



白皮书

## IBM DB2 pureScale——云及大数据世界中价优、可扩展、可靠的 OLTP

赞助商：IBM

Carl W. Olofson  
2014 年 12 月

### 关于本白皮书

---

此篇白皮书探讨了共享数据横向扩展集群的概念，论述了该集群如何提供持续可用性，以及其在可扩展事务处理支持中举足轻重的原因。与 NoSQL 数据库及 Hadoop 应用所采用的集群方法相对比，关系数据库中应用这一方法显示了关系模型的重要性，以及在大规模事务处理中持续监管模式的重要性。

随后，此篇白皮书论述了 IBM DB2 pureScale 的具体优势，该产品旨在实现服务器横向扩展架构的强大功能，令企业能够以实惠的成本开发并管理事务数据库，满足云世界的要求，应对事务数据的迅猛增长。

### 形势概况

---

#### 数据增长与 IT 预算限制

IT 经理们一直被要求少花钱多办事。他们须以相同或较少的预算处理更多的数据和日益增长的事务处理及分析查询。受保留分析流数据、机器生成数据、社交媒体数据和物联网（IoT）数据需求的推动，大数据应运而生，这无疑是雪上加霜，让这一任务更加艰巨。

随着数据量的不断增长，对吞吐量需求和在同一数据库上支持更多各种各样且更为复杂的数据库操作的需求也随之增长。此外，更高层次的事务智能执行引发更新事务与复杂查询相混合的需求。所有这些都要求更强大的处理能力。

让数据库充分扩展以满足处理能力，这牵涉到在关系型数据库管理系统（RDBMS）中使用扩展技术。如果数据库需要大量的共享存储器来处理多个复杂查询，且工作负载有大量的计算组件，则可以采用“纵向扩展”的方法，让数据库运行在具有大量插槽和大量内存的服务器上。如果旨在使得大量并发用户能够更灵活地分发事务操作或简单查询，一般会优先考虑“横向扩展”集群方法；集群的类型取决于所具体的工作负载。

## 横向扩展数据库集群

集群是一种数据库部署方法，涉及使用多个作为单个数据库服务器运行的服务器（称为“节点”）。增加节点可提高集群整体的计算能力。数据库服务器可支持一个或多个数据库，但在这两种情况下，所有节点视为一个完整的系统。

### 数据库集群与切分

RDBMS 管理的集群与“切分”不同，切分是指一个数据库被切分成多个部分（称为“碎片”），并在多个服务器上托管。凭借分片，每台服务器，无论是作为 OS 管理的集群还是网络上多个服务器的一部分，均充当单一的数据库服务器；这些服务器可以协调它们的一些操作，但不能确保跨服务器的参照完整性或作为一个单位对碎片服务器的集合实现 ACID（原子性、一致性、隔离性和持久性）等。此外，由于每个碎片均作为一个数据库服务器，所以应用程序的部署必须考虑碎片的组织。

### 集群方法

有两种基本类型的集群，分别处理不同的工作负载。它们通常被称为“无共享”和“共享存储”（我们曾称后一种方法为“共享磁盘”，但由于现今存储系统可以在内存、固态存储、固态硬盘或旋转磁盘管理持久性存储，所以“共享存储”这一术语更加精准）。

“无共享”的方法涉及到对数据库中的数据进行分区，将每个分区分到集群中的一个节点。每个节点专门管理该分区中的数据。集群与分片不同，涉及到节点之间的密切合作，它们在外表现为单一的数据库系统，跨节点分配查询，并行地执行查询，收集结果，并向应用程序返回一个结果集。由于无共享方法有利于并行查询处理（也被称为大规模并行处理，或 MPP），所以被视为更适合分析数据库的方法，如数据仓库，此类数据库拥有很少的并发更新线程，而更多的是诸多复杂、长时间运行的查询线程。（以 IBM 的情况举例：DB2 高级服务器版（Advanced Server Edition）和高级工作组版（Advanced Workgroup Edition）中的数据库分区功能（Database Partition Feature）[DPF]。DB2 的此种配置在被称为 PureData System for Operational Analytics 的集成系统包中同样可以获得。）

“共享存储”涉及到在集群节点之间分配计算工作，这些节点在一个共同的存储资源中共享数据（通常被部署为网络连接存储或存储区域网络），使用负载均衡方案在节点上分配工作负载。对于支持多更新线程且更新活动基本不变，请求简洁的数据库，如事务处理数据库，更适合使用这一方法。因此，包含并发更新控制（包括锁定管理）和共享缓冲器的技术对数据库的流畅性能和可扩展性至关重要。

### 集群实现高可用性

集群一般是为了实现高可用性。这通常是通过“活动”数据库服务器和“被动”或“待机”服务器进行配对来实现的，如果活动数据库服务器宕机，则由后者接管。在部署了无共享或共享存储集群的情况下，无需使用这种方法，因为我们通常认为除了支持可扩展性以外，对 RDBMS 的基本要求之一是必须将每个节点实施为另一节点的故障转移节点，使得此类集群方法具有内置的高可用性。然而，任何人在考虑采用集群数据库的方法时都应当调查 RDBMS 供应商在集群中处理故障转移的方式。

### 共享存储集群 RDBMS 的要求

总之，以下是用户对任何共享存储集群 RDBMS 应有的要求：

- 该数据库针对应用程序以单个单位出现，无需特殊应用程序配置或专门的 SQL。

- 内置故障转移，使数据库管理员（DBA）可以选择故障转移数据复制策略，而 RDBMS 将完成剩余工作。
- 在集群中添加或删除节点应是一项简单操作，管理开销低，并对应用程序透明。
- RDBMS 应具有启用大部分自主操作的功能，包括自动节点工作负载平衡、优化、故障转移和修复。

## DB2 pureScale

IBM 的 RDBMS、DB2 for Linux、Unix 和 Windows 提供了名为 pureScale 的配置，它可以在 x86 和 Power 系统上运行。pureScale 可使 DB2 作为共享存储集群运行，作为单个数据库系统与应用程序交互。pureScale 利用了其“近亲”大型机的架构理念，以及 IBM 的 Systems and Technology Group（STG）和 Tivoli 的技术，为客户提供高度自主，易于管理的集群数据库环境。

### DB2 pureScale 集群的工作原理

pureScale 集群中的节点共同协作，以优化其在数据库上的 SQL 事务执行。它们的数据库缓存和数据库锁由主高速缓存管理器同步，称为集群高速缓存设施（Cluster Caching Facility），利用先进的技术来确保所有节点上的数据库映像连续一致。使用传统万兆以太网的纯软件方式或（在支持它的系统上，如 Power 和 x86 系统）使用在 InfiniBand 或万兆以太网网络上工作的 RDMA（远程直接内存访问）的硬件辅助方法，均可实现这一点。如果一个节点崩溃，它们还提供透明的故障转移。甚至连高速缓存管理器也有备机，可在必要时进行故障切换。该系统采用并行日志，以防止日志 I/O 瓶颈。由此产生了一个数据库服务器节点的集合，它作为一个单位运行，操作顺畅、连续，对应用程序透明，因此不需要专门的 SQL。

### DB2 pureScale 实现横向扩展的好处

pureScale 为 DB2 用户提供按需购买、渐进扩展的系统管理方法。用户无需从现在开始为几年后的预期负载分配系统资源，而是以适宜的处理能力为现在的工作负载提供数据库，然后随着工作负载需求的扩大，向集群添加更多节点即可。节点数量可以在不中断数据库操作的情况下得到增加。pureScale 具有跨节点自动工作负载平衡的功能，并根据客户的实际情况提供近乎线性的可扩展性。

由于 DB2 pureScale 仅仅是 DB2 的一种配置，所以现有用户无需以任何方式转换数据或重新组织数据库便可以对其进行使用。现有的 DB2 应用程序的运行将不做任何改动，而且无需任何变化便可以利用 pureScale 的可扩展性。无需定义分区或创建特殊的索引。

DB2 通过以下方式实现横向扩展的收益：

- 透明、合作运行，无需特殊代码或调整变化，完全兼容任何 DB2 标准配置
- 自动故障转移，仍然无需特殊的应用程序代码（该能力还意味着，在大多数情况下，节点可以通过滚动数据库系统升级得以维护，免去针对维护的计划宕机。）
- 简洁的节点插入过程，向上扩展集群，需要极少量的管理参与，无集群中断
- 自我管理功能，在无需特殊调节和监测的情况下确保自校正、负载均衡、乃至自运行维护

## 展望未来

---

### 云计算和持续可用性

无论是在数据中心还是在云服务环境中，各企业正逐渐将云计算作为 IT 的基础架构。这意味着，为了实现云所要求的弹性的可扩展性，服务器和存储资源将可替换。就公有云而言，它往往还意味着要确保持续的可用性。从关系型数据库管理的角度来看，如 **DB2 pureScale** 这类支持横向扩展结构的关系型数据库，是满足这些需求的明智之选。

### 大数据集成

企业正在学习如何将日益增多的大数据整合至其运营当中。这不仅意味着从流媒体源抽取数据，例如来自传感器的机器产生的数据或来自数据服务的订阅源，并将该数据用于分析，而且还意味着在事务工作负载中应用该数据。这也导致事务处理更加复杂，并加大了数据库服务器处理这类事务，同时提供一致的性能，满足为每个工作负载所创建的服务水平协议的压力。通过共享存储集群等提供的灵活扩展性是提供一致性能的关键。

### 实时决策和列存储的信息管理数据库 IMDB

企业必须日益依赖于应用智能和即刻对用户可用的数据，确保智能事务处理系统产生持续的竞争力。事务处理与基于实时事务数据的复杂分析相融合是实时决策的关键要素。一些 RDBMS 供应商，包括 **IBM**，都在开发日益完善而无需牺牲性能的方式来实现这一目标。目前，**DB2** 凭借 **BLU Acceleration** 提供了利用可扩展 **OLTP** 和内存优化的列存储技术来处理复杂分析查询手段来满足这一需求。在 **DB2** 的最新版本中，列存储映像表 (**BLU Shadow Tables**)，就是列存储的加速表 (**BLU Acceleration Tables**) 即在数据库中按列存储实时事务表中的每一行，以接近实时事务的处理性能高速完成分析处理。

随着我们向企业提供实时分析处理能力的目标迈进，融合实时事务和分析处理这两个领域的技术将会持续的得到发展、改善。

### 挑战与机遇

---

RDBMS 所面临的挑战重重，其中包括满足集成云环境的要求、更有效地处理非结构化数据和来自大数据源的数据，如 **Hadoop** 和 **NoSQL** 数据库、并管理混合工作负载环境。该领域的所有供应商都在迅速地发展着自身的技术，**IBM** 肩负着并驾齐驱或超越群雄的压力。

### 结论

---

在线事务处理压力、新的数据源，以及处理更为复杂的应用程序三者方面因素的结合，特别是涉及到在线事务处理，成为固定资源关系数据库 (RDBMS) 面临的关键问题。为了解决以上面临的这些问题需要包括一个可横向扩展的基础架构，允许 **DBA** 扩大 (或缩小) 数据库服务器的处理能力，满足应用需要。从应用程序的角度来看，使用作为单个数据库服务器运行的共享存储集群是这种可扩展架构的最好体现。

在回顾这些需求及其解决方案时，企业应考量以下因素：

- 任何供应商的解决方案应提供接近线性的可扩展性，且仅需极少的管理工作。
- 集群方法应对应用和用户透明。
- 故障转移和恢复运行应自动完成，无需特殊的用户干预。
- 集群应能提供可扩展的性能，而集群访问无需数据库或 SQL 特殊条件。
- 对于现有的 DB2 用户来说，该解决方案应无需数据转换或更改应用程序。

综上考虑，企业应将 **IBM DB2 pureScale** 与任何其他产品在实施集群、可扩展的数据库配置的方面进行比较。

## 关于IDC

国际数据公司（IDC）是市场情报、咨询服务以及信息技术、电信和消费者技术市场活动方面全球首屈一指的供应商。IDC 帮助 IT 专业人士、业务主管和投资机构针对技术采购和业务战略制定基于事实的决策。1100 多名 IDC 分析师在全球 110 多个国家针对技术和产业机遇与趋势提供全球、区域及当地的专家意见。50 年来，IDC 一直致力于提供战略分析，以帮助我们的客户实现其关键业务目标。IDC 是全球领先的技术媒体、研究及活动公司 IDG 的附属公司。

## 全球总部

5 Speen Street  
Framingham, MA 01701  
USA  
508.872.8200  
Twitter: @IDC  
[idc-insights-community.com](http://idc-insights-community.com)  
[www.idc.com](http://www.idc.com)

---

### 版权声明

IDC 信息和数据的外部出版 — 凡是在广告、新闻发布稿或促销材料中使用 IDC 信息都需要预先获得相应 IDC 副总裁或国家区域经理的书面同意。此类申请均应附上所提议文件的草案。IDC 保留因任何原因拒绝批准外部使用 IDC 信息和数据的权利。

版权所有2014 IDC。未经书面许可严禁复制。

