下一代数据库的三重功能：云计算、平台即服务和开源

　　如今，越来越多的组织正在将他们的下一代数据库体系结构转移到解决方案而不是服务器上，并且接受开源数据库技术的价值。

　　2017年对于数据库技术市场来说是关键的一年，在此期间经历了几个大规模的模式转变，并且没有迹象表明很快就会停止。企业正在逐步摆脱传统的单一数据库架构，数十年来，这些数据库架构经过多代的应用，以换取更优化的、灵活的、自我管理的云数据平台战略。



　　随着对规模、安全性、性能、高可用性，以及可能最重要的成本和操作优化的需求不断增加，人们可以确定为下一代数据库架构的“2.0版本”提供动力的三个重要驱动因素：

　　(1)客户正变得以云计算为中心。

　　(2)客户越来越依赖平台即服务的数据库解决方案。

　　(3)客户正在采用开源数据库引擎。

　　想要验证企业自己的IT数据库战略是否与数据库市场的总体发展方向保持一致?企业所要做的就是确保自己采用这些原则。以下更详细地探讨它们的含义。

　　以云计算为中心的数据库时代

　　早在2014年，当重新调整微软公司的整个软件和服务战略时，首席执行官Satya Nadella提出了“移动优先，云计算优先”的口号。Nadella一直在表达微软在移动，云服务和解决方案领域所付出的努力，而整个企业市场似乎都热衷于接受类似的优先事项。

　　例如，根据麦肯锡公司的主题为“IT即服务云”和“企业云基础架构调查”，云计算作为工作负载主要环境的企业采用率预计将从2015年的10%上升到2018年的51%。麦肯锡公司还指出，企业可能会选择超大规模的云解决方案提供商提供的服务，例如亚马逊、谷歌、微软。

　　即使是非常传统的公司，他们投资拥有数十年的数据中心业务经验的大型IT组织，都会发现需要接受以云计算为中心的IT战略的优势。

　　但是，仅仅打着“云计算”旗号让企业采用是不够的，甚至可能会让人混淆不清。企业是否计划将现有服务器从本地数据中心提升或转移到云端，或者更进一步，开始为企业的数据库采用原生云计算解决方案?

　　虽然“提升和转移”(IaaS)策略可以帮助减少企业的IT管理开销，并大幅提高运营的灵活性和可用性，但是对于许多“云优先”组织来说，这只是浅尝辄止，并没有完全深入地采用云解决方案和服务。

　　从思考服务器到思考服务

　　根据调研机构Gartner公司的“2017年PaaS关键趋势”调查报告，企业客户正在采用平台即服务(PaaS)来实施越来越多的战略举措，越来越多的客户认为PaaS将成为平台交付的主要形式。

　　其原因很简单：PaaS数据库技术是自我管理和自我操作的。这些解决方案大大降低了维护数据库体系结构的运营开销，使企业可以将IT人员(尤其是数据库管理员)的时间和精力投入到以应用程序为中心的活动中，从而推动业务成功，而不是保护过于复杂的数据库基础架构。

　　数据库软件和云计算供应商正将大量的研发力量投入到PaaS数据库解决方案中，这正成为云计算领域厂商(亚马逊、微软、谷歌以及Oracle))最重要的差异化因素之一。

　　早在2014年，亚马逊公司就宣布推出一款完全可管理的关系数据库引擎Aurora，它将高端商业数据库的速度和可用性与开源数据库的简单性和成本效益相结合。此外，也许更重要的是，Aurora是AWS公司的原生PaaS数据库解决方案。自推出以来，Aurora得到了很多应用和思想共享。在亚马逊2017年度re:Invent展会期间，亚马逊公司将其云数据库产品进一步推向市场，并宣布Aurora Serverless是一个完全可管理的“抽象”数据库，而在这个数据库中，用户需要为所使用的数据库资源付费。

　　很显然，亚马逊公司在云解决方案中的创新和颠覆，PaaS数据库产品在整个行业范围内掀起了浪潮。Oracle公司是传统高端企业关系数据库软件市场的领导者之一，也是一家在研发方面投入了数十年时间研发创建全球最强大、最优化的可配置和可调的关系数据库之一的公司，该公司在OpenWorld2017年度大会推出自己的“自我驱动”数据库产品。Oracle公司首席技术官Larry Ellison描述了自治数据库云的愿景的一部分，Oracle公司关注自驱动、自扩展、自修复数据库解决方案，试图消除与调整、修补、更新和维护Oracle数据库相关的人力成本。

　　不仅仅亚马逊和Oracle公司推出了数据库产品，微软公司也推出了Azure SQL DB，这是一个完全托管的关系型云数据库服务，并尝试自动调整数据库以改进性能和保护。Azure SQL DB是微软公司的PaaS产品，结合了微软云平台SQL Server和Azure的高端性能和功能。

　　开源数据库引擎的兴起

　　商业数据库引擎(Oracle，SQL Server和DB2等)在它们所提供的功能和性能的广度上是至高无上的。事实上，这是其许可价格高的传统理由。但是，开源数据库管理系统已经成熟，成为专有商业数据库引擎极其可行的替代方案。特别是近年来，随着PostgreSQL和MySQL的进步。对于许多使用情况来说，这些开源数据库可能会与商业数据库市场的行业巨头的产品正面交锋。

　　对于那些一直关注数据库行业状况的业界人士来说，采用开源数据库引擎的热潮不应该让人吃惊。Gartner公司的一个主题为“2015年开源RDBMs状况”的报告早在2015年就预测开放源代码的数据库将被用户广泛接受：“到2018年，70%以上的新的内部应用程序将在OSDBMS上开发，现有商用RDBMS实例中的50%将被转换为正在进行中。”

　　更重要的是，如果用户将开源数据库部署到云中，那么开源关系数据库的最大缺点也许就是它们缺乏某些高端功能(与其专有的数据相比)。

　　例如，Oracle公司和微软公司都提供了直接构建到其数据库产品中的异步消息传递/排队机制。Oracle公司称之为高级队列，微软公司称之为服务代理。PostgreSQL或MySQL本身不提供类似的功能。在开始使用开源数据库技术之前，企业需要思考这个潜在的问题。

　　采用云计算

　　通过利用主要云平台提供的功能，并将这些功能作为开源数据库架构的一部分进行集成，企业可以更轻松地实现与商业数据库的功能对等。

　　例如，如果企业使用Amazon Aurora MySQL，则还可以使用亚马逊简单队列服务(它是用于微服务和分布式系统的完全托管的消息队列)和Amazon Lambda(允许用户运行无代理程序代码)作为替代方法的组合针对Oracle AQ和SQL Server Service Broker。使用Amazon SQS + Lambda，用户可以直接从Amazon Aurora MySQL对消息进行排队，就好像数据库具有内置队列服务一样。

　　同样，如果用户需要定期执行复杂的数据转换，则在微软Azure中运行MySQL数据库将允许用户使用Azure数据工厂，从而通过云端PaaS解决方案减少源数据库引擎中的缺失功能。这些云服务以及许多其他云服务可以与PaaS开源数据库引擎集成，并作为完整解决方案体系结构的一部分。

　　总而言之，现代关系型数据库的整个概念正在从单一的体系结构转向数量庞大而功能强大的数据库服务器，这些数据库服务器内置了大量的功能，以敏捷的PaaS和利用云平台作为服务的开源模型解决方案和技术。

　　这是一个明确的范式转变，可能会为今后十年的数据库技术发展奠定基础。