



综述 Linux 平台上的数据库

上海同济大学 8 卅 周杰

摘要: 本文将着重介绍 Linux 平台上常见的几种 RDBMS 的现状和发展情况, 对它们的性能进行阐述和比较, 希望能对关注 Linux 的企业用户有所帮助。

关键词: Linux DBMS

引言

Linux 作为网络服务器被广泛应用, 完成诸如 Email、WWW、FTP 及 Internet 接入这样的服务, 正因为它是在 Internet 开发的, 这些功能也就最完善、最为人们所认可。但这毕竟只是计算机应用的一个方面, 如果它要成为一个主流操作系统, 对于 OLTP、OLAP、电子商务、数据仓库、移动计算等商业智能应用的支持必不可少, 因为这些应用对企业来说是至关重要的, 这就要求在 Linux 上出现真正的企业级数据库。

由于国际风云变幻莫测, 众多企业对 Linux 的关注日益加深, 从安全角度考虑, 使用 Linux 可能是我们最好的选择。但新技术往往风险较大, 在这项技术成为绝对主流前运用它, 可能会促进企业全速发展, 也可能限制其发展。这完全取决于它在市场上的表现, 以及在企业内部如何运用它。对于 Linux 系统来说, 冒险与机遇并存。只有仔细分析 Linux, 了解它的优势与弱点, 把握其发展方向, 才能在实际运用中产生效益。

现状

目前在企业关键任务领域, 采用更多的是 BSD 或是商业 UNIX, 但这并不说明 Linux 不能作为大型数据库平台。其实 Linux 像在网络服务器领域受到好评一样, 作为数据库平台具有很大的优势。其小巧、高效的核心不输给任何 UNIX 系统, 高稳定性、低资源占有率、优异的性能曲线更不是 Windows 系统所能比的。事实上已经有大量的开发商支持 Linux, 尤其是 1998 年秋天, '顶端' 数据

库厂商如 Oracle、Informix、Sybase、IBM 都发布了 Linux 平台上的数据库产品。有意思的是 Oracle 和 Informix 在同一天宣布支持 Linux, 结果是除了微软(其实微软的早期数据库产品来自 Sybase), 所有的大型数据库厂商都有了 Linux 产品。今天 Linux 上的数据库系统如此之多, 以致于将他们一次全部列出是件很困难的事情, 但可以通过分类使大家对它们有所认识。

Linux 上的数据库分类方法大致有两种:

1. 按成本或使用版权划分

根据开发者/开发商应用版权协议的不同, 可以分为商业软件(Commercial)、共享软件(Shareware)和自由软件(GPL: 发放源代码)。表 1 列出了三类已有的数据库系统。

2. 按兼容性划分

(1)xBASE: 这是早期 DOS 中被大家熟知的'面向 PC' 的数据库系统, 它的第一代被称为 Vulcan, 运行 CP/M 下, 后更名为 dBase。通常兼容 xBASE 的数据库使用文件方式将应用数据和排序数据分别存储(如 DBF 文件), 应用程序直接读取这些文件。但这种方式过于原始, 虽然能够通过锁定机制保证多用户使用, 但当同一个文件被同时访问时, 阻塞现象仍然很严重。

(2)SQL: Ingres 的 QUEL 语言被认为是 SQL 的祖先, 但很多人认为 SQL 其实比 QUEL 差劲, 因为 QUEL 的语法比 SQL 更简单和强劲。这类数据库系统通常拥有一个中心数据库引擎, 应用程序并不直接访问数据, 而是通过引擎来完成数据的读写, 这样有效地防止了用户误操作或硬件不可预知的灾难。

表 1 Linux 下常见的 RDBMS

Commercial	Shareware	GPL
ADABAS D	GNU SQL Server	Beagle SQL
CQL++	Leap	Essentia
D3	MySQL	Gadfly
D3 ProPlus	NoSQL	jBASE
DB2 Universal Database	Paloma	MDBMS
Embedded Empress RDBMS	pql	MIMERmm
Empress DataWEB SQL)	Strong Related DataBase	mSQL (Mini
Empress Linux Suite	xsqlmenu	
INFORMIX-SE	XmySQL	
Ingres II	PrimeBase	
InterBase		
Just Logic/SO		

1. 小型数据库系统

(1)mSQL (miniSQL)。mSQL是澳大利亚Hughes公司提供的—种简单、快速的数据库引擎,支持标准SQL的一个子集,包括joins、order、distinct、null值处理等。它不像其他数据库引擎那样占用大量的系统资源,性能稳定可靠,可以称之为轻量级的数据库系统,被大量用于事务处理要求不高的网站和基本的数据库应用。mSQL拥有整整一套免费工具,支持Java、perl语言和ODBC、JDBC等web接口,提供客户机/服务器模式,支持CGI机制加快动态网页提供客户机/服务器模式,主要的数据库应用,如:varlaris、XT、//

等。此外,包括gmysql、GUI工具,其数目要超过间接说明了它的应用地位。其主要特

- 多线程支持
- 传输加密
- ODBC, Perl,C,Java,Tcl,Python的API接口
- 支持可变长度的表格式
- 提供自动增量型和timestamp型字段
- 支持所有的ODBC字段类型以及EXT、BLOB、SET、ENUM、MEDIUMINT(24位)、BIGINT(64位)等数据类型
- 支持位操作
- 拥有5个不同的优化器
- 提供大量的字符串函数和统计函数
- 对硬件要求的可伸缩性:能工作在2M--4G的内存范围内

常见的 RDBMS 的分析

关系型数据库目前仍然是市场上最流行的数据库管理系统,大多数商业应用基本上都采用C/S或B/S体系结构,架构在RDBMS上,因此Linux上的主要数据库产品也是关系型的。在对常见的几种RDBMS进行分析之前,需要明确的是一个生命力强的产业级数据库服务器应该提供这些功能:事务支持、两阶段提交、自动检查点、硬件灾难恢复、热备份以及提供相应的工具,如报表生成器、查询工具、Web网关、友好的系统管理界面等。

- 使用文件方式存储单张表
- 整型与字符串型数据可自动转换
- 可使用普通格式文件导入、导出数据库数据
- 字符串型数据可设置为大小写敏感

MySQL使用GPL版权协议,在UNIX和OS/2平台下免费,但需要的话也提供商业应用支持。当前稳定版本为3.23.10,最新为3.23.13a,主站点是<http://www.mysql.com>,可应用于Sun Solaris、Linux、FreeBSD、BSDI、IBM AIX、OSFI、HP-UX、SCO、Dec-Uinx、Irix等。

(3)PostgreSQL (Postgres95)。PostgreSQL来源于加州大学Berkeley分校Ingres数据库研究项目(此后发展为Postgres),它保持了Postgres的易扩展性、面向对象的数据建模、丰富的数据类型等显著特征,但使用SQL的扩展子集代替了PostQuel查询语言。因此,PostgreSQL是:

- 关系型的,postgres项目的研究目标之一就是显示关系型数据库本质上能够处理复杂对象和规则。PostgreSQL基本上包括了成熟的RDBMS具有的所有特性,例如声明SQL查询、查询优化、并发控制、事务处理和多用户支持。

- 高扩展性的,允许自定义操作符、字段类型、计算函数及访问方法等。

- 面向对象的,人们使用“对象-关系型(object-relational)”来描述PostgreSQL(当然大多数高端数据库厂商也声称他们的产品是对象-关系型的),支持某些面向对象的特征,如继承。

它具有如下特性:

- 处理多个并发查询的能力
- 在线备份
- 事务处理
- 多种排序方法(b-tree,hash等)
- 商业的ODBC驱动
- 使GiST实现索引

PostgreSQL由一组Internet开发者维护,是最活跃的开放源代码项目之一,有大量的第三方工具,包括PgAccess、KPGsql、GtkPGA、EARP以及guile-pg和pg-guile等Guile接口,支持C、C++、Tcl、Perl5、Python、ODBC、JDBC。其稳定版本为6.5.3,最新版本为7.0beta2,主站点为<http://www.postgresql.org/>,其商业化产品名称为IIIustra,主站点是<http://www.illustra.com/>。

(4)InterBase。在Windows下使用过DEPHI的程序

员都知道其捆绑的InterBase,2000年1月3日,Inprise/Borland公司宣布公开Interbase 6在所有平台上的源代码,成了世界上第一个开放源代码的商业数据库厂商。Interbase主要定位在关键商业计算领域、移动计算、基于Internet的应用,这次公开代码使其用户在可靠性、安全性和TOC上获益,引起了业界的震动,其先进的体系结构和复杂的技术特色必将吸引许多Linux用户。它支持ANSI SQL 92,使用UNICODE字符集,支持BLOBs数据类型、多维数组和JDBC,是一种真正的分布式跨平台的数据库环境。当前测试版为6.0,支持DEC, HP, IBM, Linux, Microsoft, NCR, Novel, SCO, SGI, SUN。

(5)GNU SQL。GNU SQL Server来自俄罗斯,是一种免费的可移动式数据库,它支持整个SQL89并拥有SQL92的某些扩展特性,提供了多用户访问和事务分离机制。目前版本0.7b5,仍然处于ALPHA测试阶段,有兴趣的朋友可以从<http://www.ispras.ru/~kml/gss/>得到关于GNU SQL的信息。

2. 企业级数据库

(1)IBM DB2。IBM DB2 Universal Database(简称DB2 UDB)作为IBM的旗舰产品,是一种完全支持Web应用的关系数据库管理系统,具有从单处理器系统到对称多处理器系统的可伸缩性,并具有先进的面向对象型关系数据库的支持,兼容image、audio、video、text等二进制大对象。DB2 UDB 5.2以后,IBM不断提高数据库在Web中的应用技术,提供了广泛的Java支持、先进的管理工具和增强的性能特征,用户可使用SQLJ以及Perl、Java、JDBC或Net.Data创建动态数据驱动的web应用。

DB2 UDB使用图形化管理工具—DB2 Web控制中心,允许用户通过Web执行管理任务。该中心提供了安装、管理和调节数据库所需要的所有工具。管理人员在一个控制点就可对多个远程DB2通用数据库进行管理。除了管理数据库实例、数据库和表空间外,控制中心也提供对于模式、表、视图、索引、用户、用户组、用户定义的类型、用户自定义函数、触发器和磁盘的管理。该中心也内置了一个调度程序,使作业可在无人值守的情况下于某一指定时间,或每隔一特定时间自动执行。

可以说DB2的数据库技术是不容置疑的,也是实际的工业标准,它足够强大以满足公司需求并具有足够的灵活性来服务于中小型业务。但在Linux领域,IBM对DB2使用的版权协议与其他操作系统下完全一样,这和Linux开放源代码的精神本身是相违背的,并不更加有利于

Linux的应用发展。当然作为商业数据库系统,这也无可厚非。

(2)Oracle。1999年Oracle宣布它正在为Linux开发Oracle 8和8.1数据库,测试版可以从它的网站<http://platforms.oracle.com/linux/>下载。Oracle宣称其Linux产品将使用户用最低的成本实现企业应用,并由此获得Oracle给开放的Linux环境带来的柔性、可度量性和可靠性。与Oracle在其他操作系统环境下一样,Oracle为Linux用户提供了一系列产品:

- WebDB--提供建设、开发、管理Web数据库应用和内容驱动的Web站点的低成本解决方案。

- Oracle Lite--用于移动应用。

- Oracle8 Server--数据库管理系统。

- Oracle8i--一个完全的集成平台,用于建设和开发internet关键应用。

- Oracle Application Server--具有可扩展的、基于工业标准的大型数据集成功能,主要用于协同工作和电子商务。

- Oracle Application--主要用于企业财务、项目管理、人力资源管理、供应链和产品制造领域。

- Oracle Developer--用于电子商务的Internet快速开发工具。

(3)Informix。Informix-SE是一个高质量、强有力、具有丰富特性的SQL数据库,主要定位在中小型数据库应用上,具有高性能、易配置、易维护、易管理的特色,使之成为Linux操作系统中广泛使用的商业级数据库。

(4)Sybase。1999年8月,Sybase宣布其旗舰产品Adaptive Server Enterprise(ASE) 11.9.2支持Linux,并准备将EAServer3.0和Replication Server12.0移植到Linux平台。在此之前Sybase已经将SQL Anywhere作为移动和嵌入解决方案增加到Red Hat Linux6.0中,同时以免费但不支持的形式提供OpenClient/C接口库,使Linux客户端能够跨平台连接到Sybase数据库。

这样Sybase从后端数据库应用到Web体系结构,在Linux上拥有了其全线产品。Sybase ASE是一种低成本、高性能的企业平台,使用户能够集中精力处理其基于

Web的关键应用,EAServer则提供给用户可扩展的、分布应用的开发环境和交叉平台支持,使得开发复杂的企业应用更加容易。

3. 数据库工具

我们知道性能优异的数据库引擎只是应用的基础,一个好的图形化管理工具无论是对专家还是一般用户都是极其重要的,不幸的是Linux下的大多数数据库系统仅仅提供命令行形式的外壳接口,很少集成像Oracle SQL * Forms或Sybase PowerBuilder这样的前端开发工具。因此除了'数据库引擎'之外,许多第三方组件也是不可少的。这些工具可分为:

- RAD(快速应用开发)工具

- 图形用户界面的查询和数据库管理工具

- 报表生成器

- Web数据库网关

由于篇幅所限不再一一列举。

发展前景

Linux有着非凡的活力,使得任何一家数据库厂商都不愿忽视它具有的潜力。于是免费的'单用户版'成为商业数据库开发商的一种策略,允许用户将其用于家庭和非商业用途,有利于开发者在低成本下实现应用。但真正的企业计算平台需要可靠的、有着成功背景的解决方案和大型关键系统的经验。大型的企业级应用不仅需要更稳定、可靠的商业数据库产品,更需要数据库厂商对Linux平台的全力支持和与Linux系统的全面结合。

让人感到可喜的是,Linux正在向着这个方向发展。一方面高端数据库厂商如IBM、Oracle、Sybase纷纷支持Linux,他们的进入为Linux平台的DBMS带来了行业标准。同时,SAP等大型商业应用软件开发商正在准备进入Linux。各种DBMS不断增加的新特性为基于Linux的电子商务、数据挖掘、OLTP和OLAP、Internet/Intranet等应用提供了稳定、高效的数据库平台和强有力的支持。随着更多的企业重视和实施Linux数据库应用项目,Linux平台上的数据库系统将走向成熟,并领导数据库发展的潮流。■