



# **EsgynDB 技术白皮书 2.8.0**

2021 年 03 月

## 版权

© Copyright 2015-2021 贵州易鲸捷信息技术有限公司

## 公告

本文档包含的信息如有更改，恕不另行通知。

保留所有权利。除非版权法允许，否则在未经易鲸捷 预先书面许可的情况下，严禁改编或翻译本手册的内容。易鲸捷 对于本文中所包含的技术或编辑错误、遗漏概不负责。

易鲸捷 产品和服务附带的正式担保声明中规定的担保是该产品和服务享有的唯一担保。本文中的任何信息均不构成额外的保修条款。

## 声明

Microsoft® 和Windows® 是美国微软公司的注册商标。Java® 和MySQL® 是Oracle 及其子公司的注册商标。Bosun是Stack Exchange的商标。Apache®、Hadoop®、HBase®、Hive®、openTSDB®、Sqoop® 和Trafodion® 是Apache软件基金会的商标。Esgyn和EsgynDB是易鲸捷的商标。

## 目录

<b>1. EsgynDB 的起源</b> .....	<b>2</b>
<b>2. NoSQL 变革和挑战</b> .....	<b>4</b>
<b>3. 传统 RDBMS 和新型 NoSQL 的完美结合——EsgynDB</b> .....	<b>6</b>
<b>4. HTAP 是什么?</b> .....	<b>8</b>
<b>5. EsgynDB 如何满足这些需求?</b> .....	<b>9</b>
5.1 流式处理.....	9
5.2 运营型工作负载.....	11
5.2.1 与 HBase 深度集成.....	12
5.3 BI/分析型工作负载.....	14
5.3.1 与 Hive 表集成.....	15
5.3.2 与 ORC/Parquet 文件深度集成.....	15
<b>6. 满足 HTAP 和未来需求</b> .....	<b>17</b>
<b>7. 总结</b> .....	<b>19</b>
<b>8. 关于易鲸捷和 EsgynDB</b> .....	<b>20</b>

## 1. EsgynDB 的起源

---

---

数据库引擎从起始到成熟需要很长时间，各式各样的、来自真实客户现场的工作负载能在不同程度上促进数据库引擎不断成熟，并在以下核心技术上持续进步：查询优化技术、事务和查询处理、支持的函数和功能。这些技术进步最终体现于满足用户 SLA 要求（规模、并发性和性能）的能力。同时，为了达到这些要求需要管理数据和工作负载、具备多租户功能，并满足安全性、一致性和可用性的要求。即使 Oracle、IBM DB2、Microsoft SQL Server 和 Teradata 已在数据库引擎领域研究了几十年，也难以满足所有用户的工作负载需求。然而，EsgynDB 架构的高灵活性和强扩展性可以最大化地满足用户的所有需求。EsgynDB 汇集了大量的专利创新，在它的发展成熟过程中，众多优秀的工程师贡献了毕生的心血。

Apache Trafodion 的鼻祖是 Tandem 的 NonStop SQL/MX（1987 年），它是一个 ANSI SQL 数据库，建立在可扩展和高可用的 MPP（Massively Parallel Processing）架构上。MX 是 Massively Extensible，代表可扩展架构。惠普公司在 NonStop SQL/MX 的基础上开发了一个旨在处理高强度企业级数据仓库负载的 MPP 数据库，这是设计 MPP 构架的初衷，而 Trafodion 直接继承该系统。因此，Apache Trafodion 能处理 OLTP、ODS、BI/分析应用程序的工作负载。

Trafodion 继承了 SQL/MX 的所有特性与核心代码，还直接继承了 NonStop SQL MX 的分支 Neoview 和其后续产品 SeaQuest（该产品线拥有数百项专利、超过 3 亿美元的投资以及超过 20 年的研发投入），因此还能处理 ODS 和 BI/分析应用程序的工作负载。2014 年 6 月，惠普公司将 Trafodion 开源。2015 年 5 月，Trafodion 成为 Apache 家族的新成员。2015 年 7 月，易鲸捷从惠普公司中剥离出来，成立了易鲸捷公司并拥有 30 名原 Neoview/SeaQuest 项目组成员，在数据库领域继承了 30 余年

的丰富的研发经验和深厚的研发实力。EsgynDB 架构在 Apache Trafodion 上，是企业服务的商业版数据库。

## 2. NoSQL 变革和挑战

---

---

NoSQL 为 Big Data 计算带来了以下前沿变革：

- **降低成本**：开源软件和标准通用硬件。
- **Schema 灵活性和 schema on read**：对于面临动态变化的成长型公司很重要。
- **增强弹性扩展**：满足大数据的高容量和速度。
- **支持半结构化和非结构化数据**：处理各类数据。
- **提高数据模型的灵活性**：处理各类工作负载。
- **并行处理用户代码**：利用 MapReduce 实现。
- **高可用模式**：适用于低廉且易出故障的通用商业服务器。

这些变革催生了几类典型的 NoSQL 解决方案，每一类解决方案都设计了自己独特的 API 接口，并仅用于解决部分特定的大数据应用场景。典型用例包括实时流和分析机器学习应用程序。更多关于这些用例的信息，请参阅《数据管理的未来》。

随着 NoSQL 和 Hadoop 越来越受欢迎，越来越多的应用程序开始迁移到这些环境。但是，当新型初创公司走向成熟阶段后，其运营型工作负载的需求不断增加，更加需要那些传统 RDBMS 所能提供的能力。虽然大型企业未面临和初创公司一样的挑战，但也想利用这些新技术的优势，并希望继续使用 SQL，原因如下：

- 企业普遍使用 SQL，这能简化开发。
- 企业已经使用了大量 SQL 相关的工具和应用程序生态系统。
- 尽管事务处理有一定开销，但很多情况下仍需要事务支持。
- SQL 引擎能高效处理 join 操作，很多情况下仍需要 join 的支持。
- 企业开发人员需要进行相对困难的应用程序开发，或者编写 MapReduce/Spark 作业，而使用 SQL 更简单。

- 预定义数据表列有很多优势，因此用户可以使用数据类型检查和约束来保证数据质量，而且无需牺牲 schema 灵活性。
- 促进了应用程序之间的元数据统一管理和规范。

## 3. 传统 RDBMS 和新型 NoSQL 的完美结合—— EsgynDB

---

---

虽然 NoSQL 运动是革命性的技术创新，但现在对传统 SQL 和 RDBMS 的需求在复苏，两者结合的新产品可以最好地满足需求。这催生了新型的强大数据库引擎 EsgynDB，它整合了 Apache HBase 原生的 NoSQL 能力、Apache ORC、Apache Parquet 和其他 Apache Hive 文件格式。

借助于 NoSQL 的实践经验，EsgynDB 的发展方向是与多种存储引擎集成，为用户提供多种数据结构/模型选项，从而支持各类工作负载和用例。EsgynDB 在 HBase 之上支持结构化的关系型抽象、Google Big Table 模型（使用原生 HBase）和采用 key-value 或列式存储（ORC/Parquet）的 Hive 表。

为了支持多数据模型和满足不同用例的需求，EsgynDB 继续与存储引擎进行集成。这些多模型支持覆盖内存、文本搜索、JSON/文件支持和图形数据库能力领域。某些集成工作早已开始，例如，与基于 Apache Geode 的 Ampool 集成、与 Apache Spark 集成（内存）、与 Apache Lucene 集成（文本搜索）、与 JSON 文档管理集成和与 JanusGraph（旧称 Titan，解决图遍历问题）集成。

多模型支持拥有可扩展数据管道架构，在生态系统中，它能轻松地与其他技术进行集成。Table UDF 基础架构允许与 Kafka 集成（流式）、与机器学习算法集成（Continuum 的 Anaconda）、与 JDBC 数据源集成（导入存储在其他数据库的数据）、与 Spark 集成（内存分析）和与 Lucene 集成（外部网站搜索）。



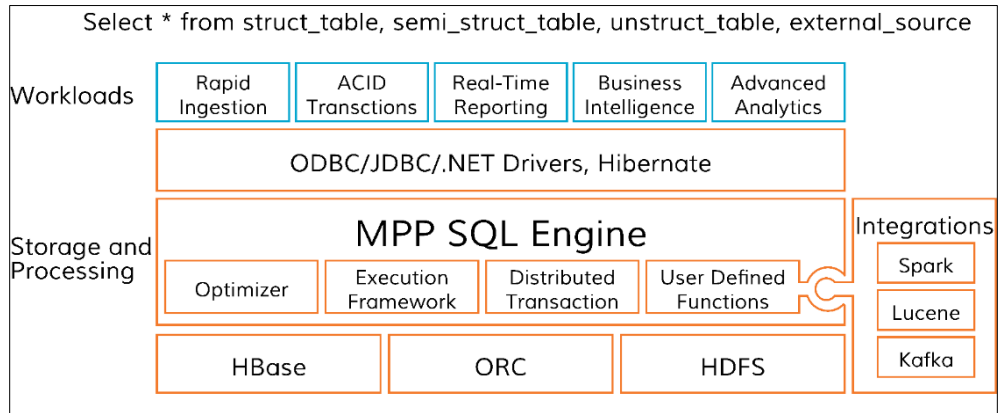


图 3-1 EsgynDB 可扩展数据管道架构

## 4. HTAP 是什么？

---

---

为了满足各种大数据需求，用户希望存在一个单一数据库引擎，它能：

- 满足所有数据模型的需求，处理所有工作负载（事务、运营、BI 和分析）。
- 将所有数据存储于单一平台，而不是存储在不同平台（不同平台使用不同的数据库处理不同的工作负载）。
- 减少数据移动和复制、降低因此产生的延迟和运营成本。
- 利用运营数据、存储在相同平台的半结构化和非结构化数据进行深度分析并挖掘潜在价值。
- 获取数据的同时生成报表和进行实时分析（无延迟），无缝集成并访问运营、历史和 Big Data。

2014 年 1 月 28 日，Gartner 在《混合事务/分析处理促进重大商业创新》报告中定义了 HTAP。维基百科将 HTAP 定义为“单个数据库同时支持 OLTP 和 OLAP，进行实时智能处理的能力”。EsgynDB 不仅支持 HTAP 工作负载，还支持多种数据模型，用于满足 Big Data 的各种应用需求。本质上，EsgynDB 是 HTAP on Big Data。

在 HTAP 领域，EsgynDB 支持多集群/多数据中心灾备、多主机扩展、灾难恢复、时间点恢复、云弹性、混合工作负载管理、多租户和安全基础架构等特性，从而提供了全面的企业级事务保护。另外，EsgynDB 在权威的 TPC-C 和 TPC-DS 基准测试中展现了业界领先的性能测试结果。

## 5. EsgynDB 如何满足这些需求?

---

---

### 5.1 流式处理

Zero Latency 公司实现了梦寐以求的愿景：实时处理高速捕捉的数据。数据批处理导致的延迟已不能被接受，许多应用程序使用流式编程模型（而不是批处理）处理注入的数据。流式处理包括两方面：一方面，保存注入的数据；另一方面，如果数据注入速度很快，处理前无法保存数据，必须进行实时处理。

HBase 擅长高速数据写入，很适合被 EsgynDB 使用，用于处理[运营型工作负载](#)。因此，EsgynDB 能满足高速注入 Big Data 的需求，并提供全面的 ACID 支持。

Kafka 作为一种缓冲，能避免大数据系统因为数据注入的瞬时增大而导致系统崩溃，这在数据实时入库保存或实时分析的场景下均可能出现。为了更好地保存并简化数据注入，EsgynDB 与 Kafka 进行集成。该集成采用了 EsgynDB 的 [Table UDF](#)，能流式地处理复杂场景或进行分析，同时，它还能利用引擎的所有 SQL 能力。

数据导入进 HBase memstore 或内存，[EsgynDB 数据注入速度](#)约为 74 $\mu$ s/行。任何需要访问数据的查询均能实现立即访问。只有那些能在 100 $\mu$ s 或 0.1ms 内处理数据的应用程序将先处理数据，再保存数据（如需要）。另外，Netflix 每天通过 3000 多个代理的 24 个集群获取 7000 亿条信息（=370 $\mu$ s/信息）；LinkedIn 通过 1100 个代理传输 1300 万条信息/秒（=85 $\mu$ s/信息）；Facebook Messenger 和 WhatsApp 每天处理 600 亿条信息。

如果您的需求大于 Netflix、LinkedIn 或 Facebook，在不使用 Kafka 的情况下，易鲸捷也能满足您高速注入数据和存储数据的需求。另外，EsgynDB 支持事务插入和更

新 (Kafka 不支持 SQL) , 还能为数据注入速度更快的场景使用 **bulk load** 进行批量加载或 mini-batch 加载。

作为通用参考, EsgynDB 的流式数据注入速度约为 50,000 行/秒/节点, bulk load 数据注入速度约为 100,000 行/秒/节点。

## 5.2 运营型工作负载

传统关系型数据库处理传统的运营型工作负载，但由于 Hadoop 能大幅降低成本、减少供应商锁定和无缝扩展以处理更多工作负载、存储更多数据，因此，使用 Hadoop 来处理工作负载的需求越来越大。

一般而言，受事务保护的运营型工作负载创造收益、接触客户、预测前景或经营业务，所以对公司而言，它是非常重要的部分。通常情况下，该工作负载在响应时间（次秒级）、事务数据完整性、用户数量、并发性、可用性和数据量方面都有严格要求。随着物联网的兴起，设备数量和种类的不断增长促进了事务和数据的增长，同时还改变了数据的类型，这些数据需要被获取并作为事务的一部分被使用。运营数据发展迅速，包括各式各样的数据格式和类型，因此，下一代运营应用程序需要使用结构化、半结构化和非结构化数据类型。例如，半结构化的日志数据和结构化的事务数据、从设备获取的物联网数据、文字信息与情感和其他分析评论、需要被搜索的文件或文本信息以及与结构化数据/元数据相关的视觉图像等。

然而，Hadoop 无法满足前述所有需求，例如，事务支持、灾备方案、强数据完整性、次秒级响应时间、运营查询优化和管理工作负载（根据优先级分配任务且并发地执行）。EsgynDB 能解决以上所有问题，它强大的 DBMS 支持应用程序和数据，减少架构在 Hadoop 上软件之间的数据延迟和冗余。

EsgynDB 提供全面的 SQL 数据库功能，支持公司利用现有 SQL 技能，提高开发人员产出和处理工作负载的效率。在处理运营型工作负载时，EsgynDB 具备以下优势：

- 与 HBase 深度集成，充分发挥 HBase 优势。
- 低延迟读写事务优化，支持事务 SQL 工作负载的高并发和快速响应。
- JNI 客户端（Type 2 JDBC 驱动）直接访问 HBase 服务，减少服务时间。

- 编译 SQL 计划缓存，减少重新编译的开销。缓存命中策略不仅是简单文字匹配，它在编译过程的多个阶段进行。如果谓词的代入值不同，导致基数有很大差异，则 EsgynDB 将为相似查询缓存出不同计划。
  - 多维访问方法 (Multidimensional Access Method) 专利技术：
    - 即使索引的前导列没有谓词，也能高效地使用索引。
    - 大幅减少使用二级索引的需求。
    - 帮助用户自动实现 salting 和 divisioning，实现数据分区。
  - EsgynDB 扩展 HBase 时为它添加了 ACID (原子性、一致性、隔离性和持久性) 事务保护，从而确保数据跨多行、表和 SQL 语句 (使用完全分布式事务管理架构) 的一致性。EsgynDB 与 HBase 深度集成，能在高并发场景时提供高扩展性、高效率的执行和事务一致性。
  - 提供高可用和数据完整性功能，对运营型应用程序至关重要。
  - 跨数据中心的全面双活分布式事务支持，扩展读/写工作负载并提供灾难恢复零事务丢失机制。
  - 对运营错误进行时间点恢复，例如，被删除的表格或错误的事务更新。
  - 弹性扩展。在不宕机和同时处理事务和查询的情况下，跨集群的节点和存储可以大幅增加或减少。下一个事务或查询将利用重新配置的计算机或存储资源
- 以上大部分功能也能处理 BI/分析型工作负载。

### 5.2.1 与 HBase 深度集成

Trafodion 宣布开源时，当时已存在基于 Big Data Hadoop 平台的用于满足 BI 和分析需求的 SQL 解决方案，但未有像 EsgynDB 一样能同时处理事务型和运营型工作负载的解决方案。为了处理这些工作负载，EsgynDB 专注于使用强大的数据库引擎、与 HBase 深度集成、提供完整可扩展的分布式事务管理基础架构并尽可能减少开销。

HBase 能处理低延迟工作负载、支持高速数据注入、有利于运营型查询的高效的键结构以及具备 schema 灵活性和从 HDFS 继承的 NoSQL 的优势。HBase 还拥有压缩、加密、高可用复制和弹性扩展等功能。

HBase 不仅为 EsgynDB 提供了用于保存结构化关系抽象的存储引擎，还提供了 Big Table 数据模型（键-列-值，即宽-列存储）处理半结构化数据（无法使用关系模型表示）。对运营型工作负载而言，EsgynDB 是 SQL 和 NoSQL 能力的完美结合。EsgynDB 不仅能在相同查询中对 Trafodion 结构化表和原生 HBase 表进行查询，还能在相同 ACID 事务中对它们进行更新。HBase 表可以定义为 EsgynDB 外表，您能通过这些表创建视图。

最后，选择 HBase 的重要原因在于 Hadoop 已经是大数据行业标准，HBase 是 Hadoop 大型生态系统的一部分，用户能利用 Hadoop 生态系统来满足各类需求。

### 5.3 BI/分析型工作负载

用户已经在 Hadoop 上运行大数据工作负载，他们过去使用 EDW 中的结构化历史数据运行 BI 查询，随后创建了为 Big Data 工作负载服务的数据湖泊，但现在他们意识到需要将 EDW 和数据湖泊结合起来。通过集成结构化历史数据、半结构化和非结构化外部 Big Data，不仅能提高决策力，而且能将部分运行在昂贵专有硬件上的工作负载转移至较低 TCO 的开源平台。这包括从这些平台上减少 ELT 过程、集成并扩展报表和分析，以集成所有企业数据并生成更高价值的分析，减少 EDW 和数据湖泊的部署。

大量用户使用 EsgynDB 处理 EDW ELT、BI 报表和分析型工作负载，所以，EsgynDB 是唯一一个有能力同时处理这些工作负载的数据库。在该领域，EsgynDB 的能力和成熟度已远超其他处理这些工作负载的 SQL-on-Hadoop 引擎。ETL 工具转换导入进 EDW 的数据，在 EsgynDB 出现之前，惠普公司已广泛地使用 EsgynDB 的底层技术处理大量数据转换，例如，使用 ELT 转换全球运营系统的数据（企业 BI 用户使用），将数据 rollup 或聚合至报表区域，以支持对 EDW 中数据执行海量并发查询并快速响应。在处理复杂 BI 和分析工作负载时，基于 EsgynDB 的数据湖泊不仅能存储和处理结构化数据，还能存储和处理未被任何 ETL 工具预处理的半结构化和非结构化数据，所以，您无需使用传统方法便能轻松处理此类工作负载。

EsgynDB 的架构使它能高效地处理 HTAP 工作负载。运营系统和分析系统可同时架构在 Hadoop 平台上，运营系统存储 OLTP 和流式应用程序的数据，数据被捕捉的同时在分析系统中进行分析，分析的结果再反馈至运营系统。由于运营系统和分析系统在同一平台上运行，闭环分析能大幅减少响应时间并增强决策力。



处理运营型工作负载需要特定功能，EsgynDB 在处理 BI 和分析工作负载时，具备以下特定功能：

- 与 ORC 文件深度集成，最大化地利用 ORC 的优势。
- 复杂优化器使用数据的统计信息，为查询生成最优计划。
- 拥有专利技术 [skew buster](#)——处理倾斜。
- 拥有专利技术 [adaptive segmentation](#)——使用查询仅需的资源，在高并发时消耗最少资源，并提高弹性。
- 数据流执行架构，能最大化地提高并发度，高效使用内存，实现快速响应。
- 拥有复杂并行数据库引擎能力，支持高并发和高吞吐量的复杂报表查询。
- 基于查询模式的多 join 策略，提供最佳性能和并发度，高效利用资源。
- 优化的 [Table UDF](#) 能力，能并行处理 C++ 或 Java 用户函数。和并行执行的 MapReduce 类似，与其他数据源和存储引擎集成。

### 5.3.1 与 Hive 表集成

EsgynDB 支持在同一个查询中同时访问以 text 和 sequence file (key-value) 格式存储的 Hive 表，您还能使用这两种格式将数据加载至 EsgynDB（速度等同于 bulk load）。Hive 表可以被定义为外表，这能提高查询的易用性。

EsgynDB 能访问 Hive 表（保存在对象存储中），包括以 ORC 或 Parquet 方式存储的 Hive 表，例如，Amazon AWS S3 或 Microsoft Azure 数据湖存储。

### 5.3.2 与 ORC/Parquet 文件深度集成

用户对 EsgynDB 有越来越多的需求，他们期望在业务数据上直接进行 BI/分析类操作，而无需进行数据移动和复制。EsgynDB 支持这类需求，虽然 HBase 并不适合这

类工作负载，但 ORC 很适合处理需要大量数据扫描的报表和分析型工作负载。

EsgynDB 与 [Apache ORC/Parquet 文件深度集成](#)，能高效地处理这类工作负载。

## 6. 满足 HTAP 和未来需求

---

---

前文分析了 EsgynDB 如何解决各种 HTAP 工作负载需求，解释了流式处理日益增长的需求和多种 Big Data 模型的应用，现在讨论企业部署 Big Data 技术的其他重要方面：

- 熟悉 Big Data 的 4V——Volume（容量）、Velocity（速度）、Variety（种类）与 Veracity（精确）：
  - **容量：** EsgynDB 架构在 Hadoop 上，拥有大规模并发处理能力并具备高可扩展性。因此，EsgynDB 擅长处理海量数据。更多关于[架构](#)和[性能](#)的信息，请参见[附录](#)。
  - **种类：** EsgynDB 支持多数据模型，并在将来支持更多数据类型。EsgynDB 不仅支持多种数据结构和模型，还支持多种工作负载，例如，OLTP、运营、BI 和分析（HTAP）。
  - **精确：** 精确与数据治理、transparency 和 lineage 息息相关，EsgynDB 能使这些流程的处理变得更简单。数据分析有利于 EsgynDB 处理半结构化和非结构化数据。分布式 ACID 事务处理、强大的灾难恢复能力与强制数据类型和检查约束使 EsgynDB 能保证数据一致性和数据质量，并在保证数据精确方面发挥了至关重要的作用。
  - **速度：** 数据速度是 Big Data 平台为 IoT 相关应用处理数据的速度。数据速度有两方面，一方面是数据注入（前文已讨论），另一方面是实时处理 HTAP 工作负载，EsgynDB 在处理高并发时，性能可以保证达到 SLA 的要求。
- 大幅降低总成本。
  - 开源软件和定价模式。
  - 在标准通用硬件而不是在专有硬件上运行。

- 使用现有 SQL 技能和工具。
- 能在云上运行（例如，AWS），拥有弹性扩展力和虚拟配置。
  - 降低管理硬件和软件资源的运营复杂成本。
  - 根据需求配置资源。例如，在销售旺季分配较多资源、在销售淡季分配较少资源。
  - 跨多个租户高效利用软件和硬件资源，这些租户可能是多个托管用户、多个企业应用程序、或者多个部门，每种租户有不同的计算和数据资源需求。
  - 向自助服务模型发展。
- 企业级管理和运营基础架构：企业能成熟地管理和运营自身的应用程序和数据部署，并期望使用相同的技能管理大数据工具。为了解决这些需求，EsgynDB 提供企业级管理和可用能力。

## 7. 总结

---

---

EsgynDB 是唯一一个能在单个平台上运行所有工作负载（事务、运营、BI 和分析）的数据引擎，同时，它还能促进结构化/半结构化/非结构化数据与 NoSQL 和 SQL 技术的集成，为您下一代应用程序提供平台。EsgynDB 能同时满足 NoSQL 和 SQL 两方面的需求，并提供最佳性能。

## 8. 关于易鲸捷和 EsgynDB

---

---

易鲸捷公司的使命是采用颠覆性的技术和产品引领大数据市场的未来，基于目前最流行的开源数据库引擎 Apache Trafodion，EsgynDB 能处理各类工作负载，例如，OLTP、ODS、BI 和分析（HTAP），它还支持流式处理、内存计算和大数据存储，与在单个系统上支持多数据模型的存储引擎深度集成。

易鲸捷的顶级产品——EsgynDB 企业版，架构在 Apache Trafodion 上，是一款专门处理 Big Data 的安全可靠的企业级 HTAP SQL 引擎。易鲸捷公司在上海、北京和贵阳设立了分公司，为 EsgynDB 提供支持、服务和培训，使 EsgynDB 能满足企业的各类需求。

更多信息，请访问 [www.esgyn.cn](http://www.esgyn.cn) 或发邮件至 [info@esgyn.com](mailto:info@esgyn.com)。

© 2016-2020 贵州易鲸捷信息技术有限公司