



现代数据环境中的 3 大转变

及其对 IT 负责人的意义

目录

问题依旧，却更难解决	3
1. 用“管道”思维代替“数据桶”思维.....	4
2. 使用基于需求的数据停放区.....	5
3. 从数据保护者转变为数据指导者	7
问题更难解决，但并非无计可施	8

问题依旧，却更难解决

为组织提供可靠的数据来协助制定更明智的决策，这是几十年来未曾根本改变的任务。尽管 IT 行业涌现出大量的技术进步和新方法，但是负责管理数据基础架构的 IT 组织今天仍然具有相同的总体使命：在数据创建时就能移动数据，使决策制定者在需要时能够访问和理解这些数据。

虽然目标一贯如此，但要想成功在企业内创建和维护一个分析事实数据源，必须克服异常艰难的阻碍。

近年来出现了大量的新数据源，它们生成了前所未有的庞大输出量，并且输出内容通常结构化程度不高甚至没有任何结构，这可能是现代数据环境中存在的最大障碍。从点击流、服务器日志和社交媒体源到机器和传感器读数，来自这些渠道的巨量数据简直令人不知所措。从经济和性能视角来看，传统企业数据仓库 (EDW) 根本无法满足这一波数据浪潮带来的需求。

这促使人们全面重新思考数据捕捉和分析策略，最终开发出新一代数据存储解决方案，这种解决方案致力于实现无架构数据捕捉、硬件可扩展性，以及将计算能力提升到接近（如果无法超越的话）数据存储本身的层级。

根据关系数据标准来看，这些更前沿的非关系解决方案尚不成熟，但是它们在最近几年内发展势头迅猛并快速成熟起来，现在可以为世界上最大、最复杂的一些企业提供支持。这一成就主要是通过补充现有企业数据仓库基础架构来实现的，但对 IT 而言，却是要管理一个更为复杂的数据生态系统。

IT 团队除了必须在当前任务中克服重重障碍，维护一个健康的数据环境之外，他们还需要确保用户可以随时从云应用程序中获取数据。许多组织都将 Google Analytics、Salesforce、Netsuite、Zendesk 等云解决方案用作其基础架构的核心部分。它们生成的数据对于组织报告至关重要。整合来自这些云解决方案的数据并使它们可供公司访问，已然成了一个对 IT 的标准要求。

当传统 EDW 不再是唯一的数据目的地时，“何时、何地、如何以及是否”将云应用程序数据引入到企业数据环境中，成了持续热议的问题。

最后，由于自助式分析已成为各种规模组织的惯常做法，越来越多的非技术用户（未接受过正式的 IT/数据培训）将进行数据发现和报告，甚至包括数据准备和高级分析。采取此举措的企业通常会看到这样的结果：极大地减轻乃至完全去除 IT 团队开发分析工具的职责。对于采用数据驱动思维模式的组织而言，尽管这一转变对组织的整体成功至关重要，但也为 IT 团队带来了新的压力，因为他们需要提供更广泛的数据访问。此外，还需确保技术满足业务部门需求，以及 IT 对安全性和管控的需求。

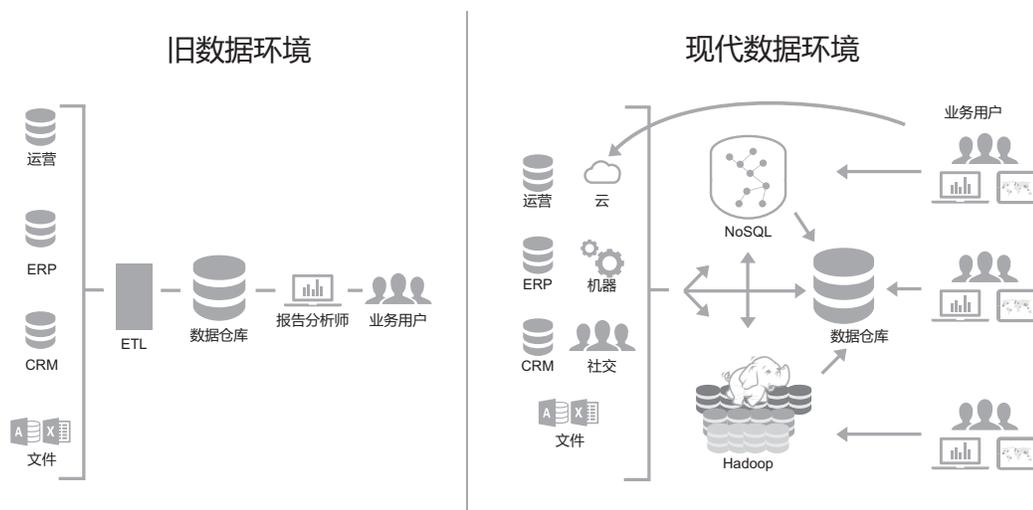
为应对这些新的挑战，许多 IT 组织都竞相采用新技术和新方法，但遗憾的是，他们并没看到，面对这些障碍，IT 团队实际上需要转变方式才能实现“从创建到使用”期间管理数据”这一目标。大数据解决方案、云数据集成以及自助式分析都用来解决更大的技术难题，但为了能够在组织中有效部署这些技术方案，IT 需要改变工作方法。

本白皮书旨在简要介绍与现代数据环境有关的 3 大思维转变，IT 负责人需要了解这些转变，才能支持组织制定数据驱动决策。

1. 用“管道”思维代替“数据桶”思维

企业数据仓库并未被弃用，只是多了一些朋友。从 Hadoop 群集到 NoSQL 数据库，关系 EDW 不再是获得正式批准的唯一数据存储位置。

这意味着，数据移动过程不再需要围绕单一位置集中执行。事实上，如今的现代数据平台更像是相互连接的神经网络，它们将组织“数据大脑”的左半球和右半球连接在一起。将此平台与当今无处不在的云解决方案相结合，基础架构和服务在几个小时内便可启动管道/ETL 项目，您只需掌握在组织内移动数据的方法 - 这是史无前例的事情。



遗憾的是，许多 IT 团队大多错过了这个机会，因为许多组织的数据管理手册都是使用去年的“万能单一数据桶”思维模式编写的。

从“数据桶”思维方式转变为“管道”思维方式，其中的诀窍是接受这样一个事实：并非组织中的所有数据问题都能从任何一个数据源中找到答案。深入探索直接从消息服务器日志中擦掉的为期五年的电子邮件资产，这种对数据粒度的需求是 EDW 可能无法满足的。而 Hadoop 部署也可能无法根据实时分析需要对销售数据提供亚秒级的查询响应。在现实世界中，许多业务问题都需要以精细的粒度快速查询一个或多个数据源，而为了完成一个数据项目，则需要不同的时段和不同的演替阶段进行这些操作。

鉴于上述要求，IT 的问题不再是“此数据进入到哪个数据桶”，而是“需要哪个数据阶段？”以及“我如何让用户轻松地在这两者之间进行移动？”。

从根本上来讲，这是关于管道的讨论，因为任何给定数据源的实际位置可能会随着项目进入不同的完成阶段而发生变化。“管道”思维模式亦从传统数据集成领域延伸到组织 workflow 领域。

业务用户如何找到其数据答案？大多数 IT 团队都知道，当前在整个组织内的电子表格中保存了许多未获批准的独立数据存储库。但是，对于这些孤立的、未经监管的本地文件数据的蔓延，IT 团队感到无力控制和管理。位于业务端的员工通常认为，围绕企业数据环境开展工作是唯一切实可行的解决方案 - 这就是他们的工作。

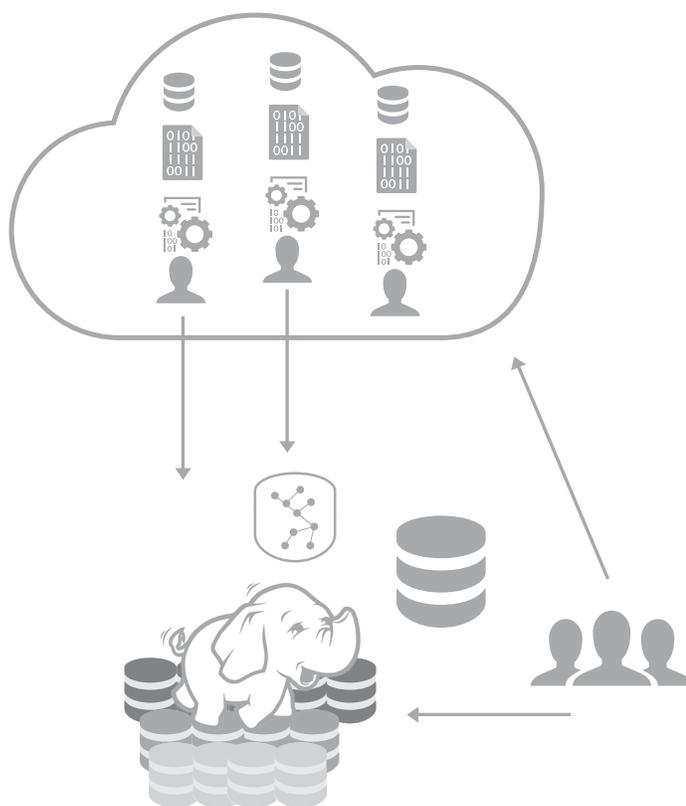
当 IT 团队全面了解业务用户查找、清理、分析和呈现数据的方式并为这一过程提供支持（同时就管控提供辅导）时，才能转变到“管道”思维模式。一旦技术团队坚定地认为，在传统批处理 IT 层面以及业务用户的即兴查询层面，数据移动比加固数据存储库更为重要，就会实现关键转变。

“ 从全球角度来看，掌握所有这些信息并能够从中发现见解越发显得至关重要。现在，我们使用我们的逻辑数据仓库和 Tableau 制作出了仪表盘，大家可以共同协作，所有这些信息也都可供访问了。

— TIM NALL, BROWN-FORMAN 首席信息官

2. 使用基于需求的数据停放区

云应用程序数据不能始终直接停放在企业数据仓库中。决策应始终反映出组织的需求。或者，更清楚地说，决策应反映出需要采取哪些措施才能使数据对组织有用。



IT 团队（甚至那些已经接受“不存在万能单一数据桶”的团队）有个常见的误解，即认为 Web 应用程序数据应自动直接整合到 EDW 中。这在表面上合乎情理，因为许多成功的企业云应用程序是销售、营销和支持基础设施的必要组成部分。这些数据具有较高的指导意义、经常被访问，且通常具有优良结构。将这些资产放置在 EDW 中的投资报酬率（ROI）（其中，每字节的价格是最高的）貌似比较合理。

但是，IT 团队需要思考的关键问题是：

- 是否可以直接从云端分析数据？换句话说，数据是否整洁，以及我们应不应该费力耗神地移动数据？
- 相关云数据的真正价值是否必须在与其他数据结合使用后才能实现？

根据这些问题的答案，IT 团队将其云数据资产放在哪里（如果他们选择将这些数据资产直接放在本地环境）可能会因组织和应用而异。

负责一定规模的 Salesforce.com 部署的任何 IT 团队都会立即告诉您，允许将没有删除重复数据的机会记录存放到 EDW 中，将会后患无穷。从不精确的佣金支出到大幅过度预计的收益数字，当处理未妥善清理的销售数据时，这些问题仅是冰山一角。

“ 最终用户获得数据后，便能快速整合自己的基础报告。最终用户是产品和应用程序变化的最直接感受者，他们可以快速发现可能需要调整的地方。

— SHARON GRAVES, GODADDY 企业数据拓展专员

同样，如果记录的准确性不能让人信服，那来自 Google Analytics、Eloqua 和 Marketo 等云数据源的网站和营销数据可能会干扰组织跟踪销售线索流程和计算客户获取成本的能力。

当必须合并数据源时，此准确性问题会变得异常糟糕。当组织需要描绘完整的客户漏斗（从第一次网站点击一直到销售决策）时，通常会遇到上述情况。虽然几乎所有 IT 团队都知道整洁数据（无论来自云或其他数据源）的重要性，但他们通常不会思考当从 Web 应用程序中下载后，这些数据会首先放在哪里。

如果某个 IT 组织拥有一个以多数据桶/管道为核心的数据环境，则当该组织将云数据集成带到本地环境时，会根据数据资产对于企业的价值来评估云数据集成。当某个云数据源相对整洁并能够在迁移后为企业提供极高的价值时，建议 IT 团队将该数据源停放在可以最快速访问的地方（关系 EDW）。

但是，当处理在被企业广泛信任之前需要繁重和/或复杂处理的 Web 应用程序数据时，IT 团队可以利用高计算速度、低字节成本的环境，例如 Hadoop 环境。借助上述第二种方法，资源可以在清理和转换数据时发挥最大作用，而不会影响 EDW 的速度。随后，处理业务的 IT 团队可以决定是否应该将清理后的数据移动到 EDW 中和/或在较大的环境中直接访问这些数据。

最后，请牢记一点：有时可能根本不需要移动云应用程序数据。许多 Web 应用程序正在为其后端存储库提供简便的接入点，这意味着业务用户可以使用自助式报告和分析工具，对实时数据执行各自的调查工作。针对这些情形，部分 IT 组织仍然希望拥有一个中间授权和管控层。而对于这些组织，部分企业友好型自助式分析解决方案还考虑到了代理连接情形。这样便可解决 IT 企业的所有需求，从基本的用户访问到高度相关的业务逻辑。鉴于此实际情况，与云应用程序数据集成相关的首要问题不是“何处”，而是“应该”。究竟是否迫切需要移动数据（以进行清理和获得其他价值等），还是可以将数据留在原处并在业务用户级别进行集成？

“ 我们之所以选择与 Tableau 合作，是因为 Tableau 实力雄厚并设有清晰的未来发展蓝图。这份蓝图与我的目标不谋而合...它帮助我按我所需、所想的
速度向前发展。

— STEVEN JOHN, AMERIPRIDE 首席信息官

3. 从数据保护者转变为数据指导者

应将公司内的自助式分析举措视为面向 IT 的握手提议，而非有敌意的接管威胁。那些首先伸出友谊之手的 IT 团队将为赢得一切胜利做好准备。

不仅现代数据环境进行了扩展，以便不只包含关系数据存储；分析环境同样得到了改善，以便包括每天都为业务用户提供巨大价值的工具。此做法与传统流程背道而驰。在传统流程中，所有分析数据都由经过严格培训的少数尖端人才进行筛选。

有效的合作关系

IT 角色

安全
数据架构
可扩展性
培训
运营中心
实现 Intranet

业务角色

具有创造性的分析工作
推动数据获取
分享专业知识
在需要时寻求帮助
传播理念
促进行动

实现



执行

结果，具有前瞻性思维的 IT 团队正在转变为其更大型组织内的数据和分析思维引领者，并且正在重新定义他们与业务团队的合作方式，从而解决数据难题。从“数据保护者”到“数据指导者”，这一 IT 思维模式的转变可能是现代数据环境中最重要的转变。此转变是决定组织能否成为真正的数据驱动型企业的一个关键因素。

重新定义 IT 与企业之间的这种关系，意味着将默认的数据管控态度从“禁止”调整为“允许”。按照传统惯例，系统会自动拒绝某个业务用户对数据的访问，除非该用户获得特批。这个新模式假设公司任何部门的业务用户都可以访问任何数据资产，只要其不违反合规性条例即可。

此态度转变的力量是巨大的。业务用户开始寻找新数据来解决新问题，这就是此转变所带来的影响。他们是否愿意这样做，取决于他们认为访问新数据将耗费多少工作量。当配备了一个易于使用的自助式分析解决方案，并因为得知 IT 团队支持广泛的数据访问而深受鼓舞时，业务部门的人员可能会成为变革的促进者。这些促进者标志着企业已经成功迈入到成为数据驱动型公司的早期阶段。

只是单单开放数据访问和分发新软件并不能解决更大的问题：真正帮助企业使用数据来制定更明智的决策。已经成功开展合作、使公司过渡到数据驱动型决策引擎的 IT 组织已实现上述目标，因为它们欣然接受了“数据指导者”这个角色。之前控制分析流程的那些 IT 团队成员通常会成为公司其他员工的指导者。他们有能力来培训其他员工学习关键功能，例如正确进行数据分析和报告等。

通过这种方式来培训业务用户，可有效将整个组织的分析报告能力提升一个量级。此外，这还可以帮助组织在自助式分析解决方案方面更快地获得更高的投资回报。然而最重要的是，使用数据来更好地全面了解业务已然成为整个组织运营精髓的一部分。

问题更难解决，但并非无计可施

从处理全新的数据存储和处理解决方案，到了解在何时、何地以及是否应该将云数据纳入到生态系统中，负责管理组织数据流的 IT 团队的责任突然变得更加艰巨。此外，自助式分析举措以及这些转变共同实现了更大的成效 - 公司处理数据的方式发生了彻底的转变。

组织需要 IT 来指导它们完成这一转变。IT 团队之所以会感觉到不确定，大多是因为他们知道问题已不在于是否将新技术应用到现有业务流程中。大数据、云技术和自助式分析模式表明，IT 团队需要从根本上改变方式，才能打造现代数据环境。

对于那些已经开始重新评估数据和分析思维模式，以适应这些转变的组织而言，障碍与机会并存。转变对于业务用户和 IT 而言都是一项艰巨的任务。当重新定义这两者之间的关系时，都会从每个阵营中传出反对的声音。但是，我们也会听到赞成的声音。

当包围着组织数据访问的坚固壁垒和流程发生改变时，IT 团队将在公司的每个角落都找到意想不到的分析工具拥护者，他们渴望成为全新数据驱动型公司的领军人物。最明智的 IT 负责人会将这些拥护者与经验丰富的数据专家配对，由专家提供指导。拥护者将通过不断演变的数据生态系统来了解他们的行进路线，IT 团队将了解到人们正在尝试解决的业务问题。对于需要依靠现代数据环境的可信度和准确性来制定决策的公司而言，这些前期步骤为在公司内培养分析和数据探索文化奠定了坚实的基础。

关于 Tableau

Tableau 致力于帮助人们查看并理解数据，不论数据规模有多大、来自什么渠道或存储在何种数据库中，用户都能轻松驾驭。从 PC 到 iPad，您可以使用各种设备对自己的数据实现快速连接、混合和可视化，感受行云流水般的使用体验。用户无需掌握编程技能，就能创建和发布可自动更新数据的营销仪表盘，与同事、团队、高层领导、合作伙伴或客户分享实时的分析洞见。

其他资源

获取免费试用版

[+ Tableau Online](#)

[+ Tableau Desktop](#)

相关白皮书

[现代商业智能和分析评估指南](#)

[受规管的大规模自助式分析](#)

[重新定义 IT 在现代商业智能领域中的角色](#)

[适用于企业的 Tableau：IT 概述](#)

