## 容错机制

Spark streaming 程序是周而复始不间断运行的，但是网路异常、CPU负荷过高、内存溢出、硬盘读写异常断电等都会导致程序出现异常。

为了避免或减少故障所造成的影响，而且能让运行的系统从故障中恢复过来，使流数据处理能继续正常下去，所以就有了容错机制。

容错原理：

两个概念，

热备，即热备份，系统在正常工作时，在物理上做好备用硬件运行所需要的状态更新。遇到异常时，系统把受影响的工作自动切换到备用硬件上运行，以保证继续不间断进行。

冷备，即冷备份。系统在正常工作时就做好日常数据和元素据的备份。遇到异常时，系统通常需要通过人工重启，并利用备份的数据和元数据来恢复工作。

Spark streaming的热备是针对Executor进行的。

Streaming中的冷备有两个：冷备有预写日志( Write Ahead Log, WAL), 冷备检查点(Checkpoint)。

WAL是备份与后续工作相关的接受数据及其块的元数据信息。

Checkpoint，是为了在出错后恢复而做Spark Streaming应用程序配置、DStream操作、未处理作业等元数据信息的备份。

Checkpoint:

1. Metadata Checkpoint

将流式计算的信息保存到具备容错性的存储上，如HDFS，它适用于当spark streaming 应用程序Driver所在的节点出错时能够恢复，元数据包括：

* 配置信息。创建Spark Streaming应用程序的配置信息。
* DStream操作。在Spark Streaming应用程序中定义的DStreaming操作。
* 未完成的batches。在列队中没有出列完的作业。

1. Data Checkpointing

将生成的RDD保存到外部可靠的存储当中，当出错时可以直接从检查点恢复。

Restart Receiver

Resend unacked data

Recover Block data

From log

LOG

FileSystem

Recover Block Metadata from log

2

Relaunch Jobs

Restart Spark Context

Restart Streaming Context

1,3

Restart computation from info in

checkpointed

FileSystem

4

LOG

FileSystem

LOG

FileSystem

Application Driver

Restart Executor

5

Computation

checkpointed

4.

FileSystem

Block Metadata

Written to log

LOG

FileSystem

Block Metadata

2

Jobs

3

Application Driver

Block data

Written to both memory + log

LOG

FileSystem

Executor

1. Input stream

Receiver

Streaming Context

Spark Context