阿里云
实时计算（流计算）

新功能发布记录

文档版本：20200519

为了无法计算的价值 | 阿里云

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. 阿里云文档中所有内容，包括但不限于图片、架构设计、页面布局、文字描述，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。
	该类警示信息可能会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，恢复业务时间约十分钟。
	用于警示信息、补充说明等，是用户必须了解的内容。	**注意：** 权重设置为0，该服务器不会再接受新请求。
	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。	单击**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。	在**结果确认**页面，单击**确定**。
Courier字体	命令。	执行cd /d C:/window命令，进入Windows系统文件夹。
斜体	表示参数、变量。	bae log list --instanceid Instance_ID
[]或者[a|b]	表示可选项，至多选择一个。	ipconfig [-all|-t]
{}或者{a|b}	表示必选项，至多选择一个。	switch {active|stand}

目录

法律声明.....I
通用约定.....I
1 公告.....1
2 Flink半托管.....6
2.1 2.1.3版本.....6
3 Blink独享/共享集群（原产品线）.....8
3.1 Blink-3.6.2.....8
3.2 Blink-3.6.0.....8
3.3 Blink-3.5.0-hotfix.....9
3.4 Blink-3.4.3.....9
3.5 Blink-3.4.4.....11
3.6 Blink-3.3.0.....12
3.7 Blink-3.2.3.....13
3.8 Blink-3.2.1.....14
3.8.1 Blink-3.2.1新功能发布记录.....14
3.8.2 Blink 3.2与Flink 1.5.1版本API兼容性报告.....15
3.8.3 Blink 3.0与Blink 2.0版本SQL不兼容项汇总.....17
4 发展历程.....19
5 文档最新动态.....21

1 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。

2020年4月28日21:00-2020年4月29日02:00-升级公告

2020年4月28日21:00-2020年4月29日02:00，对上海独享模式管控平台进行升级，升级期间现有运行作业不受影响，但集群扩缩容功能不可用，请知悉。

2020年4月20日-2020年4月22日-升级公告

2020年4月20日-2020年4月22日，对上海和深圳共享集群的存储服务进行版本升级，旨在为您提供更加稳定的实时计算服务。正常情况下，此次升级不会对用户的业务造成影响；特殊情况下，Blink3.2和Blink3.3版本的作业会Failover一次后恢复正常。

2019年12月24日-共享模式正式下线

实时计算共享模式已于2019年12月24日正式下线，将不再支持共享模式新项目的购买，仅支持原有项目的扩缩容、续费操作。如果您有新购需求，推荐使用实时计算独享模式或Flink云原生模式。

2019年8月27日-新增资源配置页签

开发页面右侧，新增**资源配置**页签。原控制台**基本属性**页签下的资源配置跳转链接已下线。

2019年8月21日-华东2（上海）共享集群扩容通知

2019年8月23日，华东2（上海）共享集群将新增11.53.0.0/16,11.50.0.0/16网段。如果您使用的上下游存储涉及到例如RDS、HBase等需要添加白名单的组件，请提前将11.53.0.0/16,11.50.0.0/16网段添加至现有的白名单中，防止发生连接异常。

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_4。

- 数据曲线

  新增AutoScaling相关曲线，详情请参见#unique_5。

- Timeline

  新增Timeline功能，详情请参见#unique_6。

- 属性参数

  新增AutoScale迭代的历史详情查询功能，详情请参见#unique_7。

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-blink-2.2.7版本发布

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，例如：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和LAST函数存在Retraction场景的性能。
    - 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注册源表和结果表的Connector。
    - 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JobManager和Task FailOver的处理效率。
  - 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，accumulator类型为Row(String, Long)，Agg result的类型为String，则需要实现如下代码。
		`public DataType getAccumulatorType\(\) { return DataTypes.createRowType\(DataTypes.String, DataTypes.LONG\); }`
		`public DataType getResultType\(\) { return DataTypes.String; }`

MapView构造函数修改	MapView构造函数形参类型由之前的TypeInformation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。例如`MapView map = new MapView<>(Types.STRING, Types.INT);`需要更新为`MapView map = new MapView<>(DataTypes.STRING, DataTypes.INT);`。
当参数是Long、Int时，除法和AVG返回类型改为Double	以前的除法和AVG函数返回的是入参的字段类型，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 。
在比较big decimal 和decimal数据类型的数据时，会考虑精度	用到Decimal的Job，可能会报BigDecimal类型不匹配的错误。	用到Decimal类型的Job，全局替换成带精度的声明方式，Decimal(38, 18)。

功能项	影响	解决办法
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SQL语法语义改为返回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，指定启动位点，具体操作如下：

  1. 停止待升级Job（清除state）。
  2. 开发界面单击右下角的**Flink版本**下拉箭头，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。

     如果步骤3执行不成功，需要人工介入查明原因后，进行如下操作：

     - 快速修复SQL，重复步骤1、2、3。
     - 如果无法修复SQL，回退到原有Blink版本。
     - 如果无法生成Json Plan，可以尝试设置如下参数：
       - blink.job.option.jmMemMB=4096
       - blink.job.submit.timeoutInSeconds=600

  Blink-2.0.1的UDX第三方插件安装包详情，请参见#unique_8。类似如下的异常，是因为UDX包的版本太低或者包冲突导致的。

```
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)
```

2 Flink半托管

2.1 2.1.3版本

本文为您介绍云原生2.1.3版本的重大功能变更和主要修复缺陷。

版本重大功能变更

- 支持自定义Flink Pod Label。

  详情请参见Flink Pod Label。
- 支持Flink 1.10镜像。
- 支持Flink 1.10用户包隔离功能。
- 支持Flink Configuration里使用占位符，动态填入配置信息。

  详情请参见Deployment Defaults。
- Flink 1.10 默认使用阿里自研的**Gemini State Backend**。

  在作业启动时，系统会自动添加以下配置，默认启动阿里自研的**Gemini State Backend**。

  - state.backend: com.alibaba.flink.statebackend.GeminiStateBackendFactory
  - fs.oss.impl: org.apache.hadoop.fs.aliyun.oss.AliyunOSSFileSystem
  - state.backend.gemini.snapshot.close.file: true

  > **说明：**
  > 如果需要使用您自己的State Backend，请在**state.backend**配置项中或者代码中指定**State Backend**。您可以使用**Namespaced Defaults**来设置Flink配置项默认值。
- API优化

  - DeploymentTarget：新增**spec.kind**字段。

    ```
    DeploymentTarget:
      metadata:
        ...
      spec:
        kind: kubernetes
        kubernetes:
          namespaces: vvp
    ```

  - Job：新增**job.spec.deploymentTarget.kind=kubernetes**字段。

    ```
    Job:
      metadata:
        ...
      spec:
    ```

```
deploymentTarget:
  kind: kubernetes
  kubernetes:
    namespace: vvp
```

- Deployment创建页面支持三种视图
  - Standard
  - Advanced
  - Yaml
- 扩展Deployment和Job的Status，返回更多信息
  - Job
    - live status
    - restarts count
  - Deployment
    - ClusterUnreachable
    - JobFailing
    - JobUnstable

主要缺陷修复

修复Flink1.8和Flink1.9镜像中**oss log appender**在**debug log Level**占用资源过大的问题。建议您使用新增的**1.8.1-stream3-scala_2.11**和**1.9.0-stream3-scala_2.11**镜像替换之前的镜像。

3 Blink独享/共享集群（原产品线）

3.1 Blink-3.6.2

本文为您介绍Blink 3.6.2版本的重大功能变更和主要修复缺陷。

版本重大功能变更

- 优化AutoScale功能：没有流量的作业会触发Scale Down，减少资源浪费。
- 优化资源Rescale。
- 优化SLS Connector解析数据的方法：使用FastLogGroup方法解析数据。
- 降低大规模作业拓补的内存消耗。

主要缺陷修复

- 修复Blink无法消费表格存储Tablestore源表积压数据的缺陷。
- 修复维表Cache ALL多并发导致NullPointException的缺陷。
- 修复MaxCompute Sink动态分区无法正常提交数据的缺陷。

3.2 Blink-3.6.0

本文为您介绍Blink 3.6.0版本的重大功能变更和主要修复缺陷。

版本重大功能变更

- 新增Connector：
  - 新增#unique_14。
  - 新增#unique_15。
  - 新增#unique_16。
  - 新增#unique_17。
  - 新增#unique_18。
- 新增功能：
  - 支持MQ结果表写入KEY。
  - 支持Gemini 2.0（Gemini 增强版）。

主要缺陷修复

- 修复Kafka 08版本结果表，语法检查报错：`Kafka dont have sufficient arguments`。

- 修复OTS源表因为某个分区一直没有数据，线上运行源表节点导致的Failover：`java.lang.NullPointerException`。
- 修复OTS维表以注册存储方式引用，Token过期导致的Failover：`OTSNoPermissionAccess, [Message]:You have no permission to access the requested resource, please contact the resource owner`。

3.3 Blink-3.5.0-hotfix

本文为您介绍Blink 3.5.0版本的重大功能变更和主要修复缺陷。

版本重大功能变更

- 新增#unique_20。
- 新增#unique_21。
- 升级Datahub Connector SDK版本。

主要缺陷修复

- 修复使用全局ORDER BY但语法检查没有报错的缺陷。
- 修复底层VARCHAR类型隐式转换为整型，行为错误的缺陷。
- 修复因calcite bug，数据按照时间分组进行聚合，导致计算结果异常的缺陷。
- 修复语法检查DataHub源表产生`java.lang.NoClassDefFoundError: com/aliyun/datahub/client/auth/Account`报错的缺陷。

3.4 Blink-3.4.3

本文为您介绍Blink 3.4.3版本的重大功能变更和主要修复的缺陷。

版本重大功能简介

GeminiStateBackend是Blink开发的新一代后台，它使用自研的存储引擎GeminiDB。通过线上部分作业的测试显示，GeminiStateBackend的性能可以达到NiagaraStateBackend性能的1.5倍。GeminiStateBackend的主要特性如下：

- 采用LSM + Hash索引的存储模型。其中LSM提高写性能，Hash索引存放在内存中以弥补LSM读放大的缺点。简单来说，GeminiDB将文件划分为更小粒度的单元（Page），以Page为粒度进行冲洗和压缩处理，Hash索引根据键快速定位其所在的单元，从而减少IO次数，提高读性能。
- 优化Cache策略。如果新插入的数据或压缩后的数据存在热点，GeminiDB会将其缓存在内存中。而传统的LSM实现会首先将这些数据冲洗到磁盘上，然后至少经过一次读IO才能进入缓存，导致缓存命中率下降。

- 优化数据冲洗到磁盘的策略。GeminiDB只有在内存的缓存填满后才会将部分数据冲洗到磁盘，因此，在内存充足并且压缩及时的情况下，不会产生数据文件。而传统LSM的实现中，会将数据冲洗到磁盘以便进行持久化处理。这在Blink的应用场景中是没有必要的，Blink具有原生的Checkpoint机制以保证数据的一致性，因此只需在Checkpoint时进行持久化处理。
- 支持In-memory Compaction。对于内存中数据及时进行压缩，减少写放大以及压缩的读IO。
- 使用Java消除RocksDB/Niagara中的JNI开销。
- 支持增量Checkpoint。
- 支持Local Recovery，使作业失效后快速恢复。
- 支持存储计算分离，使作业重启或者Rescale后的快速恢复，功能还在完善中。

GeminiStateBackend针对DataStream和SQL的作业配置如下：

- DataStream作业
  - API配置方式

```
StreamExecutionEnvironment env = StreamExecutionEnvironment.getExecuti
onEnvironment();
GeminiStateBackend stateBackend = new GeminiStateBackend(checkpointDir);
// Configuration for gemini
Configuration config = new Configuration();
config.setString("state.backend.gemini.heap.size", "1024mb");
// set configuration to backend
stateBackend.setConfiguration(conf);
// use gemini as state backend
env.setStateBackend(new GeminiStateBackend(checkpointDir));
```

  - 主要配置项

配置项	类型	单位	默认值	说明
state.backend.gemini.ttl.ms	LONG	ms	-1（未开启）	可选。数据存活时间。
state.backend.gemini.heap.size	STRING	支持以下字节单位： ■ 1024 ■ 1024kb ■ 1024mb ■ 1024gb	无	可选。单个GeminiDB能够使用的内存大小。 **说明：** 推荐进行配置，否则backend会根据JVM和TaskManager计算其默认值。

- SQL作业

```
# 使用Gemini。
state.backend.type=gemini
# State的TTL的大小。
state.backend.gemini.ttl.ms=129600000
# GeminiDB允许使用的内存大小，单位MB。注意：Operator配置的内存资源要包含这部分，默认为512MB。
state.backend.gemini.heap.size.mb=512
# 推荐的TaskManager的JVM参数。
blink.job.option=-yD env.java.opts.taskmanager='-XX:NewRatio=3 -XX:SurvivorRatio=3 -XX:ParallelGCThreads=8 -XX:+UnlockDiagnosticVMOptions -XX:ParGCCardsPerStrideChunk=4096 -XX:+UseCMSInitiatingOccupancyOnly -XX:CMSInitiatingOccupancyFraction=75 -Djdk.nio.maxCachedBufferSize=10240'
```

主要缺陷修复

- 修复Calc的codegen代码可能会出现NPE的缺陷。
- 修复按照proctime读取FirstRow，状态存储不需要存储完整行的缺陷。
- 修复code split缺陷：**JavaCodeSplitter**在把local variable抽取成member field时，没有替换foreach control语句中的local variable。 在distinct with filter场景中可能出现这种缺陷。

3.5 Blink-3.4.4

本文为您介绍Blink 3.4.4版本的重大功能变更和主要修复缺陷。

版本重大功能变更

Blink-3.4.4版本支持MaxCompute源表监听新分区。在DDL语句中设置**subscribeNewPartition**参数，参数默认是false。当参数值是true时，不指定partition参数，会动态监听新分区。详情参见#unique_24。

主要缺陷修复

- RDS连接性问题：Blink使用Druid链接池链接RDS，优化Druid链接池链接RDS的方式，避免链接长时间不用导致作业的Failover。
- DataHub结果表：修复因一条数据中某一列为NULL，造成作业报错**Caused by: java.lang.NullPointerException**的缺陷。
- DataHub结果表：修复将BIGINT类型数据插入至字段为TIMESTAMP类型的DataHub结果表中，产生类型转换异常的缺陷。

3.6 Blink-3.3.0

本文为您介绍实时计算公告内容，包括版本重大功能发布和产品变更介绍。

版本重大功能

- **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。从Blink3.3版本开始，Datastream作业支持手动编辑资源Plan，同时允许开启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

```
    t.taobao_bind AS taobao_bind,
    t.et_taobao_bind AS et_taobao_bind
FROM
    view_t_partner_map_taobao_bind t,
    lateral table (STRING_SPLIT(t.utdids, '\004')) AS T(utdid0);

INSERT INTO
```

  - 报错信息

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

3.7 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，本文为您介绍版本优化点和修复的问题。

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。

问题修复

- 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。

3.8 Blink-3.2.1

3.8.1 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Datasteam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。

- 支持Datastream API
  - Blink-3.2.1版本中正式加入了对Datastream API的支持。Blink-3.2.1基于开源Flink 1.5.1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Blink 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		
MongoDB		

- 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Blink 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不兼容项汇总。

3.8.2 Blink 3.2与Flink 1.5.1版本API兼容性报告

本文为您介绍Blink 3.2与Flink 1.5.1版本API兼容性测试的结果。

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、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、低）	Old API	Change	Effect
1	中	**TwoInputStreamOperator.processElement1** or **TwoInputStreamOperator.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。

3.8.3 Blink 3.0与Blink 2.0版本SQL不兼容项汇总

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] Lcom/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] Lorg/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

4 发展历程

本文为您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。受益于阿里巴巴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。实时计算团队起初仅负责支持双十一大屏展现和部分实时报表业务，在历经多年的摸索和发展后，最终成长为独立稳定的云产品团队。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：以开源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，阿里巴巴工程师最终开发了Flink SQL替换了原有的XML+组件的编程方式，使用经典的SQL模式能够完成计算和数据处理。 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：

- Flink SQL对标SQL功能，可以提高您和开发人员的技术成熟度。
- 您可以在Flink SQL运用您熟悉的SQL模型，降低您使用实时计算的难度。

5 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、体验优化介绍等。

2020年4月更新记录

时间	特征	类别	描述	使用对象	产品文档
2020.4.28	源表、结果表	新功能	介绍如何创建Oracle数据库源表和结果表。	所有	• #unique_33 • #unique_34
2020.4.24	结果表	新功能	介绍如何创建分析型数据库 PostgreSQL版结果表。	所有	#unique_14
2020.4.24	维表	新功能	介绍如何创建Phoenix5维表。	所有	#unique_15
2020.4.24	源表、结果表、结果表	新功能	介绍如何创建交互式分析Hologres源表、维表和结果表。	所有	• #unique_16 • #unique_17 • #unique_18

2020年3月更新记录

时间	特征	类别	描述	使用对象	产品文档
2020.3.21	Flink Datastream	新功能	介绍如何读取DataHub数据和Kafka数据。	所有	• #unique_35 • #unique_36
2020.3.6	源表	新功能	介绍如何创建增量MaxCompute源表。	所有	#unique_37
2020.3.5	暂停	新功能	介绍如何暂停作业。	所有	#unique_38

2020年2月更新记录

时间	特征	类别	描述	使用对象	产品文档
2020.2.26	结果表	新功能	介绍如何创建InfluxDB结果表和Phoenix5结果表。	所有	• #unique_20 • #unique_21

2020年1月更新记录

时间	特征	类别	描述	使用对象	产品文档
2020.1.16	Flink SQL	函数	介绍如何使用实时计算进制转换内置函数CONV。	所有	#unique_39
2020.1.9	OpenAPI	新接口	介绍如何使用API创建一个实时计算的实例。	所有	#unique_40

2019年10月更新记录

时间	特征	类别	描述	使用对象	产品文档
2019.10.10	新增商品模式介绍	新功能	实时计算Flink云原生版本商品模式介绍。	所有	#unique_41

2019年8月更新记录

时间	特征	类别	描述	使用对象	产品文档
2019.8.15	结果表	新功能	介绍如何创建分析性数据库MySQL版3.0结果表。	所有	#unique_42

2019年7月更新记录

时间	特征	类别	描述	使用对象	产品文档
2019.7.31	版本功能介绍	公告	Blink 3.3.0新功能说明。	所有	Blink-3.3.0
2019.7.03	结果表	新功能	介绍如何创建云数据库MongoDB版结果表。	所有	#unique_43

2019年6月更新记录

时间	特征	类别	描述	使用对象	产品文档
2019.6.26	维表	新功能	介绍如何创建Elasticsearch维表。	所有	#unique_44

2019年5月更新记录

时间	特征	类别	描述	使用对象	产品文档
2019.5.31	日志下载	新功能	介绍如何通过编辑实时计算作业参数，自定义实时计算作业日志的下载级别和下载路径。	所有	#unique_45
2019.5.31	Timeline	新功能	介绍作业拓扑图中节点的运行状态。	所有	#unique_6
2019.5.31	作业诊断	新功能	介绍作业诊断功能。	所有	#unique_46
2019.5.31	结果表	新功能	介绍如何创建云数据库Redis版结果表。	所有	#unique_47
2019.5.31	结果表	新功能	介绍如何创建云数据库RDS SQL Server版结果表。	所有	#unique_48