上海大学
SHANGHAI UNIVERSITY

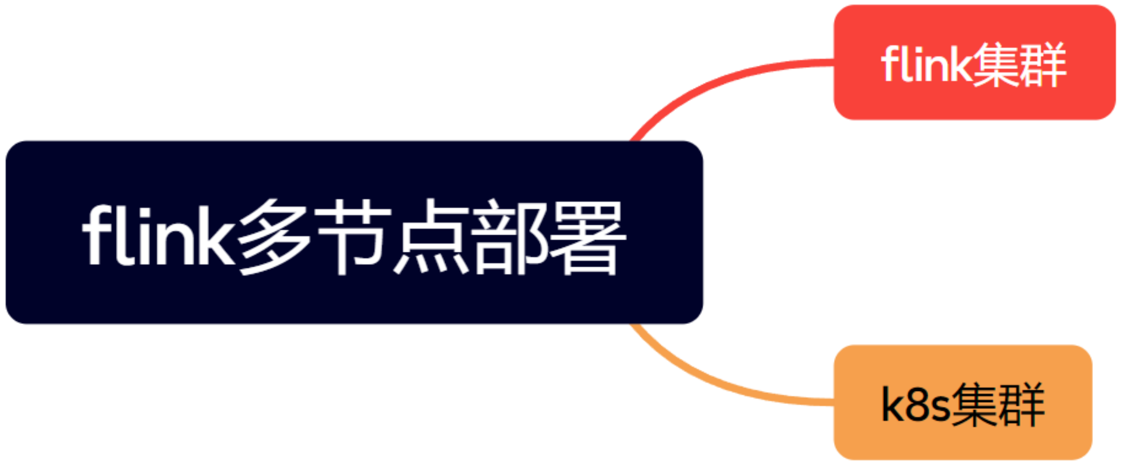
基于k8s的Flink SQL大规模数据相似度计算

汇报人：22121630 汪江豪
22120721 阮金桐

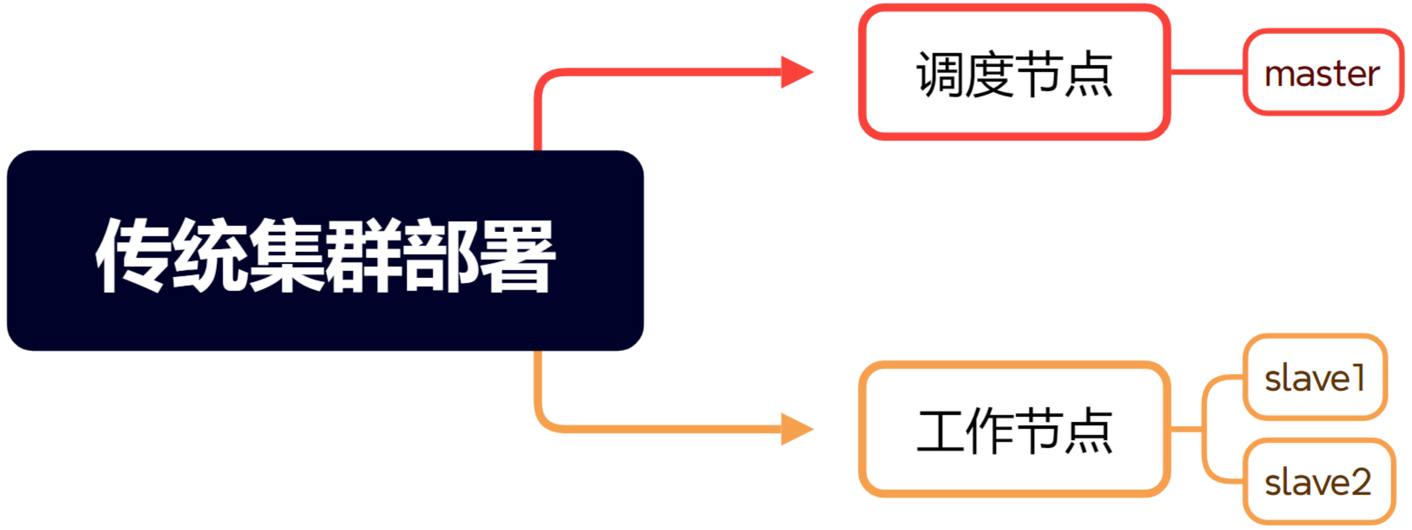
C O N T E N T

目　　录

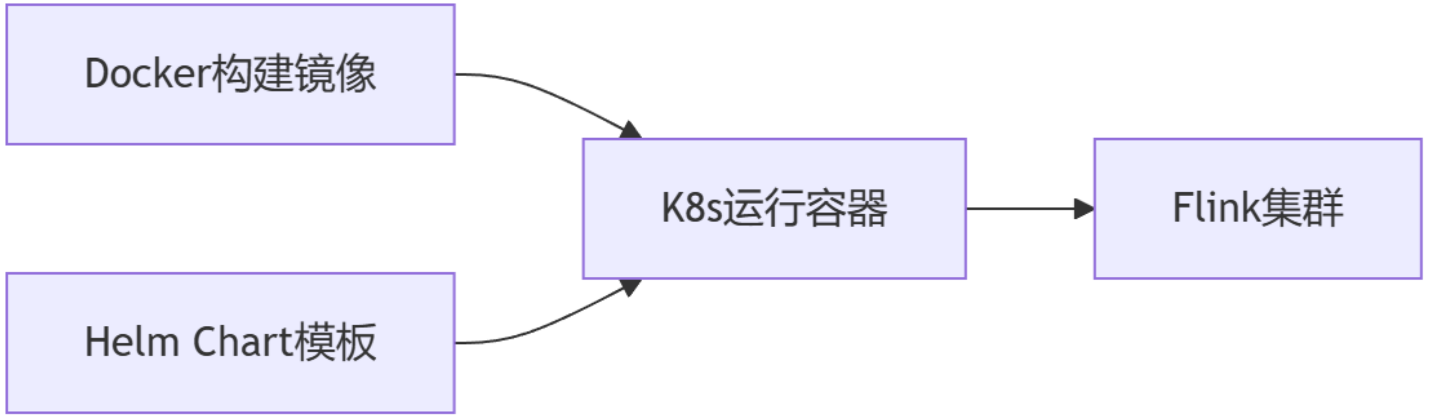
1.Flink多节点部署方式



传统集群方式
设置ssh无密码互联
修改flink/conf/flink-conf.yaml文件： obmanager.rpc.address: master taskmanager.numberOfTaskSlots: 2
修改conf目录下master文件为 master:8081
修改conf目录下worker文件为: slave1 slave2
slave1,slave2节点上设置同理 额外修改flink-conf.yaml文件中的 taskmanager.host:slave1/slave2 为自己主机名
在master节点运行start-cluster.sh，提交 jar包



依赖项



K8s是容器编排平台，用于自动化部署、扩展和管理容器化应用

Flink中的jobmanager和taskmanager作为pod运行在k8s集群中

Docker是容器化工具，可将应用及其依赖打包成轻量级、可移植容器镜像

Helm是k8s的包管理工具，类似于linux的apt

K8s集群方式部署Flink

启动k8s集群，节点加入集群

下载证书管理器
kubectl create -f https://github.com/jetstack/cert-manager/releases/download/v1.8.2/cert-manager.yaml

添加软件源
helm repo add flink-operator-repo https://downloads.apache.org/flink/flink-kubernetes-operator-1.11.0/

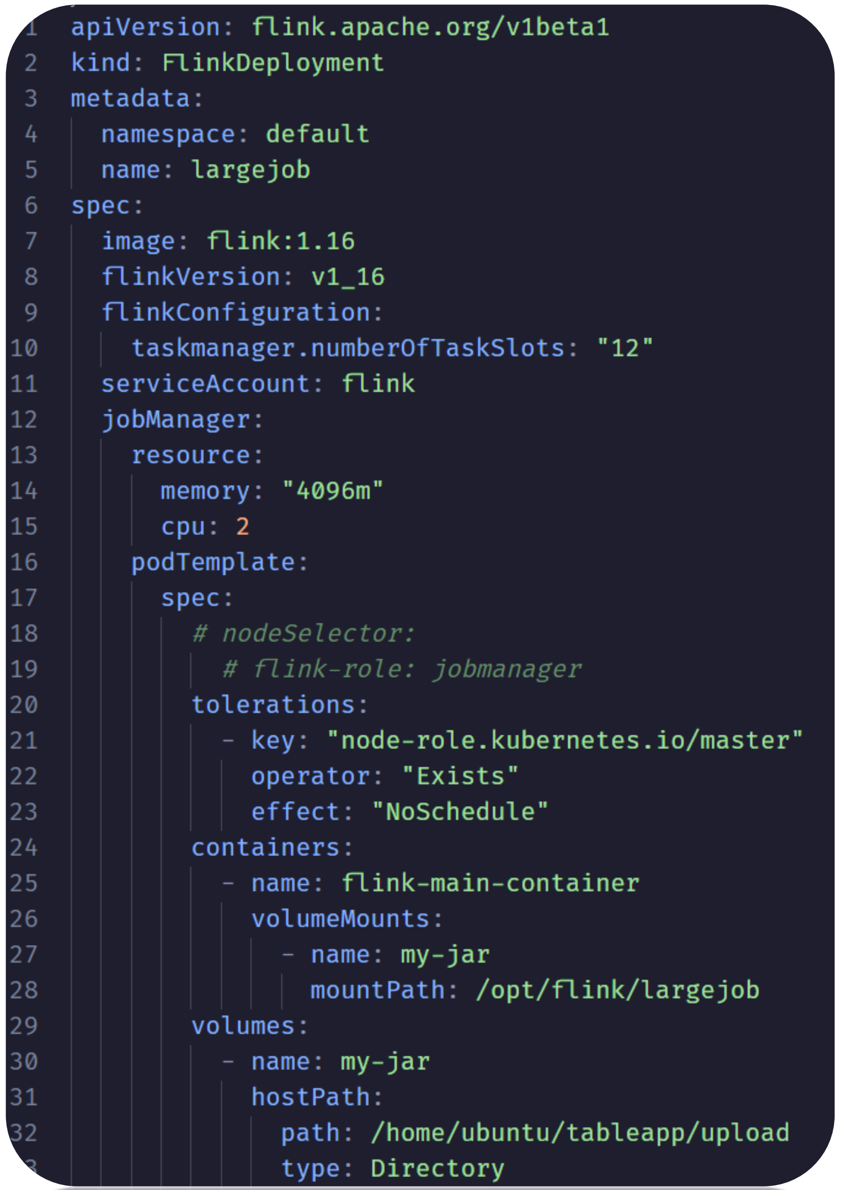
安装flink的k8s管理工具
helm install flink-kubernetes-operator flink-operator-repo/flink-kubernetes-operator

安装后operator作为pod保持运行，看到running说明安装成功
kubectl get pods

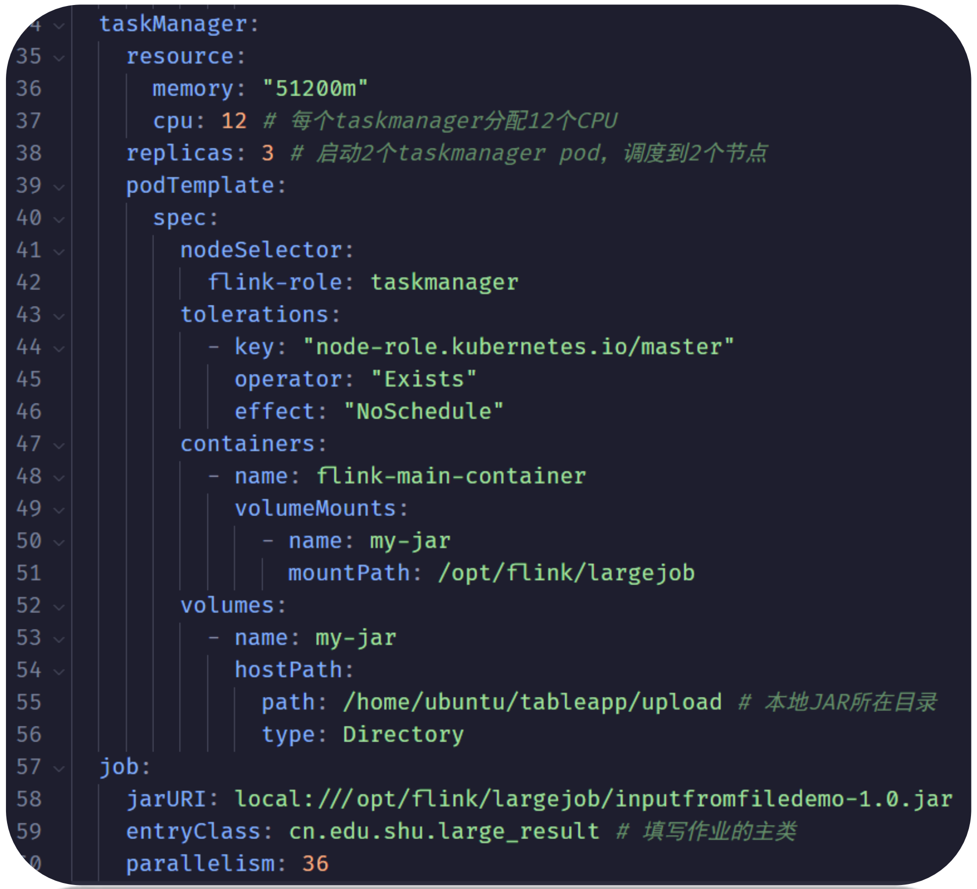
随后可通过yaml文件提交flink作业
Kubectl create –f **basic.yaml**

若要查看flink的webui，可进行端口转发，并允许外来地址访问
kubectl port-forward svc/largejob-rest 8081 --address 0.0.0.0

basic.yaml文件



```yaml
apiVersion: flink.apache.org/v1beta1
kind: FlinkDeployment
metadata:
  namespace: default
  name: largejob
spec:
  image: flink:1.16
  flinkVersion: v1_16
  flinkConfiguration:
    taskmanager.numberOfTaskSlots: "12"
  serviceAccount: flink
  jobManager:
    resource:
      memory: "4096m"
      cpu: 2
    podTemplate:
      spec:
        # nodeSelector:
          # flink-role: jobmanager
        tolerations:
          - key: "node-role.kubernetes.io/master"
            operator: "Exists"
            effect: "NoSchedule"
        containers:
          - name: flink-main-container
            volumeMounts:
              - name: my-jar
                mountPath: /opt/flink/largejob
        volumes:
          - name: my-jar
            hostPath:
              path: /home/ubuntu/tableapp/upload
              type: Directory
```



```yaml
  taskManager:
    resource:
      memory: "51200m"
      cpu: 12 # 每个taskmanager分配12个CPU
    replicas: 3 # 启动2个taskmanager pod，调度到2个节点
    podTemplate:
      spec:
        nodeSelector:
          flink-role: taskmanager
        tolerations:
          - key: "node-role.kubernetes.io/master"
            operator: "Exists"
            effect: "NoSchedule"
        containers:
          - name: flink-main-container
            volumeMounts:
              - name: my-jar
                mountPath: /opt/flink/largejob
        volumes:
          - name: my-jar
            hostPath:
              path: /home/ubuntu/tableapp/upload # 本地JAR所在目录
              type: Directory
  job:
    jarURI: local:///opt/flink/largejob/inputfromfiledemo-1.0.jar
    entryClass: cn.edu.shu.large_result # 填写作业的主类
    parallelism: 36
```

将本地数据集和jar包放在所有节点同一目录下，并修改权限为777
挂载到flink镜像中的某个目录

这里本地数据集和jar包位于/home/ubuntu/tableapp/upload

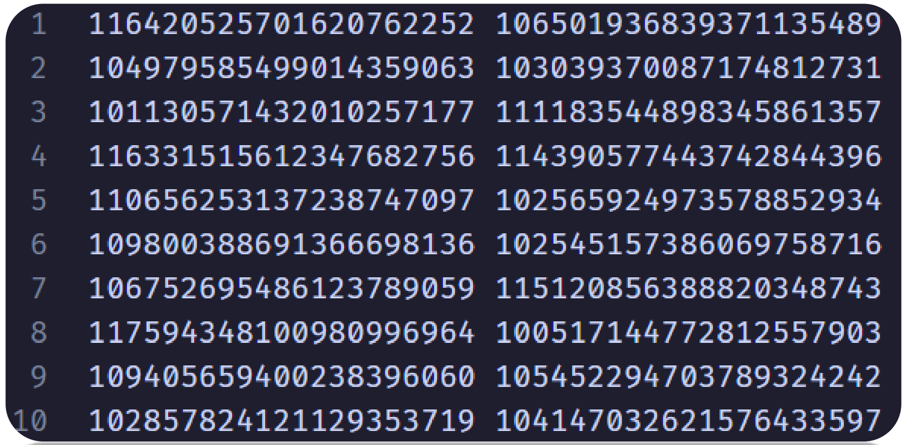
挂载目录为flink镜像中的/opt/flink/largejob

2.解法思路

解决办法

方法	配置	实际使用资源	总计耗时
Flink集群	3台4核8G 2个TM节点	2*4核8G	34m50s
K8s集群	3台16核64G	3*12核50G	8m50s

解法思路



```
116420525701620762252 106501936839371135489
104979585499014359063 103039370087174812731
101130571432010257177 111183544898345861357
116331515612347682756 114390577443742844396
110656253137238747097 102565924973578852934
109800388691366698136 102545157386069758716
106752695486123789059 115120856388820348743
117594348100980996964 100517144772812557903
109405659400238396060 105452294703789324242
102857824121129353719 104147032621576433597
```

约7万个　　约10万个

原始数据特点：
6832726行、21位数字字符串、无重复行

存储类型：
INT：10位×
BIGINT：19位×
STRING、DECIMAL性能开销大×

解决办法——转换数据集：

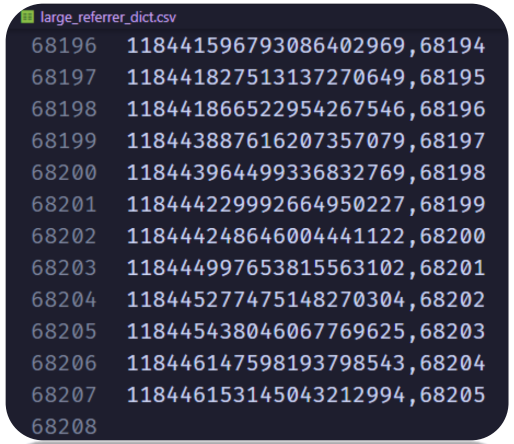
排序——两列数据分别映射到0-7万，0-10万

保存映射字典，供后续转换使用

使用INT存储

映射字典：

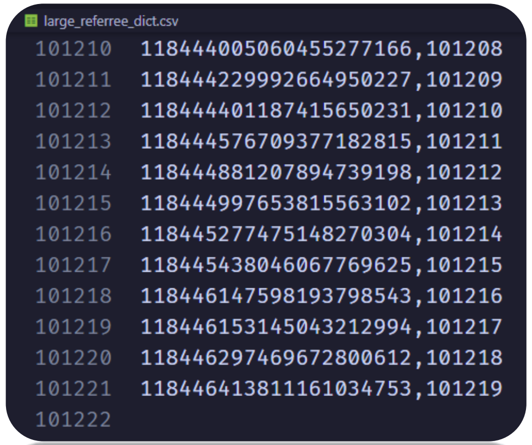
第一列



large_referrer_dict.csv

```
118441596793086402969,68194
```

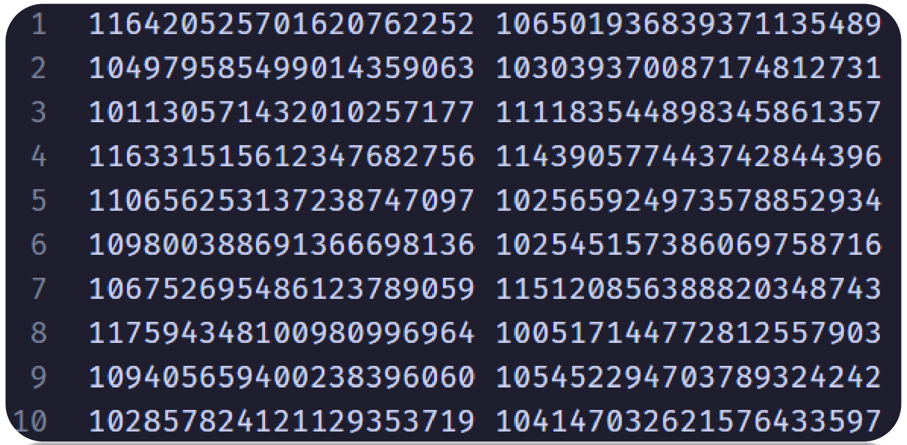
第二列



large_referee_dict.csv

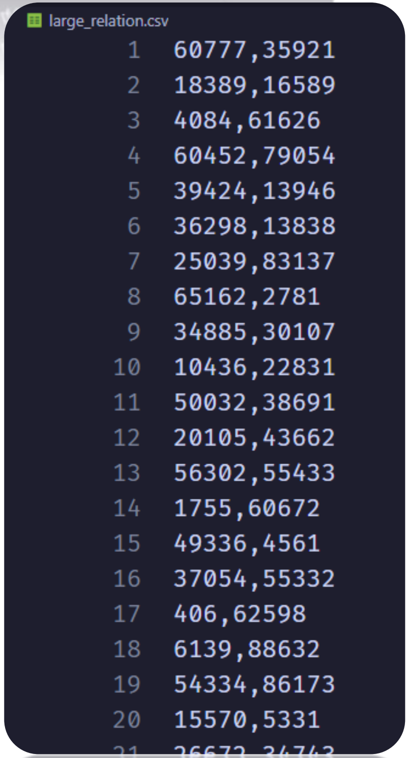
```
118444005060455277166,101208
```

原始数据集：286M



```
116420525701620762252 106501936839371135489
```

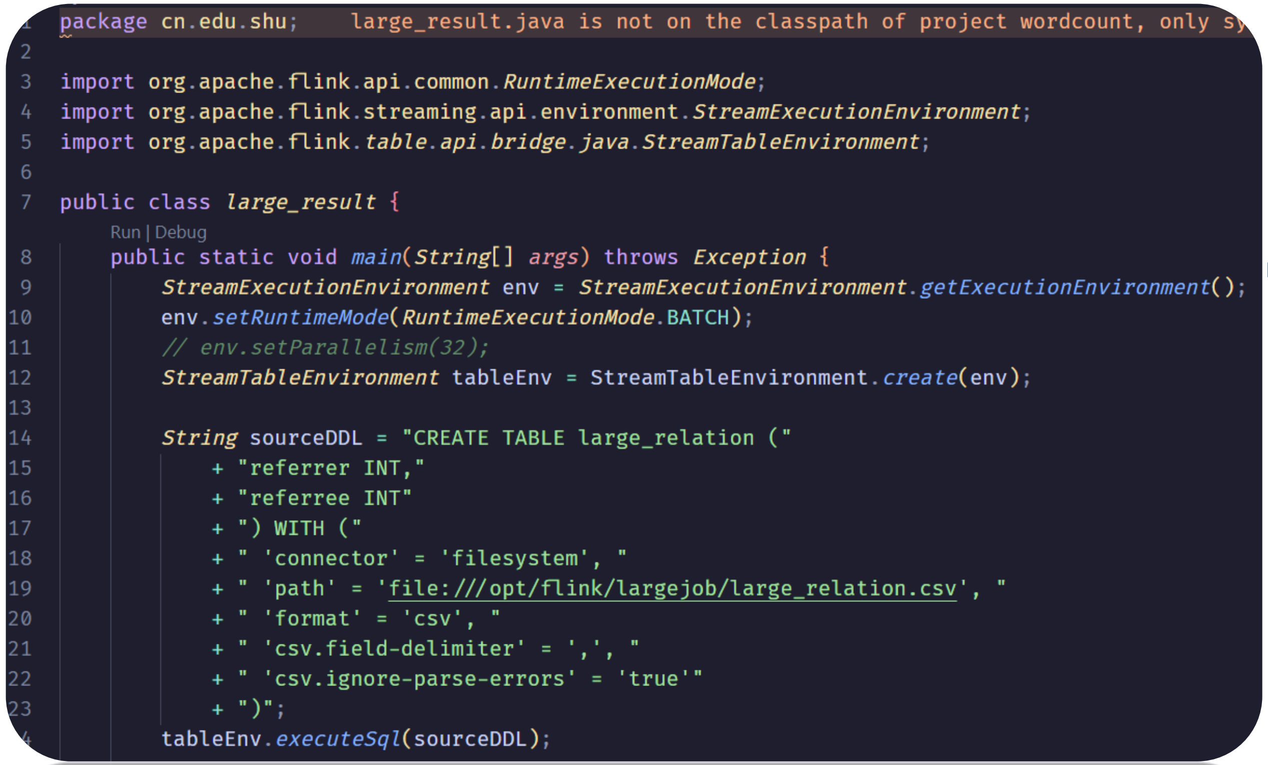
转换后数据集83M



large_relation.csv

```
60777,35921
```

JAVA代码



```
package cn.edu.shu;    large_result.java is not on the classpath of project wordcount, only sy

import org.apache.flink.api.common.RuntimeExecutionMode;
import org.apache.flink.streaming.api.environment.StreamExecutionEnvironment;
import org.apache.flink.table.api.bridge.java.StreamTableEnvironment;

public class large_result {
    Run | Debug
    public static void main(String[] args) throws Exception {
        StreamExecutionEnvironment env = StreamExecutionEnvironment.getExecutionEnvironment();
        env.setRuntimeMode(RuntimeExecutionMode.BATCH);
        // env.setParallelism(32);
        StreamTableEnvironment tableEnv = StreamTableEnvironment.create(env);

        String sourceDDL = "CREATE TABLE large_relation ("
            + "referrer INT,"
            + "referree INT"
            + ") WITH ("
            + " 'connector' = 'filesystem', "
            + " 'path' = 'file:///opt/flink/largejob/large_relation.csv', "
            + " 'format' = 'csv', "
            + " 'csv.field-delimiter' = ',', "
            + " 'csv.ignore-parse-errors' = 'true'"
            + ")";
        tableEnv.executeSql(sourceDDL);
```

在flink1.12之后，批处理模式通常比流处理模式在处理静态大数据集时效率更高：

批处理模式会自动做全局优化如(全局排序、资源复用等)。

流处理模式主要为低延迟，适合实时场景。

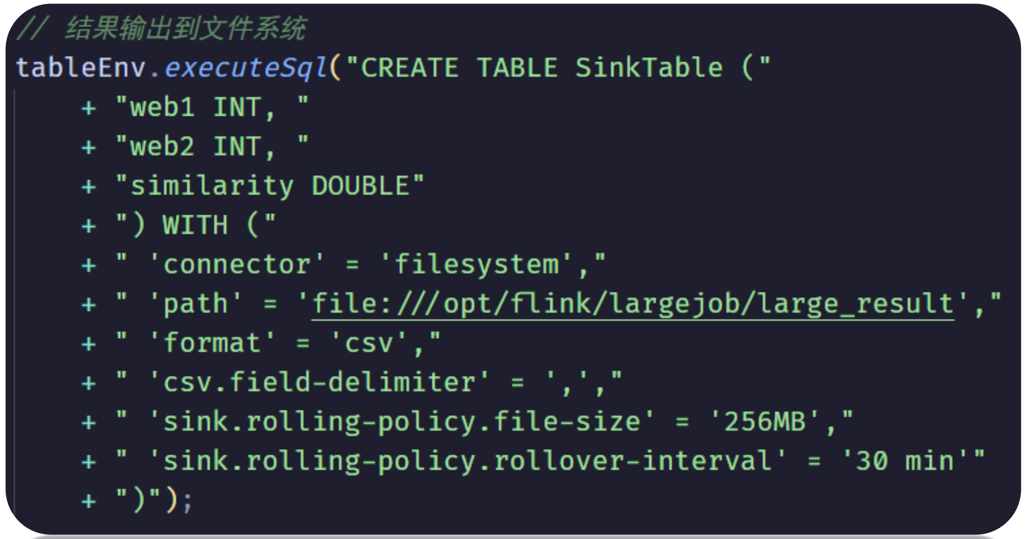
在传统集群方式中，
java代码设置并行数以提高效率

K8s集群方式，
java代码不用设置并行数，
在yaml文件中指定

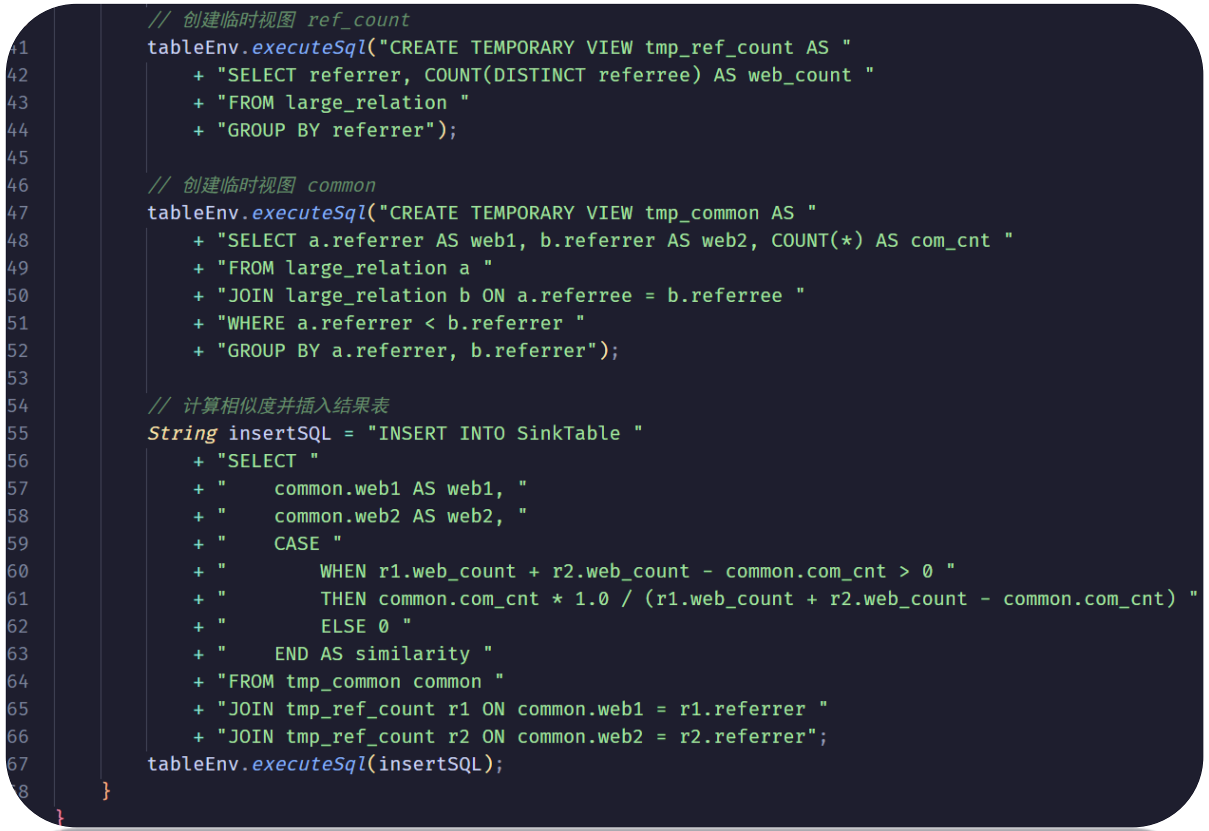
JAVA代码续

对于传统flink集群方式，path为本地数据集路径。

对于k8s集群方式，path为flink镜像中的挂载目录



```
// 结果输出到文件系统
tableEnv.executeSql("CREATE TABLE SinkTable ("
    + "web1 INT, "
    + "web2 INT, "
    + "similarity DOUBLE"
    + ") WITH ("
    + " 'connector' = 'filesystem',"
    + " 'path' = 'file:///opt/flink/largejob/large_result',"
    + " 'format' = 'csv',"
    + " 'csv.field-delimiter' = ',',"
    + " 'sink.rolling-policy.file-size' = '256MB',"
    + " 'sink.rolling-policy.rollover-interval' = '30 min'"
    + ")");
```

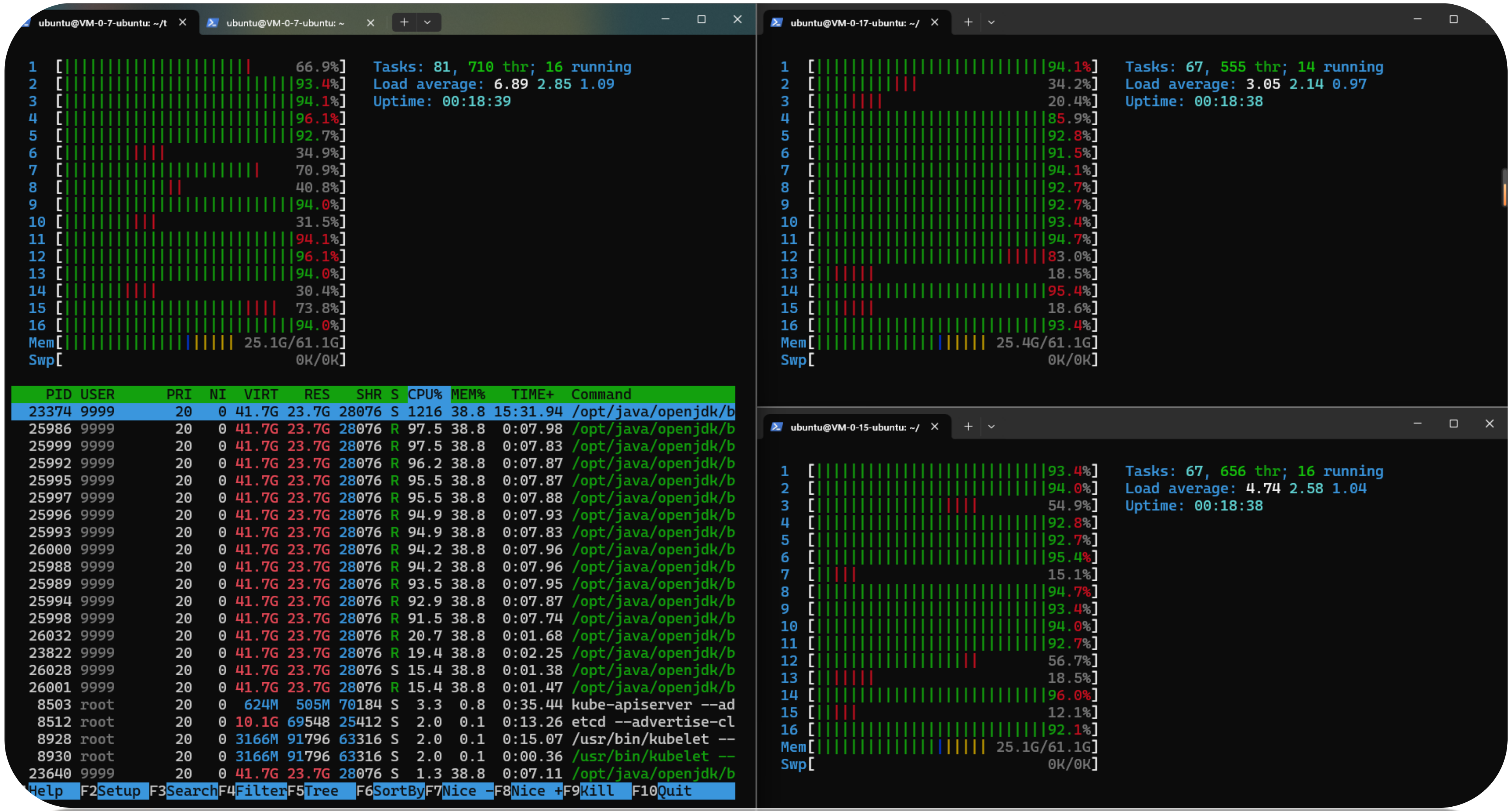


```
        // 创建临时视图 ref_count
        tableEnv.executeSql("CREATE TEMPORARY VIEW tmp_ref_count AS "
            + "SELECT referrer, COUNT(DISTINCT referree) AS web_count "
            + "FROM large_relation "
            + "GROUP BY referrer");

        // 创建临时视图 common
        tableEnv.executeSql("CREATE TEMPORARY VIEW tmp_common AS "
            + "SELECT a.referrer AS web1, b.referrer AS web2, COUNT(*) AS com_cnt "
            + "FROM large_relation a "
            + "JOIN large_relation b ON a.referree = b.referree "
            + "WHERE a.referrer < b.referrer "
            + "GROUP BY a.referrer, b.referrer");

        // 计算相似度并插入结果表
        String insertSQL = "INSERT INTO SinkTable "
            + "SELECT "
            + "    common.web1 AS web1, "
            + "    common.web2 AS web2, "
            + "    CASE "
            + "        WHEN r1.web_count + r2.web_count - common.com_cnt > 0 "
            + "        THEN common.com_cnt * 1.0 / (r1.web_count + r2.web_count - common.com_cnt) "
            + "        ELSE 0 "
            + "    END AS similarity "
            + "FROM tmp_common common "
            + "JOIN tmp_ref_count r1 ON common.web1 = r1.referrer "
            + "JOIN tmp_ref_count r2 ON common.web2 = r2.referrer";
        tableEnv.executeSql(insertSQL);
    }
}
```

在basic.yaml文件中，将数据集和jar的目录挂载好后，通过命令：

kubectl create –f basic.yaml创建任务，会自动创建镜像，执行查询

运行截图



ubuntu@VM-0-7-ubuntu: ~/t   ubuntu@VM-0-7-ubuntu: ~

```
  1  [|||||||||||||||||||||      66.9%]   Tasks: 81, 710 thr; 16 running
  2  [|||||||||||||||||||||||||||93.4%]   Load average: 6.89 2.85 1.09
  3  [|||||||||||||||||||||||||||94.1%]   Uptime: 00:18:39
  4  [|||||||||||||||||||||||||||96.1%]
  5  [|||||||||||||||||||||||||||92.7%]
  6  [||||||||||||               34.9%]
  7  [||||||||||||||||||||||     70.9%]
  8  [|||||||||||||              40.8%]
  9  [|||||||||||||||||||||||||||94.0%]
  10 [|||||||||||                31.5%]
  11 [|||||||||||||||||||||||||||94.1%]
  12 [|||||||||||||||||||||||||||96.1%]
  13 [|||||||||||||||||||||||||||94.0%]
  14 [|||||||||||                30.4%]
  15 [|||||||||||||||||||||||    73.8%]
  16 [|||||||||||||||||||||||||||94.0%]
  Mem[||||||||||||||||||||  25.1G/61.1G]
  Swp[                           0K/0K]

    PID USER      PRI  NI  VIRT   RES   SHR S CPU% MEM%   TIME+  Command
  23374 9999       20   0 41.7G 23.7G 28076 S 1216 38.8 15:31.94 /opt/java/openjdk/b
  25986 9999       20   0 41.7G 23.7G 28076 R 97.5 38.8  0:07.98 /opt/java/openjdk/b
  25999 9999       20   0 41.7G 23.7G 28076 R 97.5 38.8  0:07.83 /opt/java/openjdk/b
  25992 9999       20   0 41.7G 23.7G 28076 R 96.2 38.8  0:07.87 /opt/java/openjdk/b
  25995 9999       20   0 41.7G 23.7G 28076 R 95.5 38.8  0:07.87 /opt/java/openjdk/b
  25997 9999       20   0 41.7G 23.7G 28076 R 95.5 38.8  0:07.88 /opt/java/openjdk/b
  25996 9999       20   0 41.7G 23.7G 28076 R 94.9 38.8  0:07.93 /opt/java/openjdk/b
  25993 9999       20   0 41.7G 23.7G 28076 R 94.9 38.8  0:07.83 /opt/java/openjdk/b
  26000 9999       20   0 41.7G 23.7G 28076 R 94.2 38.8  0:07.96 /opt/java/openjdk/b
  25988 9999       20   0 41.7G 23.7G 28076 R 94.2 38.8  0:07.96 /opt/java/openjdk/b
  25989 9999       20   0 41.7G 23.7G 28076 R 93.5 38.8  0:07.95 /opt/java/openjdk/b
  25994 9999       20   0 41.7G 23.7G 28076 R 92.9 38.8  0:07.87 /opt/java/openjdk/b
  25998 9999       20   0 41.7G 23.7G 28076 R 91.5 38.8  0:07.74 /opt/java/openjdk/b
  26032 9999       20   0 41.7G 23.7G 28076 R 20.7 38.8  0:01.68 /opt/java/openjdk/b
  23822 9999       20   0 41.7G 23.7G 28076 R 19.4 38.8  0:02.25 /opt/java/openjdk/b
  26028 9999       20   0 41.7G 23.7G 28076 S 15.4 38.8  0:01.38 /opt/java/openjdk/b
  26001 9999       20   0 41.7G 23.7G 28076 R 15.4 38.8  0:01.47 /opt/java/openjdk/b
   8503 root       20   0  624M  505M 70184 S  3.3  0.8  0:35.44 kube-apiserver --ad
   8512 root       20   0 10.1G 69548 25412 S  2.0  0.1  0:13.26 etcd --advertise-cl
   8928 root       20   0 3166M 91796 63316 S  2.0  0.1  0:15.07 /usr/bin/kubelet --
   8930 root       20   0 3166M 91796 63316 S  2.0  0.1  0:00.36 /usr/bin/kubelet --
  23640 9999       20   0 41.7G 23.7G 28076 S  1.3 38.8  0:07.11 /opt/java/openjdk/b
F1Help  F2Setup F3SearchF4FilterF5Tree  F6SortByF7Nice -F8Nice +F9Kill  F10Quit
```

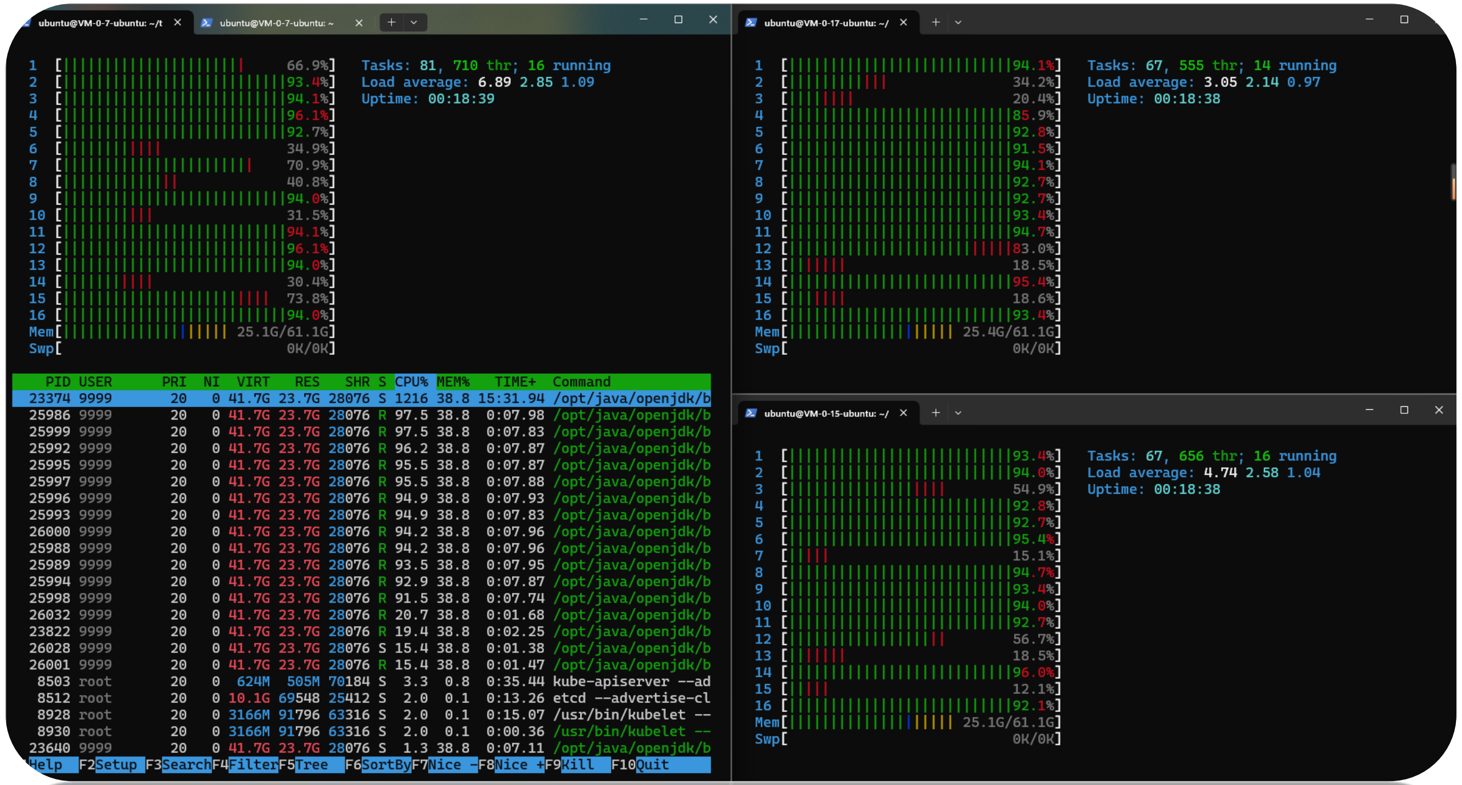
ubuntu@VM-0-17-ubuntu: ~/

```
  1  [|||||||||||||||||||||||||||94.1%]   Tasks: 67, 555 thr; 14 running
  2  [|||||||||||                34.2%]   Load average: 3.05 2.14 0.97
  3  [|||||||                    20.4%]   Uptime: 00:18:38
  4  [|||||||||||||||||||||||||||85.9%]
  5  [|||||||||||||||||||||||||||92.8%]
  6  [|||||||||||||||||||||||||||91.5%]
  7  [|||||||||||||||||||||||||||94.1%]
  8  [|||||||||||||||||||||||||||92.7%]
  9  [|||||||||||||||||||||||||||92.7%]
  10 [|||||||||||||||||||||||||||93.4%]
  11 [|||||||||||||||||||||||||||94.7%]
  12 [|||||||||||||||||||||||||||83.0%]
  13 [|||||||                    18.5%]
  14 [|||||||||||||||||||||||||||95.4%]
  15 [|||||||                    18.6%]
  16 [|||||||||||||||||||||||||||93.4%]
  Mem[||||||||||||||||||||  25.4G/61.1G]
  Swp[                           0K/0K]
```

ubuntu@VM-0-15-ubuntu: ~/

```
  1  [|||||||||||||||||||||||||||93.4%]   Tasks: 67, 656 thr; 16 running
  2  [|||||||||||||||||||||||||||94.0%]   Load average: 4.74 2.58 1.04
  3  [|||||||||||||||||||        54.9%]   Uptime: 00:18:38
  4  [|||||||||||||||||||||||||||92.8%]
  5  [|||||||||||||||||||||||||||92.7%]
  6  [|||||||||||||||||||||||||||95.4%]
  7  [|||||                      15.1%]
  8  [|||||||||||||||||||||||||||94.7%]
  9  [|||||||||||||||||||||||||||93.4%]
  10 [|||||||||||||||||||||||||||94.0%]
  11 [|||||||||||||||||||||||||||92.7%]
  12 [|||||||||||||||||||        56.7%]
  13 [|||||||                    18.5%]
  14 [|||||||||||||||||||||||||||96.0%]
  15 [||||                       12.1%]
  16 [|||||||||||||||||||||||||||92.1%]
  Mem[||||||||||||||||||||  25.1G/61.1G]
  Swp[                           0K/0K]
```

运行截图


ubuntu@VM-0-7-ubuntu: ~/t ubuntu@VM-0-7-ubuntu: ~

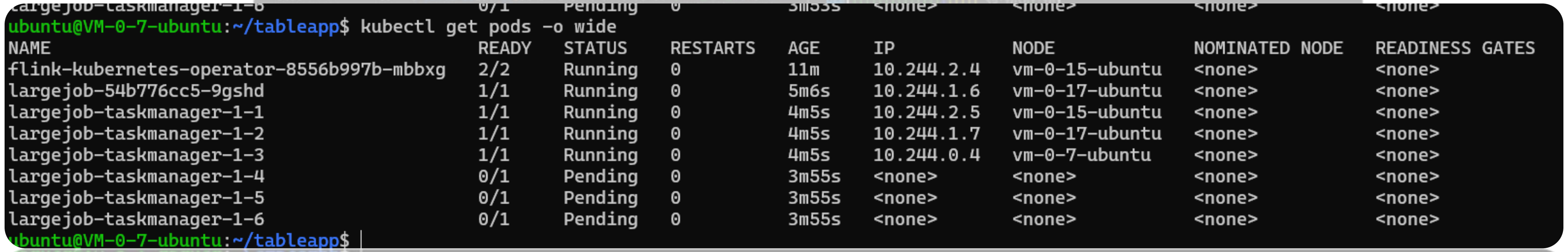
```
  1  [|||||||||||||||||||||        66.9%]   Tasks: 81, 710 thr; 16 running
  2  [|||||||||||||||||||||||||||||93.4%]   Load average: 6.89 2.85 1.09
  3  [|||||||||||||||||||||||||||||94.1%]   Uptime: 00:18:39
  4  [|||||||||||||||||||||||||||||96.1%]
  5  [|||||||||||||||||||||||||||||92.7%]
  6  [|||||||||||                  34.9%]
  7  [||||||||||||||||||||||       70.9%]
  8  [||||||||||||||               40.8%]
  9  [|||||||||||||||||||||||||||||94.0%]
  10 [||||||||||                   31.5%]
  11 [|||||||||||||||||||||||||||||94.1%]
  12 [|||||||||||||||||||||||||||||96.1%]
  13 [|||||||||||||||||||||||||||||94.0%]
  14 [||||||||||                   30.4%]
  15 [|||||||||||||||||||||||||    73.8%]
  16 [|||||||||||||||||||||||||||||94.0%]
  Mem[||||||||||||||||||||    25.1G/61.1G]
  Swp[                            0K/0K]

    PID USER       PRI  NI  VIRT   RES   SHR S CPU% MEM%   TIME+  Command
  23374 9999        20   0 41.7G 23.7G 28076 S 1216 38.8 15:31.94 /opt/java/openjdk/b
  25986 9999        20   0 41.7G 23.7G 28076 R 97.5 38.8  0:07.98 /opt/java/openjdk/b
  25999 9999        20   0 41.7G 23.7G 28076 R 97.5 38.8  0:07.83 /opt/java/openjdk/b
  25992 9999        20   0 41.7G 23.7G 28076 R 96.2 38.8  0:07.87 /opt/java/openjdk/b
  25995 9999        20   0 41.7G 23.7G 28076 R 95.5 38.8  0:07.87 /opt/java/openjdk/b
  25997 9999        20   0 41.7G 23.7G 28076 R 95.5 38.8  0:07.88 /opt/java/openjdk/b
  25996 9999        20   0 41.7G 23.7G 28076 R 94.9 38.8  0:07.93 /opt/java/openjdk/b
  25993 9999        20   0 41.7G 23.7G 28076 R 94.9 38.8  0:07.83 /opt/java/openjdk/b
  26000 9999        20   0 41.7G 23.7G 28076 R 94.2 38.8  0:07.96 /opt/java/openjdk/b
  25988 9999        20   0 41.7G 23.7G 28076 R 94.2 38.8  0:07.96 /opt/java/openjdk/b
  25989 9999        20   0 41.7G 23.7G 28076 R 93.5 38.8  0:07.95 /opt/java/openjdk/b
  25994 9999        20   0 41.7G 23.7G 28076 R 92.9 38.8  0:07.87 /opt/java/openjdk/b
  25998 9999        20   0 41.7G 23.7G 28076 R 91.5 38.8  0:07.74 /opt/java/openjdk/b
  26032 9999        20   0 41.7G 23.7G 28076 R 20.7 38.8  0:01.68 /opt/java/openjdk/b
  23822 9999        20   0 41.7G 23.7G 28076 R 19.4 38.8  0:02.25 /opt/java/openjdk/b
  26028 9999        20   0 41.7G 23.7G 28076 S 15.4 38.8  0:01.38 /opt/java/openjdk/b
  26001 9999        20   0 41.7G 23.7G 28076 R 15.4 38.8  0:01.47 /opt/java/openjdk/b
   8503 root        20   0  624M  505M 70184 S  3.3  0.8  0:35.44 kube-apiserver --ad
   8512 root        20   0 10.1G 69548 25412 S  2.0  0.1  0:13.26 etcd --advertise-cl
   8928 root        20   0 3166M 91796 63316 S  2.0  0.1  0:15.07 /usr/bin/kubelet --
   8930 root        20   0 3166M 91796 63316 S  2.0  0.1  0:00.36 /usr/bin/kubelet --
  23640 9999        20   0 41.7G 23.7G 28076 S  1.3 38.8  0:07.11 /opt/java/openjdk/b
Help  F2Setup F3SearchF4FilterF5Tree  F6SortByF7Nice -F8Nice +F9Kill  F10Quit
```

ubuntu@VM-0-17-ubuntu: ~/

```
  1  [|||||||||||||||||||||||||||||94.1%]   Tasks: 67, 555 thr; 14 running
  2  [|||||||||||                  34.2%]   Load average: 3.05 2.14 0.97
  3  [|||||||                      20.4%]   Uptime: 00:18:38
  4  [|||||||||||||||||||||||||||||85.9%]
  5  [|||||||||||||||||||||||||||||92.8%]
  6  [|||||||||||||||||||||||||||||91.5%]
  7  [|||||||||||||||||||||||||||||94.1%]
  8  [|||||||||||||||||||||||||||||92.7%]
  9  [|||||||||||||||||||||||||||||92.7%]
  10 [|||||||||||||||||||||||||||||93.4%]
  11 [|||||||||||||||||||||||||||||94.7%]
  12 [|||||||||||||||||||||||||||||83.0%]
  13 [||||||                       18.5%]
  14 [|||||||||||||||||||||||||||||95.4%]
  15 [||||||                       18.6%]
  16 [|||||||||||||||||||||||||||||93.4%]
  Mem[||||||||||||||||||||    25.4G/61.1G]
  Swp[                            0K/0K]
```

ubuntu@VM-0-15-ubuntu: ~/

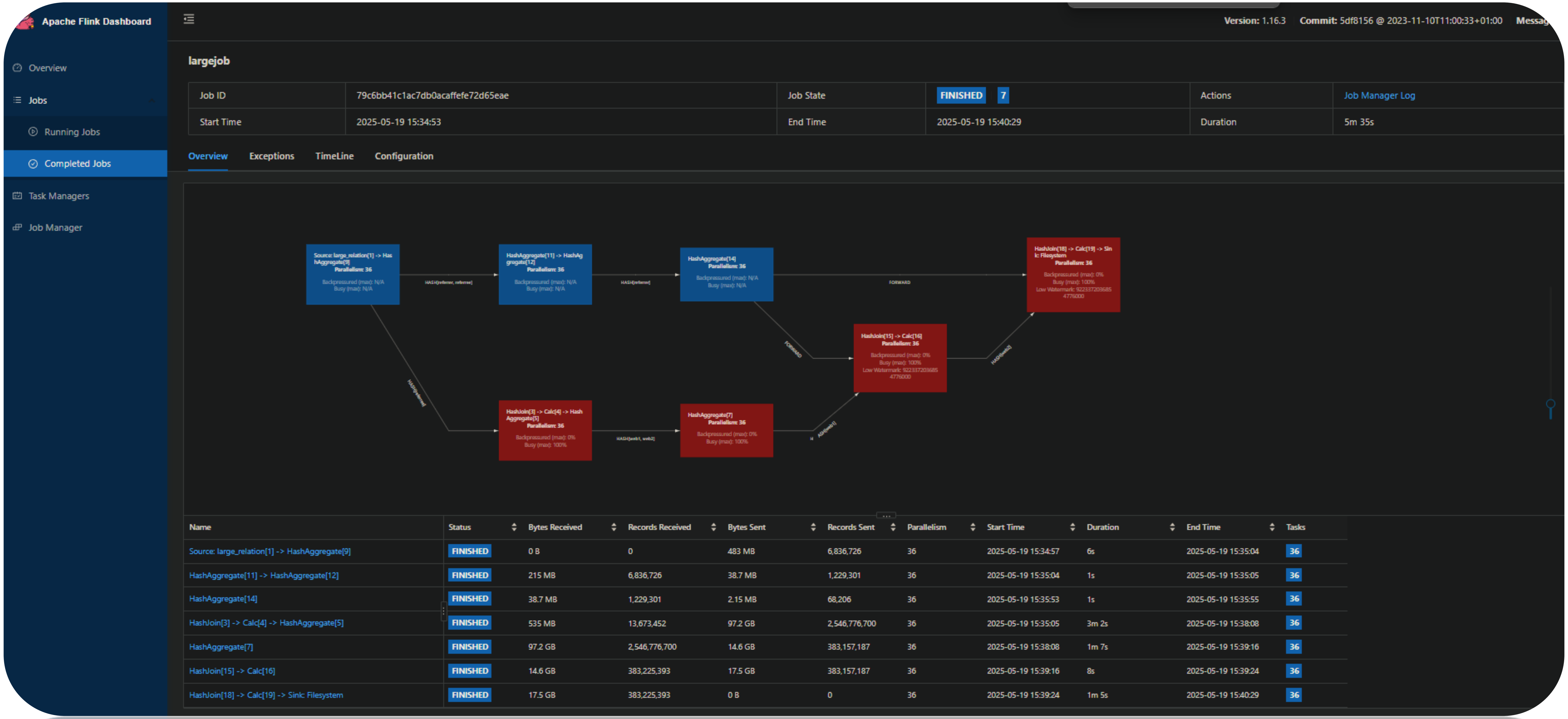
```
  1  [|||||||||||||||||||||||||||||93.4%]   Tasks: 67, 656 thr; 16 running
  2  [|||||||||||||||||||||||||||||94.0%]   Load average: 4.74 2.58 1.04
  3  [||||||||||||||||||           54.9%]   Uptime: 00:18:38
  4  [|||||||||||||||||||||||||||||92.8%]
  5  [|||||||||||||||||||||||||||||92.7%]
  6  [|||||||||||||||||||||||||||||95.4%]
  7  [|||||                        15.1%]
  8  [|||||||||||||||||||||||||||||94.7%]
  9  [|||||||||||||||||||||||||||||93.4%]
  10 [|||||||||||||||||||||||||||||94.0%]
  11 [|||||||||||||||||||||||||||||92.7%]
  12 [|||||||||||||||||||          56.7%]
  13 [||||||                       18.5%]
  14 [|||||||||||||||||||||||||||||96.0%]
  15 [|||||                        12.1%]
  16 [|||||||||||||||||||||||||||||92.1%]
  Mem[||||||||||||||||||||    25.1G/61.1G]
  Swp[                            0K/0K]
```



```
largejob-taskmanager-1-6                   0/1     Pending   0          3m55s   <none>       <none>           <none>           <none>
ubuntu@VM-0-7-ubuntu:~/tableapp$ kubectl get pods -o wide
NAME                                       READY   STATUS    RESTARTS   AGE     IP           NODE             NOMINATED NODE   READINESS GATES
flink-kubernetes-operator-8556b997b-mbbxg  2/2     Running   0          11m     10.244.2.4   vm-0-15-ubuntu   <none>           <none>
largejob-54b776cc5-9gshd                   1/1     Running   0          5m6s    10.244.1.6   vm-0-17-ubuntu   <none>           <none>
largejob-taskmanager-1-1                   1/1     Running   0          4m5s    10.244.2.5   vm-0-15-ubuntu   <none>           <none>
largejob-taskmanager-1-2                   1/1     Running   0          4m5s    10.244.1.7   vm-0-17-ubuntu   <none>           <none>
largejob-taskmanager-1-3                   1/1     Running   0          4m5s    10.244.0.4   vm-0-7-ubuntu    <none>           <none>
largejob-taskmanager-1-4                   0/1     Pending   0          3m55s   <none>       <none>           <none>           <none>
largejob-taskmanager-1-5                   0/1     Pending   0          3m55s   <none>       <none>           <none>           <none>
largejob-taskmanager-1-6                   0/1     Pending   0          3m55s   <none>       <none>           <none>           <none>
ubuntu@VM-0-7-ubuntu:~/tableapp$
```

运行截图

上海大学 SHANGHAI UNIVERSITY



Version: 1.16.3 **Commit:** 5df8156 @ 2023-11-10T11:00:33+01:00 **Messag**

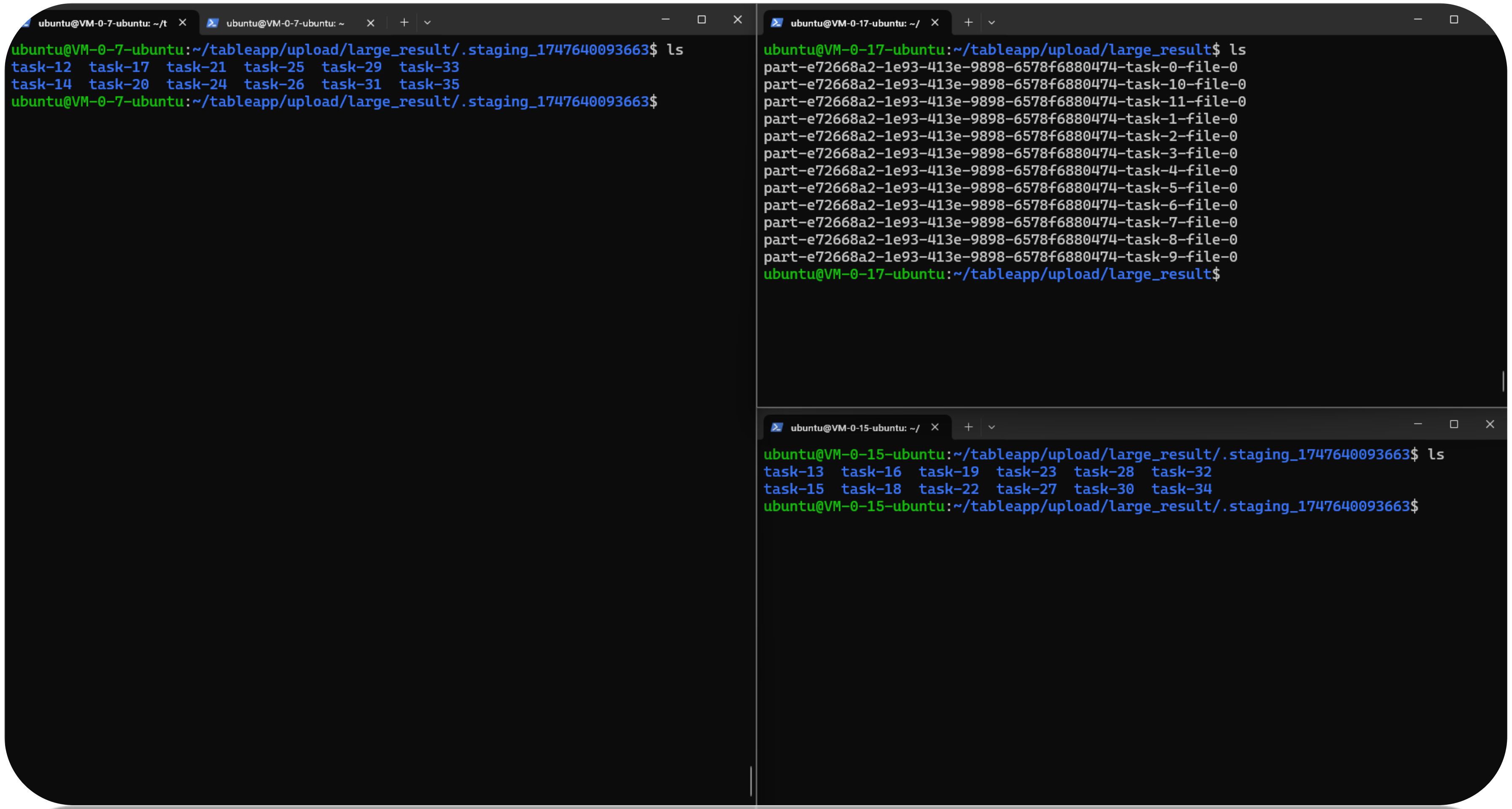
largejob

Job ID	79c6bb41c1ac7db0acaffefe72d65eae	Job State	FINISHED 7	Actions	Job Manager Log
Start Time	2025-05-19 15:34:53	End Time	2025-05-19 15:40:29	Duration	5m 35s

Overview Exceptions TimeLine Configuration

Name	Status	Bytes Received	Records Received	Bytes Sent	Records Sent	Parallelism	Start Time	Duration	End Time	Tasks
Source: large_relation[1] -> HashAggregate[9]	FINISHED	0 B	0	483 MB	6,836,726	36	2025-05-19 15:34:57	6s	2025-05-19 15:35:04	36
HashAggregate[11] -> HashAggregate[12]	FINISHED	215 MB	6,836,726	38.7 MB	1,229,301	36	2025-05-19 15:35:04	1s	2025-05-19 15:35:05	36
HashAggregate[14]	FINISHED	38.7 MB	1,229,301	2.15 MB	68,206	36	2025-05-19 15:35:53	1s	2025-05-19 15:35:55	36
HashJoin[3] -> Calc[4] -> HashAggregate[5]	FINISHED	535 MB	13,673,452	97.2 GB	2,546,776,700	36	2025-05-19 15:35:05	3m 2s	2025-05-19 15:38:08	36
HashAggregate[7]	FINISHED	97.2 GB	2,546,776,700	14.6 GB	383,157,187	36	2025-05-19 15:38:08	1m 7s	2025-05-19 15:39:16	36
HashJoin[15] -> Calc[16]	FINISHED	14.6 GB	383,225,393	17.5 GB	383,157,187	36	2025-05-19 15:39:16	8s	2025-05-19 15:39:24	36
HashJoin[18] -> Calc[19] -> Sink: Filesystem	FINISHED	17.5 GB	383,225,393	0 B	0	36	2025-05-19 15:39:24	1m 5s	2025-05-19 15:40:29	36

结果处理——生成36个子文件

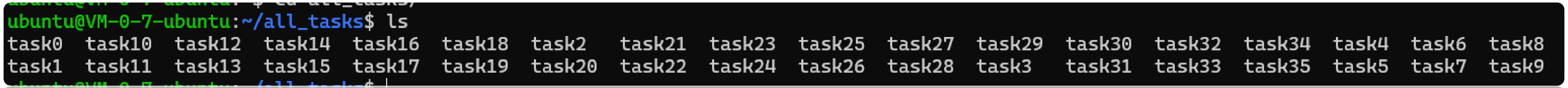


```
ubuntu@VM-0-7-ubuntu:~/tableapp/upload/large_result/.staging_1747640093663$ ls
task-12  task-17  task-21  task-25  task-29  task-33
task-14  task-20  task-24  task-26  task-31  task-35
ubuntu@VM-0-7-ubuntu:~/tableapp/upload/large_result/.staging_1747640093663$
```

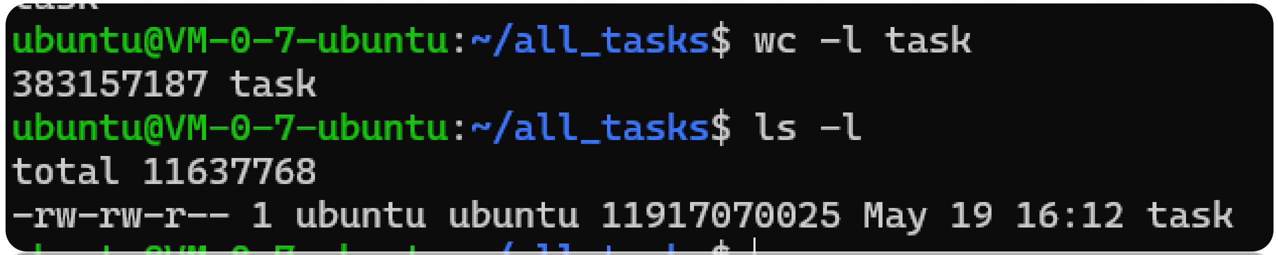
```
ubuntu@VM-0-17-ubuntu:~/tableapp/upload/large_result$ ls
part-e72668a2-1e93-413e-9898-6578f6880474-task-0-file-0
part-e72668a2-1e93-413e-9898-6578f6880474-task-10-file-0
part-e72668a2-1e93-413e-9898-6578f6880474-task-11-file-0
part-e72668a2-1e93-413e-9898-6578f6880474-task-1-file-0
part-e72668a2-1e93-413e-9898-6578f6880474-task-2-file-0
part-e72668a2-1e93-413e-9898-6578f6880474-task-3-file-0
part-e72668a2-1e93-413e-9898-6578f6880474-task-4-file-0
part-e72668a2-1e93-413e-9898-6578f6880474-task-5-file-0
part-e72668a2-1e93-413e-9898-6578f6880474-task-6-file-0
part-e72668a2-1e93-413e-9898-6578f6880474-task-7-file-0
part-e72668a2-1e93-413e-9898-6578f6880474-task-8-file-0
part-e72668a2-1e93-413e-9898-6578f6880474-task-9-file-0
ubuntu@VM-0-17-ubuntu:~/tableapp/upload/large_result$
```

```
ubuntu@VM-0-15-ubuntu:~/tableapp/upload/large_result/.staging_1747640093663$ ls
task-13  task-16  task-19  task-23  task-28  task-32
task-15  task-18  task-22  task-27  task-30  task-34
ubuntu@VM-0-15-ubuntu:~/tableapp/upload/large_result/.staging_1747640093663$
```

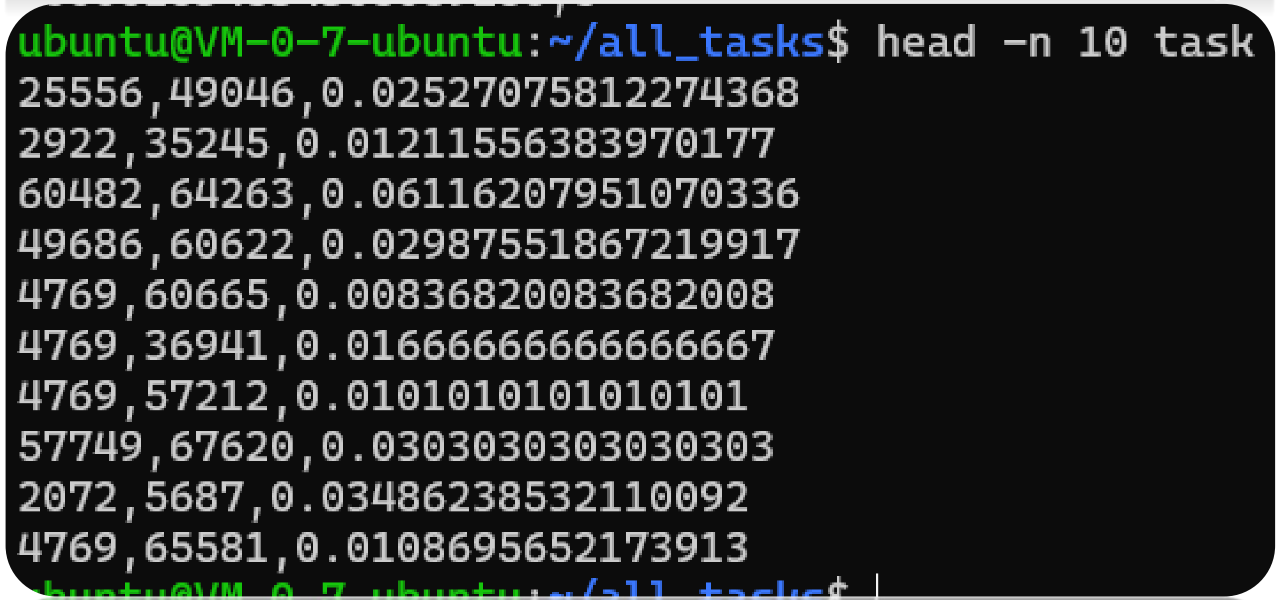
结果处理——显示结果文件信息



```
ubuntu@VM-0-7-ubuntu:~/all_tasks$ ls
task0  task10  task12  task14  task16  task18  task2   task21  task23  task25  task27  task29  task30  task32  task34  task4  task6  task8
task1  task11  task13  task15  task17  task19  task20  task22  task24  task26  task28  task3   task31  task33  task35  task5  task7  task9
```

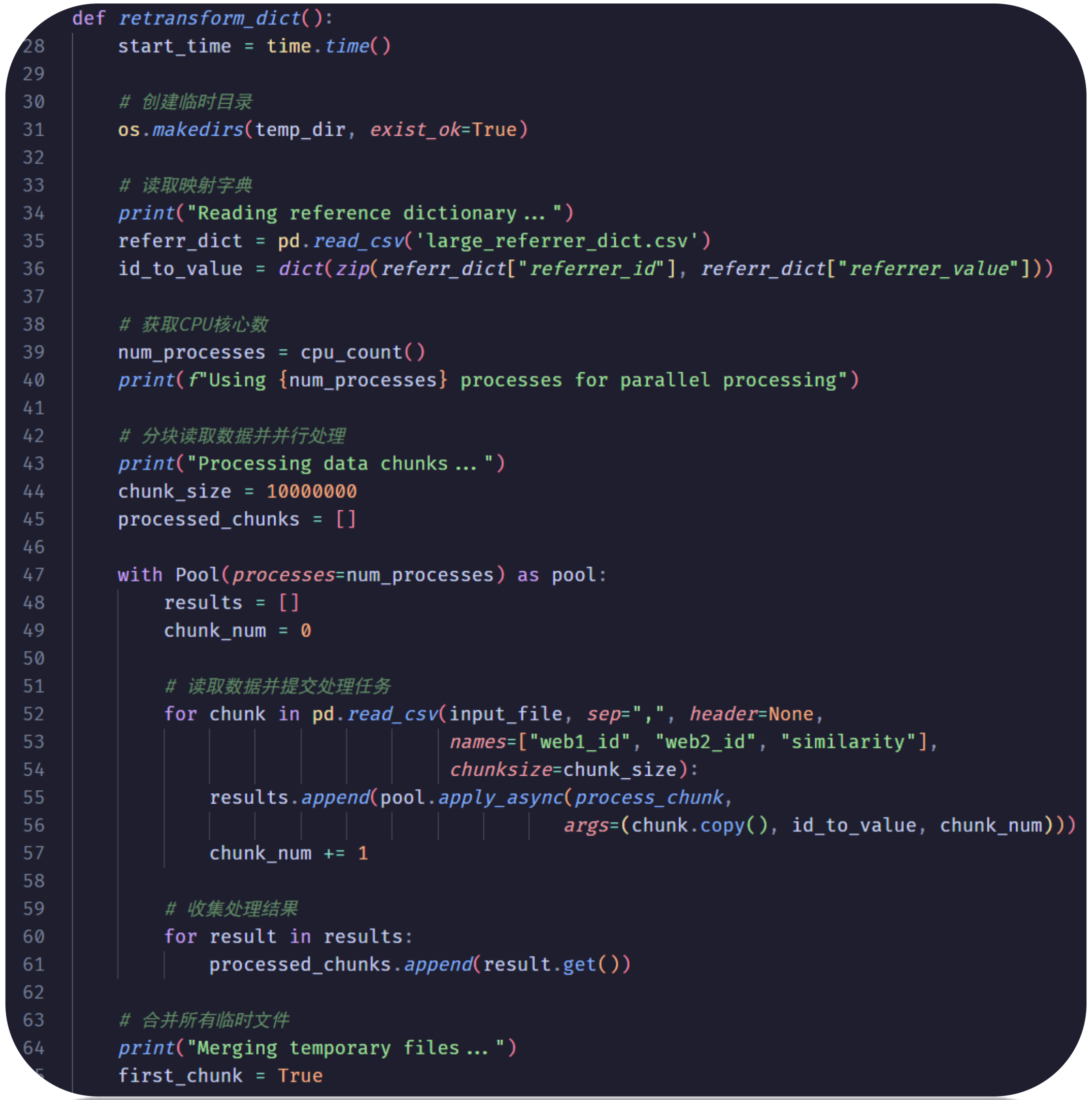


```
ubuntu@VM-0-7-ubuntu:~/all_tasks$ wc -l task
383157187 task
ubuntu@VM-0-7-ubuntu:~/all_tasks$ ls -l
total 11637768
-rw-rw-r-- 1 ubuntu ubuntu 11917070025 May 19 16:12 task
```

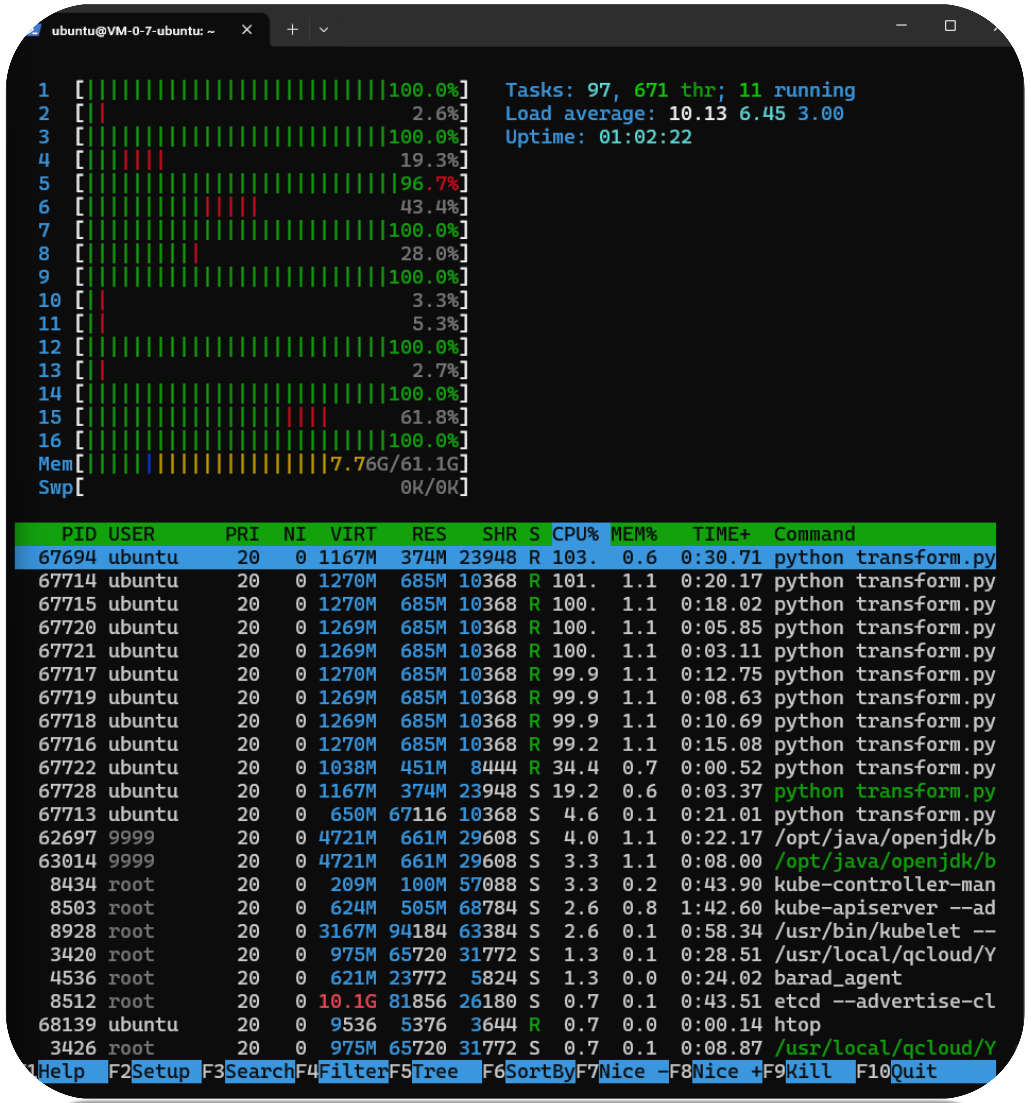


```
ubuntu@VM-0-7-ubuntu:~/all_tasks$ head -n 10 task
25556,49046,0.02527075812274368
2922,35245,0.012115563839701770
60482,64263,0.06116207951070336
49686,60622,0.02987551867219917
4769,60665,0.008368200836820080
4769,36941,0.016666666666666667
4769,57212,0.010101010101010101
57749,67620,0.030303030303030303
2072,5687,0.034862385321100920
4769,65581,0.010869565217391300
```

结果处理——Python脚本进行反映射



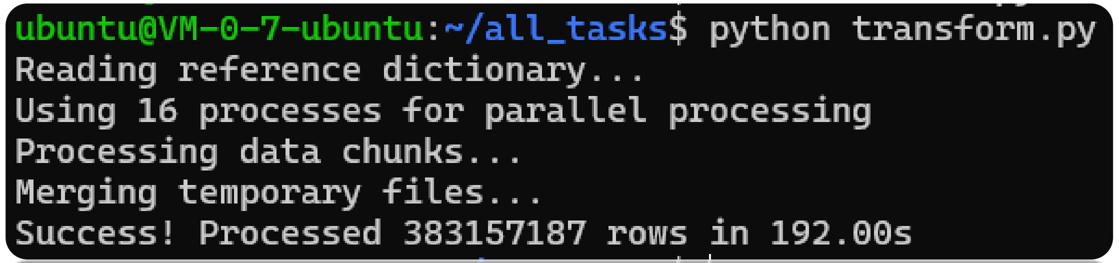
```python
def retransform_dict():
    start_time = time.time()

    # 创建临时目录
    os.makedirs(temp_dir, exist_ok=True)

    # 读取映射字典
    print("Reading reference dictionary...")
    referr_dict = pd.read_csv('large_referrer_dict.csv')
    id_to_value = dict(zip(referr_dict["referrer_id"], referr_dict["referrer_value"]))

    # 获取CPU核心数
    num_processes = cpu_count()
    print(f"Using {num_processes} processes for parallel processing")

    # 分块读取数据并并行处理
    print("Processing data chunks...")
    chunk_size = 10000000
    processed_chunks = []

    with Pool(processes=num_processes) as pool:
        results = []
        chunk_num = 0

        # 读取数据并提交处理任务
        for chunk in pd.read_csv(input_file, sep=",", header=None,
                                  names=["web1_id", "web2_id", "similarity"],
                                  chunksize=chunk_size):
            results.append(pool.apply_async(process_chunk,
                                            args=(chunk.copy(), id_to_value, chunk_num)))
            chunk_num += 1

        # 收集处理结果
        for result in results:
            processed_chunks.append(result.get())

    # 合并所有临时文件
    print("Merging temporary files...")
    first_chunk = True
```



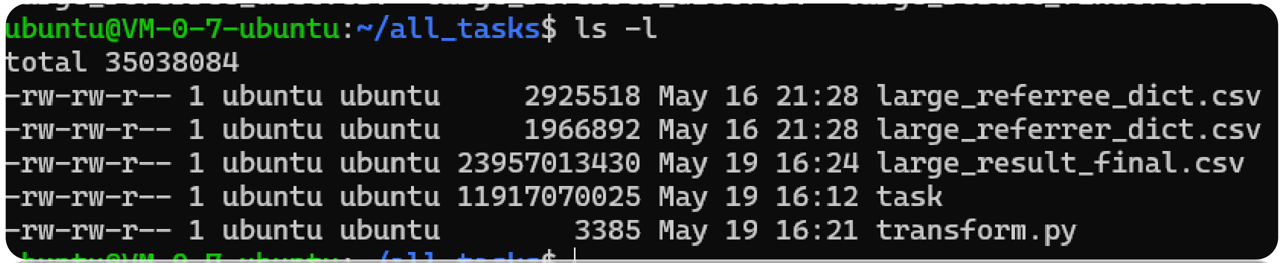
ubuntu@VM-0-7-ubuntu: ~

```
  1 [|||||||||||||||||||||||||100.0%]   Tasks: 97, 671 thr; 11 running
  2 [||                          2.6%]   Load average: 10.13 6.45 3.00
  3 [|||||||||||||||||||||||||100.0%]   Uptime: 01:02:22
  4 [|||||||                    19.3%]
  5 [||||||||||||||||||||||||||96.7%]
  6 [|||||||||||||              43.4%]
  7 [|||||||||||||||||||||||||100.0%]
  8 [|||||||||                  28.0%]
  9 [|||||||||||||||||||||||||100.0%]
 10 [||                          3.3%]
 11 [||                          5.3%]
 12 [|||||||||||||||||||||||||100.0%]
 13 [||                          2.7%]
 14 [|||||||||||||||||||||||||100.0%]
 15 [|||||||||||||||||||||      61.8%]
 16 [|||||||||||||||||||||||||100.0%]
 Mem[|||||||||||||||||||||||7.76G/61.1G]
 Swp[                          0K/0K]

  PID USER      PRI  NI  VIRT   RES   SHR S CPU% MEM%   TIME+  Command
67694 ubuntu     20   0 1167M  374M 23948 R 103.  0.6  0:30.71 python transform.py
67714 ubuntu     20   0 1270M  685M 10368 R 101.  1.1  0:20.17 python transform.py
67715 ubuntu     20   0 1270M  685M 10368 R 100.  1.1  0:18.02 python transform.py
67720 ubuntu     20   0 1269M  685M 10368 R 100.  1.1  0:05.85 python transform.py
67721 ubuntu     20   0 1269M  685M 10368 R 100.  1.1  0:03.11 python transform.py
67717 ubuntu     20   0 1270M  685M 10368 R 99.9  1.1  0:12.75 python transform.py
67719 ubuntu     20   0 1269M  685M 10368 R 99.9  1.1  0:08.63 python transform.py
67718 ubuntu     20   0 1269M  685M 10368 R 99.9  1.1  0:10.69 python transform.py
67716 ubuntu     20   0 1270M  685M 10368 R 99.2  1.1  0:15.08 python transform.py
67722 ubuntu     20   0 1038M  451M  8444 R 34.4  0.7  0:00.52 python transform.py
67728 ubuntu     20   0 1167M  374M 23948 S 19.2  0.6  0:03.37 python transform.py
67713 ubuntu     20   0  650M 67116 10368 S  4.6  0.1  0:21.01 python transform.py
62697 9999       20   0 4721M  661M 29608 S  4.0  1.1  0:22.17 /opt/java/openjdk/b
63014 9999       20   0 4721M  661M 29608 S  3.3  1.1  0:08.00 /opt/java/openjdk/b
 8434 root       20   0  209M  100M 57088 S  3.3  0.2  0:43.90 kube-controller-man
 8503 root       20   0  624M  505M 68784 S  2.6  0.8  1:42.60 kube-apiserver --ad
 8928 root       20   0 3167M 94184 63384 S  2.6  0.1  0:58.34 /usr/bin/kubelet --
 3420 root       20   0  975M 65720 31772 S  1.3  0.1  0:28.51 /usr/local/qcloud/Y
 4536 root       20   0  621M 23772  5824 S  1.3  0.0  0:24.02 barad_agent
 8512 root       20   0 10.1G 81856 26180 S  0.7  0.1  0:43.51 etcd --advertise-cl
68139 ubuntu     20   0  9536  5376  3644 R  0.7  0.0  0:00.14 htop
 3426 root       20   0  975M 65720 31772 S  0.7  0.1  0:08.87 /usr/local/qcloud/Y
1Help  F2Setup F3SearchF4FilterF5Tree  F6SortByF7Nice -F8Nice +F9Kill  F10Quit
```

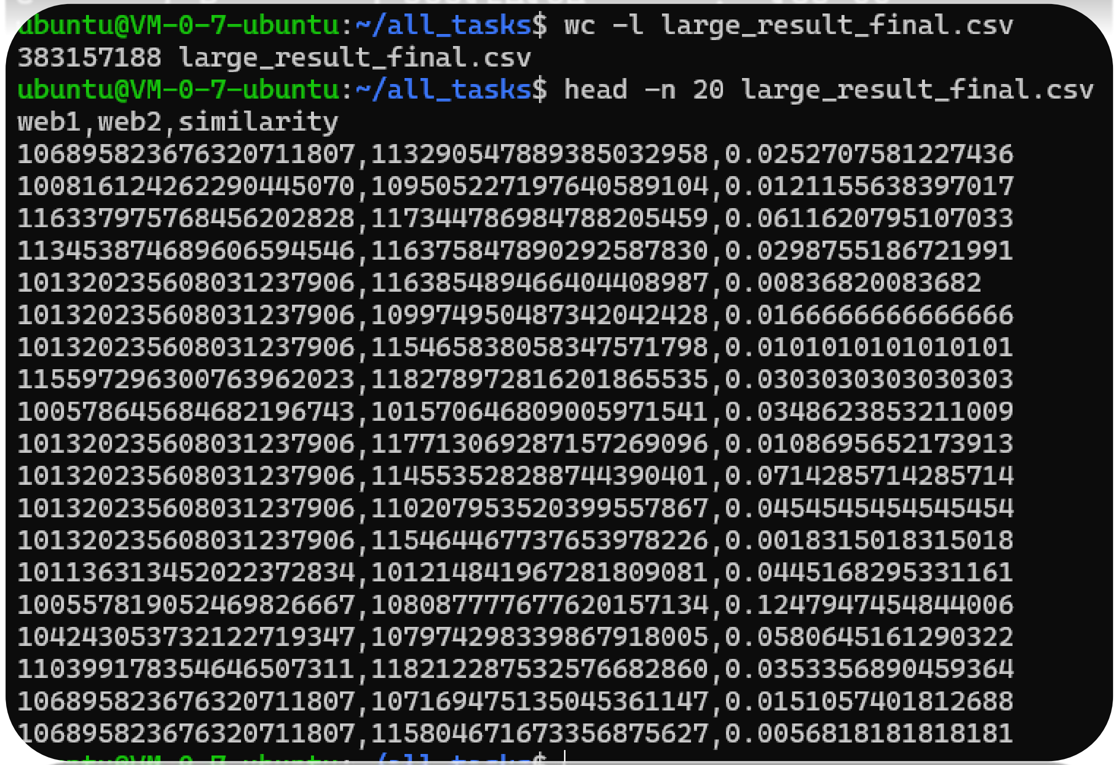
结果处理——显示最终结果



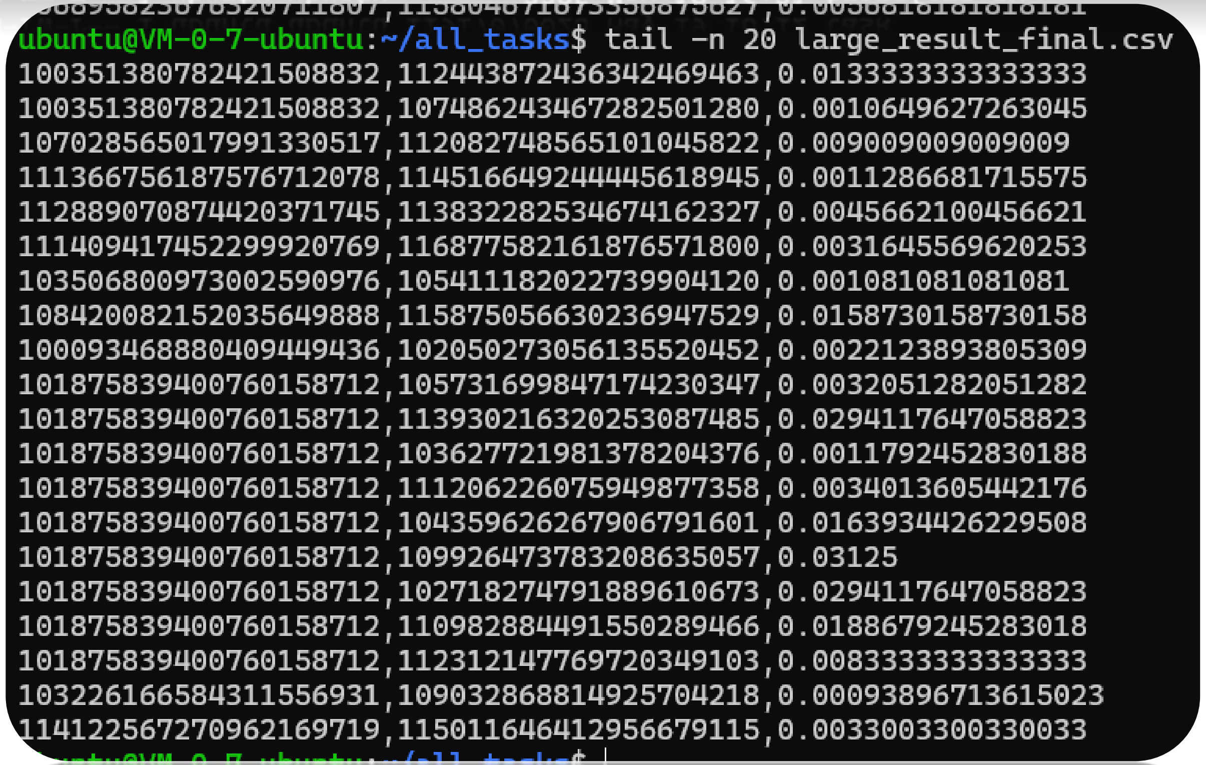
```
ubuntu@VM-0-7-ubuntu:~/all_tasks$ python transform.py
Reading reference dictionary...
Using 16 processes for parallel processing
Processing data chunks...
Merging temporary files...
Success! Processed 383157187 rows in 192.00s
```



```
ubuntu@VM-0-7-ubuntu:~/all_tasks$ ls -l
total 35038084
-rw-rw-r-- 1 ubuntu ubuntu     2925518 May 16 21:28 large_referee_dict.csv
-rw-rw-r-- 1 ubuntu ubuntu     1966892 May 16 21:28 large_referrer_dict.csv
-rw-rw-r-- 1 ubuntu ubuntu 23957013430 May 19 16:24 large_result_final.csv
-rw-rw-r-- 1 ubuntu ubuntu 11917070025 May 19 16:12 task
-rw-rw-r-- 1 ubuntu ubuntu        3385 May 19 16:21 transform.py
```



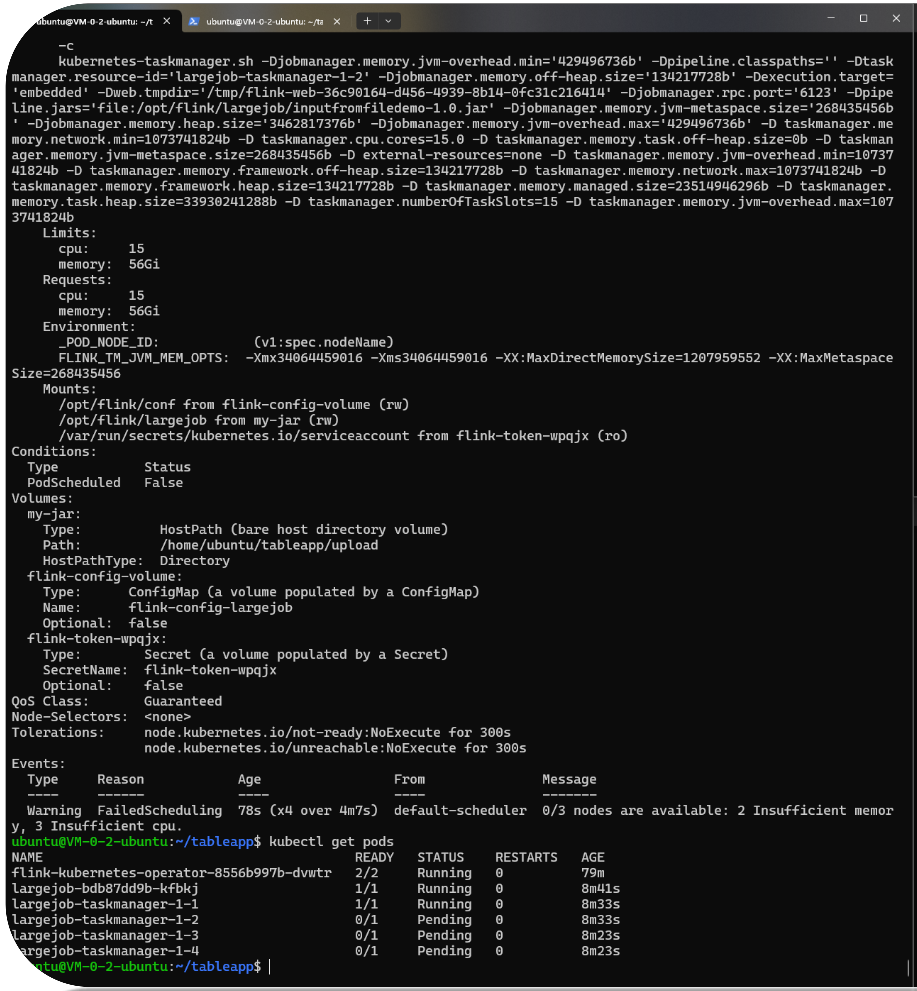
```
ubuntu@VM-0-7-ubuntu:~/all_tasks$ wc -l large_result_final.csv
383157188 large_result_final.csv
ubuntu@VM-0-7-ubuntu:~/all_tasks$ head -n 20 large_result_final.csv
web1,web2,similarity
106895823676320711807,113290547889385032958,0.0252707581227436
100816124262290445070,109505227197640589104,0.0121155638397017
116337975768456202828,117344786984788205459,0.0611620795107033
113453874689606594546,116375847890292587830,0.0298755186721991
101320235608031237906,116385489466404408987,0.00836820083682
101320235608031237906,109974950487342042428,0.0166666666666666
101320235608031237906,115465838058347571798,0.0101010101010101
115597296300763962023,118278972816201865535,0.0303030303030303
100578645684682196743,101570646809005971541,0.0348623853211009
101320235608031237906,117713069287157269096,0.0108695652173913
101320235608031237906,114553528288744390401,0.0714285714285714
101320235608031237906,110207953520399557867,0.0454545454545454
101320235608031237906,115464467737653978226,0.0018315018315018
101136313452022372834,101214841967281809081,0.0445168295331161
100557819052469826667,10808777677620157134,0.1247947454844006
104243053732122719347,107974298339867918005,0.0580645161290322
110399178354646507311,118212287532576682860,0.0353356890459364
106895823676320711807,107169475135045361147,0.0151057401812688
106895823676320711807,115804671673356875627,0.0056818181818181
```



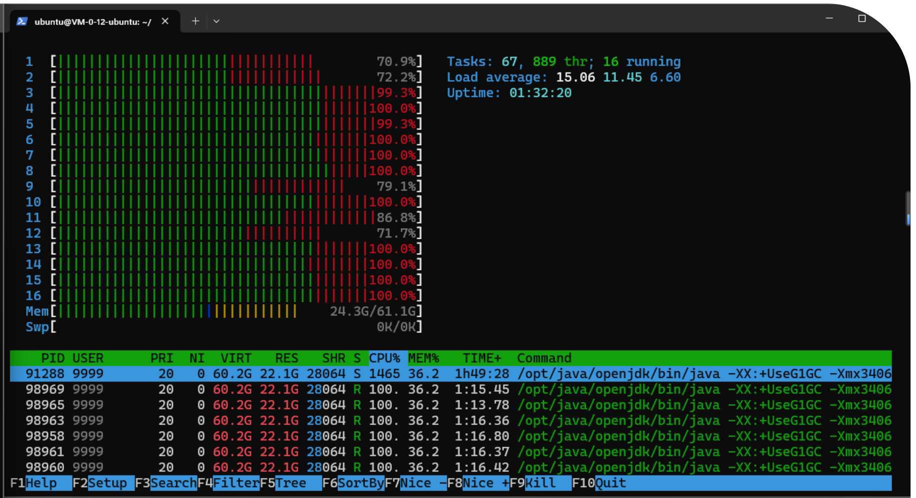
```
ubuntu@VM-0-7-ubuntu:~/all_tasks$ tail -n 20 large_result_final.csv
100351380782421508832,112443872436342469463,0.0133333333333333
100351380782421508832,107486243467282501280,0.0010649627263045
107028565017991330517,112082748565101045822,0.009009009009009
111366756187576712078,114516649244445618945,0.0011286681715575
112889070874420371745,113832282534674162327,0.0045662100456621
111409417452299920769,116877582161876571800,0.0031645569620253
103506800973002590976,105411182022739904120,0.001081081081081
108420082152035649888,115875056630236947529,0.0158730158730158
100093468880409449436,102050273056135520452,0.0022123893805309
101875839400760158712,105731699847174230347,0.0032051282051282
101875839400760158712,113930216320253087485,0.0294117647058823
101875839400760158712,103627721981378204376,0.0011792452830188
101875839400760158712,111206226075949877358,0.0034013605442176
101875839400760158712,104359626267906791601,0.0163934426229508
101875839400760158712,109926473783208635057,0.03125
101875839400760158712,102718274791889610673,0.0294117647058823
101875839400760158712,110982884491550289466,0.0188679245283018
101875839400760158712,112312147769720349103,0.0083333333333333
103226166584311556931,109032868814925704218,0.00093896713615023
114122567270962169719,115011646412956679115,0.0033003300330033
```

3.问题处理

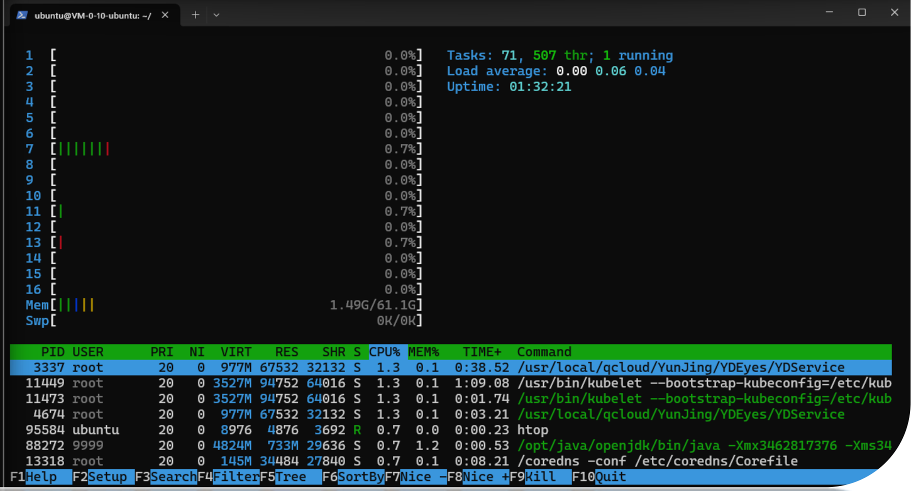
问题处理



```
ubuntu@VM-0-2-ubuntu: ~/t    ubuntu@VM-0-2-ubuntu: ~/ta
    -c
    kubernetes-taskmanager.sh -Djobmanager.memory.jvm-overhead.min='429496736b' -Dpipeline.classpaths='' -Dtask
manager.resource-id='largejob-taskmanager-1-2' -Djobmanager.memory.off-heap.size='134217728b' -Dexecution.target=
'embedded' -Dweb.tmpdir='/tmp/flink-web-36c90164-d456-4939-8b14-0fc31c216414' -Djobmanager.rpc.port='6123' -Dpipe
line.jars='file:/opt/flink/largejob/inputfromfiledemo-1.0.jar' -Djobmanager.memory.jvm-metaspace.size='268435456b
' -Djobmanager.memory.heap.size='3462817376b' -Djobmanager.memory.jvm-overhead.max='429496736b' -D taskmanager.me
mory.network.min=1073741824b -D taskmanager.cpu.cores=15.0 -D taskmanager.memory.task.off-heap.size=0b -D taskman
ager.memory.jvm-metaspace.size=268435456b -D external-resources=none -D taskmanager.memory.jvm-overhead.min=10737
41824b -D taskmanager.memory.framework.off-heap.size=134217728b -D taskmanager.memory.network.max=1073741824b -D
taskmanager.memory.framework.heap.size=134217728b -D taskmanager.memory.managed.size=23514946296b -D taskmanager.
memory.task.heap.size=33930241288b -D taskmanager.numberOfTaskSlots=15 -D taskmanager.memory.jvm-overhead.max=107
3741824b
   Limits:
     cpu:     15
     memory:  56Gi
   Requests:
     cpu:     15
     memory:  56Gi
   Environment:
     _POD_NODE_ID:           (v1:spec.nodeName)
     FLINK_TM_JVM_MEM_OPTS:  -Xmx34064459016 -Xms34064459016 -XX:MaxDirectMemorySize=1207959552 -XX:MaxMetaspace
Size=268435456
   Mounts:
     /opt/flink/conf from flink-config-volume (rw)
     /opt/flink/largejob from my-jar (rw)
     /var/run/secrets/kubernetes.io/serviceaccount from flink-token-wpqjx (ro)
Conditions:
  Type           Status
  PodScheduled   False
Volumes:
  my-jar:
   Type:          HostPath (bare host directory volume)
   Path:          /home/ubuntu/tableapp/upload
   HostPathType:  Directory
  flink-config-volume:
   Type:      ConfigMap (a volume populated by a ConfigMap)
   Name:      flink-config-largejob
   Optional:  false
  flink-token-wpqjx:
   Type:        Secret (a volume populated by a Secret)
   SecretName:  flink-token-wpqjx
   Optional:    false
QoS Class:      Guaranteed
Node-Selectors:  <none>
Tolerations:    node.kubernetes.io/not-ready:NoExecute for 300s
                node.kubernetes.io/unreachable:NoExecute for 300s
Events:
  Type     Reason            Age                From               Message
  ----     ------            ----               ----               -------
  Warning  FailedScheduling  78s (x4 over 4m7s)  default-scheduler  0/3 nodes are available: 2 Insufficient memor
y, 3 Insufficient cpu.
ubuntu@VM-0-2-ubuntu:~/tableapp$ kubectl get pods
NAME                                         READY   STATUS    RESTARTS   AGE
flink-kubernetes-operator-8556b997b-dvwtr   2/2     Running   0          79m
largejob-bdb87dd9b-kfbkj                    1/1     Running   0          8m41s
largejob-taskmanager-1-1                    1/1     Running   0          8m33s
largejob-taskmanager-1-2                    0/1     Pending   0          8m33s
largejob-taskmanager-1-3                    0/1     Pending   0          8m23s
argejob-taskmanager-1-4                     0/1     Pending   0          8m23s
ntu@VM-0-2-ubuntu:~/tableapp$ |
```



```
ubuntu@VM-0-12-ubuntu: ~/

  1  [||||||||||||||||||||||||||||||||            70.9%]   Tasks: 67, 889 thr; 16 running
  2  [|||||||||||||||||||||||||||||||||           72.2%]   Load average: 15.06 11.45 6.60
  3  [||||||||||||||||||||||||||||||||||||||||||| 99.3%]   Uptime: 01:32:20
  4  [||||||||||||||||||||||||||||||||||||||||||100.0%]
  5  [||||||||||||||||||||||||||||||||||||||||||| 99.3%]
  6  [||||||||||||||||||||||||||||||||||||||||||100.0%]
  7  [||||||||||||||||||||||||||||||||||||||||||100.0%]
  8  [||||||||||||||||||||||||||||||||||||||||||100.0%]
  9  [|||||||||||||||||||||||||||||||||||         79.1%]
  10 [||||||||||||||||||||||||||||||||||||||||||100.0%]
  11 [||||||||||||||||||||||||||||||||||||||||||| 86.8%]
  12 [||||||||||||||||||||||||||||||||           71.7%]
  13 [||||||||||||||||||||||||||||||||||||||||||100.0%]
  14 [||||||||||||||||||||||||||||||||||||||||||100.0%]
  15 [||||||||||||||||||||||||||||||||||||||||||100.0%]
  16 [||||||||||||||||||||||||||||||||||||||||||100.0%]
  Mem[|||||||||||||||||||||||||||||||      24.3G/61.1G]
  Swp[                                           0K/0K]

   PID USER      PRI  NI  VIRT   RES   SHR S CPU% MEM%   TIME+  Command
 91288 9999       20   0 60.2G 22.1G 28064 S 1465 36.2  1h49:28 /opt/java/openjdk/bin/java -XX:+UseG1GC -Xmx3406
 98969 9999       20   0 60.2G 22.1G 28064 R 100. 36.2  1:15.45 /opt/java/openjdk/bin/java -XX:+UseG1GC -Xmx3406
 98965 9999       20   0 60.2G 22.1G 28064 R 100. 36.2  1:13.78 /opt/java/openjdk/bin/java -XX:+UseG1GC -Xmx3406
 98963 9999       20   0 60.2G 22.1G 28064 R 100. 36.2  1:16.36 /opt/java/openjdk/bin/java -XX:+UseG1GC -Xmx3406
 98958 9999       20   0 60.2G 22.1G 28064 R 100. 36.2  1:16.80 /opt/java/openjdk/bin/java -XX:+UseG1GC -Xmx3406
 98961 9999       20   0 60.2G 22.1G 28064 R 100. 36.2  1:16.37 /opt/java/openjdk/bin/java -XX:+UseG1GC -Xmx3406
 98960 9999       20   0 60.2G 22.1G 28064 R 100. 36.2  1:16.42 /opt/java/openjdk/bin/java -XX:+UseG1GC -Xmx3406
F1Help  F2Setup F3SearchF4FilterF5Tree  F6SortByF7Nice -F8Nice +F9Kill  F10Quit
```



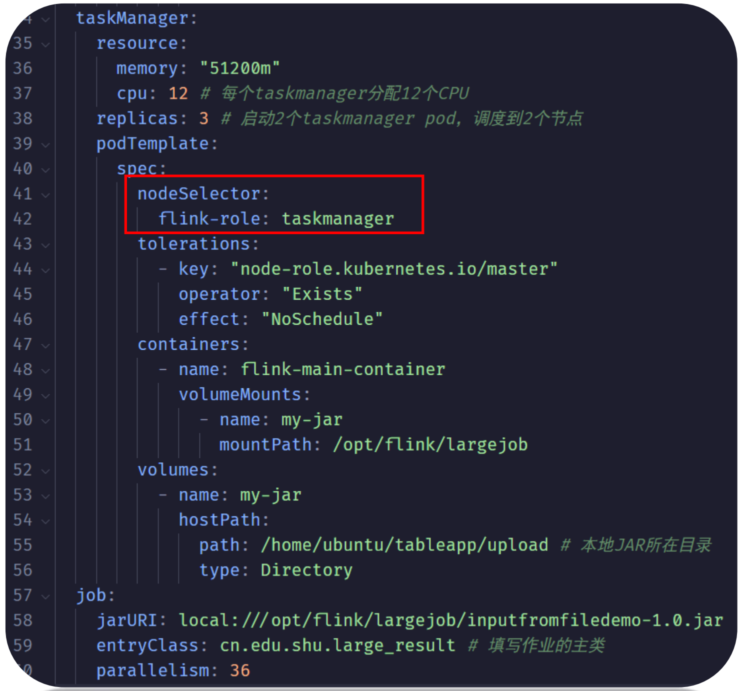
```
ubuntu@VM-0-10-ubuntu: ~/

  1  [                                             0.0%]   Tasks: 71, 507 thr; 1 running
  2  [                                             0.0%]   Load average: 0.00 0.06 0.04
  3  [                                             0.0%]   Uptime: 01:32:21
  4  [                                             0.0%]
  5  [                                             0.0%]
  6  [                                             0.0%]
  7  [|||||||                                      0.7%]
  8  [                                             0.0%]
  9  [                                             0.0%]
  10 [                                             0.0%]
  11 [|                                            0.7%]
  12 [                                             0.0%]
  13 [|                                            0.7%]
  14 [                                             0.0%]
  15 [                                             0.0%]
  16 [                                             0.0%]
  Mem[|||||                                1.49G/61.1G]
  Swp[                                           0K/0K]

   PID USER      PRI  NI  VIRT   RES   SHR S CPU% MEM%   TIME+  Command
  3337 root       20   0  977M 67532 32132 S  1.3  0.1  0:38.52 /usr/local/qcloud/YunJing/YDEyes/YDService
 11449 root       20   0 3527M 94752 64016 S  1.3  0.1  1:09.08 /usr/bin/kubelet --bootstrap-kubeconfig=/etc/kub
 11473 root       20   0 3527M 94752 64016 S  1.3  0.1  0:01.74 /usr/bin/kubelet --bootstrap-kubeconfig=/etc/kub
  4674 root       20   0  977M 67532 32132 S  1.3  0.1  0:03.21 /usr/local/qcloud/YunJing/YDEyes/YDService
 95584 ubuntu     20   0  8976  4876  3692 R  0.7  0.0  0:00.23 htop
 88272 9999       20   0 4824M  733M 29636 S  0.7  1.2  0:00.53 /opt/java/openjdk/bin/java -Xmx3462817376 -Xms34
 13318 root       20   0  145M 34484 27840 S  0.7  0.1  0:08.21 /coredns -conf /etc/coredns/Corefile
F1Help  F2Setup F3SearchF4FilterF5Tree  F6SortByF7Nice -F8Nice +F9Kill  F10Quit
```

跑flink作业时发现空闲节点
TaskManager节点只有一个

原因：
k8s对于JobManager和TaskManager随机分配
该节点被JobManager调度占用，不进行计算

Master节点被自动打上污点，TM不会分配到该节点；或master节点资源不足，TM的pod一直处于pending状态

解决办法

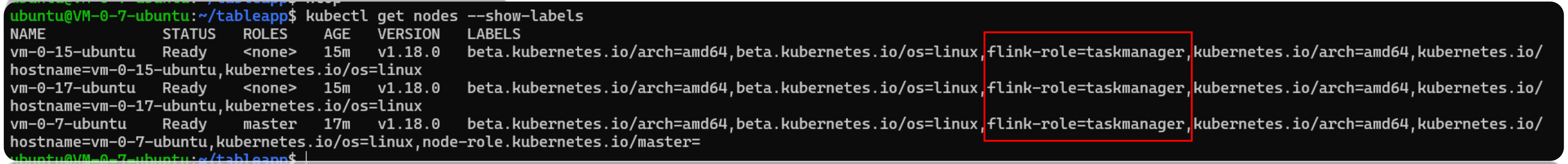


```
taskManager:
  resource:
    memory: "51200m"
    cpu: 12 # 每个taskmanager分配12个CPU
  replicas: 3 # 启动2个taskmanager pod, 调度到2个节点
  podTemplate:
    spec:
      nodeSelector:
        flink-role: taskmanager
      tolerations:
        - key: "node-role.kubernetes.io/master"
          operator: "Exists"
          effect: "NoSchedule"
      containers:
        - name: flink-main-container
          volumeMounts:
            - name: my-jar
              mountPath: /opt/flink/largejob
      volumes:
        - name: my-jar
          hostPath:
            path: /home/ubuntu/tableapp/upload # 本地JAR所在目录
            type: Directory
job:
  jarURI: local:///opt/flink/largejob/inputfromfiledemo-1.0.jar
  entryClass: cn.edu.shu.large_result # 填写作业的主类
  parallelism: 36
```

移除master节点污点限制
kubectl taint node vm-0-9-ubuntu node-role.kubernetes.io/master:NoSchedule-

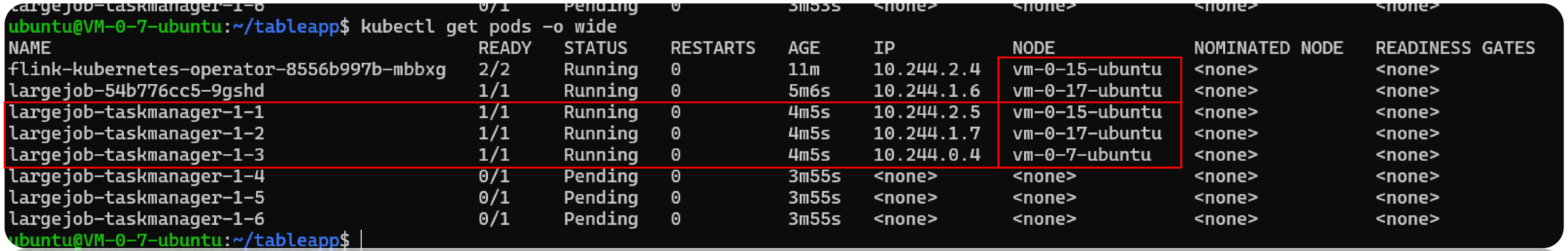
若想强制选择pod运行的节点
可以在yaml文件中添加节点选择标签，
只有当节点标签为taskmanager时才分配为TM

为节点打上标签
kubectl label nodes vm-0-14-ubuntu flink-role=taskmanager
显示节点标签
kubectl get nodes --show-labels

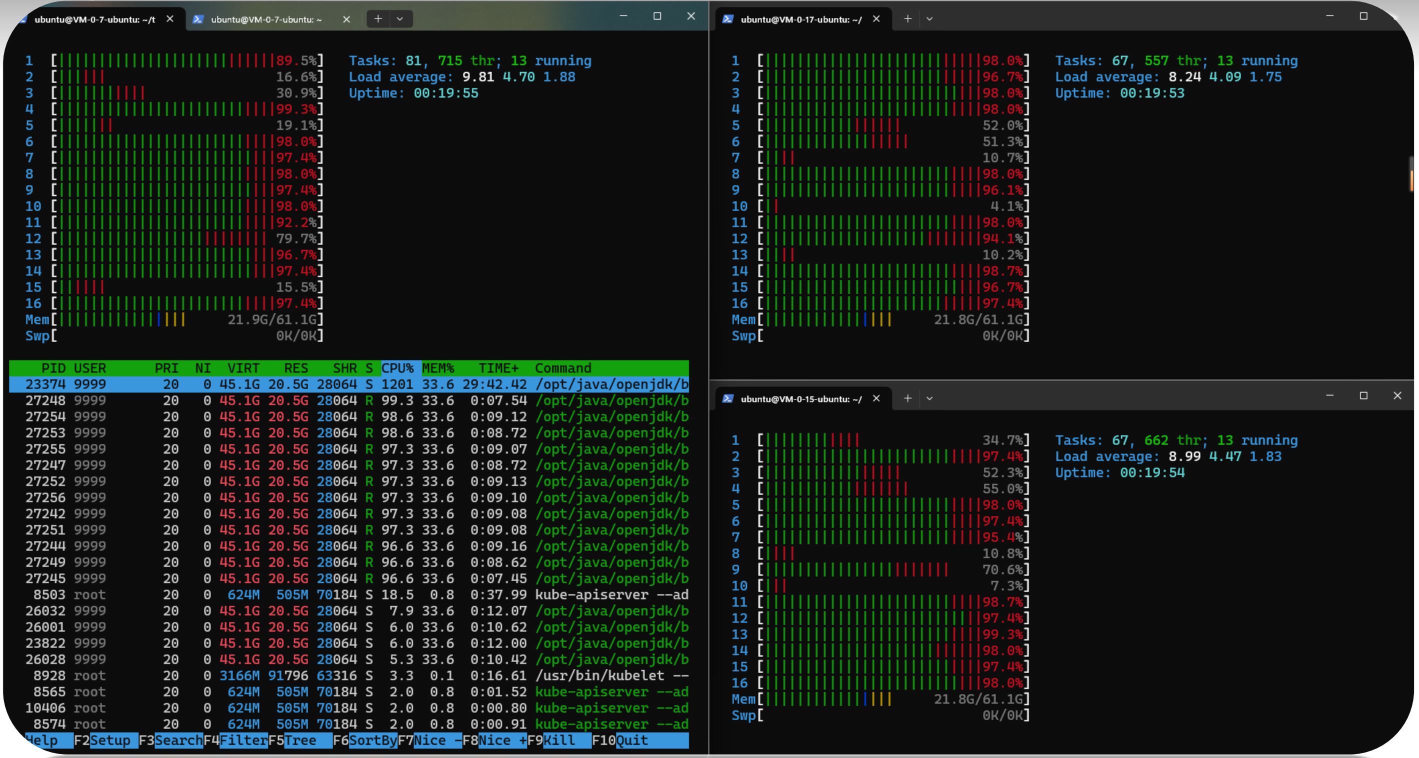


```
ubuntu@VM-0-7-ubuntu:~/tableapp$ kubectl get nodes --show-labels
NAME             STATUS   ROLES    AGE   VERSION   LABELS
vm-0-15-ubuntu   Ready    <none>   15m   v1.18.0   beta.kubernetes.io/arch=amd64,beta.kubernetes.io/os=linux,flink-role=taskmanager,kubernetes.io/arch=amd64,kubernetes.io/
hostname=vm-0-15-ubuntu,kubernetes.io/os=linux
vm-0-17-ubuntu   Ready    <none>   15m   v1.18.0   beta.kubernetes.io/arch=amd64,beta.kubernetes.io/os=linux,flink-role=taskmanager,kubernetes.io/arch=amd64,kubernetes.io/
hostname=vm-0-17-ubuntu,kubernetes.io/os=linux
vm-0-7-ubuntu    Ready    master   17m   v1.18.0   beta.kubernetes.io/arch=amd64,beta.kubernetes.io/os=linux,flink-role=taskmanager,kubernetes.io/arch=amd64,kubernetes.io/
hostname=vm-0-7-ubuntu,kubernetes.io/os=linux,node-role.kubernetes.io/master=
ubuntu@VM-0-7-ubuntu:~/tableapp$
```

让TaskManager跑在3个节点上



```
largejob-taskmanager-1-6                  0/1     Pending     0          3m53s   <none>       <none>           <none>           <none>
ubuntu@VM-0-7-ubuntu:~/tableapp$ kubectl get pods -o wide
NAME                                      READY   STATUS      RESTARTS   AGE     IP           NODE             NOMINATED NODE   READINESS GATES
flink-kubernetes-operator-8556b997b-mbbxg 2/2     Running     0          11m     10.244.2.4   vm-0-15-ubuntu   <none>           <none>
largejob-54b776cc5-9gshd                  1/1     Running     0          5m6s    10.244.1.6   vm-0-17-ubuntu   <none>           <none>
largejob-taskmanager-1-1                  1/1     Running     0          4m5s    10.244.2.5   vm-0-15-ubuntu   <none>           <none>
largejob-taskmanager-1-2                  1/1     Running     0          4m5s    10.244.1.7   vm-0-17-ubuntu   <none>           <none>
largejob-taskmanager-1-3                  1/1     Running     0          4m5s    10.244.0.4   vm-0-7-ubuntu    <none>           <none>
largejob-taskmanager-1-4                  0/1     Pending     0          3m55s   <none>       <none>           <none>           <none>
largejob-taskmanager-1-5                  0/1     Pending     0          3m55s   <none>       <none>           <none>           <none>
largejob-taskmanager-1-6                  0/1     Pending     0          3m55s   <none>       <none>           <none>           <none>
ubuntu@VM-0-7-ubuntu:~/tableapp$ |
```



ubuntu@VM-0-7-ubuntu: ~/t
ubuntu@VM-0-7-ubuntu: ~

```
  1  [|||||||||||||||||||||||||89.5%]   Tasks: 81, 715 thr; 13 running
  2  [|||||                    16.6%]   Load average: 9.81 4.70 1.88
  3  [||||||||||               30.9%]   Uptime: 00:19:55
  4  [|||||||||||||||||||||||||99.3%]
  5  [||||||                   19.1%]
  6  [|||||||||||||||||||||||||98.0%]
  7  [|||||||||||||||||||||||||97.4%]
  8  [|||||||||||||||||||||||||98.0%]
  9  [|||||||||||||||||||||||||97.4%]
  10 [|||||||||||||||||||||||||98.0%]
  11 [|||||||||||||||||||||||||92.2%]
  12 [||||||||||||||||||||||||| 79.7%]
  13 [|||||||||||||||||||||||||96.7%]
  14 [|||||||||||||||||||||||||97.4%]
  15 [|||||                    15.5%]
  16 [|||||||||||||||||||||||||97.4%]
  Mem[||||||||||||||||     21.9G/61.1G]
  Swp[                          0K/0K]

    PID USER      PRI  NI  VIRT   RES   SHR S CPU% MEM%   TIME+  Command
  23374 9999       20   0 45.1G 20.5G 28064 S 1201 33.6 29:42.42 /opt/java/openjdk/b
  27248 9999       20   0 45.1G 20.5G 28064 R 99.3 33.6  0:07.54 /opt/java/openjdk/b
  27254 9999       20   0 45.1G 20.5G 28064 R 98.6 33.6  0:09.12 /opt/java/openjdk/b
  27253 9999       20   0 45.1G 20.5G 28064 R 98.6 33.6  0:08.72 /opt/java/openjdk/b
  27255 9999       20   0 45.1G 20.5G 28064 R 97.3 33.6  0:09.07 /opt/java/openjdk/b
  27247 9999       20   0 45.1G 20.5G 28064 R 97.3 33.6  0:08.72 /opt/java/openjdk/b
  27252 9999       20   0 45.1G 20.5G 28064 R 97.3 33.6  0:09.13 /opt/java/openjdk/b
  27256 9999       20   0 45.1G 20.5G 28064 R 97.3 33.6  0:09.10 /opt/java/openjdk/b
  27242 9999       20   0 45.1G 20.5G 28064 R 97.3 33.6  0:09.08 /opt/java/openjdk/b
  27251 9999       20   0 45.1G 20.5G 28064 R 97.3 33.6  0:09.08 /opt/java/openjdk/b
  27244 9999       20   0 45.1G 20.5G 28064 R 96.6 33.6  0:09.16 /opt/java/openjdk/b
  27249 9999       20   0 45.1G 20.5G 28064 R 96.6 33.6  0:08.62 /opt/java/openjdk/b
  27245 9999       20   0 45.1G 20.5G 28064 R 96.6 33.6  0:07.45 /opt/java/openjdk/b
   8503 root       20   0  624M  505M 70184 S 18.5  0.8  0:37.99 kube-apiserver --ad
  26032 9999       20   0 45.1G 20.5G 28064 S  7.9 33.6  0:12.07 /opt/java/openjdk/b
  26001 9999       20   0 45.1G 20.5G 28064 S  6.0 33.6  0:10.62 /opt/java/openjdk/b
  23822 9999       20   0 45.1G 20.5G 28064 S  6.0 33.6  0:12.00 /opt/java/openjdk/b
  26028 9999       20   0 45.1G 20.5G 28064 S  5.3 33.6  0:10.42 /opt/java/openjdk/b
   8928 root       20   0 3166M 91796 63316 S  3.3  0.1  0:16.61 /usr/bin/kubelet --
   8565 root       20   0  624M  505M 70184 S  2.0  0.8  0:01.52 kube-apiserver --ad
  10406 root       20   0  624M  505M 70184 S  2.0  0.8  0:00.80 kube-apiserver --ad
   8574 root       20   0  624M  505M 70184 S  2.0  0.8  0:00.91 kube-apiserver --ad
Help  F2Setup F3SearchF4FilterF5Tree  F6SortByF7Nice -F8Nice +F9Kill  F10Quit
```

ubuntu@VM-0-17-ubuntu: ~/

```
  1  [|||||||||||||||||||||||||98.0%]   Tasks: 67, 557 thr; 13 running
  2  [|||||||||||||||||||||||||96.7%]   Load average: 8.24 4.09 1.75
  3  [|||||||||||||||||||||||||98.0%]   Uptime: 00:19:53
  4  [|||||||||||||||||||||||||98.0%]
  5  [||||||||||||||||         52.0%]
  6  [|||||||||||||||||        51.3%]
  7  [||||                     10.7%]
  8  [|||||||||||||||||||||||||98.0%]
  9  [|||||||||||||||||||||||||96.1%]
  10 [||                        4.1%]
  11 [|||||||||||||||||||||||||98.0%]
  12 [|||||||||||||||||||||||||94.1%]
  13 [||||                     10.2%]
  14 [|||||||||||||||||||||||||98.7%]
  15 [|||||||||||||||||||||||||96.7%]
  16 [|||||||||||||||||||||||||97.4%]
  Mem[||||||||||||||||     21.8G/61.1G]
  Swp[                          0K/0K]
```

ubuntu@VM-0-15-ubuntu: ~/

```
  1  [|||||||||||||            34.7%]   Tasks: 67, 662 thr; 13 running
  2  [|||||||||||||||||||||||||97.4%]   Load average: 8.99 4.47 1.83
  3  [||||||||||||||||         52.3%]   Uptime: 00:19:54
  4  [|||||||||||||||||        55.0%]
  5  [|||||||||||||||||||||||||98.0%]
  6  [|||||||||||||||||||||||||97.4%]
  7  [|||||||||||||||||||||||||95.4%]
  8  [||||                     10.8%]
  9  [||||||||||||||||||||||   70.6%]
  10 [|||                       7.3%]
  11 [|||||||||||||||||||||||||98.7%]
  12 [|||||||||||||||||||||||||97.4%]
  13 [|||||||||||||||||||||||||99.3%]
  14 [|||||||||||||||||||||||||98.0%]
  15 [|||||||||||||||||||||||||97.4%]
  16 [|||||||||||||||||||||||||98.0%]
  Mem[||||||||||||||||     21.8G/61.1G]
  Swp[                          0K/0K]
```

THANK YOU!

汇报人：22121630 汪江豪
22120721 阮金桐