Flink学习路线

Flink练习环境准备

1. 源码编译
2. 本地运行
3. mysql+es+redis+kafka+zk本地搭建
4. hbase+hive+yarn有条件可以虚拟机搭建3台机器的小集群

Flink Basic API 练习

1. keyBy,groupBy,LambdaFunctions(reduce,filter),TransformationFunctions(map,RichMapFunction)
2. 如何用keyBy返回key的list
3. 读本地文件+hdfs文件
4. 快速创建DataSet and DataStream
5. 计数器Accumulators & Counters(new IntCounter(),addAccumulator(),getAccumulatorResult())

DataStream API

1. Source
    - kafka
    - mysql
    - 自定义source
2. Sink
    - kafka(kafka数据分流会用)
    - redis(实时去重复)
    - es(数据分析,报表展示)
    - hive(实时数据仓库)
    - hbase,mysql等等

3. Time
    - Event Time/Processing Time(这俩公司比较常用)，Ingestion Time
4. Watermarks
    - 原理(assignTimestampsAndWatermarks,Periodic Watermarks,Punctuated Watermarks)
    - 什么时候指定时间戳？是不是在获取到source之后立即指定时间戳?
    - 如何自定义时间戳，解决kafka多个partition时间交叉使用影响水印的问题
    - 如何处理fixed amount of lateness
5. State
    - KeyState,OperatorState(面试Tip:Kafka Connector 使用,支持在并行化情况改变的时候对并行算子状态（state）进行重分布)
    - TTL (Time-To-Live),什么时候触发TTL更新
    - 如何用State做实时去重复（天级全量实时去重复，TopN等等)
    - managed state和raw state区别，（面试Tip：并行化改变的时候该用哪一个？）
        - ValueState<T>
        - ListState<T>
        - ReducingState<T>
        - FoldingState<T, ACC>（未来会去掉）
        - MapState<UK, UV>
    - 存取状态(state)的方法
        - Even-split redistribution(均分重分布)
        - Union redistribution(联合重分布)
        - CheckpointedFunction
            - snapshotState无论何时要用到检查点，都要调用

initializeState用户自定义函数初始化的时候被调用

GetCheckpointLock()，source's context中获得一个锁(lock)保证输出原子化

- Asynchronous I/O for External Data Access

  AsyncFunction的实现以分派请求

  一个callback，用于取得operation的结果并交给AsyncCollector

    AsyncCollector在第一次调用AsyncCollector.collect时完成,所有后续的collect调用会被忽略

  像应用transformation那样在DataStream上应用异步I/O

  AsyncFunction不能多线程调用

DataSet API

- 后续更新（spark支持jobserver服务化调度，flink还不支持）

Table API & SQL

- 后续更新(业务复杂的场景会比较喜欢sql,后续会详细说为什么用sql好)

数据类型& 序列化

- 后续更新

管理执行

- 后续更新

项目实战（PPT晋升star方法论,有助于你晋升加薪噢）

S(SITUATION)业务状况和背景

  让面试官知道你在做什么

T(TASK)你有什么挑战和任务

  让面试官知道你的任务很有挑战，让他做都做不出来

A(ACTION)你做了哪些行动

你方案和思考(已经要多讲你的思考)很前卫，有创新

R(RESULT)结果如何

结果很完美，圆满完成任务，方案也很赞

祝愿大家升值加薪