

Flink State 最佳实践

本文主要分享与交流 Flink 状态使用过程中的一些经验与心得，当然标题取了“最佳实践”之名，希望文章内容能给读者带去一些干货。本文内容首先是回顾 state 相关概念，并认识和区别不同的 state backend；之后将分别对 state 使用访问以及 checkpoint 容错相关内容进行详细讲解，分享一些经验和心得。

State 概念回顾

我们先回顾一下到底什么是 state，流式计算的数据往往是转瞬即逝，当然，真实业务场景不可能说所有的数据都是进来之后就丢掉，没有任何东西留下来，那么留下来的东西其实就是称之为 state，中文可以翻译成状态。

在下面这个图中，我们的所有的原始数据进入用户代码之后再输出到下游，如果中间涉及到 state 的读写，这些状态会存储在本地的 state backend（可以对标成嵌入式本地 kv 存储）当中。

<https://ververica.cn/developers/flink-state-best-practices/>