配置参数	默认值	描述
checkpointInterval	30	连续检查点之间的时间间隔(以秒为单位)
maxFileSize	1623195647	每个数据文件的最大大小(以字节为单位)。一旦文件达到这个大小(或一旦写入下一个事件将达到该大小),该文件保存关闭并在那个目录下创建一个新的数据文件。如果将此值设置为高于默认值,Channel 仍然使用默认值为最大值
minimumRequiredSpace	524288000	Channel继续操作时每个卷所需的最少的空间 (以字节为单位)。如果任何一个挂载数据目录 的卷只有这么多空间剩余, Channel 将停止操作 来防止损坏和避免不完整的数据被写入。这个 参数的最小值是 1048576 (1MB)
keep-alive	3	每次写入或读取应该等待完成的最大的时间周期(以秒为单位)

File Channel, 作为 Flume 主要的持久化 Channel, 通常整体上表现了 Agent 的性能。可以通过配置对 Channel 的几个方面进行调整。File Channel 有 capacity和 transactionCapacity 参数,这些和 Memory Channel 中的参数完全相同,但是它们的默认值更高,分别是 10000000 和 1000。

前面已经讨论过,File Channel 可以往多个磁盘写数据,但是 Channel 没有直接意识到这一点。不同磁盘可以挂载到不同的挂载点,Channel 可以配置以循环的方式往这些目录写数据。一个数据目录 Channel 总是会追加到一个文件,尽管 Channel 会按照需要从所有的文件中读取数据。因为多个 Source 可以写入到一个 Channel (并且多个线程可以写入一个 Source),Channel 将以不同的线程并行写到不同的数据目录(很多情况下,每个 Source 和 Sink 至少运行一个线程,例如 Avro Source 运行不止一个线程),因此并行磁盘的使用会有更好的性能。所以,如果你分配给 File Channel 更多起作用的磁盘,那么 Channel 的性能有可能有所改善。

File Channel 将逗号分隔的一系列数据目录作为 dataDirs 参数的值。默认值只能用来测试不建议用于生产环境使用。即使一个磁盘或有限数量的磁盘,每个磁盘可以使用多个数据目录以获得更好的性能,但是能提高性能的目录数量,随着磁盘的不同有所差异。

File Channel 定期写出到检查点以快速重启和恢复。checkpointDir 参数的值指定了检查点写入的目录。如果 Channel 正在写入检查点时停止,检查点可能是不完整或损坏的。

88