

企业数据集成开发平台

SharkBuilder

产品白皮书

目 录

目 录	I
1. 前言	1
2. SHARKBUILDER 产品概述	1
2.1. SHARKBUILDER简介	1
2.2. SHARKBUILDER定位与价值	2
2.2.1. SharkBuilder产品定位	2
2.2.2. SharkBuilder产品价值	3
2.3. SHARKBUILDER产品功能架构	3
2.4. SHARKBUILDER特点	5
2.4.1. 参数化、配置化开发	5
2.4.2. 多种不同自动化程度的开发方式	6
2.4.3. 基于元数据的脚本自动生成	6
2.4.4. 一站式开发平台	7
3. 应用场景或案例	8
3.1. 某国有银行新一代数据仓库数据区迁移案例	8
4. 联系我们	8

1. 前言

大数据时代下，各企业都在进行数据线和信息管理系统建设。通过将企业生产经营中产生的数据进行集成整合，同时运用合适的查询、分析及数据挖掘工具对数据进行处理，将信息数据变为辅助决策的知识，提供给企业决策者运筹帷幄。在建设过程中，我们需要进行大量的数据集成开发，而数据集成开发无论采用何种开发方式或者是工具，我们都脱离不了数据的元数据进行。

元数据是描述数据的数据（data about data），描述数据的各类属性，数据的元数据不仅包括我们传统的数据的定义、结构、位置、数据关系、字段mapping关系，还包括数据的处理时间、控制条件和数据生成程序及实现方法等。大部分企业进行信息化建设过程中，都建立了元数据管理系统，去构建企业完整、一致的元数据知识库，去辅助数据开发和数据管理。企业元数据管理系统中元数据信息与企业数据实际情况越贴合一致，则支撑效果越好。但是实际建设中，往往会出现元数据与实际脱节的情况，造成的原因主要是：

- 1，数据开发环节多，周期长，在各环节交互时容易产生信息的不一致；
- 2，元数据涉及面比广，比较杂，如果未与开发结合，则需要投入大量精力进行登记元数据，由于信息繁琐，登记过程中容易不全或者存在问题；
- 3，开发规范未良好确定，元数据只是作为参考使用，设计人员和开发人员等对元数据不重视，造成元数据信息缺失严重；

这种情况往往会不断恶化下去，逐渐元数据变不可用，也无人再去参考元数据系统，这样企业元数据信息与数据实际情况完全脱节。

基于上述背景，我们认为在企业数据线建设进行数据开发时，需求开发过程与元数据紧密结合，互相驱动是必要的，以此我们提出了基于元数据驱动的数据集成开发模式，并以此研发工具进行支撑。

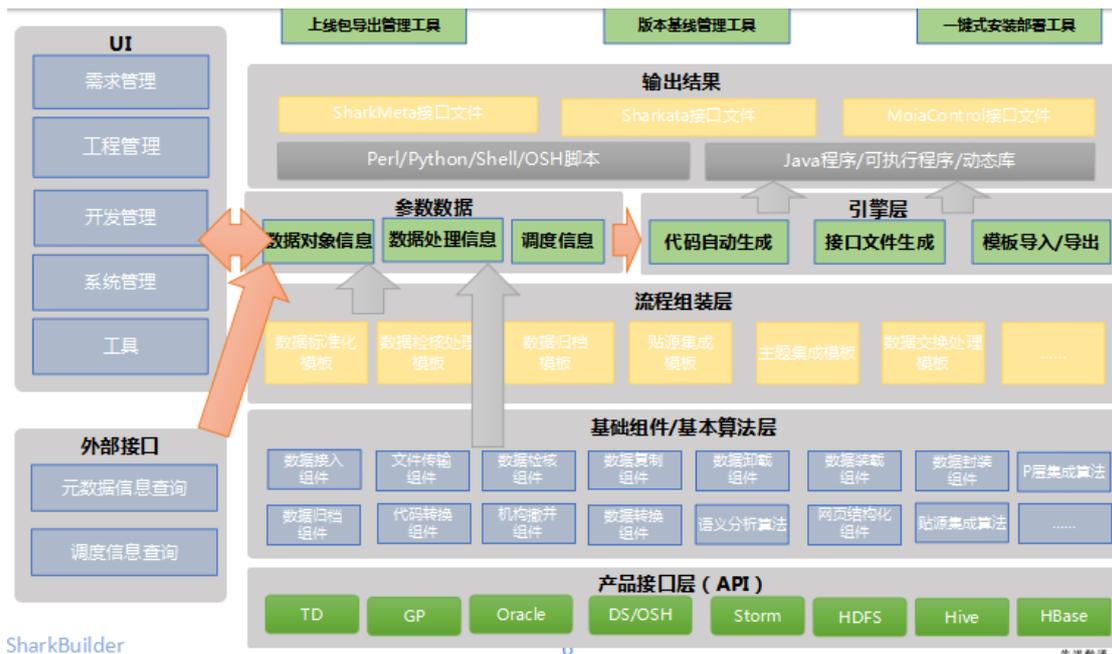
2. SharkBuilder产品概述

2.1. SharkBuilder简介

企业数据集成/数据应用开发管理平台 (SharkBuilder) 是一款用于支持企业数据集成需求开发管理的软件，是Sharkata产品族之一。它通过参数化、可视化

的机制实现数据集成的开发和管理，提供包括数据处理程序/脚本、数据的元数据信息、执行调度信息的一站式开发管理平台。

本产品提供规范、统一的企业数据集成开发管理，基于数据对象的视角，通过对数据对象的属性进行参数配置开发的方式，实现数据处理和数据流程的开发，在提升开发效率和开发质量的同时，保证企业的元数据、数据运行信息一致，提升对数据线的整体管理水平。



使用SharkBuilder可以完成以下工作：

- 1、提供一致的数据集成需求开发流程规范管控；
- 2、提供基于数据视角和面向对象的数据对象属性配置开发方式；
- 3、提供图形化的数据流程开发；
- 4、提供基于配置信息的自动化脚本生成、基于公共组件的参数配置和用户自定义开发的多种脚本开发方式；
- 5、提供通过配置数据信息自动生成产生数据对应脚本和程序；
- 6、提供基于配置的数据来源信息及数据流程的调度信息开发方式；

提供信息与元数据管理系统、调度信息和数据资产管理系统的同步功能。

2.2. SharkBuilder定位与价值

2.2.1. SharkBuilder 产品定位

SharkBuilder定位于提供企业构建良好的数据集成开发环境和流程管理，使

得企业元数据信息与数据开发不再脱节，良好互动，解决元数据信息脱节的历史难题，真正做到基于元数据驱动的开发模式，并通过工具建立规范有序的开发平台，提升开发效率和质量。

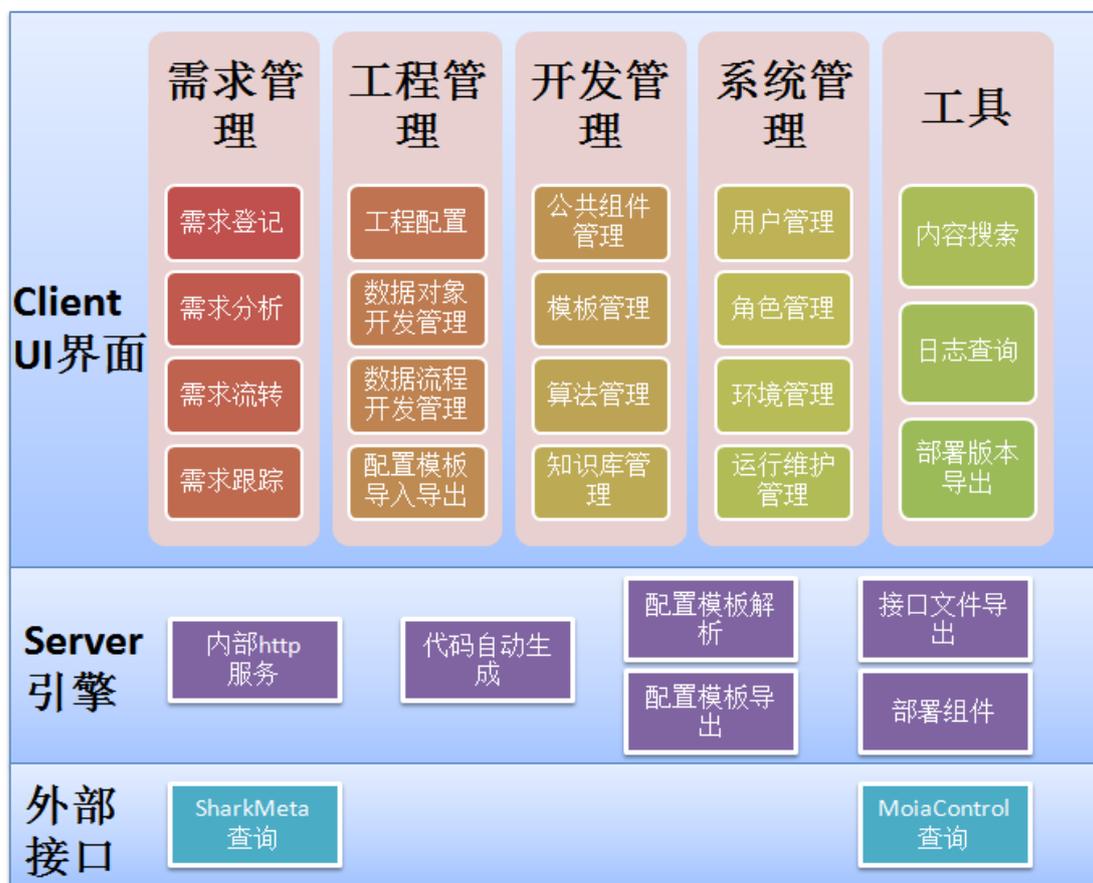
2.2.2. SharkBuilder 产品价值

基于SharkBuilder进行企业数据集成开发，首先提供参数化的开发方式，底层通过工具实现代码的自动生成，可以降低数据集成开发门槛，提高开发效率和开发质量；其次提供基于元数据驱动的数据集成全流程开发管理模式，实现元数据与开发全流程紧密结合，满足新增更新可追溯，易于维护和调整，提升总体开发管理水平；最终平台能够支撑数据及程序脚本、元数据信息、数据资产和调度信息的一站式开发，并以此为基础构建企业全局数据资产视图。



2.3. SharkBuilder产品功能架构

SharkBuilder作为数据开发平台，适合企业进行数据建设和数据开发，实现企业进行数据集成开发工作参数化、可视化，并同时同步实现元数据、调度信息和数据资产视图的配置信息的整合，在规范数据开发工作的同时提高企业数据开发效率。



SharkBuilder整个系统框架采用C/S架构，其中Client端的主要功能包括：

- 1、需求管理：提供对数据需求的登记、分配、分析和开发的过程实现全流程管理监控。
- 2、工程管理：提供面向对象式的数据开发，通过对数据对象的各个属性的开发配置，以及数据流程的开发配置，完成数据包括生成脚本、调度信息和元数据的一站式开发。最终将这些信息导出部署至对应系统中进行同步。
- 3、开发管理：提供对开发过程需要用到的组件、模板和算法进行登记和管理，并提供进行数据开发知识库的建立管理。
- 4、工具：提供对平台内容的查询，日志的查询管理和批量导入导出工具等。

其中Server端的主要功能包括内部服务，代码自动生成服务，配置模板解析和导出等。

同时提供接口向元数据和调度工具进行查询。

2.4. SharkBuilder特点

2.4.1. 参数化、配置化开发

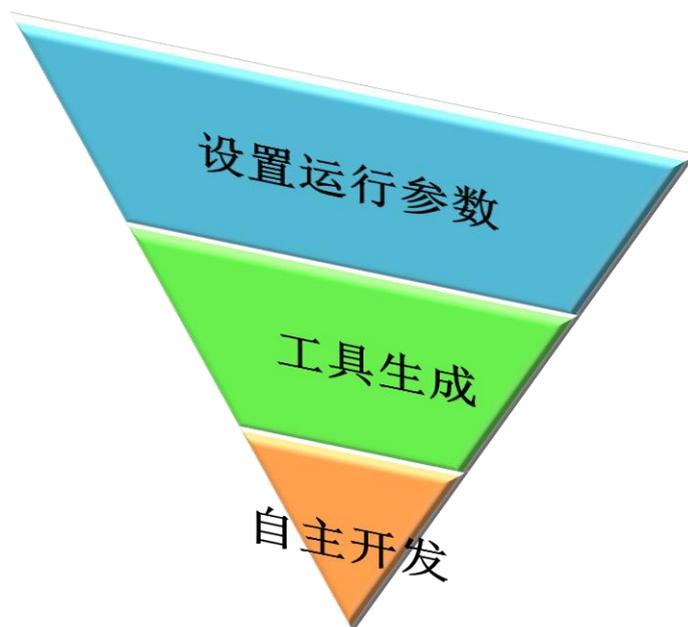
SharkBuilder与紧密结合元数据，通过将数据对象化，并将数据对象的属性设计归类为 4W1H:

- Who :数据定义
- What :数据的构成和来源关系
- When :数据的更新时间和条件
- Where :数据的位置
- How :数据的实现方法



而对于数据的开发就演变为对数据的属性的配置。

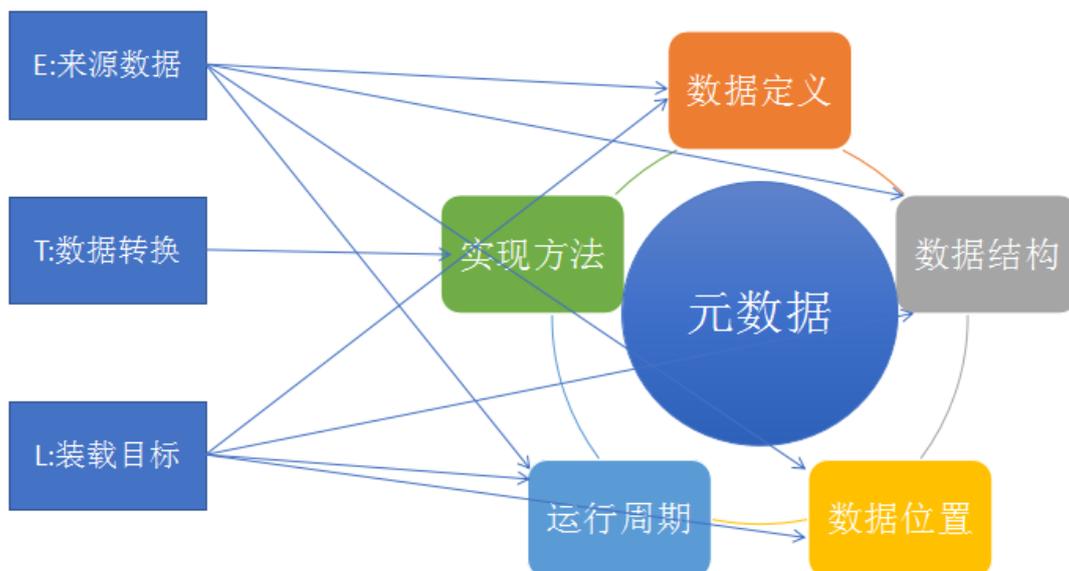
2.4.2. 多种不同自动化程度的开发方式



提供包括引用公共或自定义组件，设置运行参数的全自动化开发方式，通过数据对象属性信息，通过工具自动生成脚本程序的半自动化开发方式，以及自定义开发的手工开发方式。提高开发效率和开发质量的同时，不断完善规范，提炼组件和模板，不断促进开发向自动化方向迈进。

2.4.3. 基于元数据的脚本自动生成

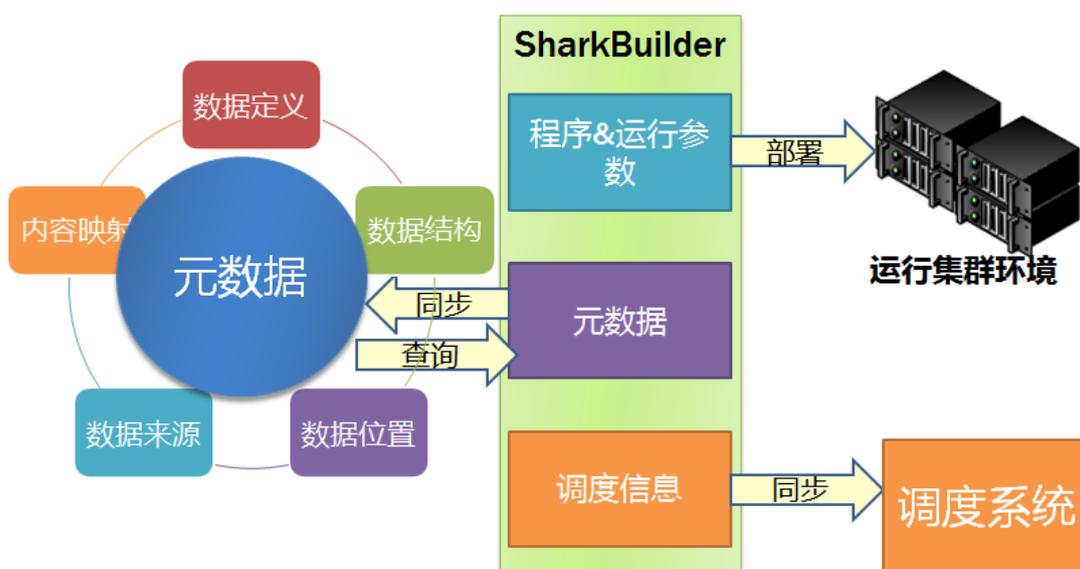
数据集成的过程解构就是一个ETL过程，以数据对象的方式已可以解构到我们对象化的数据对象元数据属性中，则ETL过程可以由元数据信息来表示，同样可以通过元数据信息生成对应的ETL脚本。



通过元数据配置，后调用脚本自动生成的开发方式，可以实现数据开发参数化、配置化，在规范开发同时提升开发质量，并可以屏蔽底层技术平台差异，支持数据处理在不同技术平台之间的快速迁移。

2.4.4. 一站式开发平台

SharkBuilder的开发主要围绕数据对象的属性信息的配置进行。而数据对象的属性信息分别涵盖了数据的元数据信息，数据关联作业的程序、运行参数及调度信息，数据定义信息。

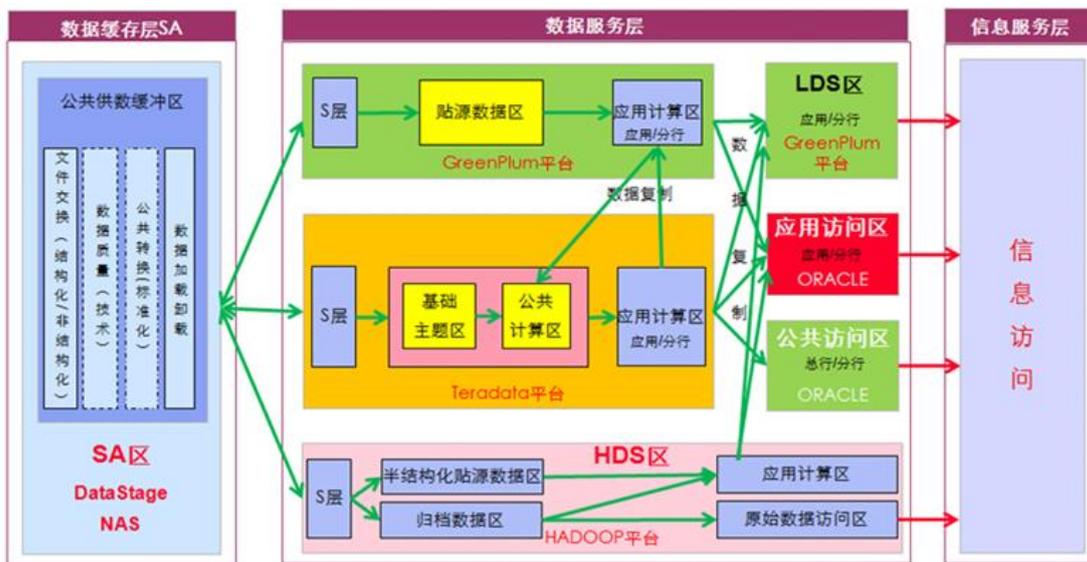


当完成数据对象的开发配置后，数据的脚本程序，数据的元数据、数据关联作业及调度信息相应都完成了，只需要向相应的系统进行对应的信息同步工作，

即实现数据所关系的系统的相关开发配置工作。

3. 应用场景或案例

3.1. 某国有银行新一代数据仓库数据区迁移案例



该银行采用多技术平台建设新一代企业级数据仓库。仓库中存在多个技术平台，包括TD、GP、Hadoop、Oracle、NAS。

其中基础主题区和公共计算区初期部署在TD平台上实现，在过程中出于对成本和资源快速准备的综合考虑，将基础主题区逐步迁移至GP平台中。其中在保证数据处理的一致的前提下，需要迁移工作的快速平稳进行。并且在迁移期间，面临着原有数据处理进行调整和变更的情况。

该行通过本产品来实现基础主题区和公共计算区数据的开发，通过数据的设计文档，填写产品的模板，最终导入生成脚本，在GP上实现原TD上的数据处理。用3个月时间实现初期规划的数据迁移，并在后续半年内完成整体的数据迁移。

4. 联系我们



北京先进数通信息技术股份公司

<http://www.adtec.com.cn>

电话： 010-68700075, 010-68700009

邮件支持：sharkata@adtec.com.cn

微信公众号：[adt-bigdata](https://www.adt-bigdata.com)

公司简介

北京先进数通信息技术股份公司（以下简称先进数通）是中国金融信息技术的重要参与者和领先企业。公司总部位于北京中关村科技园区，在上海、广州、成都、武汉、厦门、深圳、重庆设立了分公司，办事处及服务网点遍及全国 30 余个中心城市，形成了强大的支持服务网络。

自 2000 年成立以来，先进数通始终致力于以金融信息化为核心的软件解决方案、IT基础设施建设和IT运维服务。历经多年的稳健发展，业务已涵盖银行、保险、证券及企业等多个领域，为约 400 家金融企业及其他用户提供IT解决方案和专业服务，包括五大国有商业银行、股份制商业银行、区域银行及邮储银行、政策性银行、保险、证券、企业等。

秉持“诚信为本、艰苦奋斗、荣辱与共”的核心价值观，凭着稳健、进取的核心团队，领先的技术研发能力和自主知识产权，丰富的金融行业积淀和共创价值的合作伙伴，先进数通专注耕耘十余载，在金融前置与渠道应用、数据管理与商业智能应用、企业大数据平台、数据中心网络建设、行业云平台、企业网络及IT运维服务等领域打造出领先的综合竞争力，不断为客户提供高业务价值的产品和服务。

- ◇ 中国软件自主可靠企业核心软件品牌、中国软件行业（金融领域）优秀企业奖
- ◇ 最具投资价值方案商TOP50、十佳大数据方案商、十佳银行行业方案商
- ◇ 中国方案商百强企业、银行行业方案商十强
- ◇ 十大金融科技杰出企业奖、金融科技企业创新奖、十大金融科技企业用户信赖产品奖