

## Apache Hadoop

# 推动物联网 (IoT) 的发展：四个客户使用案例



## 目 录

概述	3
物联网（IoT）数据管理平台	3
案例研究 1：降低互联汽车的停修时间	4
案例研究 2：提供基于使用行为而定的保险和车载资讯系统	5
案例研究 3：在工业环境中实现预测性维护	6
案例研究 4：医疗保健和帕金森氏症研究的革命性突破	7
Cloudera 公司简介	8

## 概述

研究预测，到 2020 年，全球通过物联网（IoT）连接的设备将达到 300 亿台，<sup>1</sup> 使得物联网（IoT）有可能成为近年来最具颠覆性的技术进步之一。企业机构正在开始利用物联网来推动其关键业务目标的实现，包括监控和提升产品性能、提高内部效率、改善客户体验、开发新的营收和基于使用量的业务模式。

然而，物联网还将推动数据量的爆炸性增长，这些数据需要被处理、存储、管理、分析和（在某些情况下需要实时），从而实现业务价值。每分钟从众多数据来源将生成数以百万计的数据事件，物联网将创造出比大多数信息领导者目前所熟悉的大得多的数据量和数据种类，这就需要对其信息基础设施进行现代化改造来实现其价值。

要从物联网中获取价值，需要有效地管理不仅是静态数据而且有动态数据。事实上，物联网的成功部署将取决于企业机构对于所有这些快速变化的大批量数据中获得洞察力的能力。而要取得实时洞察力意味着动态数据处理过程中（摄取、处理和服务）每一步都是至关重要的。

## 物联网数据管理平台

今天，全球领先的企业机构正在采用基于 Apache Hadoop™ 的 Cloudera 企业版作为用于存储、管理、处理以及更重要的是从所有物联网数据中驱动分析的数据管理和分析平台。

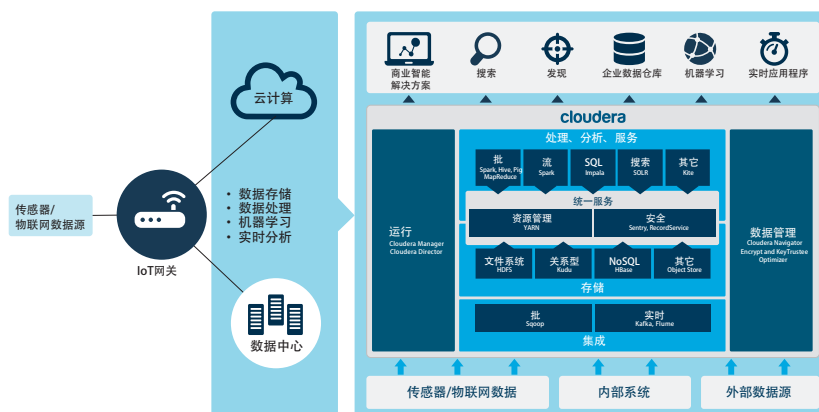


图 1: Cloudera 企业版 - 物联网数据管理平台

## 当今的世界都依靠数据在运行。这对您和您的业务意味着什么呢？

借助 Cloudera 企业版，企业机构可以轻松地将来自多个数据源的数据集成到一个单一、统一、安全的平台上 - 通过活动日志、事务数据、客户数据、外部数据等将物联网传感器数据流结合和关联在一起。任何类型的数据（无论是结构化、非结构化还是半结构化数据）都可以在不改变其格式的条件下加载到 Cloudera 企业版中，从而可以确保数据完整性，并提供完整的分析灵活性。

Cloudera 企业版提供了一个可扩展的具有弹性的企业级数据管理平台，通过该平台可以轻松获取、存储，处理和分析动态数据和静态数据。企业机构可以使用诸如 Apache Kafka 和 Apache Flume 等工具轻松地从各种传感器中获取批处理和流式传输数据。同时，Cloudera 企业版还支持使用 Apache Spark 和 Spark Streaming 的批处理和实时处理数据，由 Apache HBase 和 Apache Kudu 等存储选项支撑。此外，根据不同的使用案例，Cloudera 企业版提供了各种不同的分析选项，包括使用 Apache Impala（孵化中）等工具的 SQL 分析，机器学习库以及与领先的商业智能（BI）解决方案的集成。

<sup>1</sup> 来源：IDC，Worldwide Internet of Things Forecast Update，2016-2020 年。

此外，因为 Cloudera 企业版可以在工业标准硬件、设备或云计算环境中运行，每 TB 数据量的平均存储和处理成本比传统数据管理机制低十倍。

在汽车、制造、公用事业、工业自动化、保险、医疗保健、电信和科技等多个行业中，众多的领先企业机构已经采用 Cloudera 企业版作为他们物联网的数据管理平台，该平台每秒采集和处理数百万条事件，从而获得切实可行的业务洞察力。

事实上，Cloudera 正在为行业中最具吸引力的物联网使用案例提供着支撑，包括互联汽车和车载资讯系统、联网家庭的预测性维护、工业物联网、智能城市、基于使用行为而定的保险以及医疗保健物联网。下文将予以具体介绍。



在 Cloudera 数据中心基础之上构建的远程诊断物联网平台已经将车队维修的成本平均降低了 30% 至 40%。

## 案例研究 1: 使用物联网 (IoT) 降低互联汽车的停修时间

使用案例: 互联汽车 / 预测性维护

客户: 一家领先的汽车制造商

地理区域: 北美

### 概述

该客户是北美地区领先的汽车制造商之一，其 2015 年的营业收入超过 100 亿美元，提供各种类型的商用载重车辆、公共汽车、防务车辆和发动机等产品。

### 业务驱动因素

#### 车辆性能的实时视图

- 为了真正了解载重车辆性能以及提高车辆正常运行时间，该汽车制造商希望实时捕获和分析超过 18 万辆载重车辆的详细数据，包括从速度、制动和发动机性能到燃油经济性、地理位置、怠速时间、潜在故障等各种性能数据。
- 该汽车制造商还想了解其生产的载重车辆在各种条件下的性能，以及在发动机故障之前发现发动机故障，以尽量减少故障和车队停修时间。

### 解决方案

#### 基于 Cloudera 的实时远程诊断解决方案

- 该汽车制造商开发了一种远程诊断工具，利用 Cloudera 企业版实时收集其载重车辆上的车载资讯和地理位置数据。
- 安装在车辆上的各种传感器可以用于采集和传输各种数据，包括速度、发动机性能、燃油效率、冷却液温度和制动磨损等。这些数据以 15 至 60 秒的时间间隔进行采集，并将其传输至 Cloudera 企业版中以进行高级分析。
- 然后，工程师和分析人员使用各种分析工具和机器学习功能及早地发现发动机问题，预测维护问题，并在载重车辆故障之前解决问题。

### 影响

#### 提升客户体验以及减少车队停修时间

- **减少停修时间:** 该解决方案使车队所有者能够通过智能手机或平板电脑实时轻松地监控其车队的性能，预先安排维护计划，并减少高达 30% 的计划外维修和车辆停修时间。
- **降低维护成本:** 通过对其车辆的实时监控，车队所有者能够将每英里的车队维护成本降低高达 80% (从每英里 0.12 美元到 0.15 美元降低至每英里 0.03 美元)。
- **启用新服务:** 该汽车制造商现在可以向车队所有者提供分析即服务 (analytics as a service)，并计划建立一个在线门户以集成车载资讯系统数据、附加的 GPS 数据和零部件库存信息。此服务功能将使车队所有者能够快速找到具有其所需零部件库存的最近的经销商服务地点，以及拥有可用技术人员和维修间隔的服务地点。



## 案例研究 2：提供基于使用行为而定的保险和车载资讯系统

使用案例：汽车车载资讯系统 / 基于使用行为而定的保险

客户：一家大型的欧洲汽车保险公司

地理区域：欧洲

### 概述

一家大型的欧洲汽车保险公司使用 Cloudera 和车载资讯系统来实时收集、存储和分析客户车辆上安装的黑匣子设备的数据。该公司使用这些信息根据每一名驾驶员的驾驶模式调整保险费率和免赔额，在提高利润和市场份额的同时降低风险。

### 业务驱动因素

#### 提高竞争力，减少欺诈性理赔和提供增值服务

在激烈的竞争压力以及不断提升的每宗理赔的平均损失成本的情况下，汽车保险公司希望有效地推出新的服务项目，向他们的客户提供基于使用行为而定的保险（UBI）。此外，该公司需要一套基于车载资讯系统的解决方案，以帮助他们减少欺诈性理赔，并向客户提供各种新的增值服务，包括道路救援和事故援助、被盗车辆追踪、远程诊断燃料管理等。

这种新的业务洞察力已经显著减少各种欺诈性理赔，并将总体理赔数量减少了 30%。

### 解决方案

使用 Cloudera 企业版和英特尔体系架构推出了一套基于车载资讯系统使用的保险解决方案

- Cloudera 公司联合英特尔公司一起，帮助保险公司推出了一套基于车载资讯系统使用的保险，允许该保险公司可以个性化定制保险责任范围以及根据客户的实际驾驶习惯向客户收取保险费。
- 在投保客户的车辆上安装了一个黑匣子，可以持续记录相关的信息，例如 GPS 位置、驾驶速度、驾驶距离和时间，加速和制动类型（例如，快速或平滑）和转弯习惯。
- 使用 Cloudera 企业版来收集、存储和分析大量的车载资讯系统数据，实时从黑匣子设备以及其他外部数据源收集，例如天气数据、路况报导和与其他车辆之间的碰撞事故。这些数据被提取、清理、转换和聚合到 Apache Impala（孵化中）和 Apache HBase 表中。
- 然后，保险公司进一步分析这些数据，创建出一份针对每一位客户驾驶习惯的个性化记录。通过重点分析每一位客户驾驶倾向的个体特征，保险公司将提高准确预测客户提出理赔的几率和成本的能力，可相应地调整保险费率和免赔额。

### 影响

#### 提高驾驶安全性和降低保险理赔

- **减少 30% 的理赔：**该保险公司现在可以根据司机的现实生活习惯和驾驶条件进行更一致和更准确的预测，并相应调整保险费率。这种新的业务洞察力可以显著减少欺诈性理赔，并将所有理赔的数量减少了 30%。
- **吸引并保留低风险（更安全）的司机：**通过向更安全的司机提供更低成本的保费激励，该保险公司已经能够增加其用户群，并同时提高他们的盈利能力。
- **鼓励良好的驾驶习惯：**凭借 Cloudera 企业版提供的高级分析功能，该保险公司可以为他们的客户提供门户网站以学习和提高他们的驾驶技能，并鼓励高风险驾驶员改变他们的驾驶习惯。



## 案例研究 3：在工业环境中实现预测性维护

使用案例：工业物联网 / 预测性维护

客户：一家领先的工业自动化公司

地理区域：北美

### 概述

北美地区一家领先的全球工业自动化公司通过工厂管理软件、运动控制传感器和机器安全组件完善其产品，帮助制造商提高效率，增加设备和生产运行时间。

### 业务驱动因素

#### 通过实时运营监控以提高效率

- 该公司需要能够实时监控、存储和分析来自多个工厂的数千个不同制造系统的数据，以便进行预测分析并提高运营效率。
- 由于客户每月每家工厂生成将近 50 GB 到 1 PB 的数据，包括测量工厂设备速度、受力、温度、压力和每分钟转速等传感器数据，该公司正在寻找一个可以处理巨大数据量的可扩展数据管理平台，其中包括静态数据和动态数据。

### 解决方案

#### 一套在 Cloudera 企业版上运行的基于云计算的互联工厂解决方案

- 通过基于 Cloudera 企业版并在 Microsoft Azure 上运行的工业物联网解决方案，该公司能够充分利用 PB 级物联网数据的预测分析和机器学习功能，帮助其客户减少生产停机时间并提高竞争力。
- 通过一种构建在 Hadoop 之上的专有诊断和分析应用程序，用户能够连续监控关键机器数据信号，例如扭矩、速度和位置误差，并确定移动部件的机械磨损以支持部件级别的维护。
- 通过在云计算环境中运行 Cloudera 企业版，可以轻松地提供数据和分析即服务，使他们能够提供实时应用程序、机器学习和商业智能（BI）解决方案，从而帮助客户更高效地管理其生产资源。凭借该公司开发的物联网解决方案，客户现在可以通过平板电脑或移动设备毫不费力地实时监控其机械设备的健康和功能，以确保平稳运营。

### 影响

#### 预测性维护和工业资产实时视图

- **实现实时预测性维护：**通过对来自生产系统的时间序列传感器数据进行处理和分析，该公司现在可以在工厂员工肉眼发现机械磨损和设备问题之前对其进行有效的检测，以便他们能够在发生故障之前采取行动。
- **致力于零停机时间：**通过将来自工厂车间的传感器数据流与来自内部和外部系统的结构化数据相结合，该公司工程师现在能够在影响生产之前发现和解决新出现的问题。
- **降低成本：**减少停机时间可以防止数百万美元的生产损失。此外，当需要进行维护时，机器学习功能可以帮助该公司匹配派遣合适的技术人员或工程师更快地解决问题，使他们能够在全球范围内扩展服务，而无需在额外资源方面投入数百万美元的成本。

减少停机时间可以防止数百万美元的生产损失；并且，当需要进行维护时，机器学习功能可以帮助该公司匹配派遣合适的技术人员或工程师更快地解决问题。





## 案例研究 4：医疗保健和帕金森氏症研究的革命性突破

使用案例：医疗保健物联网和可穿戴设备

客户：Michael J. Fox 帕金森氏症研究基金会

地理区域：北美

### 概述

据预计，全球大约有 700 万至 1000 万人患有帕金森症，这是一种神经退行性疾病，其发病率仅次于阿尔茨海默病。迈克尔 •J• 福克斯帕金森氏症研究基金会（Michael J. Fox Foundation for Parkinson's Research, MJFF）致力于通过积极资助的研究议程，寻找帕金森氏症的治疗方法，确保进一步改善帕金森氏症患者的治疗疗法。迈克尔 •J• 福克斯帕金森氏症研究基金会（MJFF）是由演员迈克尔 •J• 福克斯（Michael J. Fox）于 2000 年成立，该基金会已成为全球帕金森症研究最大的私人资助者，至今已投入的研究资金高达 4.5 亿美元。

### 业务驱动因素

#### 提升帕金森氏症的数据收集和管理功能

- 帕金森氏症研究中面临的挑战之一是有限的可用数据 - 数据来自于非常小的样本，数据收集不够频繁，并且不够客观。
- 帕金森氏症的主观测量方法仍然大致与 200 年前相同。通常，用于研究的数据收集来自于患者可能每三个月一次访问医生的诊所，而且主要来自患者的自我评估和记录。
- 在去诊室就医期间，患者会向医生陈述他们的症状，医生会要求他们做一些活动，然后由医生根据患者的情况进行主观评估。

### 解决方案

#### 在 Cloudera 上构建的英特尔可穿戴技术的解决方案

- 英特尔公司已经与迈克尔 •J• 福克斯帕金森氏症研究基金会合作，旨在通过一项多阶段研究，使用可穿戴设备和大数据分析平台（基于 Cloudera）监控帕金森氏症患者的症状，收集数据并检测参与者数据中的模式。
- 使用可穿戴设备每周 7 天每天 24 小时实时收集和传输客观的体验数据。
- 该合作项目建立了一个基于 Cloudera CDH 的大数据分析平台，用于收集、存储、管理和驱动对来自可穿戴设备的所有数据流的分析。该平台部署在经优化的英特尔架构的云基础设施上，使科学家能够专注于帕金森氏症的研究而不是基础计算技术。
- 该基础设施可以管理每天每位病人高达 1GB 的数据，并且随着数据和样本量的增长，该数据量可不断扩展。

### 影响

#### 针对帕金森氏症治疗所需数据的革命性突破

- **更好地了解帕金森氏症：**通过使用可穿戴技术，迈克尔 •J• 福克斯帕金森氏症研究基金会现在可以轻松地实时收集和汇聚体验数据（包括动作迟缓度、震颤强度、睡眠质量等），以便更好地了解帕金森氏症的临床进展情况，以及跟踪其与分子变化的关系。
- **改进数据访问和质量：**研究人员以前只有数量非常少的数据点，凭借这个新的平台，现在可以每秒收集来自数千名患者的 300 多个读数，获得大量的数据来侦测模式，探索新的研究发现。
- **为研究人员创建一个集中式数据存储库：**目前，该平台作为一种集中式存储库，拥有与帕金森氏症相关的最大量的数据集。在不久的将来，该平台可以存储其它类型的数据，例如患者、基因组和临床试验数据。此外，该平台可以支持其它高级技术，例如机器学习和图形分析，从而提供更准确的预测模型，研究人员可以使用这些模型来检测疾病症状的变化情况。

研究人员以前只有数量非常少的数据点，现在可以每秒收集来自数千名患者的 300 多个读数，获得大量的数据来侦测模式，探索新的研究发现。



## 关于 Cloudera

Cloudera 公司提供了现代化的数据管理和分析平台。全球领先的企业机构都信赖 Cloudera 公司可以帮助他们通过利用 Cloudera 企业版产品，以解决他们所面临的最具挑战性的业务问题。Cloudera 企业版是 Cloudera 公司开发的基于 Apache Hadoop 的最快速、最便捷、最安全的数据平台。我们的客户可以有效地捕捉、存储、处理、分析大量的数据，使其能够利用先进的分析技术以迅速、灵活、更低成本的方式推动业务决策。为了确保客户取得成功，我们将为客户提供全面的支持、培训和专业服务。

欲了解更多信息，请访问 Cloudera 公司网站：[cloudera.com](http://cloudera.com).

---

[cloudera.com](http://cloudera.com)

电话：+86-21-62369001

地址：上海长宁区延安西路 2299 号世贸商城 26 楼 2612 室