

做出合理的选择

数据库技术已经成为计算机科学技术中发展最快的领域之一，也是应用最广的技术之一。当今，数据库管理系统已经进入到百家争鸣的时代。面对琳琅满目的数据库产品，应用开发者应该如何进行选择呢？

著名的数据库学者，图领奖获得者Michael Stonebraker曾在2010年提出“One size does not fit all”的观点，即在大数据时代，不可能存在某一类数据管理系统能够同时支持对结构化数据、半结构化数据、非结构化数据的所有应用。传统关系数据库管理系统（如Oracle）一统天下的局面将不复存在。此后，各种各样的数据管理系统如雨后春笋般不断涌现。这些系统在功能、性能和易用性三个衡量指标上表现各不相同。

- 功能：衡量系统的数据处理和数据分析功能，如，支持事务处理，保证事务ACID属性；支持复杂查询，如连接查询、聚集查询等；支持搜索功能等；
- 性能：衡量系统数据存储和处理性能，如可扩展性、可伸缩性、高可用性、系统每秒能处理的查询数（Queries Per Second, QPS）、每秒能处理的事务数（Transactions Per Second, TPS）、响应时间（Response Time, RT）以及系统同时能处理的请求数（系统并发）等；
- 易用性：衡量基于数据管理系统做应用开发的难易程度。

目前，不存在三个指标都能做到极致的数据管理系统。因为，在数据管理系统设计之初，总需要在指标上做一些折中和取舍。例如，为了追求系统性能和易用性，牺牲部分系统功能，如不支持事务处理和连接查询等。权衡之后的数据管理系统虽然无法适用于所有的应用领域，但是能在某一应用领域中表现最优。

不同的权衡和取舍造就了如今数据管理系统的百花齐放、百鸟争鸣的新局面。这也给应用开发者带来了无限的困扰——如何才能为某一应用选择合适的数据管理系统？通常，数据库选型需要从应用类型、数据类型、功能需求、性能需求以及易用性五方面进行综合考虑。

- 应用类型方面：首先明确应用的类型，如OLTP事务处理应用、OLAP数据分析应用，根据应用类型缩小选型范围。比如，如果应用属于OLTP类则只需在OLTP数据管理系统中进行选型；
- 数据类型：然后明确应用支持的数据类型，如结构化数据、非结构化数据、半结构化数据，进一步缩小选型范围。比如，数据类型为非结构化数据则只需在非关系数据库管理系统中进行选择；
- 功能需求：然后分析应用的数据管理功能需求，如需要事务处理功能，基于地理位置搜索功能等，根据应用需求确定选型范围；
- 性能需求：然后分析应用的性能需求，如需要支持可扩展性、高并发量等，根据性能需求确定选型范围
- 易用性：最后比较数据库管理系统的开发效率和运维成本，确定最终的选择。

做出合理的选型通常需要综合考虑多方面因素，而这些都是建立在应用开发者对各种数据管理系统理解和应用实践基础之上。本小节只是给出了一个较为通用的选型策略。