1.1	在 master 上删除一条记录时出现的故障	112			
4.1.2	主键重复	114			
4.1.3	在 master 上更新一条记录, 而 slave 上却找不到	115			
4.2	特殊情况: slave 的中继日志 relay-log 损坏	116			
4.3	人为失误	118			
4.4	避免在 master 上执行大事务	119			
4.5	slave_exec_mode 参数可自动处理同步复制错误	120			
4.6	如何验证主从数据一致	121			
4.7	binlog_ignore_db 引起的同步复制故障	123			
4.8	MySQL5.5.19/20 同步一个 Bug	124			
4.9	恢复 slave 从机上的某几张表的简要方法	126			
4.10	如何干净地清除 slave 同步信息	127			
第 5 章	性能调优	129			
5.1	表设计	129			
5.2	字段类型的选取	133			
5.2.1	数值类型	134			
5.2.2	字符类型	139			
5.2.3	时间类型	141			
5.2.4	小技巧: 快速修改表结构	148			
5.2.5	pt-online-schema-change 在线更改表结构	152			
5.2.6	MySQL5.6 在线 DDL 更改表测试	158			
5.3	采用合适的锁机制	161			
5.3.1	表锁的演示	161			
5.3.2	行锁的演示	164			
5.3.3	InnoDB 引擎与 MyISAM 引擎的性能对比	166			
5.4	选择合适的事务隔离级别	168			
5.4.1	事务的概念	168			
5.4.2	事务的实现	169			
5.4.3	事务隔离级别介绍	171			
5.5	SQL 优化与合理利用索引	177			
5.5.1	如何定位执行很慢的 SQL 语句	177			
5.5.2	SQL 优化案例分析	178			
5.5.3	合理使用索引	188			
5.6	my.cnf 配置文件调优	198			
5.6.1	per_thread_buffers 优化	198			
5.6.2	global_buffers 优化	200			
5.6.3	Query Cache 在不同环境下的使用	201			
5.6.4	tuning-primer.sh 性能调试工具的使用	205			
5.6.5	72 GB 内存的 my.cnf 配置文件	208			
5.6.6	谨慎使用分区表功能	211			
5.7	MySQL5.6 同步复制新特性详解	213			
第 6 章	备份与恢复	223			
6.1	冷备份	224			
6.2	逻辑备份	224			
6.2.1	mysqldump 增加了一个				

重要参数	225
6.2.2 取代 mysqldump 的新工具 mydumper	226
6.2.3 逻辑备份全量、增量 备份脚本	229
6.3 热备份与恢复	230

第三部分 高可用集群管理篇

第 7 章 目前流行的 4 种高可用 架构	236
7.1 采用 MySQL 自带的 Replication 架构	237
7.1.1 Keepalived+MySQL Replication 架构的搭建 演示	237
7.1.2 MMM+MySQL Replication 架构的搭建演示	241
7.2 Heartbeat+DRBD+MySQL 架构的搭建演示	249
7.3 红帽 RHCS 共享存储架构的 搭建演示	254
7.3.1 安装过程	257
7.3.2 红帽 RHCS 集群的维护	265

7.4 MySQL 高可用集群 HA 解决 方案的测试评估	267
--	-----

第 8 章 批量管理服务器

8.1 开源工具 pssh 的使用方法	270
8.2 自己编写的 SSH 服务器批量 管理工具	273

第四部分 监控篇

第 9 章 性能监控

第 10 章 服务监控

10.1 Nagios 搭建与维护	283
10.2 MySQL 数据库的监控脚本	288

第五部分 项目案例

第 11 章 项目案例讲解

11.1 数据碎片整理方案	292
11.2 用户信息表水平拆表方案	296
11.3 阿里巴巴中间件 Cobar 水平拆表方案	299

第一部分

MySQL5.5 新特性篇

第 1 章 MySQL5.5 介绍

第 2 章 半同步复制

第 1 章

MySQL5.5 介绍

MySQL 是一个中小型的关系型数据库管理系统，由瑞典 MySQL AB 公司开发，目前属于 Oracle 公司。由于它具有性能高、成本低、可靠性好等特点，近几年已经成为最流行的开源数据库，被广泛地应用在 Internet 上的中小型网站中。而且，随着 MySQL 的不断成熟，现在它也逐渐用于更多大规模的网站和应用了，比如，维基百科、Google 和 Facebook 等。非常流行的开源软件组合“LAMP”中的“M”指的就是 MySQL。

这几年，MySQL 的版本在不断变更，可以说是有了翻天覆地的变化，在之前的 4.0 版本中，没有存储过程、触发器、函数、事件，对 CPU 多核的支持也不好，在经历了 5.0 和 5.1 两个过渡版本后，到 5.5 版，其性能和功能上已经得到了很大的改善，主要体现在 CPU 多核处理上有了很大提高，宕机恢复时间减少，可快速创建索引，并具有半同步复制功能等方面，目前这个版本很稳定。2013 年春节，5.6 版本已经出了 GA 版，但这里不推荐直接将其用于生产环境中，因为该版本还有许多未知 Bug 在修复，建议 1 年后再开始应用，不要充当小白鼠，本书中也只会针对该版部分功能上的新特性进行介绍。

本章主要讲解的是 MySQL5.5 和 InnoDB 的一些增强性能，这些增强性能极大地提高了系统和 MySQL 的性能。下面将详细介绍每一个关键的增强性能及其实现过程。

为了不误导读者，保证全文的准确性，下面的内容会结合 MySQL5.5 官方手册《14.4. New Features of InnoDB 1.1》来讲解，帮助大家认识 MySQL5.5 中一些较为重要的改变，其中也许会有疏漏的地方，不到之处请大家访问 <http://dev.mysql.com/doc/refman/5.5/en/innodb-5-5.html>，参考相关的英文文档。

1.1 性能上的显著改变

1.1.1 MySQL5.5 默认存储引擎的调整

在 MySQL5.1.X 之前的版本中，默认的存储引擎是 MyISAM，每个 MyISAM 在磁盘上会将数据存储成三个文件。第一个文件是表结构文件，它的名字以表的名字开始，其扩展

名会指出文件类型，为 .frm 文件存储表定义；第二个文件是数据文件，其扩展名为 .MYD (MYData)；第三个文件是索引文件，其扩展名是 .MYI (MYIndex)。

MyISAM 存储引擎的特点是表级锁、不支持事务和全文索引，适合一些 CMS 内容管理系统作为后台数据库使用，但是在高并发、重负荷生产系统上，表锁的特性显得有些力不从心，并且如果系统出现宕机、MySQLd 进程崩溃，MyISAM 引擎表很容易受到损坏，这时不得不用外部命令 MyISAMchk 去修复它。

从 MySQL5.5.X 开始，默认的存储引擎变更为 InnoDB Plugin 引擎。

InnoDB 给 MySQL 提供了具有提交、回滚和崩溃恢复能力的事务安全 (ACID 兼容) 存储引擎。InnoDB 锁定在行级，并且也在 SELECT 语句提供一个与 Oracle 风格一致的非锁定读。这些特色增加了多用户部署和性能。之所以没有在 InnoDB 中扩大锁定的需要，是因为 InnoDB 中的行级锁定适合非常小的空间。InnoDB 也支持 FOREIGN KEY (外键)。在 SQL 查询中，你可以自由地将 InnoDB 类型的表与其他 MySQL 表的类型混合起来，甚至在同一个查询中也可以混合。

InnoDB 是为处理巨大数据量时拥有最大性能而设计的。它的 CPU 效率可能是任何其他基于磁盘的关系数据库引擎所不能匹敌的。

InnoDB 存储引擎被完全整合到了 MySQL 服务器中，不过，为了能在主内存中缓存数据和索引，InnoDB 存储引擎会维持它自己的缓冲池。InnoDB 会在一个表空间中存储它的表和索引，表空间可以包含数个文件 (或原始磁盘分区)。这与 MyISAM 表不同，在 MyISAM 表中，每个表是被存在分离的文件中的。另外，InnoDB 表可以是任何尺寸，即使是在文件尺寸被限制为 2 GB 的操作系统上。

InnoDB 存储引擎应用在众多高压力、高并发的大型数据库站点上。比如，著名的 Internet 新闻站点 Slashdot.org 运行在 InnoDB 上。Myrix, Inc. 在 InnoDB 上存储着超过 1 TB 的数据，还有一些其他站点在 InnoDB 上处理平均每秒 800 次插入 / 更新的负荷。

随着 InnoDB 存储引擎的崛起，更多公司为其提供了 patch 补丁，使其性能发挥至极致，包括 Google 公司、Percona 公司、Sun Microsystems 公司等，在 MySQL5.1.X 版本里，你可以自己选择是否加载打过补丁后的 InnoDB Plugin (版本 1.0.X)，InnoDB Plugin 较 Built-in 版本新增了很多特性，在后面会一一介绍，这里不再叙述。我之前用的 MySQL5.1.43 二进制版中，InnoDB Plugin 已经包含在其 /usr/local/MySQL/lib/plugin/ 目录下，如图 1-1 所示。

```
total 17808
drwxr-xr-x 2 mysql mysql 4096 Jun 1 2010
drwxr-xr-x 3 mysql mysql 4096 Jun 1 2010
-rw-r--r-- 1 mysql mysql 11686678 Jun 1 2010 ha_innodb_plugin.so
-rwxr-xr-x 1 mysql mysql 1054 Jun 1 2010 ha_innodb_plugin.ln
lrwxrwxrwx 1 mysql mysql 25 Jun 10 2010 ha_innodb_plugin.so -> ha_innodb_plugin.so.0.0.0
lrwxrwxrwx 1 mysql mysql 25 Jun 10 2010 ha_innodb_plugin.so.0 -> ha_innodb_plugin.so.0.0.0
-rwxr-xr-x 1 mysql mysql 6390758 Jun 1 2010 ha_innodb_plugin.so.0.0.0
```

图 1-1 InnoDB Plugin 补丁

修改 my.cnf 配置文件，采用 InnoDB Plugin，添加如下内容：


```
ignore_builtin_innodb
plugin-load=innodb=ha_innodb_plugin.so;
innodb_trx=ha_innodb_plugin.so;
innodb_locks=ha_innodb_plugin.so;
innodb_lock_waits=ha_innodb_plugin.so;
innodb_cmp=ha_innodb_plugin.so;
innodb_cmp_reset=ha_innodb_plugin.so;
innodb_cmpmem=ha_innodb_plugin.so;
innodb_cmpmem_reset=ha_innodb_plugin.so
```

重启 MySQL 服务后，登录“MySQL>select @@innodb_version;”验证是否成功，这个是 MySQL5.1 的配置方法。

而在 MySQL5.5.X 版本里，你可以省去上面的那些操作步骤，直接修改 my.cnf 配置文件，添加如下内容即可：

```
innodb_file_per_table=1
innodb_file_format=barracuda
innodb_strict_mode=1
```

登录到 MySQL 里，执行下面命令验证是否成功。

```
MySQL> select @@version;
+-----+
| @@version |
+-----+
| 5.5.19    |
+-----+
1 row in set (0.00 sec)
```

```
MySQL> select @@innodb_version;
+-----+
| @@innodb_version |
+-----+
| 1.1.8            |
+-----+
1 row in set (0.00 sec)
```

关于 InnoDB Plugin 配置介绍，参见 MySQL5.5 手册：

Because the InnoDB storage engine has now replaced the built-in InnoDB, you no longer need to specify options like `--ignore-builtin-innodb` and `--plugin-load during startup`.

To take best advantage of current InnoDB features, we recommend specifying the following options in your configuration file:

```
innodb_file_per_table=1
innodb_file_format=barracuda
innodb_strict_mode=1
```

在 MySQL5.1 里，Built-in InnoDB 文件格式是 Antelope，在 MySQL5.5 里 InnoDB Plugin 文件格式要调整为 Barracuda。Barracuda 格式支持表压缩功能，TRUNCATE TABLE 的速度比以前要快。