

管理数据，满足合规性 & 提高生产率

探索发掘

- 在统一的元数据储存库内，轻松追踪、分类并且定位数据

优化

- 对工作负载的瞬时洞察和优化指导，取得最佳Hadoop成果

审计

- 对Hadoop全面的审计历史记录保留在同一个地方

追踪

- 将数据的上下游继承关系可视化，以验证可靠性

保护

- 通过性能优化的加密以及密钥管理来保护敏感数据的安全

管理生命周期

- 通过集成元数据策略，来定义并自动化复杂的数据生命周期活动

“当客户信任我们，把他们的钱汇给他们的亲朋好友之时，我们就是顾客金融信息的监护人。我们需要确保自己合乎规定，有正确的监控和审计机制，而这正是Cloudera Navigator可以发挥作用的地方”。

西联汇款

Hadoop的集成数据管理与监管

Cloudera Navigator

大数据对于企业来说，意味着一个在战略、运营和客户互动上实现信息驱动的一个重大机遇：拥有存储与分析任何种类和规模数据的能力，并且将其开放给更多用户和分析工具。但是，这个机会也会带来数据管理上的挑战：大量的业务用户想自助访问可发掘数据；管理员需要知道数据是怎么被用来优化分析性能的；安全团队需要看见数据的访问方式以及它们是怎么满足合规性的。一个企业大数据平台必须有能力在整个机构内解决数据管理以及合规性需求，但同时不能牺牲大数据本身的灵活性和优势。

Cloudera Navigator是唯一为大数据和Apache Hadoop提供集成数据管理和监管的解决方案。Cloudera Navigator作为Cloudera 企业版的核心，获得了大量用户的信任，获得了对数据前所未有的可视性。它提供必要的审计和数据保护，来满足最严苛的规范限制。有了唯一的全面的元数据基础作支撑，Cloudera Navigator自动将各个平台上的所有技术元数据，以及企业内用户定义的业务元数据整合成一个单一的、可搜索的储存库。至此，Cloudera Navigator可以从四个基础方面提供有效的数据管理：

自助数据发掘 & 分析

企业用户可以毫不费力地找到并且信任最重要的数据。

- 通过直观性的全文本搜索和SQL访问，在唯一的统一元数据储存库中发掘和探索数据。
- 基于业务情境和分类来定位数据集，与自动化技术情境相结合 – 使其容易找到类似的、相关的数据。

活跃数据的优化

数据库管理员可以迅速获取洞察来优化最关键的工作负载。

- 快速分析现有的SQL日志获取全面的视角，包括哪些查询最为关键、哪些数据被最为经常的访问、以及数据是如何来被使用的。
- 通过智能优化指导来提高Hadoop的性能和效率。
- 通过可兼容的识别来减少工作负载的研发时间，从而在Hadoop上快速取得成功。

合规性监管 & 保护

安全团队可以追踪、了解并且保护对敏感数据的访问。

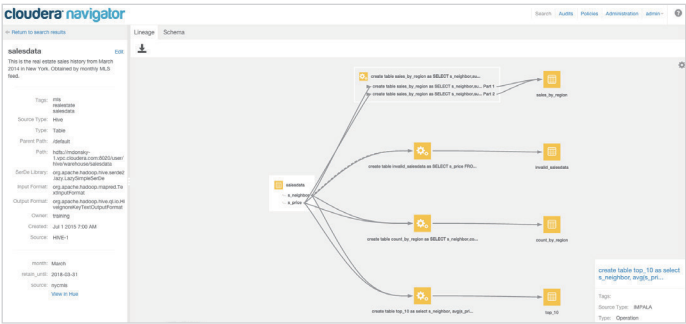
- 自动保存一个完整的审计记录，追踪每一次访问尝试，细化到用户ID、IP地址以及完整的查询文本。
- 追踪数据的使用方式以及在列层级和视觉沿袭上的变化，从而快速确认一个数据集的来源以及它对下游工件的影响。
- 通过高性能加密和Navigator 加密以及Navigator 密钥受托者的密钥管理来保护所有数据
- 与领先企业的元数据、沿袭、以及SIEM应用进行集成，可即用。

Hadoop 规模数据生命周期自动化

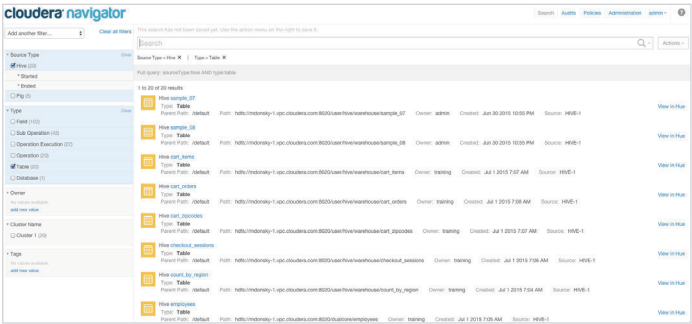
数据管家可以有效管理并且实施关键的生命周期政策，而且免于风险

- 自动化关键管理以及综合处理活动 - 比如元数据的分类，数据归档以及保存，甚至调用合作伙伴产品作为额外的数据准备及转化 – 通过灵活的策略引擎
- 通过唯一内置的备份及灾难恢复来确保业务连续性
- 通过无缝集成合作伙伴工具，来管理不只是Hadoop的数据生命周期

| 特色 | 功能 | 益处 |
|------------|---|---|
| 探索 & 挖掘 | | |
| 统一的技术性元数据 | 将所有来自HCatalog、HDFS、Sqoop等平台的元数据集中于同一个可搜索的界面 | 保持对不断增长的数据的控制，用以持续精准的提取 |
| 全面的业务元数据 | 带有定制化业务情境、标签以及关键/值组的文件、图表以及个人列表 | 面向不同的相关用户（包括业务分析师、数据科学家以及数据管家）将数据进行有意义的分类，释放数据的价值 |
| 元数据搜索 | 跨所有业务和技术元数据搜索寻找资产（比如时间超过7天以上的数据；特定用户创建的表格；包含敏感信息的文件等） | 通过一个共有的界面来找到多样化的数据 |
| 审计 & 访问控制 | | |
| 审计配置 | HDFS、Impala、HBase以及Sentry的自动配置审计跟踪；设置阈值 | 通过几下简单的点击，确保所有必要审计数据都被成功捕捉 |
| 审计仪表板 | 通过一个简单的、可查询的界面，将数据访问可视化并提供摘要 | 一个单一的平台来快速识别异常值和安全漏洞 |
| 第三方集成 | 集成审计信息，用于全球安全信息和事件管理(SIEM)系统 | 将Hadoop的数据集成到整个基础架构范围的报告和流程 |
| 数据继承关系 | | |
| 直观的可视化 | 在一个浅显易懂的图形文件、表单个列级查看数据继承关系，掌握上游数据源以及下游走向 | 快速识别数据集的来源，及其对下游操作的影响 |
| 加密 & 密钥受托者 | | |
| 透明的加密 | 保护所有数据却毫不影响性能 | 即便是最敏感的数据也可用于分析，免于风险 |
| 企业密钥管理 | 通过一个单独的、集中化解决方案来管理加密密钥和安全资产 | 与现有HSM以及密钥管理策略进行集成，以满足合规性需求 |
| 优化器 | | |
| 活跃数据优化 | 分析所有数据管理系统的工作负载，以优化峰值性能及合理化工作负载 | 将正确的工作负载在正确的时间放在正确的地方，获得最佳Hadoop成果 |
| 元数据策略 | | |
| 元数据策略管理 | 基于到达或预定的间隔，引发特定数据集的执行动作（比如元数据的自动归类） | 轻松设置，监控并且实施数据管理策略，并且与第三方工具集成 |



查看上游和下游记录



搜索指定条目的元数据