

10年砥砺 梦想成真 (关于 IBM 通用数据库 UDB 的问答)

如果说数据库技术是现代信息系统的能源,那么数据库管理系统自身的核心又是什么呢?多年来很多企业 and 专家都在探索这个问题。IBM 公司同时也是在数据库技术发展中做出贡献最大的企业之一。10年前 IBM 制定了一项名为“恒星爆发(Starburst)”的研究计划,当时这还仅仅是朝着理想数据库方向探索的纯研究项目,随着“恒星爆发”研究过程的日益深入、研究成果的日益积累,它自身日益强大的凝聚力把 IBM 遍布全球的研究力量全部集合在了一个宏伟的数据库远景技术之下,现在随着 IBM 通用数据库(IBM UDB)的正式面世,“恒星爆发”终于揭开了神秘面纱,这项计划所引发的一场深刻的数据库技术革命正在全球展开,企业、专家和用户期待已久的通用数据库终于变成了现实。

John McPherson 博士是 IBM“恒星爆发”研究计划的关键人物之一,他的正式职务是 IBM Toronto 实验室 DB2 公共技术开发部经理,管理着包括编译器、数据通信、交易管理、OLAP 和并行查询处理在内的多个研究项目。在加入 Toronto 之前,McPherson 曾经担任过 IBM Almaden 研究中心探索数据库系统部经理,和一群举世公认的数据库技术精英们共同进行对象关系型系统、异构多媒体、先进工作流程和数据挖掘等方面的研究工作,如今他们的一系列研究成果正在强烈地冲击着市场。

近日 McPherson 博士从多年潜心修行的幕后走上了台前,在一次接受国外 IT 媒体的采访中详细介绍了 IBM 通用数据库技术发展的全过程,从中可以深切地感受到在 IBM 数据库技术成功背后一批杰出人才的执着追求。

问:对大多数用户来说,从研究开发工作发展到新产品的全过程是一个十分神秘的过程。IBM DB2 UDB 是如何完成这一过程的呢?

McPherson: IBM DB2 UDB 包含着许多数据库技术精英的聪明才智和杰出贡献。在很多情况下,开发新产品的工作是首先从我们研究部门开始的,IBM 拥有

世界上最引人关注的数据库研究机构,只要稍作回顾,不难发现象关系型数据库、基本查询优化器技术 SQL 和许多我们今天已经视为理所当然的那些基本的东西,都是由 IBM 研究部门发明的。

问:“恒星爆发”研究成果形成了 IBM DB2 UDB 的核心。你能不能谈谈在“恒星爆发”研究计划背后的一些考虑?

McPherson:好的。我们得先回到 80 年代中期,那时候我们已经提供了几个关系型数据库:系统 R(IBM 的原型关系型 DBMS)、SQL/DS、DB2/MVS,但我们仍然强烈地感觉到应该认真回顾历史、重新考虑系统的结构性能和如何支持今后不断变化着的应用。那时候在数据库圈子里有两种截然不同的观点,一种观点是应该建立“艾达”数据库—即将任何人可能需要的功能特性都放到 DBMS 中;另一种观点则是:“不,你不能预先想到人们将来需要什么,我们应该建立一种可扩展的数据库,使人们可以很容易地往其中加入更多的功能。”在“恒星爆发”研究计划中,我们采纳了后一种观点,朝可扩展数据库方向努力。刚开始时,我们的重点研究方向只是如何支持新数据类型的问题。

问:OS/2 扩展版 DBMS 是否也是出自“恒星爆发”研究计划?

McPherson:“恒星爆发”研究计划最初的部分工作形成了 OS/2 扩展版 DBMS 的核心,那时候我们只是刚刚开始着手构建,并没有完全重写查询,因此没有许多后来在 DB2 公共服务器 2.0 版中出现的功能,不过我们当时有一个查询图模型(QGM)的早期版本,从这个早期版本出发,我们继续开发查询重写技术,后来这一技术把产品推到了艺术级水准。

到 1991 年来至 1992 年初,“我们必须开发一个多平台可移植数据库”这一点开始变得明朗起来,因为那时并没有 AIX 平台上的数据库,迫使我们开发一个非 IBM 平台的数据库,由此我们想到了可移植性,因此,可移植性成为“恒星爆发”研究计划的另一个主要目标。大约从那个时候开始,研究部门开始与 IBM 软件

解决方案部形成一种默契的全作关系,双方密切合作的结果是建立 DB2 公共服务器 2.0 版和 DB2 PE,可以说 OS/2 和 DB2 PE 确实改变了我们在 IBM 工作的方式。

问:我们曾经听说过数据库技术研究机构(数据库技术研究机构),这个研究机构到底是怎么回事?它现在还在活动吗?

McPherson:是的,数据库技术研究机构还存在,虽然现在很少能听到有关它的消息,但它的重要性丝毫没有降低。数据库技术研究机构和所有研究组之间建立起密切的合作关系,Pat Selinger 是数据库技术研究机构的主任,Hamid Pirahesh 是新近任命的经理。数据库技术研究机构的目标有二:一是帮助将技术从研究部门转化到产品部,二是促进研究项目与产品部所面临问题之间的密切联系,这实际上是信息的双向流动。数据库技术研究机构跨越了各部门的边界,在将技术转化到产品部时,能够有效建立起各部门之间的信任和理解。

问:当 IBM 在收购了一个公司以后,例如 Lotus,情况会有怎样的改变?把他们的技术结合进来需要进行怎样的转换?

McPherson: Lotus 自己有一个研究组织,当然我们研究部门的人员也要与他们一起工作,但我们并不是去告诉他们应该做些什么或改变他们做事情的方式,而是渗透到他们正在进行的工作当中。Almaden 有很多数据库事物处理管理、并行管理和数据复制方面的专家, Lotus 现在正在与 Almaden 合作以便提高 Lotus Notes 的可伸缩性,这是 IBM 研究部的专家和 Lotus 合作的一个很好实例。自从 IBM 发展了一系列重要合作关系以来我们一直是这样做的。

问:你怎样确立研究规划?数据库技术研究机构是不是一个指导委员会?

McPherson:我们根据规划与数据库技术研究机构协同工作,同时我们也独立自主地工作,这决定于研究工作本身。10 或 15 年前,我们常常在做一些工作后就不了了之,现在我们的工作程序是建立规划、协同工作,然后传递给产品部门,并帮助产品部门开发产品。IBM 很重视最基础的探索性研究工作,这些研究的工作程序是:由技术负责人设立规划——这些负责人通常都是能够胜任工作的博士——对于这样一些世界级

的研究人员,自上而下的沉重管理毫无意义。

问:怎样把这些研究设想移植到产品中去?用户的要求又是怎样与研究规划相协调?

McPherson:所有研究部门和产品部门的人员每年都集聚一到两次共同讨论研究。去年 12 月我们聚在一起确定了 B2 第六版的研究重点,在这期间组成了二级委员会,深入研究在这些重点方面的用户要求,然后把这些重点按优先次序排列起来,以便能够把注意力集中在关键要求上。比如在 DB2 UDB 中,一个优先考虑的方面是在事务处理 OLAP,我们把注意力集中在星型连接查询、OLAP 处理和多维 Cube 及 roll-up 功能上。

问:当某一项技术进入产品开发阶段,随后又进入发布阶段时,这项技术已经发生了很大的改变,某些东西可能在实验室里表现非常好,但在实际应用中并不是那么好。对研究部来说,与产品保持同步是不是很难?

McPherson:这是一个很重要的话题。当我们做“恒星爆发”研究计划时是从零开始的,完全是从空白起步。对于“恒星爆发”研究计划不仅要研究,同时还想建立一个用于研究和开发的试验台或工作台,因此我们选择建立一个可伸缩数据库,它的最大优势是可以做试验,我们已经把它应用到 DB2 公共服务器 2.0 版中,现在又把它用到了 DB2 UDB 中,它使得我们能够在研究和产品中都拥有同样的代码基,我们没有时间先建立经典原型然后再进一步发展,而在实验室中就必须为将来要交付的产品建立原型,所以在这种意义上,“恒星爆发”研究计划是 IBM 的一个重要发展。IBM 每一个人的工作台都是我们的代码库,这使得 Hawthorne 实验室与 San Diego 超级计算机中心能够合作进行超大规模数据库支持的研究工作,Almaden 能够进行异构分布式数据库的工作。

问:这是被人们称为的大蒜(Garlic Project)计划吗?

McPherson:是的,这正是大蒜计划。他们正在把对象扩展放进我们的代码基中。在我们开发 5.0 版的同时,有一个叫做 Object Strike Force 的小组进行对象扩展、抽象数据类型和其他一些工作。这个小组里有 Almaden、Toronto 和 STL 的研究人员。我们计划要在今年年底前公布其中的一部分新特性,并有能力做到这一点,因为我们是从正确的代码基开始的。

问:你是否正在同许多 β 版用户一起工作?

McPherson:是的。通常我们对 β 用户是很有选择性的,因为 β 项目牵涉到IBM和用户两方面的许多开发工作。性能调整是 β 测试项目的一个重要内容。例如,我们已经有一个运行DB2 PE的重要用户,这个用户在DB2 UDB和PE上运行同样的查询基准测试程序时发现DB2 UDB运行得更快些,这是新优化器最惊人的特性之一。因此我们在DB2 UDB中为PE用户彻底更换了编译器,这是一个非常有趣的新的“恒星爆发”编译器。“恒星爆发”研究计划的一个难于抓住的方面是——我们的目标不再集中在任何一种查询上面,验证你是否达到预期目标的唯一方式是运行大量的查询——真正的用户查询,用真正的数据。 β 版用户在达到这一目标过程中起着非常重要的作用。

问:DBMDS就像一个非常复杂的动物。你是在什么时候决定要重写产品,而不是继续增加新特性和功能的?

McPherson:在某种意义上,我们总是在重写,但什么时候决定要重写产品的主要部分呢?这常常是在有了某种新范例以后——当有了某些打破我们最初所作的所有假设的新特性时。

今天的关键问题是要持续不断地向基础结构投资,基础结构的改变要求我们不能只着眼于低级别用户的利益。当你增加一个新功能时,很容易只把注意力集中在准备开发的新功能上,例如新的数据类型。但我们一直努力提醒自己:“我们以前做过这件事,上次就很费劲,这次也不会轻松,所以真的应该花时间把基础结构搞好,使其具有可伸缩性。”这就是使我们保持警觉的动力——希望向基础结构投资,并且意识到这不是我们最后一次打算改变特殊性能和功能。建立好基础结构可能不是把事情弄好的最快捷的方法,但在增加某些新特性并能很好地使用时,你就会真正体会到搞好基础结构的好处了。

问:查询重写似乎是“恒星爆发”研究计划的基本部分。你们是怎样开展工作的?你打算怎样继续使用这项技术呢?

McPherson:我们在“恒星爆发”研究计划中发展出来并已用于DB2 UDB的一项技术就是QGM,这是一种

能够提供通用查询表示方式和易于操作的内部表达查询的方式。在QGM之前,我们使用的是一种查询分析树表达。

查询重写可能从深层操作数据结构,把查询从一个表达工转换成另一个等价的表达形式,以便得到一个适合于优化器的表达形式。

问:是在SQL中重写吗?

McPherson:是。这两种方式我们都可以用。我们可以用SQL把查询转换成QGM;然后可以在任何一点上把查询再从QGM转回到SQL,这是一种很有威力的机制。随着时间的不断推移,我们将加入越来越多的查询转换。在DB2 UDB中就有了扩展性更强的子查询连接转换。

问:在使用户对所得到的性能满意和让他们使用新功能之间似乎要仔细地权衡考虑。查询重写是不是一个必须注意不要超出用户能力太远的例子?

McPherson:好的,这是查询重写的一个精彩方面。这个问题也是我们对优化器技术一直非常注意的方面,我们不要求用户知道写查询的最佳方法,只需用户写出所需要的查询,我们来替他进行转换。事实上大多数查询都不是由用户编写的,而是通过工具产生到DB2当中,用户只需点击一下界面便可建立查询。

问:你能否对为什么查询重写有助于可伸缩性这个问题稍做解释?

McPherson:这样说吧,你作为用户,可以按你所想的方式写查询,我们的数据库来实施这个查询。数据库能够了解你所使用的系统,对每一个平台,产生的查询计划可能相去甚远,但你不必为此担心,你看到的是完全相同的应用、完成相同的查询。对于我们来说,可伸缩性不仅意味着数据库拥有一个可以支持任意数量的节点、任意数量的处理器、任意容量的内存、任意数量的磁盘和磁盘配置的体系结构,还意味着用户不必考虑如何优化查询。这就是为什么优化器技术是我们的可伸缩性的关键。使优化过程看起来自动化了——这是我们孜孜已求的目标。

(来稿时间:1998年1月)