

Data Integration Features Comparison Matrix

Talend Open Studio

vs

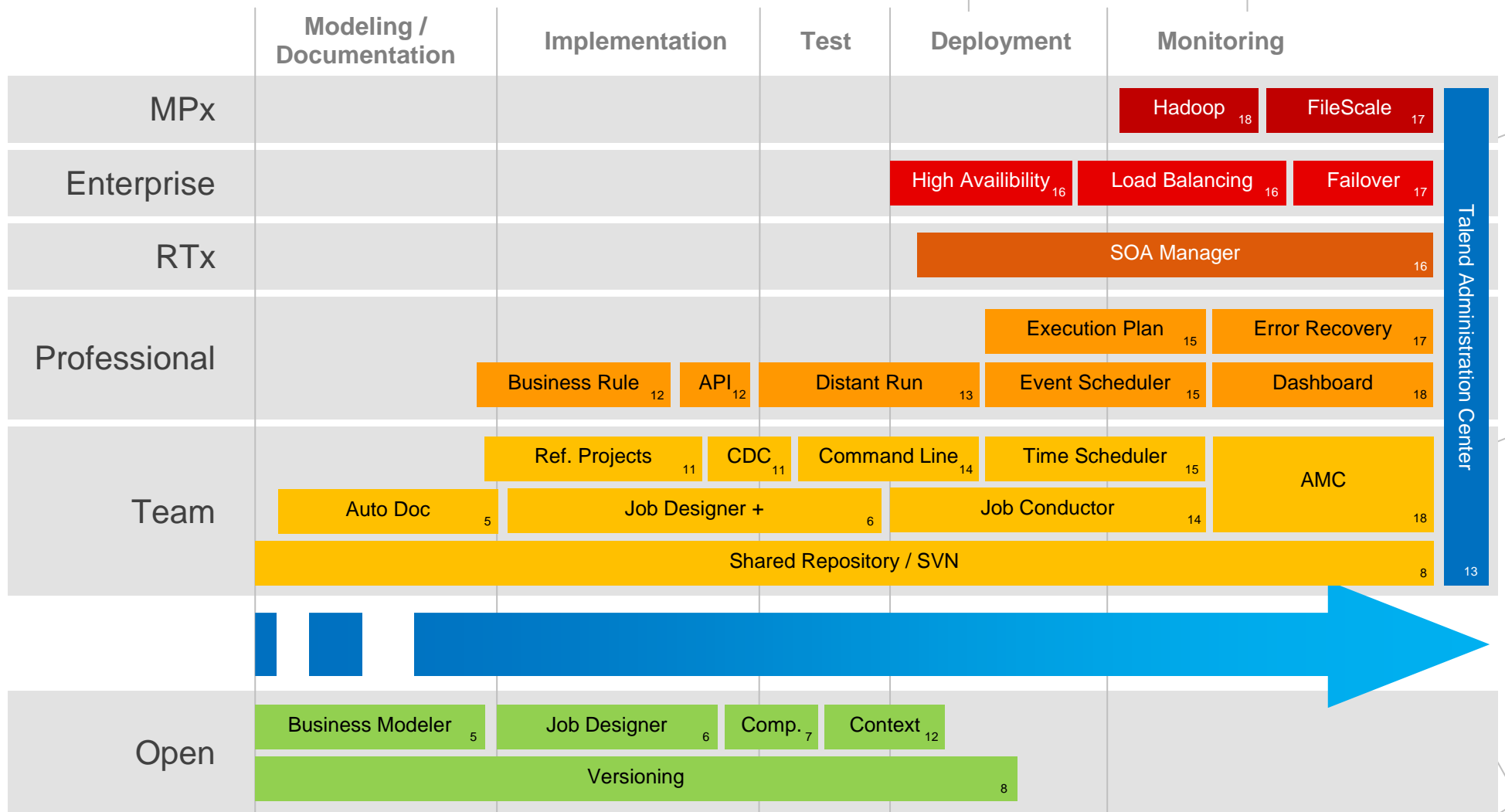
Talend Integration Suite

Team, Professional, RTx, Enterprise & MPx editions

<http://cn.talend.com/products-data-integration/matrix.php>

Q1 2011

Feature details



许可证类型与补偿

开源 GPL 许可证

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx

GNU 通用公共许可协议是免费软件的一种执照，它为 GNU 项目下免费软件的分配设立了法律条款。

这个GNU GPL 许可证的创立是为了保证用户的下列权利。

- 为任何使用目的执行软件而不受到限制的权利;
- 分析软件机能并按照需要对其进行调整的权利。

如果该软件修改内容的作者决定共享这个软件，他或者她也必须按照该 GPL 许可证行事。

该 GPL 许可证的全文请见链接：

<http://www.opensource.org/licenses/gpl-2.0.php>

访问源代码

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx

Talend Open Studio 的源代码可在如下网址访问：

<http://www.talendforge.org/trac/tos/>

为了达到透明与一致的目的，Talend 还在商业版本中为请求该源代码的顾客提供全部需要的工具。

订阅许可证

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx

“企业”版包含了增值的功能(见下文)与服务，提高了开源产品的性能;这些版本按照商业许可证模式分销。

Talend 定价模式实际落实了价格的透明性和可预估性:

它的价格不是立基于数据规模或者对CPU或连接器更多的潜在需求之上，而是立足于投入的开发人员（Studio）的数量，产品功能的级别（即选定的版本）和订阅的期限之中。

这种订阅的方法保证了对您投资的回报：每年订阅许可证的数量可以增加也可以减少，用以配合项目内容和配备人员的发展变化。

Talend 解决方案的部署，维护和支持都更为低廉，比起同等的专有解决方案要便宜50%到80%之多。

补偿协议

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx

因为开源软件是协作开发，共同努力的结果，所以最终代码融合了不同资源贡献的代码。如果来自各种贡献的整合代码没能小心地管理控制，最终软件的使用可能会损害原版贡献者的权利。因而最终用户可能会因为侵权而受到司法起诉和经济诉讼，即使这种侵权并非有意为之。Talend 为其订阅客户提供补偿条款。Talend 借此确保为用户提供法律上和经济上的保护，即便发生了Talend 代码侵犯第三方权利的情况。

支持与文档

社区支持: 论坛，Bugtracker...

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx

Talend 用户社区由成百上千的专业人士组成，极其活跃。社区的主要贡献包括：

- 新版本的测试和品质保证，
- 请求新的功能，
- 产品翻译和本地化，
- 论坛支持和交流，
- 新组件、连接器、作业、模型、其他插件的开发和共享。

Talend Exchange使社区成员能够发布他们自己的插件用以和其他用户共享。这些贡献大多数都会在经过我们的内部的研发团队彻底的测试改良之后，最终整合到产品当中。

另外，Talend 也贡献了许多关键的开源项目，同时还是 Eclipse 和 Apache 基金会的成员。

SLAs企业级支持

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx

我们技术专家每天在和我们的研发团队交流，订阅 Talend 支持服务，分享他们的内部经验。建立这些服务的目的在于确保方案的效用，保证订阅客户的安全和从容。服务共有三个等级：白银，黄金以及白金。其中每一个等级都关系到回应报出的 bug 花费的初始时间，提供补丁需要的回应时间等等。

<http://cn.talend.com/professional-support/support.php>



用户指南，参考指南

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx

Talend Open Studio 的 PDF 文档可以免费下载，同时拥有英文和法文两个版本。共有两个使用指南：用户指南（User Guide，276 页）和组件参考指南（Reference Guide）。下载请至：

<http://www.talend.com/resources/documentation.php>

您还可以在 Amazon 购买印刷版的使用指南：

http://www.amazon.com/Talend-Open-Studio-Guide-Version/dp/1453632077/ref=sr_1_1?ie=UTF8&s=books&qid=1296040035&sr=1-1

此外您可以观看我们的 Tutorials，这也是价值丰富的信息来源。

访问链接请点击：<http://www.talendforge.org/tutorials/menu.php>

建模/ 文档

业务建模专家

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx



业务建模专家并非一款可执行工具（如同Microsoft Visio）。

但它可以帮助您将数据集成流程所需要的相关文档和技术元素全部都组织在一个业务友好型的结构视图中，使不同的团队（设计，开发，测试，生产等）可以在共同的模型基础上工作，使用共享的工具。

例如，业务用户可以使用业务模型来表达他们的数据集成需求。

IT 开发部门和操作人员可以凭借这些模型来更好地理解这些业务需求并将他们转译成一个个的技术进程（作业）。

每一次技术实施（作业执行）过程完成以后，业务模

型都可以很轻松地得到更新，同时便于其他的相关人员跟踪查看开发的进展状况。

DBAs 可以使用业务模型来分享配置数据库连接所需要的元数据，从而系统架构师们能够对数据集成所要求的条件形成宏观概念。

企业在数据管理或集成项目的初始阶段就需要选定其业务模型来确保最后的成功，因而设计业务模型是一个企业最佳实践的一部分。因为业务模型通常可以协助侦测和快速地解决项目遇到的瓶颈和缺陷，因而采用他们有利于限制超支水平并/或者减少前期投资。

自动文档

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx

这个功能可以按照需要为你所有的作业生成详细的技术文档。这个文档汇集了作业元数据（作者，版本，状况，更新日期，等等），图形化作业视图，以及作业中使用的所有组件的参数。文档使用便于互动的格式（HTML / XML），应用方便。

通过添加个性化评论，文档内容可以轻松地得到扩充。

高级自动文档

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx



通过高级自动建档功能，每个作业的每个版本的技术文档都可以自动生成（见前段）：当您保存作业的时候它的文档会得到更新并储存在 **Repository** 存储库中，因此，用户都可以自动分享使用该文档。

高级自动建档功能也可以自定义文档的图形化显示，因而可以添加您自己的 logo 和公司名称，或者通过自定义的 CSS 来更改颜色。

作业设计专家

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx



作业设计专家为您提供可视化，可执行功能的视窗，是执行集成进程的真正场所，配备了一整套图形化的组件和连接器工具栏。

集成的不同进程可以通过简单的拖拽组件和连接器来完成。在可视化作业设计界面上，拖拽放置组件和连接器，在他们之间通过简单描画建立关系，并设置他们的属性。

无论数据的集成、转移或同步进程需要什么样的源数据和目标数据，作业设计专家都可以通过内容丰富完备的组件库来对这些资源或目标进行访问。组件和连接器涵盖了对于数据，数据管理和数据流序列所能实现的全部类型的任务和操作。其中连接器可以访问，读写所有的源数据和目标数据，从而实现数据的集成，转移和同步。而在选择每一个组件的时候，各种参数都可以在一个视图上集中配置，或者从元数据管理专家

家 (Repository) 中导入。

复杂组件都有专门的可视化图形界面，或者内建的向导，来帮助用户建立他们的作用。

为了始终能够方便理解一个作业的设计，作业的线图能够分为若干Subjob (次级作业)，然后可以进一步设置为子作业与父作业关系，从而建立一个作业执行的序列。编译用组件以及不同类型的关系可以帮助用户为不同进程的执行排出顺序。而内建的console (控制台) 则可以让用户在 Studio 上迅速监测到执行的状态，检查并跟踪执行的性能。

组件

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx



Talend提供的原生的开源连接器不受IT环境限制，既可满足technical需要，也可以回应business需求。它可以用无与伦比的速度将相异多样的数据结构桥接起来，连接器对异构数据广泛地连接功能足以实现所有数据、应用以及数据库之间的互动，而这种广博的连接功能，也使得 Talend方案可以不断丰富，日趋完美。超过550个组件可以免费使用，其中60% 都有 Talend 社区设计并开发。

外部开发的连接器与组件可以通过Talend Exchange (<http://talendforge.org/exchange/>) 共享。大量提交的组件都经由Talend验证并改善，并最终原生统一于 Talend 方案之中，得到方案支持。

请参见<http://www.talendforge.org/components> 获得所支持连接器的详尽的清单。

ETL 支持

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx

ETL（提取，转换，加载）是 Talend 数据集成解决方案使用的默认模式。它本质在于一行近挨一行数据流形式处理数据。这个模式经专门调整，用于异质的环境中，它可以实现源系统和目标系统中任何技术的整合（web 服务，文件，数据库，MOM业务应用，等等）。

ETL 模式还可以用于批处理和实时处理中。ETL 进程可以并行运转，以进一步地加速其执行。

Talend 独特的架构生成独立的各种进程，可以部署于任何服务器之上（无论公司外部或者内部），因此不受任何执行引擎的限制。另外，ETL 各种进程可以在尽可能靠近数据的地方执行，将存取时间和带宽消耗最小化，并消除瓶颈。

在同一个作业中，这个方法可以和 ELT 结合（见下段），获得最高性能，不受任何架构限制。

ELT 支持

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx

Talend 数据集成解决方案还支持 ELT 模式（提取，加载，转换），其本质在于直接在目标数据库的 DBMS 中以集合操作（使用 Union, Except 和 Intersect 运算）来处理数据。

这个模式专为同质环境（一个数据库）使用。它享有得天独厚的设备资源优势，尤其推荐在“数据仓库设备”的环境中（如 Teradata, Netezza, 等等）处理极大规模的数据时使用。

在同一个作业中，这个方法可以和 ETL 结合（见前段），达到最高性能，不受架构限制。

版本管理

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx



您可以通过 Talend 原生的手动版本管理功能来对 Talend Studio 中的 items 进行版本管理。

主次版本号（major & minor）可以在作业创建时自动设置，并随着时间增长和更新次数次第增加。在 Talend Studio 的设计专家 perspective（透视图）中，您可以直接访问专门的版本控制面板来进行管理。

Studio 中创建的所有 items 的版本都可以这样来管理，这包括：业务模型，作业，例程，元数据，文档，等等。

版本管理总体上是最佳实践计划的一部分，其目标是促进 item 的重复利用率，并在需要时候回复到前面的开发阶段。

共享 Repository 存储库

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx



共享 Repository 存储库（或称元数据管理专家）用于在中央存储库中集中整合所有的项目信息和企业元数据，供集成流程中所有的相关人员共享。

在 Studio 方面，用户得到授权，并根据由 Talend Administration Center 规定的权限访问工程项目。

共享 Repository 存储库实现了集成项目所有参与人员之间的团队合作，有利于存储共享他们所有的 Talend item，这包括：业务模型，作业（进程），Joblet，例程，元数据定义（例如对源系统或目标系统的连接配置），等等。

共享存储库是在一个产业标准化的资源管理器（Subversion）之上运行的，可以存储和管理所有版本的 item。

自动锁定系统可以保证正在设计的作业能够被有效锁定，其他用户不能在此时对其进行改动。

从 4.0 版开始，我们充分发挥的 subversion 版本管理的作用，可以处理不同的 branch，数据 check in /check-out，手动或自动 commit 提交，评论加注等。

数据查看器（Data Viewer）

this feature is included in:

Open

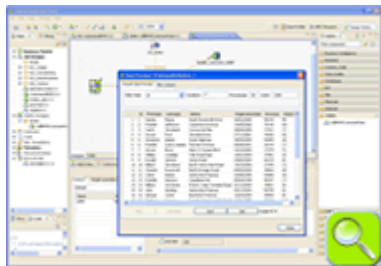
Team

Professional

RTx

Enterprise

MPx



在使用 Talend 开发作业的时候，您或许需要查看不同源系统或目标系统的内容（文件，数据库等）。数据查看器可以深入到这些系统的数据当中进行钻取，而不必再经过连入它们通常所需要的应用，比如：txt & csv 文件的 Notepad，查看数据库表的 SQL query browser，XLS 文件的 MS Excel，html 浏览器等。

不再需要使用多种工具，也不再需要浏览系统才能钻取到数据分布的位置。数据查看器使用特定的源与目标路径设置，直接访问实际的数据。

因为可以直接在 Studio 中使用，数据查看器可以节省您大量的时间，无论任何组件，您只需右键单击即可。这也十分便于您在开发您的集成进程时查看您源系统和目标系统中的数据，同时还能不再受数据格式（Excel，数据库表，CSV 等）的限制。

向导

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx

动态 Schema

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx

动态 schema 可以通过未知的 column 结构和数量来设计作业。根据开发人员的选择，动态的 column 可以使用 Pass-through 模式直接映射到目标中。这种功能主要应用于 scenario 重现，或者多个 column 的一对一简单映射。这一功能使得设计这些类型的作业简单上手。例如，一名开发人员需要迁移整个数据库及其上百个表格，他无需了解全部的表格结构，只使用单一的一个作业就可以完成！

冲击分析 (Impact Analysis)

this feature is included in:

Open

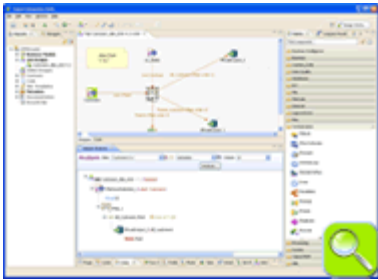
Team

Professional

RTx

Enterprise

MPx



冲击分析功能帮助您理解每一个改变所可能造成的后果。该功能可以通过元数据管理专家来使用。您可以对任何元数据（数据库，文件等）的任意列进行冲击分析。分析的结果以图形化互动形式的报告显示，通过这个报告，您可以跟踪分析的列，并看到在整个作业中所有应用于该列的操作，从源系统一直到目标系统。您可以将此报告导出生成HTML 文件。

Data Lineage

this feature is included in:

Open

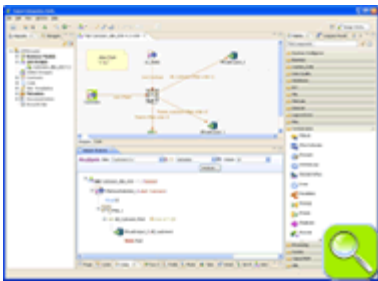
Team

Professional

RTx

Enterprise

MPx



Data Lineage 功能帮助您理解改变的位置。该功能通过数据管理专家使用，可以应用于任何元数据（数据库，文件等）的任意列上。Data lineage 分析的结果会以报告形式从作业的目标端组件开始回溯，直到源系统端组件为止，追踪出一条改动的整个演进过程。您可以将该报告导出生成 HTML 文件。

作业比较

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx



作业比较功能可以鉴定作业的两个版本或不同作业之间的异同。

作业比较功能完全嵌入 Talend Integration Suite Studio 之中。比较的结果会生成html 或 xml 格式的可视化互动报告，标注出所有不同。

在本例中，比较报告显示了 tFileInputDelimited 组件的设置中，delimiter 分隔符这一栏在比较的两个作业中定义的内容不同：3.2版本中，分隔符为“\t”，而4.2版本则为“\n”。

Joblets

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx



Joblets 功能是指将作业的一部分（或一个Subjob）因子化为一个相对固定的Joblet 组件。当您预计需要再次使用作业的一些组件组成的部分，或者需要把这一部分因子化的时候，只需选中这些组件，并在目录上点击“Refactor to Joblet”项即可。

作业设计的过程自动得到了简化。因为需要使用的组件可以打包成单一的一个 Joblet 组件。这个Joblet 组件可以通过 Palette组件栏上专门的 Joblet 文件夹实现分享，从而可以在任何作业中使用。

Joblet 的出现大幅度地简化了复杂，冗繁作业的维护。

另外，配套的“冲击分析（Impact Analysis）”机制有利于找出哪些作业使用了特定的 Joblet。

引用项目

this feature is included in:

Open

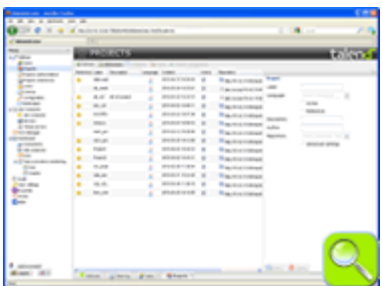
Team

Professional

RTx

Enterprise

MPx



引用项目（Reference Projects）可以避免项目之间对item（如作业，例程，文档，元数据等）不断复制（复制—粘贴）的需要。

“仆”项目与一个或多个“主”项目按参照引用的关系联系在一起，由此从一个或多个父项目中继承item。

在仆项目中，来自于主项目的资源以只读模式出现：它们只可使用和执行，但不能更改。

因为主仆项目之间建立了强大的连接关系，一旦有人在主项目中修改某个item，所有的仆项目都能得到相应更新。

引用项目分析一个项目中所有的冗余item（作业，模板，元数据）以便于其它项目使用。这一功能重新利用了30%的item，提高了使用率。这30%的item 对所有的数据集成项目往往是通用的，因而可以大幅度减少相关的维护需要。

变化数据捕捉

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx



数据仓库的实现，涉及到从一个或多个数据库提取数据，并搬运到一个或多个目标系统中来进行分析。但是，这也意味着要占用大量的时间和资源来提取和搬运大规模的数据。

实时捕捉且只捕捉有变化的数据的能力被称为**变化数据捕捉（CDC）**。只捕捉变化可以降低系统之间的数据交换量，从而缩短ETL的用时。

Talend CDC架构采用的是“发布—订阅”模式。发布者捕捉数据变化，并提供给订阅者（即作业）。订阅者使用从发布者那里得到的数据变化。这一功能实时侦测数据中发生的变化，将变化的数据立即发送给订阅作业，从而省去了ETL或数据集成操作过程中加载和更新数据需要的时间。

Talend 变化数据捕捉功能最常用的模式包括：触发，以及Redo（重做）log。哪个模式可用取决于要使用何种类型的数据库。

业务规则

this feature is included in:

Open

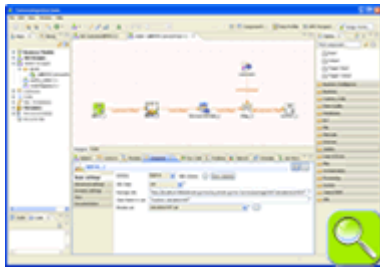
Team

Professional

RTx

Enterprise

MPx



业务规则通常有业务用户按照规格文档定义，然后由技术人员释译并实现。

Talend Integration Suite内嵌了一个业务规则引擎，帮助用户配置自己的业务规则。用户因而能够定义自己的市场划分标准（年龄，地区等），并通过Excel spreadsheet，或经由Talend Administration Center上的 Drools Guvnor 界面直接建立相应的业务规则。

Drools Guvnor 界面使业务专家可以使用图形化的编辑器来直截了当地建立和编辑规则，控制对规则和其它功能的访问，管理规则版本和及时修改。规则可以通过业已开发的作业进行测试和命名。（浏览下一个幻灯片）

API

this feature is included in:

Open

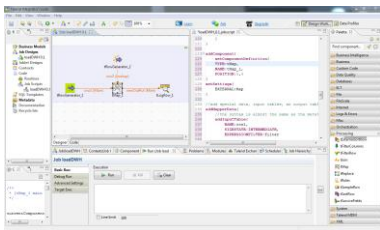
Team

Professional

RTx

Enterprise

MPx



Sophisticated but simple to grasp scripting language which enables programmatic job generation

Fully featured dedicated editor including syntax highlight, code completion aid & warning system

Automatic reverse engineering from an existing Job design to produce the corresponding script

测试

上下文管理

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx



上下文几乎使所有的组件或作业参数都可以外在化。因而用户或可以在作业运行中实时定义参数，或可以对测试和生产阶段使用不同的参数，抑或更多。

上下文可以更加不同的环境类型按照需要来定义（开发，测试，生产等），而上下文创建的数量则没有限制。

用户可以在设计阶段或运行阶段等任何时候切换上下文，使用业已定义的设置。

参数值可以在设计和测试阶段通过一个对话框来更改。另外，您也可以使用专门的参数加载组件动态地取代任何参数值。

远程运行

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx



远程运行功能可以通过Studio在任意服务器上远程执行作业。

这在需要测试作业的时候极为实用，例如：

- 使用与生产环境类似的配置。
- 使用任何种类的操作系统。
- 使用所要求的系统。

这就免除了复杂部署流程所带来的麻烦。

目标系统可以在运行过程中通过Studio动态选取。同时，在远程执行模式中，所有常规的调试，追踪和实时统计功能均可使用。

Talend Administration Center

this feature is included in:

Open

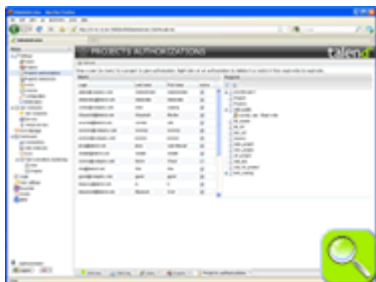
Team

Professional

RTx

Enterprise

MPx



订阅内容包括一个Studio（根据用户数量或可更多）部分和一个安装于服务器上，并通过基于web的界面**Talend Administration Center**来管理的软件部分。

所有的Studio均可以用**远程**模式连接在Talend Administration Center中定义的工程项目，不再局限于使用**本地**连接方式中。

Talend Administration Center是一款轻量级应用（在浏览器中使用，不需要部署），可以帮助集成项目的经理管理用户，工程，用户权限，许可证，等等。

工程项目的授权可以轻松地在不同用户当中分派（支持LDAP目录），角色不同，项目访问权限各异。这包括：无权（No permission），只读，读写，等等。

在用户被授权访问的工程项目范围内，他们可以与其他用户在Studio中直接分享repository存储库中各种item（如Jobs，业务模型，数据库连接元数据等）。如需了解更多信息，请查看后面的幻灯片。

随着您所订阅的Talend Integration Suite的版本的变化，数量繁多的附加plug-in插件会出现在左侧的导航面板中，供您取用。这些插件包括Dashboard，SOA管理器，Server 管理器等等。

Job Conductor

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx



Job Conductor协调执行多个数据集成作业，在中央执行界面上，按照需求，或者根据时间安排（Team Edition功能）或是基于事件（Professional Edition功能）来启动所有的作业。

Job Conductor运作的基石是“JobServers”，或者作为代理安装在执行作业的每一个服务器上的小应用程序。

代理安装以后，您可以使用Job Conductor来实时监控您所有的硬件资源（可用CPU，RAM，HD等），从而在最适用的服务器上，将作业的执行分布于网格中，而其原生的JMX支持可以监控逾40多个指数。因此轻松一次点击就可以将任何作业部署到任何服务器上。

Command Line

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx



用作业设计专家开发的集成进程可以通过**Command Line**模块在Talend Studio GUI之外部署，升级并执行。

Talend Command Line模块提供一整套的命令行选项，开发人员，管理员可以凭借它们轻松的进行批量操作。

Command Line几乎可以使用所有通过Talend Studio和Talend Administration Center使用的作业管理功能。比如，这些功能包括：更新作业属性，将项目等级提升为生产级别，导入/导出作业或作业组，等等。

有了Command Line功能，您可以轻松快捷地开展大量复杂的作业部署和执行工作，甚至涵盖了他们的依存关系，和执行使用的元数据。

原生的命令行帮助文件提供详尽的可用命令清单，以及对应的简短功能描述。

时间调度器

this feature is included in:

Open

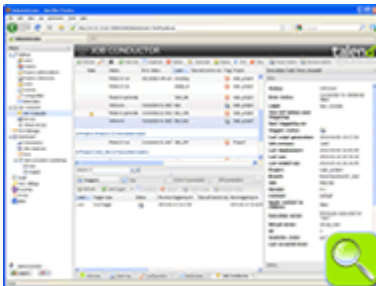
Team

Professional

RTx

Enterprise

MPx



时间调度器（Time-based scheduler）可以在所定义的时间和日期执行作业（当月第一个星期一，每个星期二，等等），或者在一段时间内对作业进行周期性规律执行。一个 Task则集中作业执行所需要的所有信息（工程名称，作业名称，作业版本，服务器，等等）。

任务按照调度安排得以触发，因而作业会在指定的时间和指定的服务器上部署并执行。一套方便实用的status（状态）系统可以通过Job Conductor直接监测触发状态和执行进行的成功或失败的情况。

从Professional Edition开始，又添加了基于事件或文件的调度功能。（见事件调度器幻灯片）

事件调度器

this feature is included in:

Open

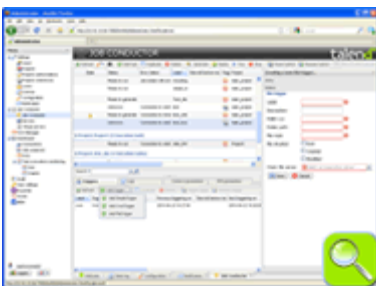
Team

Professional

RTx

Enterprise

MPx



事件调度器是为实现实时集成而对基于时间的调度功能的扩张。

时间监听器可以让进程的执行按照需求，或基于事件来触发一次执行。

所谓事件既可以是针对文件的操作如文件出现，消失，文件修改，也可以是基于SQL的“wait for”条件应用。一旦预计的时间得到确认，一次执行任务即会触发，作业的部署和执行随之进行。

您可以给任何任务轻松添加新的事件触发器，从而扩展自动化执行的产业化规模。

执行计划

this feature is included in:

Open

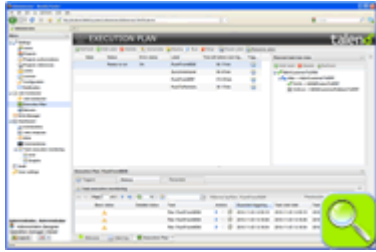
Team

Professional

RTx

Enterprise

MPx



执行计划功能可以直接通过Job Conductor将各种作业的执行任务排列，编排起来，轻松实现错误恢复（error recovery）。执行计划是一个面向任务的功能，可以勾勒出不同任务之间的依存关系，编排执行顺序。

任务依存关系在主任务层级视图中定义，每一个任务都可以有一个次级任务。

您也可以调度和触发多个执行计划。仅仅通过Job Conductor的视窗，这些执行计划就可以使用所有环境定义的执行参数。

SOA管理器/ Web 服务

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx



SOA管理器提供了一个基于web的完全图形化的环境，以服务（Web服务）的方式暴露一个或多个集成作业，这些作业因而能够使用 SOAP binding（PRC或基于文档），在异构的应用和系统之中抑或之间自动部署。一套专用的WSDL向导有利于生成WSDL描述器来把作业作为Web服务进行暴露，并且在消费Web服务的同时找出匹配的UDDI条目。

SOA管理器以优化的请求池和队列管理系统为基础，配备了先进的管理所接受的请求的功能。用户自定义的常用服务池可以即时处理大量最需要的请求，同时队列管理器（queue manager）则处理其它附加请求，从而缓冲吞吐量，以实现异步数据处理。

SOA管理器将在Talend Integration Suite Studio中设计的作业暴露在以Web服务为中心的环境中，从而建立了横跨多个应用和系统的互动。

均衡加载（Load Balancing）

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx



Grid Conductor（网格控制器）通过Job Conductor访问，确保最优化使用执行网格，从而最大限度的提高集成进程的扩展性和可用性。Grid Conductor按照虚拟服务器的定义运转，不受系统类型的限制（包括CPU，OS等）。

任务会分派给Grid Conductor的虚拟服务器执行，而不再依赖单一的一个执行服务器。

通过对所有执行服务器上可用资源的持续监测，Grid Conductor可以保证所有作业都在触发时顺利执行，同时可用资源得到充分利用，并避免了传统的单一服务器方法所引起的各种瓶颈。

这缓解了大量作业同时执行时，或者无专用服务器可用时所带来的资源优先级的困扰。当执行需要的资源不可用时，Grid Conductor还可以自动执行失败转移（fail-over）。

高可用性

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx



高可用性是通过部署多个Job Conductor和多个作业执行服务器实现的。

另外，通过建立起数据库簇（database cluster），您可以确保failover的执行，并防止执行被打断的情况发生。

失败转移 (Failover)

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx

FileScale

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx

Benchmark Results

Configuration

- 2 CPU Xeon 5520 Quad Core / 2.26GHz / 24 GB RAM
- Sun StorageTek 2540 Fiber Channel: 12 HD SAS 136Gb

Results

	1	10	100	300	550
# rows	6 million	60 million	600 million	1.8 billion	3.3 billion
data volume	740 MB	7.4 GB	74 GB	225 GB	415 GB
sorting time	6 sec	1 min 02 sec	21 min 30 sec	2 hr 01 min	5 hr 17 min

Independent benchmark made by Sun Micro

Talend Integration Suite

MPx拥有多服务器，多CPU以及多核的架构优势，代码和相互独立的次级进程可以并行执行，从而充分利用这种架构。这种大规模并行功能最大程度的提升了企业服务器的性能，并增加了可用处理器的数量，从而极大缩短了处理时间。

Talend Integration Suite

MPx的FileScale技术独树一帜，是兼顾了服务器硬件架构的高端技术，不但可以让您的硬件物尽其用，还可以发挥low-level排序算法的最佳

性能。FileScale技术实行bulk模式，处理大规模，甚至极大规模的文件。这种技术全面利用其执行架构的优势，摆脱了JVM或是执行引擎的限制，已经彻底革除了传统数据集成架构的特有缺陷。

FileScale

技术为数据处理运用了前沿的算法，注重提高给数据排序和转置的效力。它采用MapReduce映射化简架构，可以分解任意的数据处理操作，形成若干微粒进程，达到最快最好的处理目的。参见Sun Microsystem workbench：http://blogs.sun.com/aja/entry/talend_s_new_data_processing

Hadoop

this feature is included in:

Open

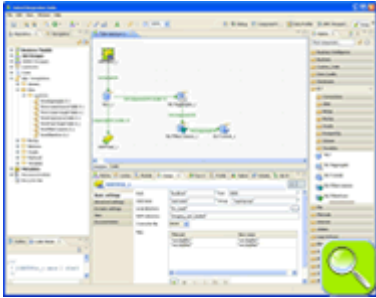
Team

Professional

RTx

Enterprise

MPx



Apache Hadoop是开源的Java软件框架，支持数据密集型的分布式应用。它采用MapReduce架构，使得各种应用能够使用低廉服务器组成的大规模网格，来处理数以千计的节点（node）和达到PB字节的数据。Talend Integration Suite MPx包含了对Hadoop的原生支持，因此可以扩展到任何级别，支持所有复杂的数据类型，所以企业能够充分利用他们的Hadoop簇，处理最极端规模的数据和复杂的数据转换。

Palette组件栏提供一套专用的组件帮助读写HDFS以及Hive系统，并提供ELT和SQLtemplate功能。

监测

AMC

this feature is included in:

Open

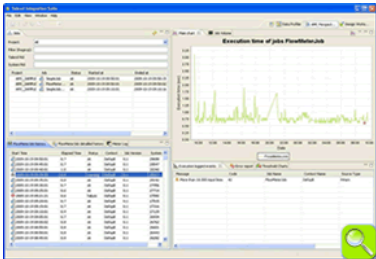
Team

Professional

RTx

Enterprise

MPx



Talend Activity Monitoring Console是一款便利的图像化界面，是一个中央监测工具。

它各种细致的监测功能可以整合所收集到的log信息，理解底层Job的互动，预防突发性错误，并支持系统管理方面的决策。

Activity Monitoring

Console只通过一个中央控制台就可以监测作业执行的事件（成功，失败，警告，等等），执行时间，数据规模。

该工具既可以独立使用，也可以完全融合在Studio中。

Dashboard

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx



Dashboard是Web版的Activity Monitoring Console，可以通过Web浏览器轻松使用。

Dashboard提供性能的线图和状态指数，任何相关人员都可以借此查看任何集成进程执行情况的当前和历史状态。

它各种细致的监测功能可以整合所收集到的log信息，理解底层Job的互动，预防突发性错误，并支持系统管理方面的决策。

错误恢复

this feature is included in:

Open

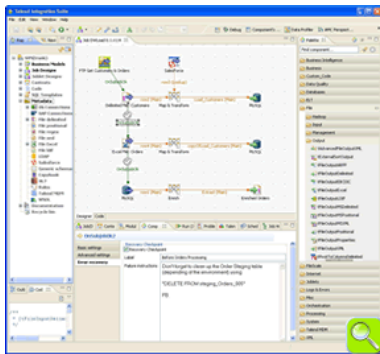
Team

Professional

RTx

Enterprise

MPx



如同备份与修复操作一样，作业执行的进程可能要耗费大量的时间。Talend Integration Suite Studio提供了在设计作业的过程中设置恢复点的功能。

一旦执行失败，任何进程都可以从某一个恢复点重新执行。作业开发人员还可以通过使用“on-failure”功能，专门针对一些具体的错误情况，设计并贯彻执行具体的错误管理方案。

恢复点可以在某些数据流之间进行合理设置（设置在连接组件的触发器上）。其目的是为了在作业执行的进程因为失败需要重新进行时，尽量节省要投入的时间和精力。

因为有了错误恢复点功能，任何进程都可以在失败之前最近的恢复点重新开始（当然也可以是失败之前的其它任意恢复点），从而不再需要回到作业执行进程的最开始。