

# Flink 内存模型和管理

dafeihou@hillstonenet.com

2022 年 12 月 8 日

## 说明

本文基于 Flink 版本 1.14。

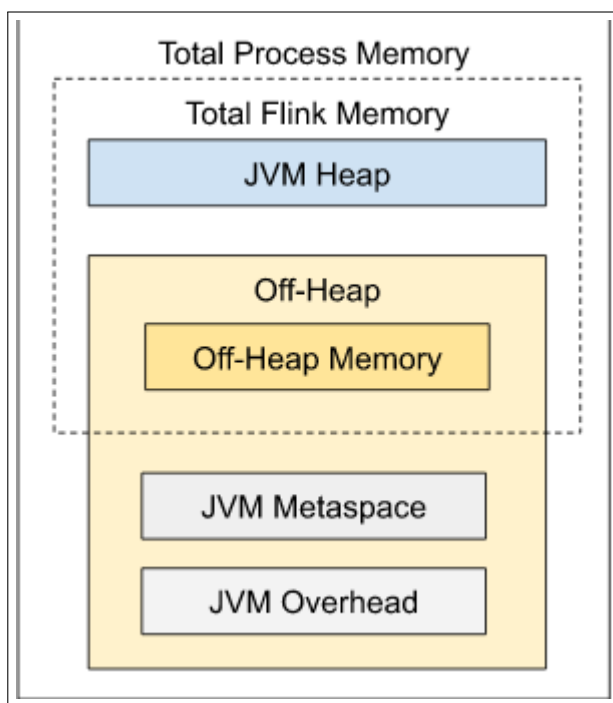
## 内存模型和配置

### 内存模型

参考: [Set up Flink's Process Memory](#)

Flink JVM 的内存总量包括 Flink 自身消耗的内存和运行进程消耗的 JVM 内存, FLink 自身的内存又包括 JVM Heap(堆内存) 和 Off-heap(对外内存) 内存。

即:



$$\begin{aligned}
 TotalFlinkMemory &= JVMHeap \\
 &\quad + OffHeap \\
 TotalProcessMemory &= TotalFlinkMemory \\
 &\quad + JVMMetaspace \\
 &\quad + JVMOverhead
 \end{aligned}$$

注: 每个组件的作用会在后续说明

手动管理 Flink 内存的方式比较简单，原生提供如下参数可供修改：

类型	Taskmanager	Jobmanager
Total Flink Memory	taskmanager.memory. flink.size	jobmanager.memory.flink.size
Total process Memory	taskmanager.memory. process.size	jobmanager.memory.process.size

## 详细的内存参数选项

参考：[Memory Configuration](#)

tips：Flink 虽然提供了如下参数来修改内存数据，但是官方建议是尽可能不要使用复杂的配置策略，在大多数情况下，用户应该只配置：

- taskmanager.memory. process.size 调整 total process 内存
- taskmanager.memory. flink.size 调整 total flink 内存
- taskmanager.memory. managed.fraction 根据百分比调整其他内存组件比值

一些比较详细的配置如下：

key	default value	类型	描述
jobmanager.memory.enable-jvm-direct-memory-limit	false	boolean	是否启用 jobmanager 的直接内存限制
jobmanager.memory.flink.size	null	memory	配置 jobmanager 的总 flink 内存大小
jobmanager.memory.heap.size	null	memory	配置 jobmanager 的堆内存大小，推荐最小值 128mb
jobmanager.memory.jvm-metaspace.size	256mb	memory	配置 jobmanager 的 jvm metaspace 大小
jobmanager.memory.jvm-overhead.fraction	0.1	float	为 jvm overhead 保留的 total process memory 百分比，即预留的一部分堆外内存
jobmanager.memory.jvm-overhead.max	1gb	memory	保留的最大值
jobmanager.memory.jvm-overhead.min	192mb	memory	保留的最小值
jobmanager.memory.off-heap.size	128mb	memory	此参数包含全部的堆外内存分配情况
jobmanager.memory.process.size	null	memory	jobmanager 的总进程内存大小，即上文 total process memory。在容器化部署中，这应该设置为容器内存。

key	default value	类型	描述
taskmanager.memory.flink.size	null	memory	total flink memory 大小
taskmanager.memory.framework.heap.size	128mb	memory	TaskExecutors 的框架堆内存大小, 并不分配给 slot
taskmanager.memory.framework.off-heap.batch-shuffle.size	32mb	memory	用于阻塞式 shuffle 和读取 *; 需 要 比 taskmanager.memory.framework.off-heap.size 小, 可以 为大规模的批处理加大该配置
taskmanager.memory.framework.off-heap.size	128mb	memory	TaskExecutors 的框架堆外存大小, 并不分配给 slot
taskmanager.memory.jvm-metaspace.size	256mb	memory	TaskExecutors 的 JVM metaspace 大小
taskmanager.memory.jvm-overhead.fraction	0.1	float	为 jvm overhead 开销保留的堆外内存
taskmanager.memory.jvm-overhead.max	1gb	memory	jvm overhead 最大值
taskmanager.memory.jvm-overhead.min	192mb	memory	jvm overhead 最小值
taskmanager.memory.managed.fraction	0.4	float	如果未明确托管内存大小, 则使用 total flink memory 的百分比做托管内存
taskmanager.memory.managed.size	null	memory	TaskExecutors 的托管内存大小, 用于排序、哈希表、中间结果的缓存、RocksDB 的状态后端。
taskmanager.memory.network.fraction	0.1	float	total flink memory 的一部分, 用于网络 buffer 等
taskmanager.memory.network.max	1gb	memory	最大
taskmanager.memory.network.min	64mb	memory	最小
taskmanager.memory.process.size	null	memory	TaskExecutors 的总进程大小, total process memory, 容器化部署时应和容器分配的内存一致。
taskmanager.memory.task.heap.size	null	memory	TaskExecutors 的 task heap memory, 如果未指定, 将是 TotalFlinkMemory - 框架堆内存 - 框架堆外内存 - 托管内存 - 托管堆外内存 - 网络内存。