

阿里云 实时计算（流计算）

产品简介

文档版本：20190816

法律声明

阿里云提醒您在阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。
2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。
3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。
4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的“现状”、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。
5. 阿里云网站上所有内容，包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含“阿里云”、“Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。
6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。

通用约定

格式	说明	样例
	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	 禁止： 重置操作将丢失用户配置数据。
	该类警示信息可能导致系统重大变更甚至故障，或者导致人身伤害等结果。	 警告： 重启操作将导致业务中断，恢复业务所需时间约10分钟。
	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	 说明： 您也可以通过按Ctrl + A选中全部文件。
>	多级菜单递进。	设置 > 网络 > 设置网络类型
粗体	表示按键、菜单、页面名称等UI元素。	单击 确定 。
<code>courier</code> 字体	命令。	执行 <code>cd /d C:/windows</code> 命令，进入Windows系统文件夹。
<code>##</code>	表示参数、变量。	<code>bae log list --instanceid Instance_ID</code>
<code>[]或者[a b]</code>	表示可选项，至多选择一个。	<code>ipconfig [-all -t]</code>
<code>{ }或者{a b}</code>	表示必选项，至多选择一个。	<code>swich {stand slave}</code>

目录

法律声明.....	I
通用约定.....	I
1 什么是阿里云实时计算.....	1
2 业务流程.....	4
3 应用场景.....	8
4 产品生态.....	11
5 产品安全.....	12
6 产品模式.....	14
6.1 产品模式概述.....	14
6.2 独享模式与共享模式的区别.....	16
6.3 独享模式公测转商业化.....	18
7 服务介绍.....	20
7.1 文档说明.....	20
7.2 法律说明.....	20
7.3 服务限制.....	21
8 发展历程.....	23
9 公告.....	25
10 Flink Forward Asia 2019 门票说明.....	35
11 文档最新动态.....	36
12 新功能发布记录.....	40
12.1 Blink-3.2.1.....	40
12.1.1 Blink-3.2.1新功能发布记录.....	40
12.1.2 Flink 3.2与Flink 1.5.1版本API兼容性报告.....	41
12.1.3 Blink 3.0与Blink 2.0版本SQL不兼容项目汇总.....	43
12.2 Blink-3.2.3.....	44
12.3 Blink-3.3.0.....	45

1 什么是阿里云实时计算

阿里云实时计算（Alibaba Cloud Realtime Compute）是一套基于Apache Flink构建的一站式、高性能实时大数据处理平台，广泛适用于流式数据处理、离线数据处理等场景。

产品特点

- 强大的实时处理能力

阿里云实时计算集成诸多全链路功能，方便您进行全链路实时计算开发，包括：

- 强大的流计算引擎。

1. 阿里云实时计算提供Flink SQL（详情请参见[#unique_4](#)），支持各类错误场景的自动恢复，保证故障情况下数据处理的准确性。
2. 支持多种内置函数，包括：字符串函数、日期函数、聚合。
3. 精确的计算资源控制，高度保证公共云用户作业的隔离性。

- 关键性能指标为开源Flink的3到4倍。数据计算延迟优化到秒级甚至亚秒级。单个作业吞吐量可做到百万（记录/秒）级别。单集群规模为数千台。
- 深度整合各类云数据存储。阿里云实时计算可以直接读写包括数据总线DataHub、日志服务LOG、云数据库RDS版、表格存储TableStore、分析型数据库MySQL版在内的各类数据存储系统，无需进行额外的数据集成工作。

- 托管的实时计算服务

不同于开源或者自建的流式处理服务，阿里云实时计算是完全托管的流式计算引擎。阿里云可针对流数据运行查询，无需预置或管理任何基础设施。在阿里云实时计算，您可以享受一键启用的流式数据服务能力。阿里云实时计算天然集成数据存储、数据开发、数据运维、监控报警等功能，方便您以较小成本试用和迁移流式计算。同时，实时计算提供完全租户隔离的托管运行服务。从最上层工作空间，到最底层执行机器，提供高度有效的隔离和全面防护，让您放心使用实时计算。

- 低廉的人力和集群成本

大量优化的SQL执行引擎，提供比原生Flink作业更高效且更廉价的计算作业。在开发成本和运行成本方面，阿里云实时计算均要远低于开源流式框架。例如，项目预算时您需要考虑如下成本：

- 编写一个复杂业务逻辑下Flink作业Java代码的人力成本。
- 针对、作业的调试、测试、调优、上线工作成本。
- 后续长期用于Flink、Zookeeper等开源软件的运维成本。

如果使用阿里云实时计算服务，上述问题交由阿里云平台承担，您可以专注于业务。

产品定位

- 目前实时计算适用应用场景

- 实时的网络点击PV、UV统计。
- 统计交通卡口平均时间段内（例如平均每5分钟）的车流量。
- 水利大坝的压力数据的统计和展现。
- 网络支付中涉及金融盗窃固定行为规则的告警。

- 目前实时计算无法实现的场景

- Oracle存储过程无法使用实时计算替换。实时计算无法从功能上完全替换掉Oracle存储过程，两者面向问题领域不一致。
- Spark作业无法无缝迁移至实时计算。Spark中涉及实时计算的部分，可以通过改造，完成从Spark至实时计算的迁移。完成迁移后您可以省去运维Spark和开发Spark等工作的成本。
- 实时计算无法实现多条复杂规则引擎的告警功能。如果单一数据存在多条复杂规则的告警，在系统运行的同时，告警本身也会发生变化。这类场景建议使用规则引擎系统解决，实时计算主要针对的不是此类问题。

当前实时计算对外接口定义为Flink SQL加UDF。实时计算提供服务于流式数据分析、统计、处理等应用场景的一站式开发工具。面向的用户包括数仓开发人员、数据分析师等。您通过编写Flink SQL，即可完成自身流式数据分析业务，不需要参与底层代码开发。

基本概念

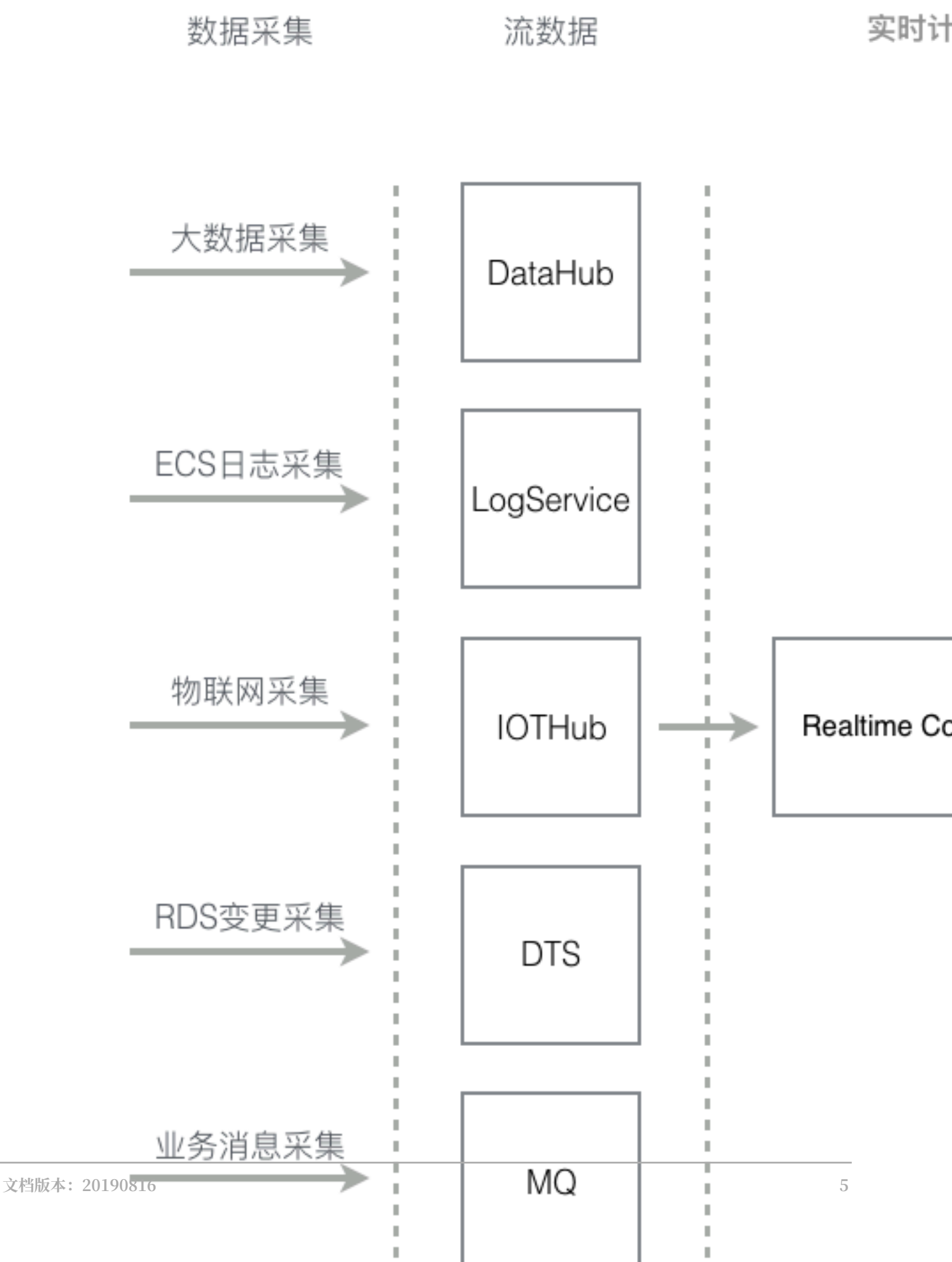
概念	说明
计算集群（Compute Cluster）	计算集群是承载实时计算产品计算任务的分布式集群系统，基于YARN模式。根据集群的形态不同，实时计算分为独享模式和共享模式，详细见 产品模式概述 。

概念	说明
开发界面（WebConsole）	实时计算提供了一套完整在线IDE开发工具，一站式集成数据存储、数据开发、数据运维、监控报警等功能，辅助您进行业务开发。
项目空间（Project）	项目空间是实时计算最基本的业务组织单元，是您管理集群、作业、资源、人员的基本单元。您可以新建项目，也可以以子账号身份加入其它项目空间。  说明： 实时计算的项目空间，通过阿里云RAM主子账号，支持多人同时操作。
实时计算单元（CU）	在实时计算中，作业的实时计算单元为CU。1 CU描述了1个实时计算作业最小运行能力，即在限定的CPU、内存、I/O情况下对于事件流处理的最小能力。1个实时计算作业可以指定在1个或者多个CU上运行。 当前对实时计算单元（CU）运行能力的定义：1 CU=1 CPU + 4G MEM。其处理能力约为： · 对于简单业务，比如单流过滤、字符串变换等操作，1 CU每秒可以处理10000条数据。 · 对于复杂业务，比如join、窗口、group by等操作，1 CU每秒可以处理1000到5000条数据。  说明： 实时计算单元（CU）的具体处理能力也与您的业务实际情况相关。

2 业务流程

本文为您介绍阿里云实时计算业务流程系统架构。

实时计算业务流程系统架构图如下。



1. 数据集采集

广义的实时数据采集指使用流式数据采集工具将数据实时地采集并传输到大数据Pub/Sub系统。Pub/Sub系统将为下游实时计算提供源源不断的事件源去触发流式计算作业的运行。阿里云大数据生态中提供了针对不同场景领域的流式数据Pub/Sub系统。阿里云实时计算天然集成上图中诸多Pub/Sub系统，能够集成各类流式数据。



说明：

例如：您可以直接使用实时计算对接SLS的LogHub系统，以做到快速集成并使用ECS日志。

2. 流式计算

流数据作为实时计算的触发源驱动实时计算运行。一个实时计算作业至少使用一个流数据作为数据源。同时，对于复杂的业务场景实时计算还支持和静态数据存储进行关联查询。



说明：

例如，针对每条DataHub流式数据，实时计算将根据流式数据的主键和RDS中数据进行关联查询（即join查询）。

3. 实时计算集成

阿里云实时计算将计算的结果数据直接写入目的数据存储，从而最大程度降低全链路数据时延，降低数据链路复杂度，保证数据加工的实时性。为了打通阿里云生态，阿里云实时计算天然集成了OLTP（RDS产品线等）、NoSQL（OTS等）、OLAP（ADB等）、MessageQueue（DataHub、ONS等）、MassiveStorage（OSS、MaxCompute等）。

4. 数据消费

流式计算的结果数据进入各类数据存储后，您可以运用个性化的应用，操控结果数据。如使用数据存储系统访问数据，使用消息投递系统接受信息，或者使用告警系统生成异常结果数据警报。

数据链路情况

对于上图的数据链路，部分数据存储由于和实时计算模型不能一一匹配，需要使用其他类型的流数据做中转，说明如下：

· DataHub

DataHub提供了多类数据（如日志、数据库BinLog、IoT数据流等）上传到DataHub的工具、界面，以及开源、商业软件的集成。请参见[DataHub使用文档](#)对DataHub做更详尽了解。

- LogService

LogService（日志服务）是针对日志类数据的一站式服务。LogService提供了诸多针对日志的采集、消费、投递、查询分析等功能。请参见[LogService采集方式](#)，了解如何使用日志进行流式数据采集。

- IoTHub

IoTHub（物联网套件）是阿里云专门为物联网领域的开发人员推出的。其目的是帮助开发者搭建安全的数据通道，方便终端（如传感器、执行器、嵌入式设备或智能家电等等）和云端的双向通信。使用IoTHub规则引擎可以将IoT数据方便投递到DataHub，并利用实时计算和MaxCompute进行数据加工计算。请参见[IoT设置规则引擎](#)，了解如何将IoT数据推送到DataHub。

- DTS

DTS（数据传输）支持以数据库为核心的结构化存储产品之间的数据传输。DTS是一种集数据迁移、数据订阅及数据实时同步于一体的数据传输服务。使用DTS的数据传输功能，可以方便您将RDS等BinLog解析并投递到DataHub，并利用实时计算和MaxCompute进行数据加工计算。请参见[MySQL 到 DataHub 数据实时同步](#)，了解如何使用DTS进行流式数据采集。

- MQ

阿里云MQ服务是一套完整的消息云服务。阿里云MQ服务基于高可用分布式集群技术，搭建了包括发布订阅、消息轨迹、资源统计、定时（延时）、监控报警等功能。

3 应用场景

实时计算广泛应用于大数据实时化的场景，从简单到详细，可以按照已有流处理系统、部门场景和技术领域进行划分。

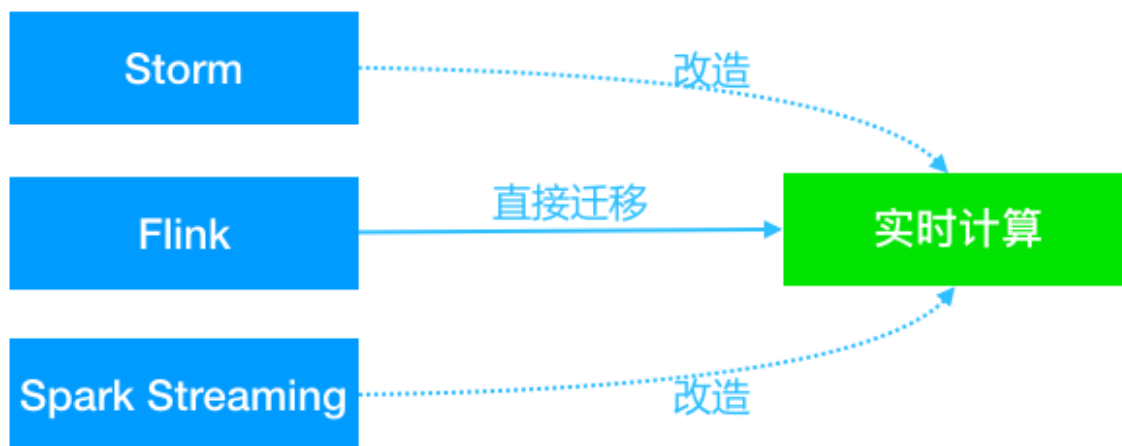


说明：

更多场景案例请参见[阿里云实时计算产品案例&解决方案汇总](#)。

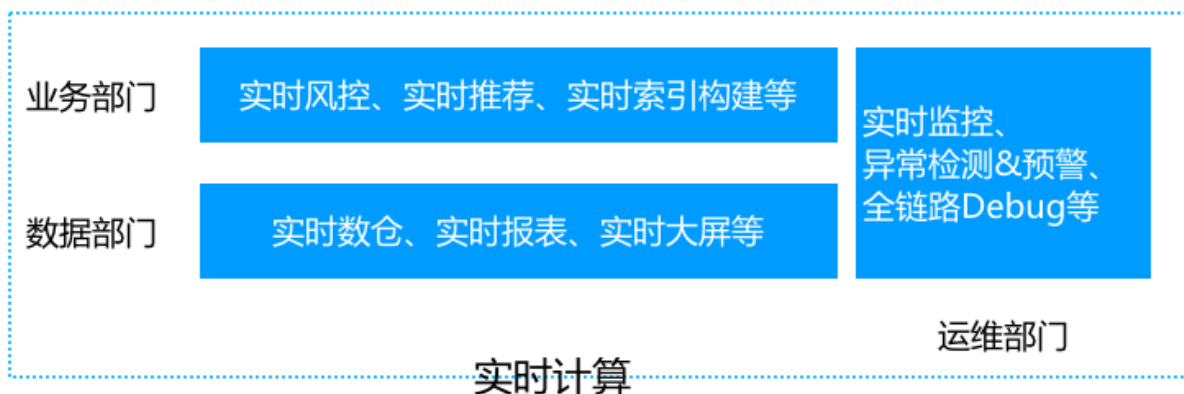
已有流处理系统

最基本的，如果您本地已经有Flink、Storm、Spark Streaming系统，一般这部分是可以直接迁移到实时计算产品上。具体的迁移细节请与实时计算团队联系。



按照部门场景划分

从企业部门职能的角度可以将实时计算的划分为以下场景：



- 业务部门：实时风控、实时推荐、搜索引擎的实时索引构建等。
- 数据部门：实时数仓、实时报表、实时大屏等。

- 运维部门：实时监控、实时异常检测和预警、全链路Debug等。

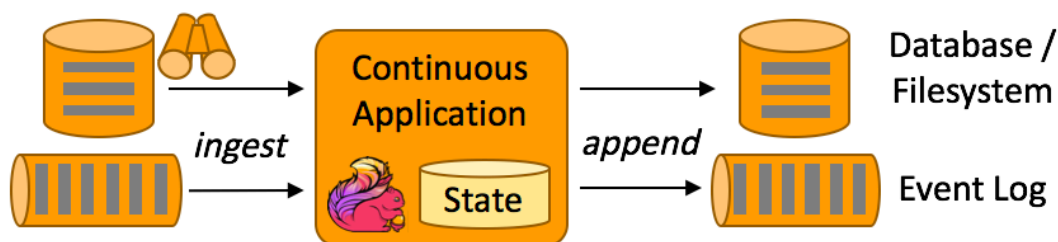
按照技术领域划分

从技术领域来看，实时计算主要用于以下场景：

- 实时ETL&数据流

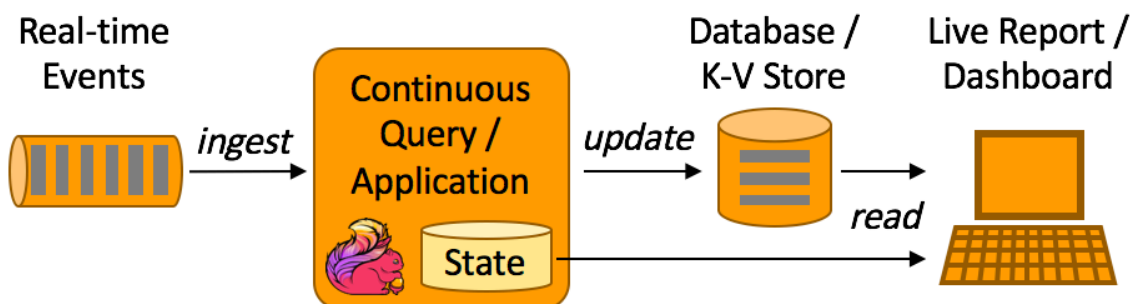
实时ETL&数据流的目的是实时的把数据从A点投递到B点。中间可能会加上一些数据清洗和集成的工作，比如实时构建搜索系统的索引，实时数仓中的ETL过程等。

Real-time Events



- 实时数据分析

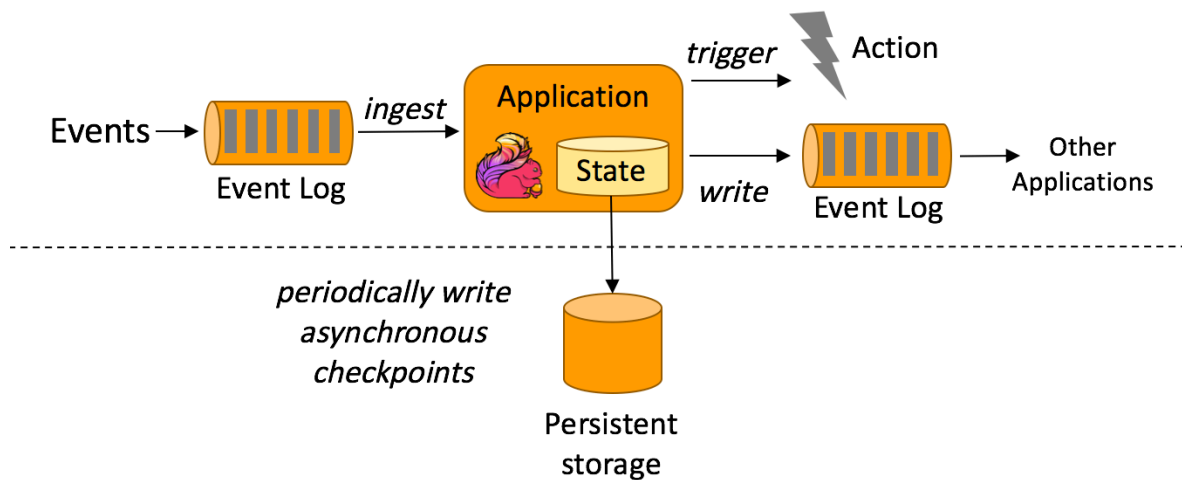
数据分析指的是根据业务目标从原始数据中抽取对应信息并整合的过程。比如查看每天卖的最好的十种商品，仓库平均周转时间，文章平均点击率，推送打开率等等。实时数据分析则是上述过程的实时化，一般最终体现为实时报表或实时大屏。



- 事件驱动应用

事件驱动应用是对一系列订阅事件进行处理或作出响应的系统。事件驱动应用往往还会依赖内部状态，比如点击欺诈检测，风控系统，运维异常检测系统等。当用户的行为触发某些风险控制点

时，系统会捕获这个事件，并根据当前行为和用户之前的行为进行分析，决定是否对用户进行风险控制。



说明:

以上技术领域场景流程图来自<https://flink.apache.org/usecases.html>。

4 产品生态

实时计算支持丰富的上下游生态。

实时计算当前支持的上下游生态

- 数据源表
 - [#unique_9](#)
 - [#unique_10](#)
 - [#unique_11](#)
 - [#unique_12](#)
- 数据结果表
 - [#unique_13](#)
 - [#unique_14](#)
 - [#unique_15](#)
 - [#unique_16](#)
 - [#unique_17](#)
 - [#unique_18](#)
 - [#unique_19](#)
 - [#unique_20](#)
 - [#unique_21](#)
 - [#unique_22](#)
 - [#unique_23](#)
- 数据维表
 - [#unique_24](#)
 - [#unique_25](#)
 - [#unique_26](#)
 - [#unique_27](#)



说明:

若有其它上下游对接需求，请联系实时计算产品技术支持。

5 产品安全

实时计算支持整体全链路实时计算的安全，包括账号安全，业务安全以及数据安全。

账号安全分为实时计算账号安全和数据存储账号安全：

- 实时计算账号安全

实时计算账号当前仅支持阿里云账号体系（包括登录用户名+密码或签名密钥）。传输链路使用HTTPS协议，保证全链路的用户账户安全。实时计算账号安全详情，请参见[账号管理](#)。

- 数据存储账号安全

针对将数据存储保存到连接账号的问题，实时计算提供基于RAM或STS的两种连接方式。避免您因为账户信息丢失导致的业务信息泄露。数据存储账号安全详情，请参见[角色授权](#)。

业务安全

实时计算的业务安全部分主要包含项目隔离安全和业务流程安全：

- 项目隔离安全

实时计算对不同的项目进行了严格的项目权限区分。不同用户或项目之间是无法进行跨越项目权限的访问、操作项目内的所有子产品实体。项目级别的资源隔离能够保证您与其他用户的资源使用情况相互之间不受干扰。



说明：

例如，一个用户的作业在运行期间，由于数据量的突增，提升了这个作业对CPU使用率。这种情况下，由于实时计算在底层使用虚拟化技术进行资源隔离，保证该用户的作业CPU使用率增加，不会影响到其他用户作业的CPU使用情况。

- 业务流程安全

实时计算对于流式计算开发进行了严格的流程定义，区分了数据开发和数据运维。保证了整体业务流程的完整和安全性。

- 提供代码版本

支持代码版本回滚和对比。方便您进行代码追溯、比对、排错。

- 提供IDE单机调试容器

避免代码线下运行影响线上真实数据。您可以自行构造线下输入表、维表、输出表相关数据，不影响线上生产作业。

- 提供发布流程

发布流程避免线下代码改动直接影响生产运行。线下代码调试完成后，通过上线作业将作业提交到数据运维系统。已运行的实时计算作业并不会直接使用新代码，需经过您确认后，停止已运行作业并使用新代码重新运行。从流程上保证发布的严谨性。

数据安全

数据安全分为实时计算系统数据安全和业务数据安全，阐述如下。

- 系统数据安全

实时计算系统保证自身数据安全：

- 使用HTTPS访问方式，确保传输链路的安全。
 - 数据存储使用AES高强度加密方式连接信息，避免敏感信息泄露。
 - 实时计算系统通过全面且深入的攻击测试，并有阿里云安全团队为实时计算提供安全服务。

- 业务数据安全

实时计算本身不负责存储用户的业务数据。具体业务数据安全交由不同的阿里云存储系统保证。请您参考不同的数据存储的安全模型以及最佳安全实践做更多了解。

6 产品模式

6.1 产品模式概述

阿里云实时计算支持共享模式和独享模式2种产品模式。独享模式是基于共享模式的补充，具备更加丰富的功能。



说明：

共享模式和独享模式详细的区别对比参见[#unique_30](#)。

共享模式

不同用户共享计算集群的物理资源（网络、磁盘、CPU或内存等），通过账号管理、CGroup（Control Groups）等方式进行资源隔离和安全管理。基于账号安全、业务安全和数据安全方面的考虑，共享模式不提供自定义函数功能。



说明：

自定义函数详情，请参见[#unique_31](#)。

独享模式

· 独享模式优势

独享模式是指在阿里云ECS上单独为用户创建独立计算集群，单个用户独享计算集群的物理资源（网络、磁盘、CPU或内存等），与其他用户的资源完全独立。与共享模式相比，独享模式具有以下优点：

- 丰富的硬件

实时计算独享集群可充分复用阿里云在CPU：MEM配比、GPU或FPGA等硬件层面的各类优化，为您解决各类硬件适配问题。

- 用户间的隔离

如果使用ECS独享集群，您能够使用专有网络VPC和独享计算资源。既能满足您对专网专用、资源独享的需求，也能够与您的开发平台连通，满足您的业务需求。

- 支持自定义函数

独享模式在网络及物理机层面与其他用户完全的隔离，支持自定义函数等更底层的API，满足您的业务需求。

- 丰富的功能

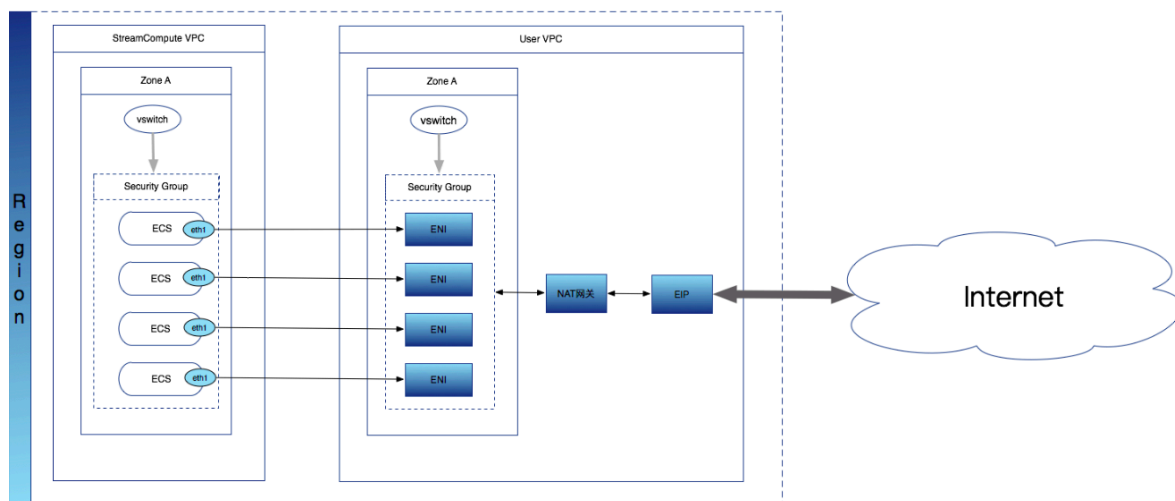
■ Data Lake场景下的ETL：通过Flink SQL+UDF的方式，使ETL任务开发更加便利。

■ 异构数据源计算：支持从异构数据源读取数据做分析。例如，远程从OSS读取数据归档日志，并JOIN HBase中高危IP，完成网络攻击分析等。

■ 支持更加丰富的数据存储：例如，[#unique_33](#)和[#unique_34](#)。

· 独享模式系统网络架构

独享模式系统网络架构说明如下：



- 实时计算独享模式为全托管模式。您所购买的ECS托管在实时计算VPC下，暂不提供用户登陆接口。如果您需要登陆ECS，请提交工单。
- 为了访问VPC内服务，集群创建时需要在您账号下申请弹性网卡。通过弹性网卡，您可访问其VPC下所有资源。
- 如果需要访问公网，可在弹性网卡上绑定NAT网关及弹性公网IP，具体操作步骤请参见[#unique_35](#)。
- 若您需访问VPC内其他安全组的服务，请配置相应安全组的规则。



说明：

不访问公网时，弹性网卡不会产生任何额外费用。

6.2 独享模式与共享模式的区别

本文为您介绍实时计算独享模式和共享模式之间的区别。

独享模式与共享模式区别列表



说明：

独享模式实时计算集群仅能访问同一VPC、同一Region下的上下游存储资源。若需访问其他VPC下的资源，请配置安全组的端口号，通过高速通道等方式打通网络。

	独享模式	共享模式
自定义功能开发	支持UDF和API等功能，作业开发更加灵活。	不支持

网络类型	支持阿里云VPC网络	不支持
机器类型选择	购买界面支持机器类型选择。支持GPU、FPGA（待上线），可为后续的AI做准备。	不支持
支持数据源类型	· 数据总线（DataHub） · 日志服务（Log Service） · 消息队列（MQ） · 消息队列（Kafka）  说明： 更多实时计算数据源表介绍，请参见#unique_37。	· 数据总线（DataHub） · 日志服务（Log Service） · 消息队列（MQ）  说明： 更多实时计算数据源表介绍，请参见#unique_37。
支持数据输出类型	· 分析型数据库（AnalyticDB） · 数据总线（DataHub） · 日志服务（Log Service） · 消息队列（MQ） · 表格存储（Table Store） · 数据库（RDS/DRDS） · 高性能时间序列数据库（HiTSDB） · 云数据库（HybridDB for MySQL） · Kafka · 云数据库（HBase） · ElasticSearch（ES）  说明： 更多实时计算数据源表介绍，请参见#unique_38。	· 分析型数据库（AnalyticDB） · 数据总线（DataHub） · 日志服务（Log Service） · 消息队列（MQ） · 表格存储（Table Store） · 数据库（RDS/DRDS） · 高性能时间序列数据库（HiTSDB） · 云数据库（HybridDB for MySQL）  说明： 更多实时计算数据源表介绍，请参见#unique_38。

支持地域	· 华东1（杭州） · 华东2（上海） · 华南1（深圳） · 华北2（北京） 如果对其它地域有开通需求，请联系阿里云技术支持。  说明： 独享模式支持地域介绍，请参见服务限制。	· 华东2（上海） · 华南1（深圳） · 华北2（北京）  说明： 共享模式支持地域介绍，请参见服务限制。
形态	每个用户在独立的ECS小集群上，搭建独立的实时计算引擎。	大集群，共同使用公共资源池。
隔离性	隔离性强。使用ECS隔离网络、安全组等资源，与其他用户的资源相互独立。	隔离性较弱。无法隔离网络等资源。
面向客户群体	具备较强开发技术，对开发的灵活性和可控性具备需求的大数据团队。	仅需实现流式业务，对开发过程无特殊需求的团队或个人。
成本	相对共享模式，独享模式成本较高。高可用模式下每个用户需单独承担额外的3台Master机器的费用。	仅需支付计算服务费用
集群管理和运维	需要具备集群管理、网络配置等技能。	无需管理集群

6.3 独享模式公测转商业化

本文为您介绍独享模式公测转商业化方案。

经历了四个月的邀测后，实时计算独享模式已经成功完成商业化上线。商业化后，您将在独享模式享受到更专业的支持服务和可用性SLA保障。

- 您可以通过以下方案完成公测转商业化：

- 代金券的方式

试用服务目前已经关闭，需要试用的新用户可以联系您公司对接的阿里云销售，通过代金券的方式购买实时计算独享模式来试用。若没有对接的阿里云销售，欢迎联系实时计算团队帮助您解决。

- 集群扩容或者迁移

想续使用的用户，请联系实时计算团队帮助您进行集群的扩容或者迁移，整个过程中支持和技术团队全程跟进，减少对您业务的影响。

- 如果您暂时不考虑继续使用独享模式，也请您联系实时计算团队，清除您已有的业务和数据，保证您信息的安全。公测阶段的设备将于2018年11月12号全部收回，请各您在此之前联系实时计算团队，实时计算钉钉号21772761。可以咨询您对接的阿里云销售。

7 服务介绍

7.1 文档说明

本文对实时计算文档内容做总体说明，并介绍本文档适用对象。

文档主旨

本文档主要对阿里云实时计算产品使用 and 开发所涉及的相关内容进行统一的文档说明，为您进行大数据开发工作提供指导。文档主要内容包括产品简介、产品定价、准备工作、快速入门和详细的开发使用指南，提供常见应用场景的解决方案和最佳实践，并针对常见问题提供FAQ方便您排查问题。

适用对象

本文档适用于以下对象：

- 实时计算开发人员
- 数据仓库工程师
- 数据分析人员
- 相关技术选型决策方

7.2 法律说明

阿里云提醒您在阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档的，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本文档内容或提供给任何第三方使用。
2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。
3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。
4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以大数据集成服务平台的“现状”、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠

性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。

5. 阿里云网站上所有内容，包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含“阿里云”、“Aliyun”、“AliCloud”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）详情请参阅：[阿里云产品服务协议](#)。
6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。
7. 为了保证用户的权益，我们提供了阿里云流计算服务等级协议相关条款。详情请点击：[阿里云流计算服务等级协议](#)。

7.3 服务限制

本文介绍实时计算所支持的服务范围和相应的限制，包括CU处理能力的限制、项目创建的限制。

- 实时计算当前支持地域如下。实际支持地域以实时计算官方购买页面为准。
 - 共享模式:华东2（上海）、华南1（深圳）、华北2（北京）。
 - 独享模式:华东1（杭州）、华东2（上海）、华南1（深圳）、华北2（北京）。
- 实时计算当前在我们内部压测场景下，一个CU的处理能力大约如下：
 - 对于简单业务，比如单流过滤、字符串变换等操作，1CU每秒可以处理10000条数据。
 - 对于复杂业务，比如join、窗口、group by等操作，1CU每秒可以处理1000到5000条数据。
- 实时计算共享模式暂不提供UDF功能。如果需要支持UDF，请使用实时计算独享模式。

- 实时计算对整个Project下属的Task、Task版本、IDE打开Task页面数量均有不同限制。包括：
 - 单个Project下允许最多创建作业的个数为100。
 - 单个Project下最多允许文件夹的个数为50，深度最大不超过5层。
 - 单个Project下最多允许UDF/JAR个数为50。
 - 单个Project下最多允许数据存储注册的个数为50。
 - 单个Task最多允许历史保存版本数为20。
- [实时计算控制台](#)仅支持Chrome浏览器访问。



说明：

请您仔细评估上述限制对于各您业务的影响情况。如有疑问，请与实时计算产品及时沟通，确认产品是否满足您的业务需求。

8 发展历程

本文介绍阿里云实时计算的发展历程。

阿里云实时计算在原有Flink系统基础上，提供一整套的开发平台，完整的流式数据处理业务流程。受益于阿里大数据多年的技术和业务沉淀，在您使用实时计算时可以享受到阿里巴巴集团前沿的计算引擎能力，业务上可规避阿里巴巴集团多年在流式大数据的试错和教训，让您可以更快、更轻松地实时化大数据处理流程，助力业务发展

- 起源：脱胎于双十一实时大屏业务

实时计算脱胎于阿里巴巴集团内部双十一实时大屏业务，在阿里巴巴集团内部从最开始支持双十一大屏展现和部分实时报表业务的实时数据业务团队，历经多年的长期摸索和发展，到最终成长一个独立稳定的云计算产品团队。实时计算定位将阿里巴巴集团本身沉淀多年的实时计算产品、架构、业务能够以云产品的方式对外提供服务，助力更多中小企业实时化自身大数据业务。

- 萌芽阶段：开源Flink作为基础

最初阿里巴巴集团支撑双十一大屏等业务同样采用的是开源的Flink作为基础系统支持，并在上面开发相关Flink代码。这个时期的实时业务处于萌芽阶段，规模尚小。数据开发人员使用Flink原生API开发流式作业，开发门槛高，系统调试难，存在大量重复的人力工作。

- 发展阶段：基于Flink的API开发

阿里巴巴集团的工程师针对这类大量重复工作，开始考虑进行业务封装和抽象。工程师们基于Flink的API开发出大量可复用的数据统计组件。例如，实现了简单过滤、聚合、窗口等等作为基础的编程组件，并基于这类组件提供了一套XML语义的业务描述语言。基于这套设计，流式计算用户可以使用XML语言将不同的组件进行拼装描述，最终完成一整套完整的实时计算处理流程。基于XML+Flink组件的编程方式，从底层上避免了用户大量的重复开发工作，同时亦降低了部分使用门槛。但我们的数据分析人员仍然需要熟悉整套编程组件和XML描述语法，这套编程方式离分析人员最熟悉的SQL方式仍然差距甚远。

- 现阶段：Flink SQL开发完成

任何技术的发展一定遵循 小众/创新 到 大众/普及 的成长轨迹。而从小众到大众，从创新到普及的转折点，一定在于技术的功能成熟和成本降低。阿里巴巴工程师开始思考如何更大程度降低数据分析产品门槛，从而普及到更多的用户。得益于用户群对关系型数据库几十年的沉淀。阿里工程师最终开发一套Flink SQL替换了原有的XML+组件的编程方式，使用经典的SQL模式去计算和处理数据。这套系统成为今天实时计算的核心计算引擎(Flink)。当前这套系统以单机群数千

台机器规模。日均消息处理数千亿级，流量近PB级别，成为阿里巴巴集团最核心的流式计算集群。

Flink SQL的优势

- 可以对标SQL功能。从而提高开发人员的技术成熟度。
- 可以利用用户熟悉的SQL模型。可极大降低用户上手使用实时计算的门槛。

9 公告

本文为您介绍实时计算公告内容，包括版本更新、功能更新、产品活动等。

2019年5月30日-实时计算3.0.0以上版本新功能

- 运行信息

新增Vertex相关信息查询功能，详情请参见[#unique_46](#)。

- 数据曲线

新增AutoScaling相关曲线，详情请参见[#unique_47](#)。

- Timeline

新增Timeline功能，详情请参见[#unique_48](#)。

- 属性参数

新增AutoScale迭代的历史详情查询功能，详情请参见[#unique_49](#)。

2019年5月29日-北京、上海地区实时计算共享模式停止售卖

2019年3月14日-共享模式默认版本升级为2.2.7

自3月14日起，实时计算共享模式默认版本升级为[2.2.7](#)，此次升级仅影响新增作业，不影响原有运行作业。

2019年1月24日-版本发布blink-2.2.7

`blink-2.2.7` 是**blink-2.x**系列中最新稳定版本，在**blink-1.x**版本（最新稳定版本为**blink-1.6.4**）进行了全面的升级，采用了自主研发的新一代存储Niagara作为Statebackend的底层存储，优化了SQL的性能，增加了一系列新功能。

- 主要特性

- SQL

- 新增Window Emit机制，可以控制window结果的输出策略，如：1小时窗口，每1分钟输出一次。
- 双流join在功能上支持了miniBatch，针对不同场景优化了retraction处理和state存储结构，提高了性能。
- AGG 支持 Filter 语法，可以只聚合满足条件的行。
- 对Local-global AGG进行优化。
- 重构了SQL的optimize阶段，解决了SQL编译时间过长的若干问题。
- 增加了对SortedMapView中key的多种数据类型的支持：Boolean、Byte、Short、Int、Long、Float、Double、BigDecimal、BigInt、byte[]和String。
- 优化了min/max/first/last等在有retraction场景时的性能。
- 新增若干标量函数，例如时区相关的解析、格式化、转换函数：TO_TIMESTAMP_TZ、DATE_FORMAT_TZ和CONVERT_TZ。
- 对SQL，Connector模块的错误信息进行了归类，并对每种类型设计了相应的ERROR_CODE。

- Connector

- 支持用户自定义的tableFactory方式注册source和sink插件。
- 支持用户通过udtf方式直接解析数据源类型。
- 支持读取和写入kafka。
- 支持写入到ElasticSearch。

- Runtime

- 通过Blink Session机制，统一了用户提交Job、获取执行结果等行为。
- 开放了调度插件机制，允许计算模型根据需求自定义调度逻辑。
- 在有限流的情况下，通过避免不必要的全图重启，提高了JM和Task FailOver的效率。

- StateBackend

- 使用NiagaraStateBackend替换RocksdbStateBackend，具备更好的读写性能。
- (Experimental) NiagaraStateBackend支持计算存储分离，支持failover过程中state秒级恢复

· 与blink1.6.4不兼容的语法

功能项	影响	解决办法
TableFunction接口修改	所有使用自定义TableFunction的用户	更新代码，实现新的getResultType接口
ScalarFunction接口修改	所有使用自定义ScalarFunction的用户	实现新的getResultType接口
AggregateFunction接口修改	所有使用自定义AggregateFunction的用户	实现新的getAccumulatorType和getResultType接口。例如 ACC类型为Row (String, Long), Agg result的类型为String 则需要实现： <pre> public DataType getAccumulatorType \(\) { return DataTypes.createRowType\ (DataTypes. String, DataTypes .LONG\); } public DataType getResultType\(\) { return DataTypes. String; } </pre>
MapView构造函数修改	MapView构造函数形参类型由之前的TypeInfo变更为DataType。所以在自定义udaf中声明了MapView的job都会受影响。	更新代码，按DataType去构造MapView。比如MapView <pre> map = new MapView<>(Types.STRING, Types. INT);得更新为MapView map = new MapView<>(DataTypes.STRING, DataTypes.INT); </pre>
当参数是long、int 时，除法和AVG返回类型改为了double	以前的除法和AVG返回的是参数类型，现在是double，会导致类型不匹配的错误。例如：除法/AVG的结果直接写入结果表，可能会报结果表与query的字段类型不匹配的错误。	在除法/AVG 的结果上强制加上CAST。

功能项	影响	解决办法
BigDecimalTypeInfo现在在比较时会考虑精度	用到Decimal的 job，可能会抛BigDecimal 类型不匹配的问题。	用到Decimal 类型的job，全局替换成带精度的声明方式，Decimal(38, 18)
NULL与字符串的比较语义	1.x版本NULL与字符串比较返回true，在2.x版本后遵循了数据库语义改为返回false。	所有NULL与字符串比较的地方，例如：WHERE SPLIT_INDEX(shop_list, ':', 1) <> shop_id，如果SPLIT_INDEX(shop_list, ':', 1)返回了NULL，在1.x版本上where条件会返回true，在2.x上面会返回false将数据过滤。

· 如何升级到blink-2.x

使用blink-1.x版本的job升级到blink-2.x，需要进行数据回溯升级，数据回溯是指用户根据业务需要，在启动job时候指定启动位点，具体操作如下：

1. 停止待升级job（清除state）；
2. 开发界面点击右下角的测试版本，修改job的blink版本为blink-2.2.7，上线job；
3. 启动修改后的job并指定启动位点；
4. 若步骤3不成功，人工介入查明原因后，做如下操作：
5. 可以快速修复sql，重复步骤1，2，3；
6. 无法修复sql，回退到原有blink版本；
7. 如遇上无法生成json plan的情况，可以尝试设置如下参数：

- blink.job.option.jmMemMB=4096
- blink.job.submit.timeoutInSeconds=600



说明：

blink-2.0.1的udx第三方插件安装包，请参见[#unique_50](#)：

如果遇到类似如下的异常，是因为udx包的版本太低或者包冲突导致：

```
code:[30016], brief info:[get app plan failed], context info:[
detail:[java.lang.NoClassDefFoundError: org/apache/flink/table/
functions/aggfunctions/DoubleSumWithRetractAggFunction
at java.lang.ClassLoader.defineClass1(Native Method)
at java.lang.ClassLoader.defineClass(ClassLoader.java:788)
```



```
at java.security.SecureClassLoader.defineClass(SecureClassLoader.  
java:142)  
at java.net.URLClassLoader.defineClass(URLClassLoader.java:467)  
at java.net.URLClassLoader.access$100(URLClassLoader.java:73)
```

2018年5月29日-版本发布

· 公告内容

实时计算为了给用户带来更好的开发交互体验,将于明日19:00上线新版本。更新内容如下。

总览

新功能

- Job overview 页面增加SQL物理执行计划展示。
- 为window算子增加每分钟丢弃数据量的metric汇报。
- 支持In、exist、not in、not exist, 并转换为semi-anti join执行。
- 使用英文显示Job launch 的错误信息。
- 所有的sink插件支持ignoreDelete参数。
- ads支持ak和sk参数来替代userName/passWord。
- 支持[引擎版本切换](#)

· Bug修复

- first_value/last_value第一个字段为null的异常。
- OTS维表获取pk字段的bug。
- 智能配置根据metrics判断failover信息不准确的bug。
- cast比较时出现ClassCastException的bug。
- LogicalTableFunctionCall错误的继承结构引起的分段优化错误。
- TopN与OVER语法的规则匹配问题。
- connector open阶段增加并发度检查时抛出IndexOutOfArrayLength异常的bug。
- 阿里云sts过期导致的failover问题。
- 同一个container内同时存在多个sink时可能会出现的插入异常。
- rds mysql sink不支持插入关键字的bug。

2018年5月18日-公测用户转正式商用最后一次通知

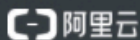
实时计算已于2018年4月30日完成正式商用，但发现部分有使用需求的公测用户仍未付费购买。为确保您能继续使用实时计算，请务必在5月15日前完成付费并绑定到原有项目上，否则将视为放弃使用实时计算，我们将释放其占有的全部计算资源，并清空项目空间。因此造成的一切损失由客户自行承担。

2018年4月23日-公测用户转正式商用

实时计算已于2018年4月1日正式商用，公测用户有一个月的缓冲期，须在4月内付费并绑定到原有项目上。如果您还有继续使用实时计算的需求，请务必在4月内未进行购买并绑定，否则将视为放弃使用实时计算，我们将释放其占有的全部计算资源，并清空项目空间。另外，购买时可以根据目前CU的使用量来购买。比如，公测申请了50个CU，但实际使用了20个CU，那么购买只用购买20个CU即可。如果目前使用了50个CU，只购买了10个CU，那么会导致您的作业随时有被强行停止的风险。您如果有任何问题，欢迎及时反馈。

2018年4月20日-实时计算首次大优惠，包年8.5折

包年8.5折。

 阿里云 短信服务

流计算

配置参数

区域

华东2

华北2

华南1

不同地域之间的产品内网不互通；订购后不支持更换地域，请谨慎选择

购买时长

10 CU

计算能力: 1CU可以处理瓶颈1000条数据/秒

购买时长

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

66

67

68

69

70

71

72

73

74

75

76

77

78

79

80

81

82

83

84

85

86

87

88

89

90

91

92

93

94

95

96

97

98

99

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626

627

628

629

630

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741

742

743

744

745

746

747

748

749

750

751

752

753

754

755

756

757

758

759

760

761

762

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778

779

780

781

782

783

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

810

811

812

813

814

815

816

817

818

819

820

821

822

823

824

825

826

827

828

829

830

831

832

833

834

835

836

837

838

839

840

841

842

843

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859

860

861

862

863

864

865

866

867

868

869

870

871

872

873

874

875

876

877

878

879

880

881

882

883

884

885

886

887

888

889

890

891

892

893

894

895

896

897

898

899

900

901

902

903

904

905

906

907

908

909

910

911

912

913

914

915

916

917

918

919

920

921

922

923

924

925

926

927

928

929

930

931

932

933

934

935

936

937

938

939

940

941

942

943

944

945

946

947

948

949

950

951

952

953

954

955

956

957

958

959

960

961

962

963

964

965

966

967

968

969

970

971

972

973

974

975

976

977

978

979

980

981

982

983

984

985

986

987

988

989

990

991

992

993

994

995

996

997

998

999

1000

1001

1002

1003

1004

1005

1006

1007

1008

1009

1010

1011

1012

1013

1014

1015

1016

1017

1018

1019

1020

1021

1022

1023

1024

1025

1026

1027

1028

1029

1030

1031

1032

1033

1034

1035

1036

1037

1038

1039

1040

1041

1042

1043

1044

1045

1046

1047

1048

1049

1050

1051

1052

1053

1054

1055

1056

1057

1058

1059

1060

1061

1062

1063

1064

1065

1066

1067

1068

1069

1070

1071

1072

1073

1074

1075

1076

1077

1078

1079

1080

1081

1082

1083

1084

1085

1086

1087

1088

1089

1090

1091

1092

1093

1094

1095

1096

1097

1098

1099

1100

1101

1102

1103

1104

1105

1106

1107

1108

1109

1110

1111

1112

1113

1114

1115

1116

1117

1118

1119

1120

1121

1122

1123

1124

1125

1126

1127

1128

1129

1130

1131

1132

1133

1134

1135

1136

1137

1138

1139

1140

1141

1142

1143

1144

1145

1146

1147

1148

1149

1150

1151

1152

1153

1154

1155

1156

1157

1158

1159

1160

1161

1162

1163

1164

1165

1166

1167

1168

1169

1170

1171

1172

1173

1174

1175

1176

1177

1178

1179

1180

1181

1182

1183

1184

1185

1186

1187

1188

1189

1190

1191

1192

1193

1194

1195

1196

1197

1198

1199

1200

1201

1202

1203

1204

1205

1206

1207

1208

1209

1210

1211

1212

1213

1214

1215

1216

1217

1218

1219

1220

1221

1222

1223

1224

1225

1226

1227

1228

1229

1230

1231

1232

1233

1234

1235

1236

1237

1238

1239

1240

1241

1242

1243

1244

1245

1246

1247

1248

1249

1250

1251

1252

1253

1254

1255

1256

1257

1258

1259

1260

1261

1262

1263

1264

1265

1266

1267

1268

1269

1270

1271

1272

1273

1274

1275

1276

1277

1278

1279

1280

1281

1282

1283

1284

1285

1286

1287

1288

1289

1290

1291

1292

1293

1294

1295

1296

1297

1298

1299

1300

1301

1302

1303

1304

1305

1306

1307

1308

1309

1310

1311

1312

1313

1314

1315

1316

1317

1318

1319

1320

1321

1322

1323

1324

1325

1326

1327

1328

1329

1330

1331

1332

1333

1334

1335

1336

1337

1338

1339

1340

1341

1342

1343

1344

1345

1346

1347

1348

1349

1350

1351

1352

1353

1354

1355

1356

1357

1358

1359

1360

1361

1362

1363

1364

1365

1366

1367

1368

1369

1370

1371

1372

1373

1374

1375

1376

1377

1378

1379

1380

1381

1382

1383

1384

1385

1386

1387

1388

1389

1390

1391

1392

1393

1394

1395

1396

1397

1398

1399

1400

1401

1402

1403

1404

1405

1406

1407

1408

1409

1410

1411

1412

1413

1414

1415

1416

1417

1418

1419

1420

1421

1422

1423

1424

1425

1426

1427

1428

1429

1430

1431

1432

1433

- 总览界面

- 增加新版本功能提示。
- CU使用率超过90%时增加扩容提示。
- 增加剩余可使用时间，少于30天时提示续费。
- 增加项目信息，包括管理员、用户、地域、创建时间等信息，管理员可以直接在总览页面管理项目。
- 增加新手指导部分

- 开发界面

- 文件树

- 重新设计了底层逻辑
 - 支持拖拽移动文件夹
 - 支持拖拽移动文件
 - 支持非弹窗快速修改文件夹名称
 - 右键菜单根据窗口边界自动适应
 - 增加快捷操作列表

详情请参看[#unique_51](#)

- 数据源

- 重写了数据源逻辑
 - 数据源数据抽样与血缘关系数据抽样打通

- 运行引擎

- Blink版本的下拉列表宽度自适应，不再需要拖拽宽度

- 代码界面

- 支持错误提示精确到字符
 - 所有不可进行操作（调试、运维）等按钮增加不可操作原因的提示
 - 增加沉浸式全屏编辑模式
 - PlanJSON编辑页面重新设计支持一键切换到代码模式，优化了现有展示方式，不再会被遮挡。
 - 作业版本增加对比、回滚、删除按钮，增加每次上线原因的提示
 - 文件对比JSON与配置部分重写高亮效果
 - 语法检查错误提示支持高亮

- 调试

- 支持顺序抽样功能
 - 调试数据预览与结果固定表头

详情请参看[#unique_52](#)

- 运维界面

- 运维列表

- 增加了资源消耗的排序和显示
 - 批量操作模式下禁止了无关操作的可能性
 - 运行模式筛选方式改变
 - 新作业提交的提示方式修改
 - 作业启动时增加透明升级部分的操作逻辑

更多详情请参见[数据运维](#)

- Vertex列表

- 重新设计了Vertex拓扑的布局方式，目前的排列更加规则整齐
 - Vertex列表固定了ID和Task状态在首尾两列
 - 增加切换视图选项，可以将Vertex Name和数据展示方式进行切换

- Vertex详情

- 重新设计了拓扑结构，按照RecIn和RecOut来判断拓扑是否有错误
 - 重命名了之前不合理的Tab名称
 - SubTask List 移除了TaskContainer的跳转，增加查看历史的提示
 - 历史曲线时间选择器固定在头部

- 数据曲线

- 重写了数据曲线逻辑，增强了拓展性
 - 重新设计了曲线自动刷新时的loading效果，及曲线的配色体系
 - 增加WaterMark曲线
 - 曲线的tooltip目前不再会在y方向被遮挡了
 - 固定时间选择器在用户始终可见的位置

- Failover

- 最近一次的FailOver详情增加高亮效果
 - FailOver历史列表优化，错误信息在展开后可见，增加高亮。

- JobManager

- 增加Attempt List及直接跳转到Resource的方式
 - 增加当前的Metric列表，不再需要弹窗显示。

- 血缘关系

- 重新设计了血缘关系的展示方式

- 数据抽样不再需要右键点击，在数据表直接点击即可抽样

更多详情请参看[#unique_54](#)

- 告警设置

- 优化了交互，用户现在可以更容易的添加告警规则了

- 属性参数

- 将原有Configuration部分移入属性参数

- 合并历史列表展示

2018年4月12日-RDS连接报错解决方案

添加实时计算白名单

近期实时计算集群扩容，对于已经在实时计算中注册过的RDS用来说没影响，但有些用户可能未注册数据源而直接在sql中以代码的形式访问RDS，对于这部分用户，有可能出现无法访问RDS的错误。这是因为RDS系统出于自身安全的考虑，禁止了外部访问，所以需要这些用户在RDS系统中授权实时计算访问。如何添加RDS的白名单，请见[#unique_55](#)。请在添加RDS白名单之后，进行数据源的注册，这样实时计算就能在扩容时自动添加白名单，避免出现链接错误。

10 Flink Forward Asia 2019 门票说明

本文为您介绍Flink Forward Asia 2019购票说明以及注意事项。

会议通票包含内容

- 会议通票包含所有主题演讲、专题演讲以及厂商共建专场（不包含VIP闭门会议）。
- Apache Flink China 社区定制大礼包（卫衣 + Flink 技术书籍 + 马克杯 + 社区专刊）。
- 会议期间展区参观或参与互动，各种小礼品的搜集。



说明：

会议通票不包含培训，如有培训需求，请购买VIP套票。

VIP套票包含内容

- 会议通票所有权益。
- 可参加VIP闭门会议。
- 为期1.5天的Apache Flink线下培训。

退票和退款政策

如果您需要退票，请在购票后的5天内提交退票申请。



说明：

2019年11月10日后，无法提交退票申请，请妥善安排时间。

关于发票开具

请在购票的订单页面，费用 > 发票中正确填写相关信息，以便发票开具。

- 正在开票处理：开票订单生成中。发票尚未打印，可以单击作废取消申请。
- 申请邮寄：信息已经记录，发票尚未打印，可以单击作废取消申请。
- 发票打印中：发票已经在打印。
- 已开发票：发票已经邮寄，可在发票详情中查看邮寄进度。



说明：

从申请邮寄至已开发票状态，大约需要3个工作日。

特别说明

- 请正确填写购票基本信息，大会电子门票将在会前一个月以短信形式发送至您的手机。
- 遇到极端情况会议取消，主办方责任仅限于退回已支付的票款。

11 文档最新动态

文本为您介绍实时计算文档的最新动态，包括新增功能介绍、体验优化介绍等。

时间	特征	类别	描述	使用对象	产品文档
2019.3.25	计算列方法介绍	新功能	介绍计算列的概念和使用场景。	所有	#unique_58
2019.3.25	Flink SQL概述	产品介绍	介绍实时计算Flink SQL的主要内容。	所有	#unique_59
2019.3.25	Flink SQL开发指南概述	产品介绍	介绍Flink SQL开发指南的主要内容。	所有	#unique_60
2019.3.25	Flink SQL作业开发	新功能	介绍如何开发实时计算作业	所有	#unique_61
2019.3.25	Flink SQL作业上线	新功能	介绍如何上线实时计算作业。	所有	#unique_62
2019.3.25	Flink SQL作业启动	新功能	介绍如何启动实时计算作业。	所有	#unique_63
2019.3.25	Flink SQL作业本地调试	新功能	介绍如何进行Flink SQL作业本地调试。	所有	#unique_64
2019.3.25	Flink SQL作业线上调试	新功能	介绍如何进行Flink SQL作业线上调试。	所有	#unique_65
2019.3.25	autoconf自动配置调优	新功能	介绍Flink SQL作业autoconf自动配置调试功能。	所有	#unique_66
2019.3.25	autoscale自动配置调试	新功能	介绍Flink SQL作业autoscale自动配置调试功能。	所有	#unique_67

时间	特征	类别	描述	使用对象	产品文档
2019.3.25	Flink Datastream 开发指南概述	产品介绍	介绍Flink Datastream 开发指南的主要内容。	所有	#unique_68
2019.3.25	Flink Datastream 作业提交	新功能	介绍如何提交Flink Datastream 作业。	所有	#unique_69
2019.3.25	Flink Datastream 作业开发	新功能	介绍如何开发Flink Datastream 作业。	所有	#unique_70
2019.3.15	计算列方法介绍	新功能	介绍计算列的概念和使用场景。	所有	#unique_58
2019.3.15	watermark 方法介绍	新功能	介绍 watermark 的概念和使用场景。	所有	#unique_71
2019.3.15	实时计算产品概述	产品介绍	为您介绍实时计算的特点、定位以及基本概念。你	所有	#unique_72
2019.2.23	API参考	新功能	介绍如何使用API对实时计算项目进行操作。	所有	API参考
2019.1.21	Flink SQL概述	体验优化	介绍Flink SQL的基本概念和使用方法。	所有	#unique_59
2019.1.16	创建 MaxCompute (ODPS) 结果表	新功能	介绍如何创建实时计算 MaxCompute (ODPS) 结果表	所有	#unique_74

时间	特征	类别	描述	使用对象	产品文档
2019.1.16	创建 MaxCompute (ODPS) 维表	新功能	介绍如何创建实时计算 MaxCompute (ODPS) 维表	所有	#unique_75
2018.12.12	创建自定义结果表	新功能	介绍如何创建实时计算自定义结果表	所有	#unique_76
2018.12.04	注册分析型数据库 (AnalyticDB)	新功能	介绍如何使用存储注册的方式连接分析型数据库 (AnalyticDB)，以及连接过程中的常见问题。	所有	#unique_77
2018.12.04	高性能 Flink SQL 优化技巧	体验优化	介绍有助于提升系统性能的 Flink SQL 推荐写法、推荐配置、推荐函数。	所有	#unique_78
2018.12.04	EMIT 语法	新功能	EMIT 语句说明	所有	#unique_79
2018.12.04	MULTI_KEYVALUE	新功能	MULTI_KEYVALUE 函数说明	所有	#unique_80
2018.11.20	独享模式与共享模式的区别	体验优化	介绍实时计算独享模式与共享模式的区别，包括支持数据类型、机器类型、地域等方面的区别。	所有	#unique_30
2018.10.26	DDL 语句 创建 ES 结果表	新功能	新增创建 ES 结果表的说明	所有	#unique_81

时间	特征	类别	描述	使用对象	产品文档
2018.10.25	查看 AccessID、 AccessKey信 息	体验优化	查看 AccessID、 AccessKey信 息说明	所有	如何查看 AccessID、 AccessKey信 息
2018.10.23	实时计算 共享 模式自动续费	新功能	新增实时计 算 共享模式 下，对实例的 自动续费功能 的开启、修改 以及关闭。	所有	#unique_82

12 新功能发布记录

12.1 Blink-3.2.1

12.1.1 Blink-3.2.1新功能发布记录

本文为您介绍Blink-3.2.1版本的核心功能以及新版本对于Datastream和SQL的兼容性。

版本核心功能介绍

Blink-3.2.1是Blink正式开源后第1个基于开源代码的正式版本。Blink-3.2.1版本核心功能介绍如下：

- Job AutoScale

Blink-3.2.1版本中加入了AutoConfig和AutoScale相结合的自动调优功能。自动调优功能能够在作业运行过程中，根据作业运行状态以及输入流量，自动调整各个算子的并发和资源，从而保证作业的低延时。在Blink-3.2.1版本中，该功能处于公测阶段，欢迎您使用。试用过程中如果遇到问题，请提交工单。

- Datastream API支持

- Blink-3.2.1中正式加入了对Datastream API的支持。Blink-3.2.1基于开源Flink1.5.1分支开发，Blink-3.2.1中Datastream API接口与开源Flink1.5.1版本兼容性对比请参见[Flink 3.2与Flink 1.5.1版本API兼容性报告](#)。
- Datastream Connector新增

对接系统	Source	Sink
Kafka	兼容	兼容
HBase	不兼容	
JDBC		
RDS<MySQL>		
ES		
MangoDB		

- SQL Connector新增

Blink-3.2.1在Blink2.0版本基础上新增了如下Connector。

对接系统	Connector类型
ES	DIM
MongoDB	SINK
Redis	DIM
Redis	SINK
RocketMQ(4.2.0)	SOURCE
SQL Server	SINK

兼容性

- Datastream兼容性

请参见[Flink 3.2与Flink 1.5.1版本API兼容性报告](#)。

- SQL兼容性

请参见[Blink 3.0与Blink 2.0版本SQL不兼容项目汇总](#)。

12.1.2 Flink 3.2与Flink 1.5.1版本API兼容性报告

校验范围

- flink-clients
- flink-core
- flink-java
- flink-java8
- flink-optimizer
- flink-scala
- flink-scala-shell
- flink-streaming-java
- flink-streaming-scala
- flink-yarn
- flink-connectors
- flink-filesystems
- flink-formats
- flink-metrics

- flink-queryable-state
- flink-state-backends

兼容性详情

- flink-core

总计方法个数：6126；不兼容数量：1。

序号	严重程度（高、中或低）	Old API	Change	Effect
1	中	GenericCsvInputFormat.supportsMultiPaths()	该override方法从该类中删除，返回的默认值发生变化，新的默认值不支持multiPath。	所有GenericCsvInputFormat子类涉及multiPath的功能会受到影响。

- flink-connector-elasticsearch

总计方法个数：14；不兼容数量：1。

序号	严重程度（高、中或低）	Old API	Change	Effect
1	中	ElasticsearchSink	父类从ElasticsearchSinkBase<T>变为了ElasticsearchSinkBase<T,org.elasticsearch.client.Client>。	子类不兼容。

flink-json

总计方法个数：34；不兼容数量：1。

序号	严重程度（高、中或低）	Old API	Change	Effect
1	中	JsonSchemaConverter	类名重命名成了JsonRowSchemaConverter。	子类不兼容。

flink-streaming-java

总计方法个数：3031；不兼容数量：4。

序号	严重程度（高、中或低）	Old API	Change	Effect
1	中	TwoInputStreamOperator.processElement1 or processElement2	void返回值变成了TwoInputSelection。 TwoInputStreamOperator中新增endInput1和endInput2 abstract方法。	不兼容所有的TwoInputStreamOperator实现。
2	中	OneInputStreamOperator类	添加endInput() abstract方法。	不兼容所有的OneInputStreamOperator实现。
3	中	StreamOperator类	添加requireState abstract方法。	不兼容所有的StreamOperator实现。
4	中	OneInputStreamOperator类	添加了endInput() abstract方法。	不兼容所有的OneInputStreamOperator实现。

12.1.3 Blink 3.0与Blink 2.0版本SQL不兼容项目汇总

SQL语法变更

- 语法变动

[Over Agg] The window rank function without order by

- 行为变动

- [Division to double] | 隐式类型转换，从2.2版本开始生效。
- [Decimal] DDL decimal type default precision changed to (10, 0) | Decimal的默认精度变更，从2.2版本开始生效。
- [CEP] pattern不可以贪婪匹配结束。例如，不支持pattern (a b+)。可以通过将(a b+)转换为(a b+ c)，并把c定义为not b的方式，解决不能以贪婪匹配结束的问题。
- [CEP] within语句不支持动态窗口。

开发接口变更

- 重构和语义变更

[StreamTableSink#emitDataStream returns values changes from void to StreamTableSink]

- Class位置挪动

- [Parser 继承 com.alibaba.blink.streaming.connectors.common.source.SourceCollector]
- [Class not found] com/alibaba/blink/exceptions/NotEnoughParamsException
- [Class not found] com/alibaba/blink/exceptions/UnsupportedTableException
- [Class not found] org/apache/flink/table/sources/BatchExecTableSource
- [Class not found] org/apache/flink/table/functions/aggfunctions/DoubleSumWithRetractAggFunction
- [Class not found] org/apache/flink/table/functions/Monotonicity
- [Class not found] com/alibaba/blink/cache/Cache
- [Class not found] org/apache/flink/table/row/GenericRow

- 实现变更

- [Method not found] com.alibaba.blink.table.api.RichTableSchema.getColumnTypes
- [Method not found] org/apache/flink/table/types/DataType.of
- [Verification] java.lang.VerifyError: class com.koubei.blink.connector.sls.CustomTableFactory overrides final method setClassLoader
- [Class not found] com/aliyun/odps/OdpsException

Connectors

- [ODPS] ODPSTableSink stream mode do not support overwrite
- [ODPS] Only batch mode support overwrite

12.2 Blink-3.2.3

为了给您带来更好的开发交互体验，实时计算发布了Blink-3.2.3版本。

优化

- 优化Vertex Topology页面与作业资源配置界面展示不一致问题。
- 优化3.2.1版本数据曲线页面Task ID的ID名称不展示问题。

- 优化运维界面TaskManager.log日志中乱码问题。
- 优化3.2.1版本GC日志覆盖上线调试数据问题。
- 优化独享集群AutoScale需要增加作业参数才能开启问题。
- 优化作业运行代码为LEFT JOIN，运维界面Vertex拓扑图中显示为INNER JOIN问题。
- 优化代码编辑有误时不定位至具体的一行问题。

BUG修复

- REGEXP_EXTRACT函数参数为null或正则表达式不合法时，返回null，与实际不符。
- 开发代码中使用反斜线的正则表达式后，使用用分号编译报错。
- 上线调试运行结果打印至taskmanger.out中，与实际不符。
- LEFT JOIN维表，JOIN时无维表标识不报错。
- Time类型的数据写入ADS结果表出现异常。
- 3.2.1版本MaxCompute维表使用TIMESTAMP类型进行JOIN，作业运行报错。
- 3.2.1和3.2.2版本使用minibatch作业参数，作业运行报错。
- CEP语法检测正常，作业运行报错。
- 云数据库HBase结果表的maxRetryTimes参数未生效。
- Kafka源表消费在Kafka端未展示对应的Group ID。
- 只写入一个VARBINARY类型的数据至MQ，运行报错。

12.3 Blink-3.3.0

本文为您介绍实时计算公告内容，包括版本重大功能简介和产品变更。

版本重大功能简介

- 小CU模式
 - 新增初始plan设置CU预期：生成plan时，作业的初始并发度会根据用户设置的CU预期进行缩放。
 - 新增小CU模式资源配置：当用户设置的CU预期很小，每个Vertex缩小成1个并发都无法放下时，会触发小CU模式，自动资源设置进行压缩，同时将所有Vertex调度到一个TaskManager上，减少框架资源开销。
- DataHub
 - 写入Datahub支持通过hashFields的配置，使相同列值的记录写入同一个下游Shard。
 - 修复Blink 3.2版本读取closed shard的Datahub数据Failover，且任务不能恢复的问题。

产品变更

• Auto Scale变更

- Autoscale最大资源设置语义修改：Max CU限制从作业整体资源的上限调整为Plan的资源上线，以解决Max CU限制，导致作业无法启动的问题。Max CU限制调整后，作业实际使用的资源可能会超过Max CU。
- Autoscale添加可选作业参数配置，可以手动关闭scale down功能，保证作业稳定性。配置项：`healthmanager.resource.scale.down.enabled`（资源scale down开关）和 `healthmanager.parallelism.scale.down.enabled`（并发度scale down开关）。
- 新增支持Datastream作业手动设置资源，开启Autoscale。Blink3.3版本开始，Datastream作业可手动编辑资源Plan，同时允许可以开启Autoscale。Datastream的Autoscale目前仅处于试用阶段。

• FirstRow on Rowtime

支持您按照Rowtime字段读取FirstRow，去重后保持time字段的Rowtime属性，即您可以在去重节点后继续使用Window。

• SQL文件大小写变更

该改动将忽略SQL文件中的大小写（大小写不敏感），导致部分作业在编译时报错。该错误表明您在同一段代码中，使用大小写来区分变量或标识符。

- 示例

```
4      t.taobao_bind AS taobao_bind,  
5      t.et_taobao_bind AS et_taobao_bind  
6  FROM  
7      view_t_partner_map_taobao_bind t,  
8      lateral table (STRING_SPLIT(t.utdids, '\004')) AS T(utdid0);  
9  
10 INSERT INTO
```

- 报错信息

```
5  org.apache.flink.table.api.ValidationException:  
6  *****  
7  ERR_ID:  
8      SQL-00120001  
9  CAUSE:  
10     SQL validation failed:  
11     From line 3, column 9 to line 3, column 69: Duplicate relation name 'T' in FROM clause  
12  ACTION:  
13     Please see descriptions above. If it doesn't help, please contact customer support for  
14     this.  
15  DETAIL:  
16  .....
```