

ESG 实验室评审

InterSystems IRIS 数据平台：实现快速商业洞见的统一、高效数据平台

日期：2018年3月 作者：高级IT验证分析官 Kerry Dolan

摘要

这一份ESG实验室评审记录了我们为InterSystems公司的IRIS 数据平台进行性能测试的结果。我们在概念验证环境中测试了IRIS数据平台的性能，并将其与其他三种常用数据库进行对比。测试重点关注于金融服务行业常用的查询语句运行的性能。

挑战

事务型数据库是很多收益驱动的业务过程所必需的，同时数据分析能提高竞争优势。但企业往往会受困于传统的关系型数据库管理系统（RDBMS）模型难以适用于他们已有的数据类型和想要寻求的知识。非关系型数据库（NoSQL）和新型数据库（NewSQL）能提供有用功能，但同时使用多个数据库意味着需要构建不同的架构，而这些数据库架构都有性能、可扩展性和灵活性方面的要求。此外，企业一直希望提高数据库性能从而加速商业洞见；内存数据库的处理速度很快，但处理海量数据的能力有限。上述所有需求都会增加成本。所以，毫无意外，当ESG请IT决策者列出数据库面临的挑战时，48%的受访企业提出“管理数据库大小及其增长”，35%受访企业提出“满足数据库性能需求”而34%受访企业提出“支持基于云端的数据库”¹。

图1 数据库面临的挑战



48%的受访企业提出
“管理数据库大小及其增长”



35%受访企业提出
“满足数据库性能需求”



34%受访企业提出
“支持基于云端的数据库”

来源：企业策略集团

大多数企业最终选择了使用多个数据库、增加基础架构和人力成本。对于企业所有的数据库需求----事务型工作、分析型工作，实时数据、历史数据、自然语言处理和商业智能---如果企业可以减少实现需求的步骤，它们就能节省成本。同时数据库必须提供高性能，数据持久性、可扩展性且易于管理等特点。

¹来源：ESG研究调查，企业数据库发展趋势，2017年1月

这份ESG白皮书是受InterSystems公司委托完成的，在ESG许可下发布。

©2018企业策略集团股份有限公司。版权所有。

解决方案 : InterSystems IRIS 数据平台

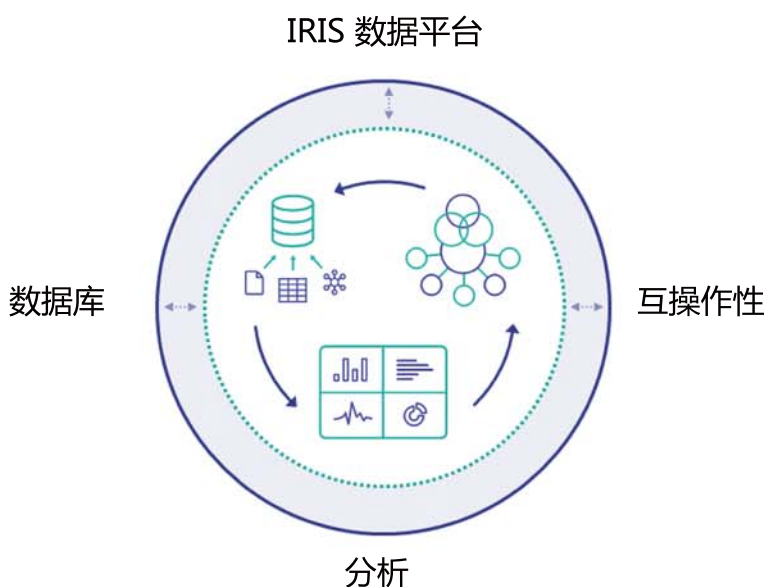
InterSystems IRIS 数据平台旨在帮助企业构建能支持多重事务处理和分析性用例的富含数据的智能应用程序。

InterSystems IRIS 是一种统一的实时平台，它将强大的数据管理、分析和互操作性能力融入单个解决方案。它被设计用于在云端、本地或混合环境中大规模地并且经过优化地处理复杂工作及数据敏感的应用程序，同时提供内置安全机制、经过验证的可靠性以及完善的技术支持。

InterSystems IRIS 数据平台能在单个平台中、使用单个数据拷贝来处理当前的多种业务需求。这能帮助企业从购买和管理多个解决方案以及存储和映射多个数据拷贝的复杂情况中解脱出来。IRIS数据平台具有的功能特性包括：

- 多工作。InterSystems IRIS 数据平台能在同一个引擎中处理大规模的事务工作和分析工作，从而消除了将事务型数据经过移动/转换操作再用于分析以及部署多数据库等复杂的工作。InterSystems IRIS能在处理和检索事务数据的同时对实时数据或历史数据进行分析，从而提供对业务的更快响应和更好洞见。
- 多模型。InterSystems IRIS提供对于单一数据库的同一数据进行多种方式展示，而不是采用传统的单数据库的展示方法。它支持众多数据模型，包括关系型数据、键值数据、文档数据、对象数据和一系列schema-free的NoSQL数据。
- 互操作性。这个全面整合的开放分析平台能为数据库管理员（DBAs）提供简化的数据协调和业务流程编排能力，从而使连通数据源和业务流程变得更为简单。
- 分析。InterSystems IRIS具有实时分析和批量分析能力，包括分布式SQL处理、商业智能、预测模型运行引擎、对来自电子邮件和社交媒体的非结构化数据进行自然语言处理，同时还集成了例如Apache Spark等大数据处理框架。
- 企业级安全。IRIS集成Kerberos、轻量级目录访问协议（LDAP）、密匙管理互操作协议（KMIP）和基于角色的权限控制，能加密中转数据和静置数据，提供坚如磐石的安全。

图2 InterSystems IRIS 数据平台



来源：企业策略集团

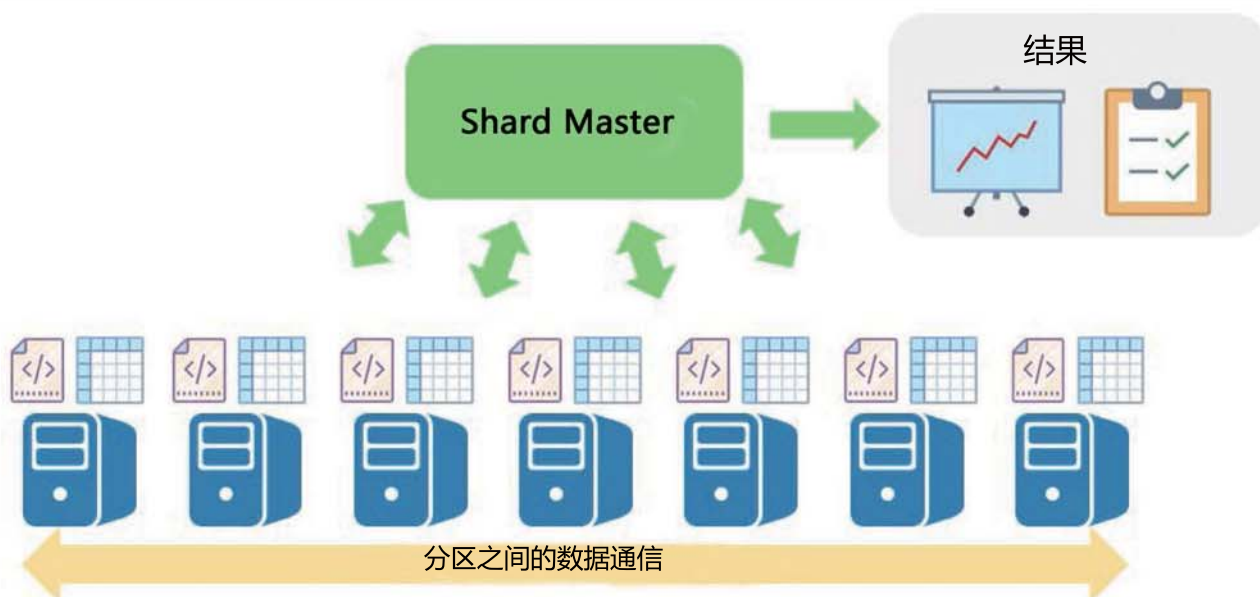
InterSystems IRIS架构

InterSystems IRIS能在快速获取和存储事务数据的同时，处理批量分析工作、实时分析工作以及符合ACID（原子性、一致性、隔离性和持久性）原则的数据分析工作；这能消除将实时数据传送到分析平台而导致的延迟。它能借助并行SQL处理实现纵向扩展，也能采用差异化方法实现横向扩展。

对于大型数据集和复杂查询，InterSystems IRIS采用一种差异化方法来进行横向扩展。通过使用商品服务器来分配查询处理，它充分利用了智能分区和分布式缓存。通过将数据库分区，工作和表格能被分配到多个节点；分区对应用程序透明可见，如果代码修改也只需做很少的改变。InterSystems IRIS使用一个智能Shard Master，它能生成用于并行执行的应用程序代码、汇集所有分区处理的结果并返回一个最终结果（具体请见图3）。每一个分区都知晓所有其他数据的位置；如果在一个分区内执行查询所需数据位于另一个分区内，那么它可以直接连通另一个分区、并可以只提取所需数据。如果采用其他解决方案，整个表格都会被传送到另一个节点，这会增加延迟和超时。智能分区管理功能使得InterSystems IRIS呈现极高性能，即使对于分割数据执行复杂连接查询任务时也一样。

每一个服务器的缓冲器都是独立的，因为分区创建了不相交的各部分。因此，横向扩展的服务器将能以线性方式增加服务器集群的存储容量。所以，当部署规模合适时，InterSystems IRIS能达到与内存数据库接近的性能，但另一方面因为它将经常使用的数据和“较少使用”的数据分别存储于缓存器和磁盘，IRIS不需要将所有数据存入内存。

图3 横向扩展性



来源：企业策略集团

ESG实验室测试

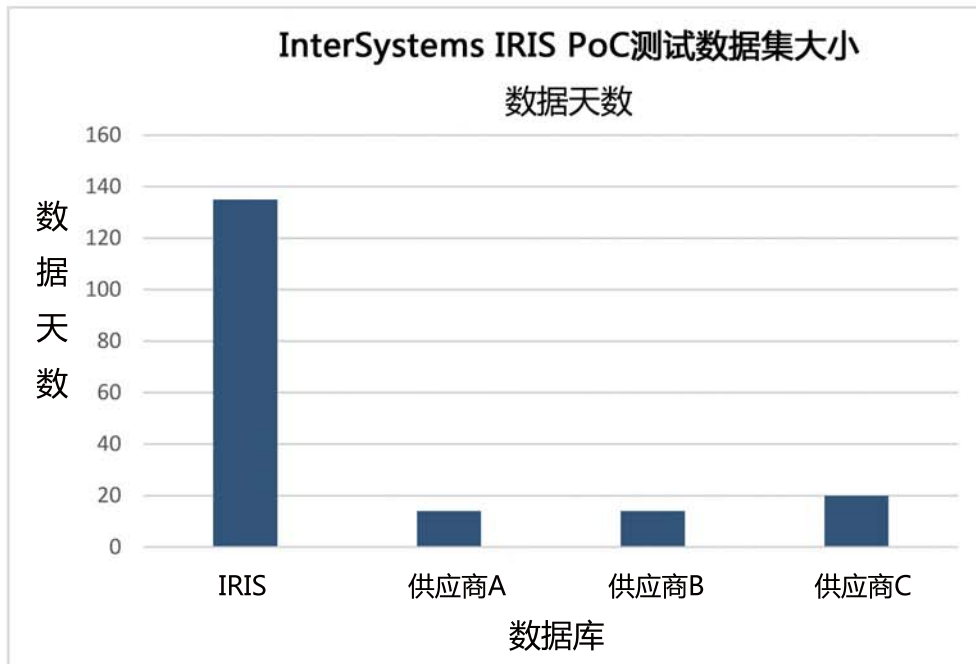
ESG实验室审核了一家金融行业公司的在数据库应用程序上进行性能概念验证性测试的结果。测试重点对比InterSystems IRIS和金融业界常用的多种其他数据库在海量金融票据中用于分析查询的性能。该应用覆盖全公司的多个关键业务；算法交易应用程序需要在严格性能约束条件下成功执行SQL查询，而来自公司各个部门的大量用户需要点对点地访问查询结果。毫秒级别的性能表现非常关键；如果不能在所需性能服务水平协议（SLAs）内完成查询，算法交易将无法执行而公司将不能成功实施其交易策略。

测试包括InterSystems IRIS和其他三种常用的数据库系统。表1显示，相比于竞争产品，配置InterSystems IRIS所需电脑架构更少（包括更少的虚拟机[VMs]、处理器核心和运行内存[RAM]），但IRIS却能对更多数据运行查询算法（以表1和图4中的数据天数、数据大小[GBs]和数据行数来衡量）。

表1 测试环境配置

数据库	所需配置	数据天数	数据集大小(GB)	数据集大小(行数)
InterSystems IRIS	4 VMs: 8 核, 96GB RAM, 400GB本地存储区	135	320	2,000,000
供应商A	8 VMs: 16 核, 256GB RAM, 400GB 本地存储区	14	33	207,400
供应商B	8 VMs: 16 核, 256GB RAM, 400GB 本地存储区	14	33	207,400
供应商C	3个专用物理服务器，每个服务器具备24核和256GB RAM，2TB 共享的存储区域网[SAN] 存储区	20	47	296,200

图4 数据集大小对比

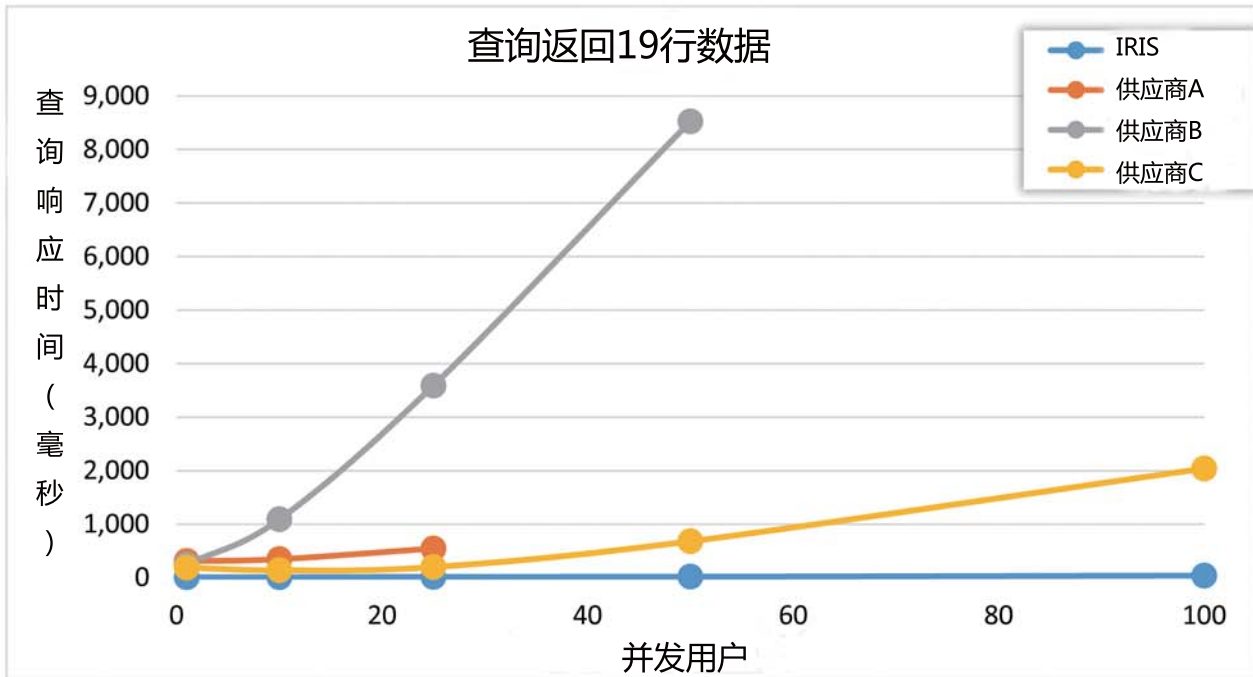


来源：企业策略集团

我们一共测试了5项查询任务。所有任务都是5到6个表格之间的复杂连接且在WHERE子句中包含8到10个断言，而查询结果数据集包含19到6504行数据、每一行包含数百字段。在并发用户数量增加时（1人、10人、25人、50人和100人），我们测量了每次查询执行的响应时间。

图5 展示了查询任务1的结果，查询结果返回了19行数据。通过测量在并发用户数量增加时的每次查询完成所需时间（毫秒）来检验数据库系统的性能。

图5 并发用户数量增加情况下，执行查询任务1的性能表现



来源：企业策略集团

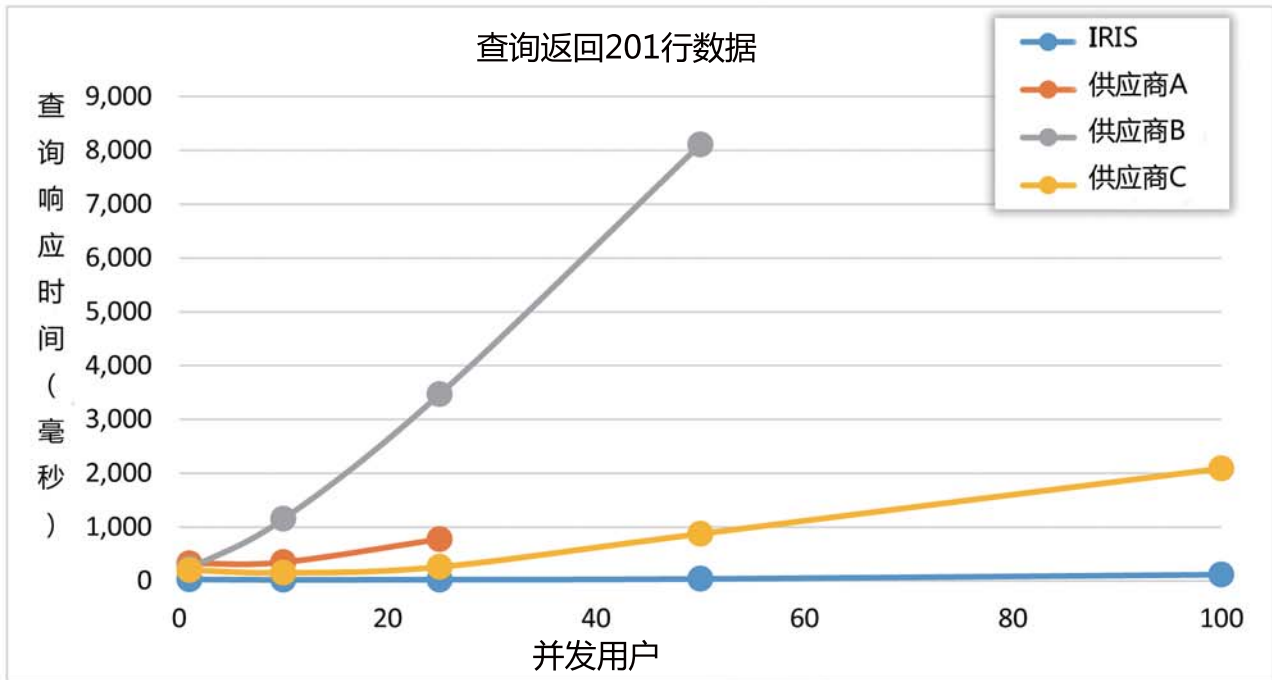
这些数字意味着什么

- InterSystems IRIS的运行需要更少基础架构，能比其他供应商数据库处理多5到8倍的数据。
- 在所有5次查询中，我们只描绘了没有发生错误的结果。供应商A只具有25个并发用户的数据，因为当并发用户数量达到50和100后、查询出现了错误；供应商B只具有25和50个并发用户的数据。这些解决方案不能处理大规模数据，这限制了其有用性。
- 对于查询任务1，在所有并发用户数量水平上，InterSystems IRIS的性能都显著快于其他数据库。InterSystems IRIS性能从8毫秒到38.97毫秒不等，在并发用户数量增至100人时只增加了4.87倍。
- 在单个用户使用的情况下，供应商A的响应时间是310毫秒（比InterSystems IRIS慢37倍）、供应商B响应时间是245毫秒（比InterSystems IRIS慢30倍）而供应商C则是188毫秒（比InterSystems IRIS慢22倍）。
- 供应商C 具有第二好的性能。在50人同时使用时，供应商C的响应时间比InterSystems IRIS慢35倍；而在100人同时使用时，其响应时间比InterSystems IRIS慢51倍。

查询任务2-4的测试获得了类似结果，InterSystems IRIS持续表现出更快的性能。

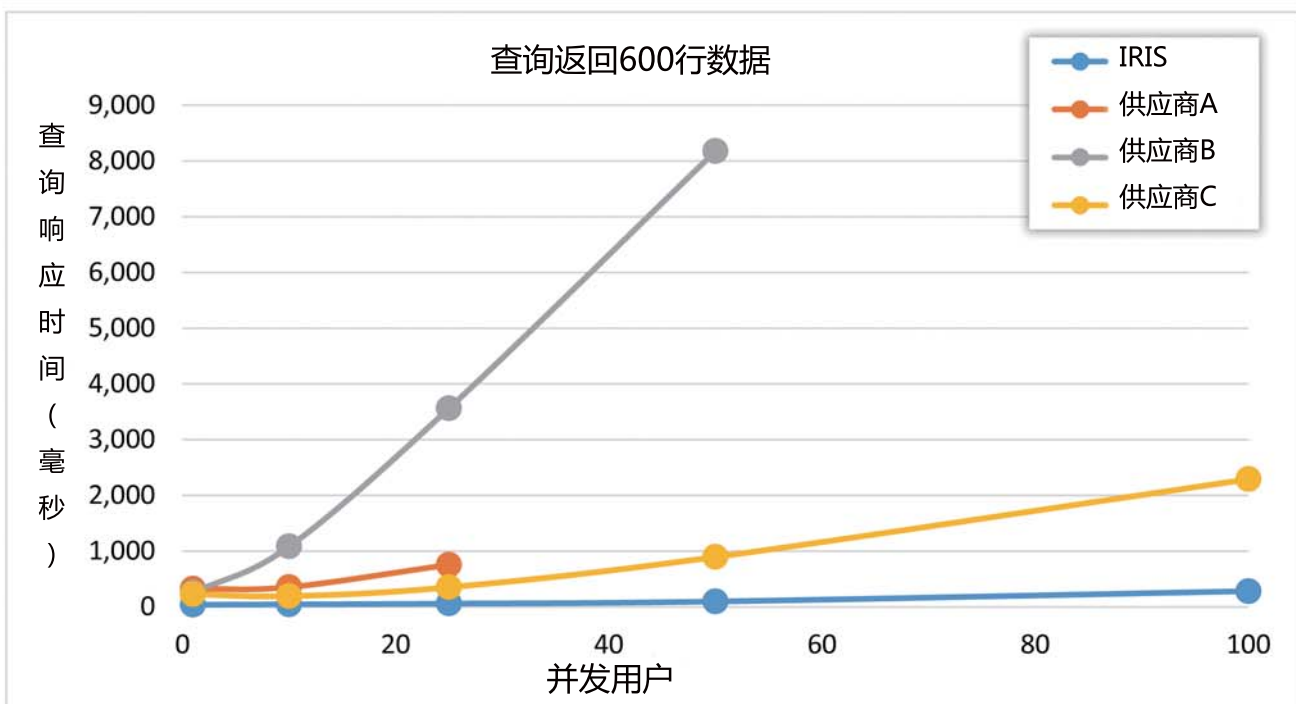
查询任务2返回201行数据，查询任务3返回600行数据而查询任务4返回984行数据；图6、7和8显示了这些结果。

图6 并发用户数量增加情况下，执行查询任务2的性能表现



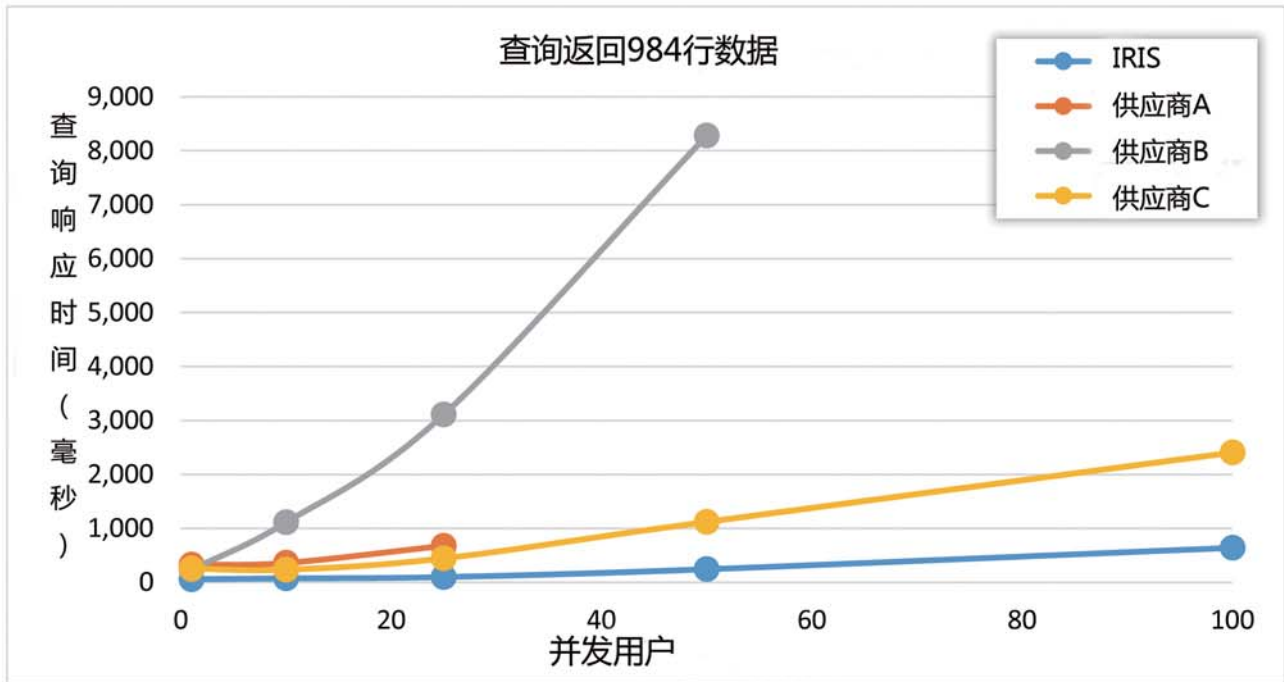
来源：企业策略集团

图7 并发用户数量增加情况下，执行查询任务3的性能表现



来源：企业策略集团

图8 并发用户数量增加情况下，执行查询任务4的性能表现



来源：企业策略集团

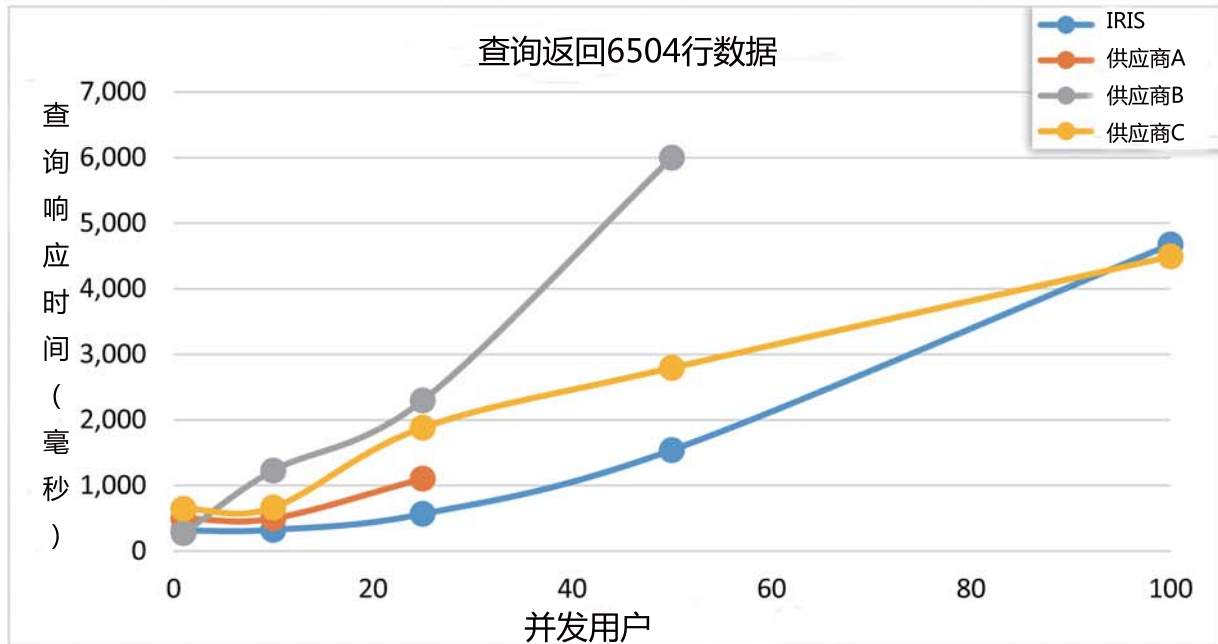
这些数字意味着什么

跟我们预计的一样，随着查询返回更多行数的数据，查询会花费更多时间。而在执行这三项查询任务时：

- 在单个用户使用的情况下，其他供应商数据库的性能比InterSystems IRIS慢3-10倍。
- 在10个并发用户使用，其他供应商数据库的性能比InterSystems IRIS慢2-59倍。
- 在25个并发用户使用，其他供应商数据库的性能比InterSystems IRIS慢3-147倍。
- 在50个并发用户使用，其他供应商数据库的性能比InterSystems IRIS慢3-218倍。
- 只有供应商C的数据库能支持100个用户同时使用，它的性能比InterSystems IRIS慢2-17倍。

图9呈现了查询任务5的结果，这次查询返回了6504行数据。在执行这个更为复杂的查询任务时，并行用户数量增加对性能的影响更大。

图9 并发用户数量增加情况下，执行查询任务5的性能表现



来源：企业策略集团

这些数字意味着什么

- 相比于其他供应商的数据库系统，InterSystems IRIS的运行需要更少基础架构，也能访问更多天数的数据。
- InterSystems IRIS的性能从单个用户使用时的312毫秒到100个用户同时使用时的4673毫秒，增加了14倍。
- 在100个并发用户使用，供应商C的表现（4492毫秒）略快于InterSystems IRIS（4673毫秒）。

为什么这很重要

根据ESG的调查，企业在接下来的一年内对技术进行投入的前三项驱动因素，其中的两项是降低成本和提升商业智能对于实时数据的分析能力以及客户自定义洞见的需求²。唯一位于这两者之前的因素是加强网络安全。与分析相关的因素成为企业优先事项毫不出人意料-----通过数据分析获取商业洞见能提升决策能力。你能访问的数据越多、你进行的查询越快返回结果，你就能做出更好的决策。

ESG实验室验证了，InterSystems IRIS对比于其他接受测试的数据库提供了持续稳定的高性能，它使用更少基础架构却能访问更多数据。虽然另一个供应商的数据库在一个单一数据点上表现出略好的性能（查询任务5，100个并发用户），考虑到相关成本和负担，添加另一个独立的解决方案很可能并不值得。InterSystems IRIS在所有查询任务中都表现出更好性能、更容易的扩展性、占用更少基础架构且没有出现任务失败。通过访问更多数据和更快返回结果，InterSystems IRIS能提供更好的商业洞见。

² 来源：ESG管理者调研结果，2018支出意向调研，2017年12月。

更重要的事实

充分利用信息能使企业获得竞争优势。金融业、医疗行业、制造业、电子通讯业和其他所有行业的公司都明白分析发展趋势有助于更好地做出决策。在你已经增加了蓝色新产品的生产量后，最终分析结果才揭示客户更喜欢红色新产品，这会让你感受到挫败。企业更多和更快地理解客户、供货商和市场变化趋势，它们就能更好地做出业务驱动的决策。挑战在于从尽可能多的数据中尽可能快地获得洞见，并将其迅速转化为实际行动。

InterSystems IRIS是一种具有良好成本效益的统一平台，能帮助企业通过构建数据驱动的、实时业务处理商业智能来改善决策能力。InterSystems IRIS是一种创新型知识平台，能通过避免使用多个事务数据库和分析数据库来加速基于实时数据的分析性查询。企业能使用同一个引擎运行事务型工作和分析型工作，无需传输、映射或转换数据；数据能同时以多种方式呈现，用户无需部署多个相互孤立的解决方案，从而获得更高性能。

ESG实验室对比验证了InterSystems IRIS相比于多个常用数据库解决方案的性能。结果证明InterSystems IRIS具有更快性能、能利用更多数据且只需要使用更少的基础架构---它能让企业以更低支出获得更快、更好的洞见。通过获取更多、更深入的洞见，企业能增加收入、改善客户体验、创造定制化产品和服务、提升接受度、降低风险、使操作更为高效和降低成本。如果你所在企业希望获得上述益处，我们推荐你考虑InterSystems IRIS。

所有商标的所有权归各自公司。本出版物包含的信息都来自企业策略集团 (ESG) 认为可靠的来源，但ESG不保证其可靠性。出版物可能包含ESG的观点，这些观点可能随时间进展而改变。本出版物版权归企业策略集团股份有限公司所有。未经企业策略集团明确同意，任何对本出版物的复制或发行，无论整体或是部分章节，无论纸质版或是电子版，或向未获授权接受的人提供本出版物，都违反美国版权法并将接受民事损害诉讼以及可能适用的刑事诉讼。如果有任何疑问，请联系ESG客户关系部 508.482.0188。

ESG验证报告的目标是向IT专业人士介绍可用于所有类型和规模的企业各种信息技术解决方案。ESG验证报告旨在提供新兴技术相关信息，不能代替做出购买决策之前的评估流程。我们的目标是探索一部分拥有具更高价值的特性和功能的IT解决方案，展示如何使用它们来解决客户的现实问题，并识别出需要改进之处。ESG验证团队专家的第三方见解是基于我们自行进行的测试以及对在生产环境中使用这些产品的客户的访谈。