

# iwos\_flink流数据处理

## Table of Contents

- [1. 部署](#)
  - [1.1. docker](#)
  - [1.2. 容错机制](#)
- [2. 应用开发](#)
  - [2.1. 说明](#)
  - [2.2. 任务类型](#)
  - [2.3. 编程语言](#)
  - [2.4. 编程模型抽象级别的挑选](#)
  - [2.5. java StreamAPI](#)
  - [2.6. python tableAPI](#)
  - [2.7. 推荐学习顺序](#)
  - [2.8. 制作模板](#)

## 1 部署

### 1.1 docker

- 查看flink-deploy

### 1.2 容错机制

- checkpoint
  - 需要在代码中配置([参考代码java](#))
  - [docs](#)
- savepoint
  - [docs](#)

## 2 应用开发

### 2.1 说明

- 微服务
- 编程语言, java, python都可以, java文档更全

### 2.2 任务类型

- 流数据实时处理
  - 单/多
- 流数据统计
  - 单/多

### 2.3 编程语言

- java
  - [项目构建设置docs](#)
- python
  - [构建pyflink](#)

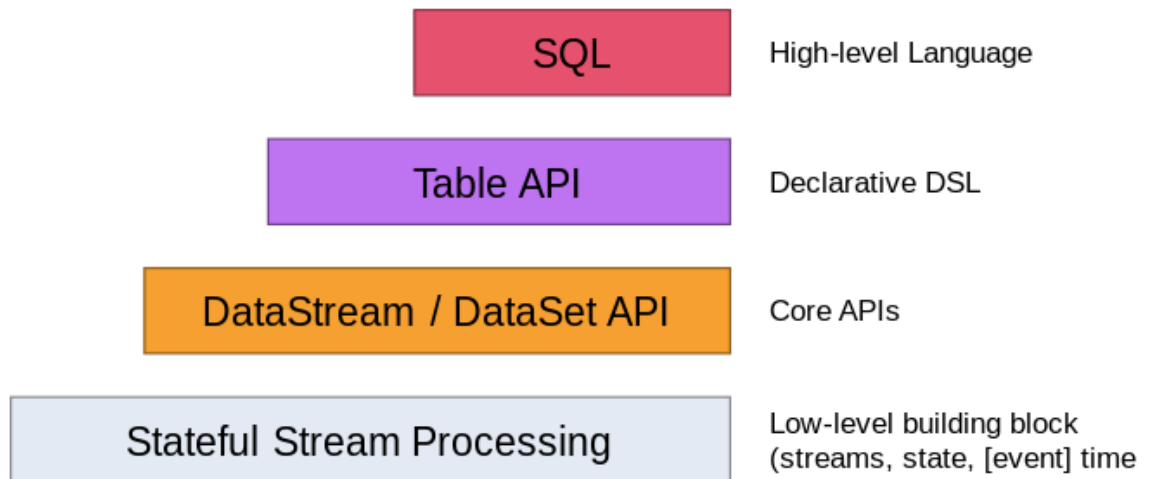
- 编程模型使用的抽象级别为:tableAPI
  - 特别说明：如果需要安装新版的pyflink必须使用上面的说明，pip安装的是1.0版本(**flink**)

## 2.4 编程模型抽象级别的挑选

•

# Levels of Abstraction

Flink offers different levels of abstraction to develop streaming/batch applications.



- 推荐: java主要使用coreAPIs(文档全，功能稳定), python使用declarativeDSL
- 推荐使用：stream(流数据)处理(flink相较于spark的优势)

## 2.5 java StreamAPI

- [项目模板](#)
- [docs](#)

## 2.6 python tableAPI

- pyflink
  1. [demo学习](#)
  2. [docs](#)
- flink
  - 有example

## 2.7 推荐学习顺序

- getting started
  - [https://ci.apache.org/projects/flink/flink-docs-release-1.9/zh/getting-started/tutorials/datastream\\_api.html](https://ci.apache.org/projects/flink/flink-docs-release-1.9/zh/getting-started/tutorials/datastream_api.html)
- StreamAPIs
  - 概览
  - Event Time
  - 算子
  - Connectors
    - Kafka

## 2.8 制作模板

- 包含内容

- 基本配置
- Checkpoint配置
- 流程
  - 环境
  - source
  - **operator** (自定义部分，对应算子)
  - sink
- Kafka Connector
- Serialization/Deserialization

Author: iws-m1

Created: 2019-11-20 三 15:15

[Emacs](#) 25.3.1 ([Org](#) mode 8.2.10)