# BLINK安装包

**Blink是阿里巴巴在flink上增强的一个版本，是flink的一个分支，未来两个分支会合并**

从下面链接下到的编译好的blink安装包

<http://fetching118.com/article/5.html>

# BLINK文档

Blink文档

<http://fetching118.com/blink_doc/concepts/programming-model.html>

对应Flink文档（参照着看，sql功能少很多）

<https://ci.apache.org/projects/flink/flink-docs-release-1.5/>

# BLINK开发及运行

首先要运行flink常驻任务

两种方式：

## 部署在yarn里（生产环境推荐这种）

配置HADOOP\_CONF\_DIR环境变量及YARN\_HOME环境变量

nohup ./bin/yarn-session.sh -n 2 -tm 2048 -jm 2048 -s 2 &

-n 2 两个manager

-tm 2048 taskmanager的内存为2048M

-jm 2048 jobmanager的内存为2048M

-s 2 双核

注意，这是一个常驻任务，保留了一个常驻的yarn-session，只需运行一次，

之后的./bin/flink run ….jar 或者./bin/sql-client embedded均会自动识别并使用该session

## 部署blink集群

相当简单，参照教程

<http://fetching118.com/blink_doc/ops/deployment/cluster_setup.html>

## 开发代码

着重看文档中开发的部分，尤其是例子，blink的tableapi对sql查询有很好的支持

<http://fetching118.com/blink_doc/dev/table/>

我的代码仓库，里面展示了查询含有“树”的单词及词频

[git@codemanager.pkpm.com:evanxuhe/flinkpkpm.git](mailto:git@codemanager.pkpm.com:evanxuhe/flinkpkpm.git)

mvn clean package -Dmaven.test.skip=true

将jar包上传至s2 /data/blink目录下

./bin/flink run ./ flink-pkpm-1.0-SNAPSHOT.jar--input hdfs://s2/xuhe

会正常打印



# Blink Sql客户端

./bin/sql-client.sh embedded

一下DDL会创建 表SqlTest与sqltest.csv的对应关系，通过表SqlTest增删改查，实际上是对csv的增删改查，反之亦然

CREATE TABLE SqlTest(

    id INT NOT NULL,

   name VARCHAR NOT NULL,

    PRIMARY KEY(id)

) WITH (

    type='csv',

    path='hdfs://s2/xuhe/sqltest.csv'

);

INSERT INTO `SqlTest` VALUES (100,’xuhe’)

将sqltest.csv从hdfs里get出来发现多了一行数据