Flink

# 简介

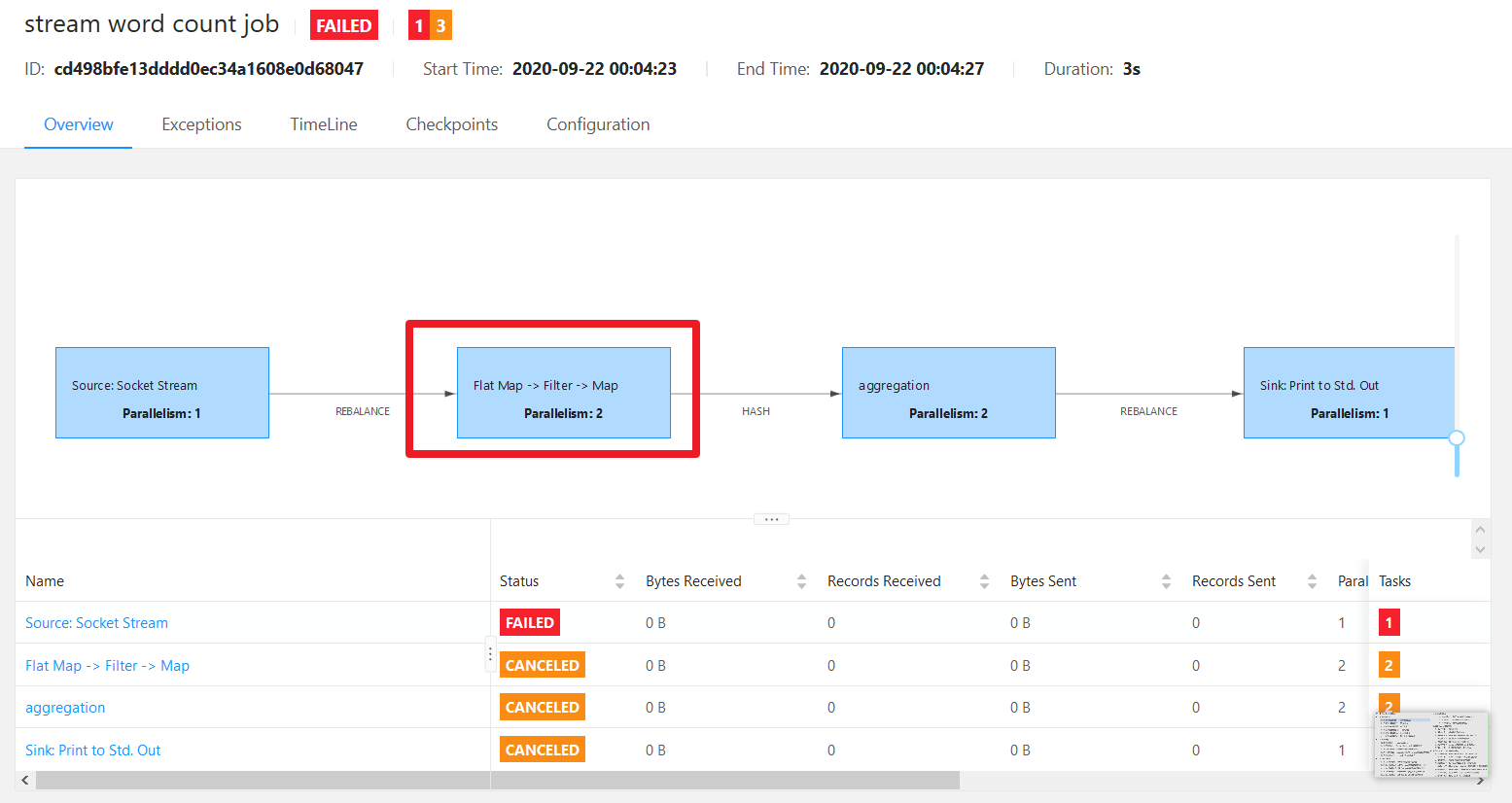
## 支持hadoop

添加支持hadoop的jar包：flink-shaded-hadoop-2-uber-2.7.5-10.0.jar

## 并行度的优先级：

单个算子并行>env>job提交指定的并行度（网页或shell）>配置文件中的配置





## 提交作业

除了页面提交，在shell界面提交：

./bin/flink run -c com.hong.wc.WordCount.Stream -p 1 xx/xx.jar –host hadoop203 –port 7777

提交后，同样也能在页面中看到运行状态

Cancel job：./bin/flink list （查看正在运行的job） ./bin/flink cancel job的id

## Yarn提交及区别

## 和spark的区别



# 运行时架构

## 任务链的理解

任务合并：one-to-one & 并行同。

逻辑视图-并行化视图-优化视图（任务合并）。

任务数和slot的关系是，整个slot负责一条链

## 各组件的理解

JobGraph 、Logical DataFlow Graph

## 什么是任务

代码中定义的每一步操作（算子，operator）就是一个任务。

算子可以设置并行度，所以每一步操作（任务）就可以有多个并行的子任务。

但是注意，keyBy() 这个算子不可以设置并行度，因为它不涉及数据操作，只是数据分组。

Flink可以将前后的任务就行合并，条件：① 1-1操作； ② 并行度一致

## 什么是Slot

Slot是TaskManager拥有的计算资源的子集。

同一个任务的并行子任务，必须运行到不同的slot上。（否则就不是并行了）

如果不同算子的任务，可以共享一个slot。

### 并行度和slot数量的关系：

1. 并行度和任务有关，就是每一个 算子拥有的并行任务数量；动态概念
2. Slot数量只跟TM的配置有关，代表TM并行处理数据的能力；静态概念（能力）

结论：

一个job需要的slot数量，就是并行度最大的算子的并行度。也习惯说是整个流的并行度

### Slot的共享组：

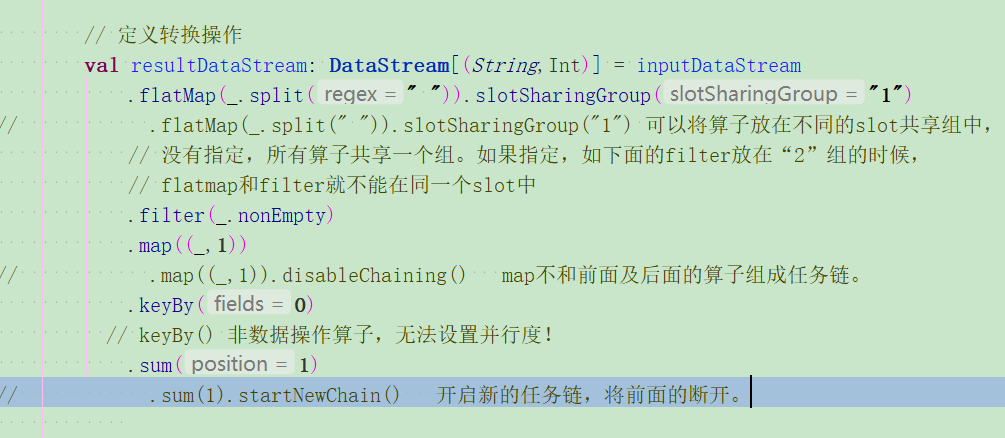
可以在算子后面调用配置，这样slot的数量的分析就有些复杂了。需要单独为共享组分配slot，这样job的总slot就是 slot组+上面常规分析slot数。

这种情景一般是用在某一个算子的计算复杂，但是默认的flink又会和其他算子进行共享slot时，可以给这些单独的算子单独分配slot组。

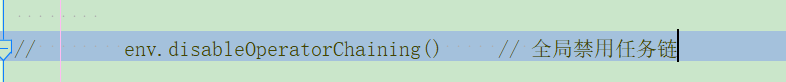
### disableChaning

见2.4.4 的代码图

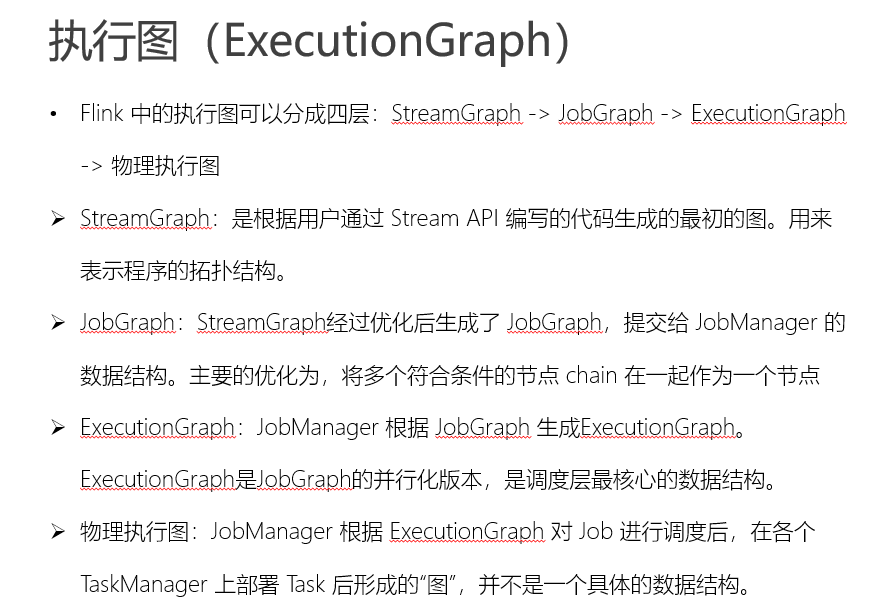
### startNewChain



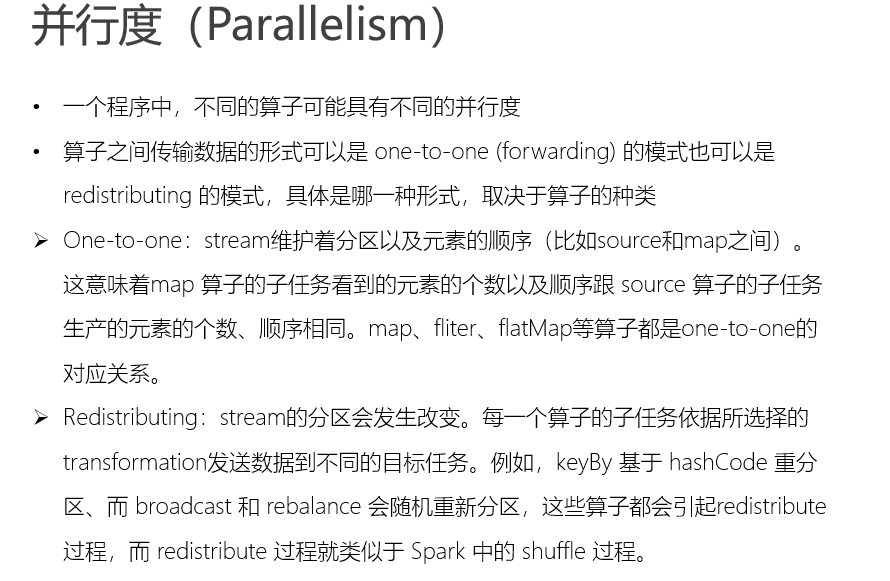
### 全局禁用任务链



## 执行图

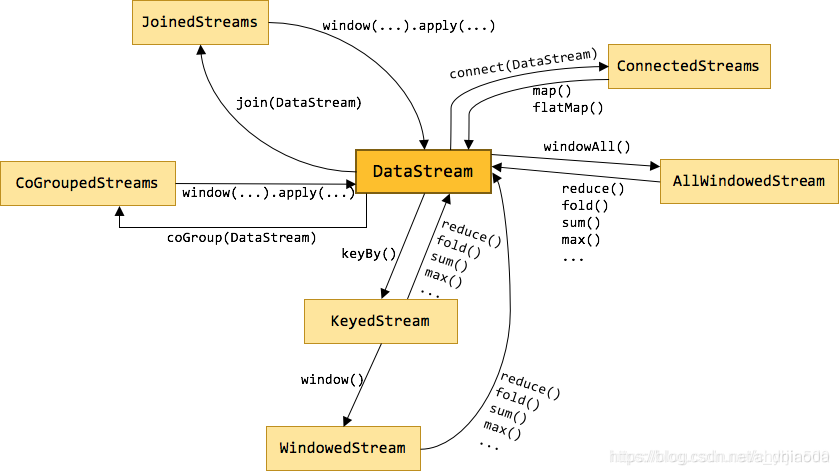


## 数据的传输形式

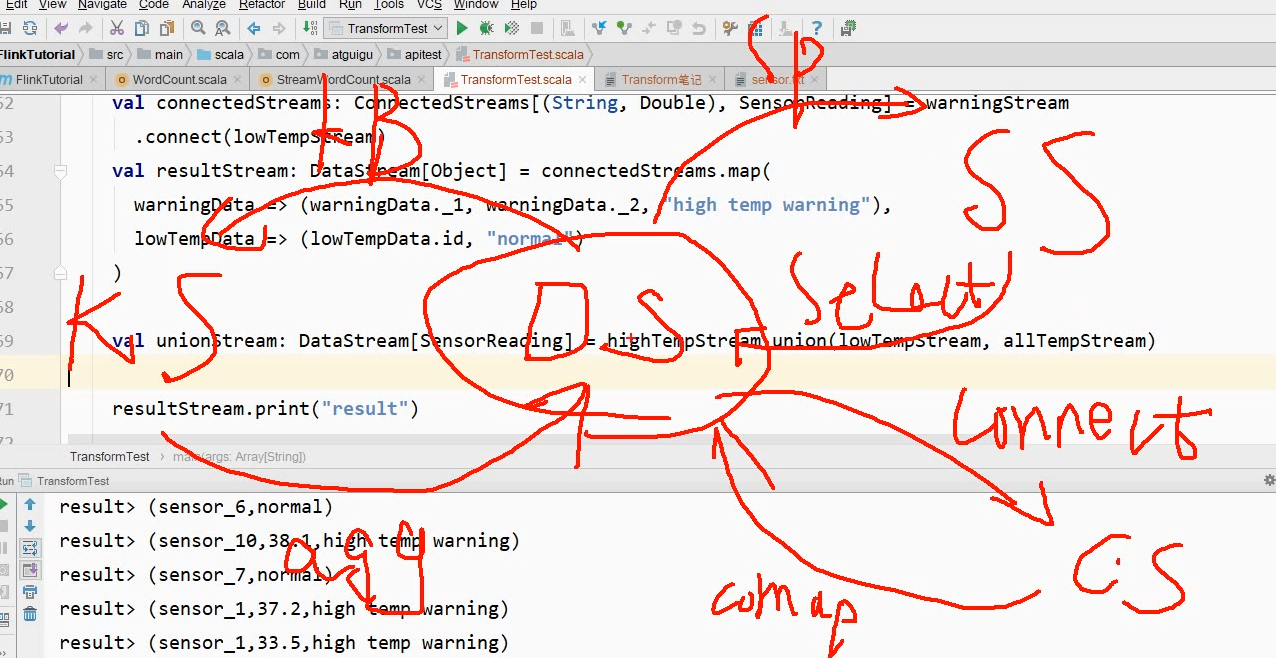


# Transform

DataStream没有滚动聚合，只有转成KeyedStream才可以。



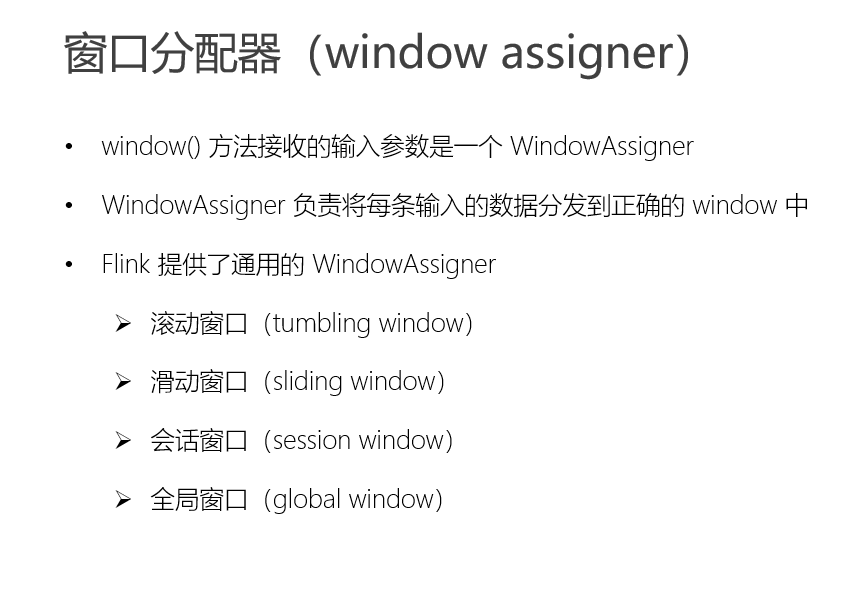
少了个split+select ：DataStream <-> SplitStream



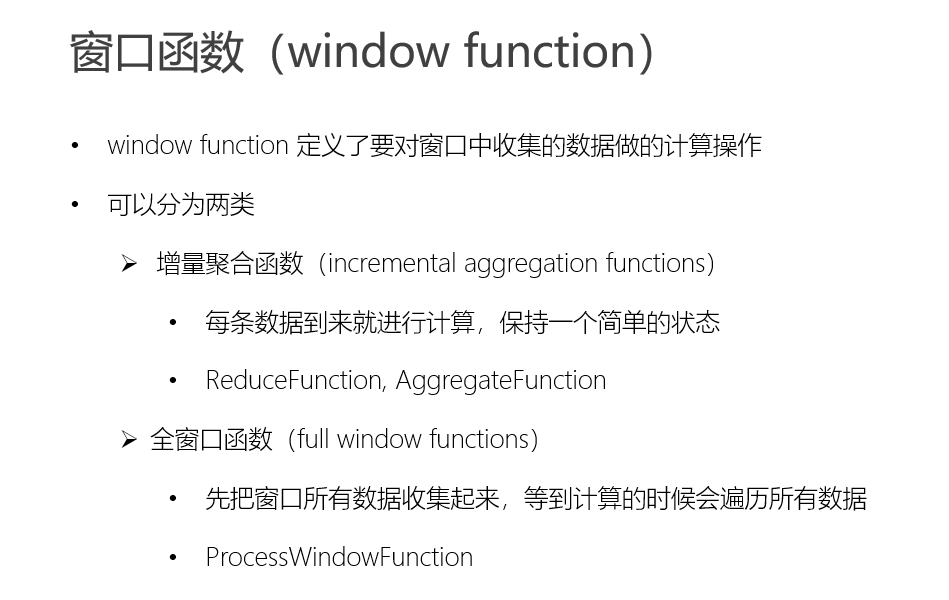
# Window API

## window assigner

1. keyBy() 后，直接调用timeWindow/countWindow。每个key对应的分区都开窗。
2. keyBy() 前，只有windowAll()，所有的数据都会进入到一个窗口，变成一个分区，不能并行处理，所以不推荐使用。



## 两种窗口函数



增量聚合函数：来一个处理一个，并将状态保存起来ReduceFunction/AggregateFunciton。

全窗口函数：可以获取整个窗口的信息。弥补增量聚合函数的不足。WindowFunction/

ProcessWindowFunction

窗口函数是基于当前窗口内的数据，是有界数据集的计算，通常只在窗口关闭时输出一次。

## Window API总览



## 其他问题

### 窗口偏移量理解

SlidingProcessingTimeWindows(Time size, Time slide, Time offset)：

由于flink的窗口时间是自然时间，如果一个窗口的大小是1小时，那么一般就是

9:00-10:00 为一个窗口，而刚好业务想统计9:05-10:05这个窗口，那么offset设置 +5min的偏移量。另外一个场景是，统一时区，如flink的时间是标准的时间，而我们是东八区，如果是Time.days(1) ，格林尼治时间0:00-第二天0:00，由于我们是早八小时，那么就是 8:00-第二天8:00的数据，统计的就不是我们这边的一天的时间，这时候偏移量设置为 -8h。

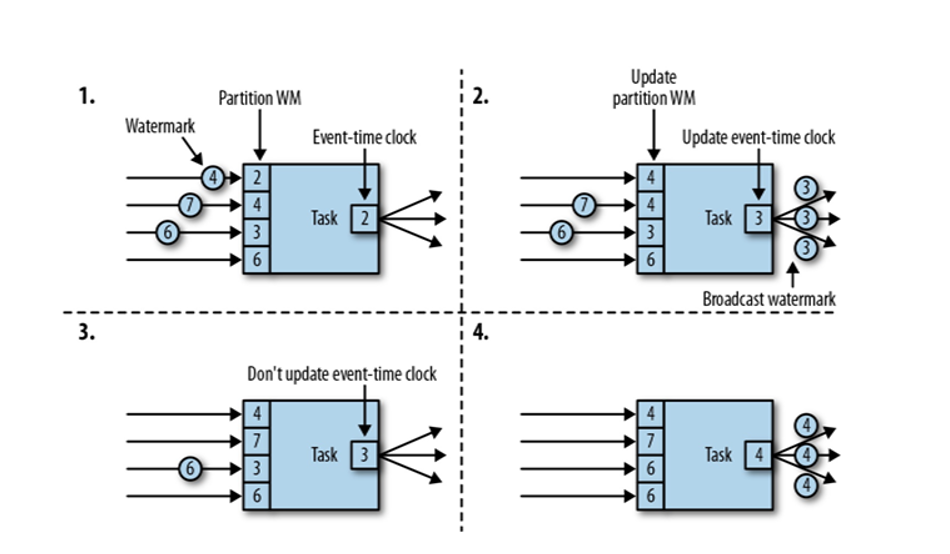
# 时间语义

## WaterMark

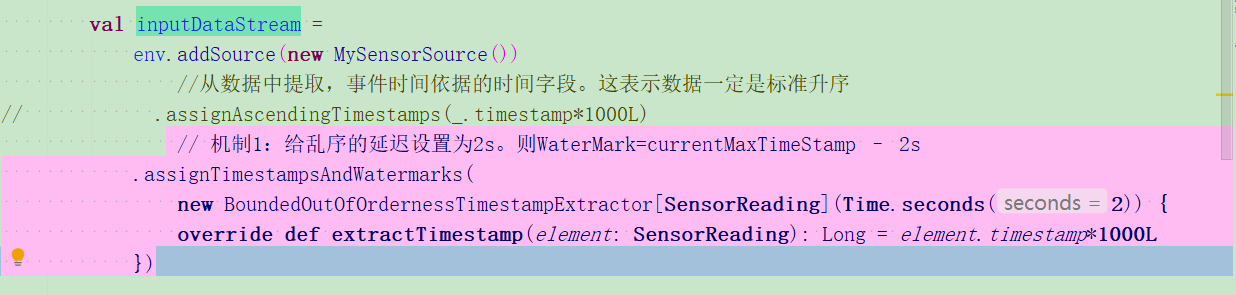
1. waterMark就是一个事件时间，代表当前时间的进展
2. WaterMark用来处理乱序数据，一般就是直接定义一个延迟时间，延迟触发窗口操作
3. WaterMark延迟时间的设置，一般要根据实际的乱序情况定，一般设置成最大乱序程度。
4. watermark 必须单调递增，以确保任务的事件时间时钟在向前推进。
5. WaterMark=currentMaxTimeStamp – lateness(允许延迟时间)，当

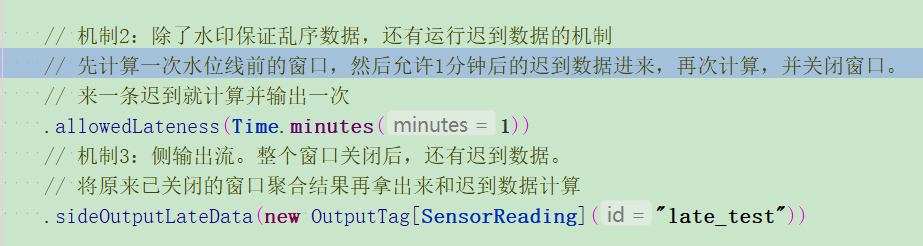
WaterMark >= WindowEndTime 时，就关闭一个窗口[ start, end)。会被任务，之后小于WaterMark时间长的数据不会再进来。

## WaterMark传递



## 乱序数据的三重保证





### 窗口的两个重要操作

* 1. 触发计算；② 清空状态（关闭窗口）

### 三种机制

WaterMark可以设置延迟时间

Window的allowedLateness方法，可以设置窗口允许处理迟到数据的时间

Window的sideOutputLateData方法，可以将迟到的数据写入侧输出流。

## 自定义时间戳

# ProcessFunction（底层API）