京东数据仓库海量数据交换工具

-Plumber开发实践

张侃

京东大数据平台部

http://weibo.com/opendayjd

关于我

张侃

大数据平台-实时数据产品研发部

微信： phoenix747

新浪微博： @pho_coder

目 录 CONTENTS

目 录 CONTENTS

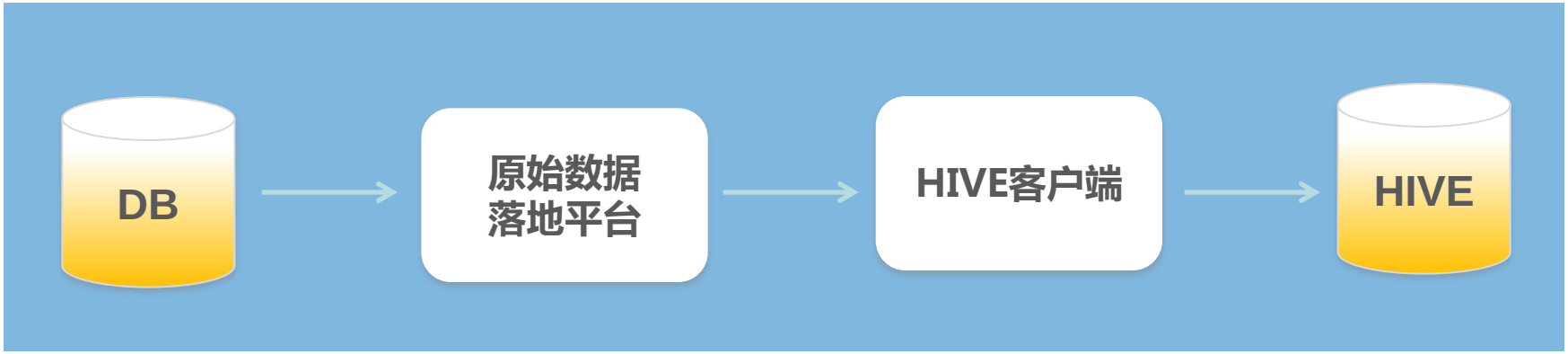
技术突破

- **海量**
  - 每日进出上TB数据
  - 每天数千数据传输任务
- **异构**
  - 结构化：mysql, sqlserver, oracle, hive
  - 非结构化: mongodb, hbase, log
- **场景复杂**
  - mysql分库分表
  - 全国各地仓库数据抽取

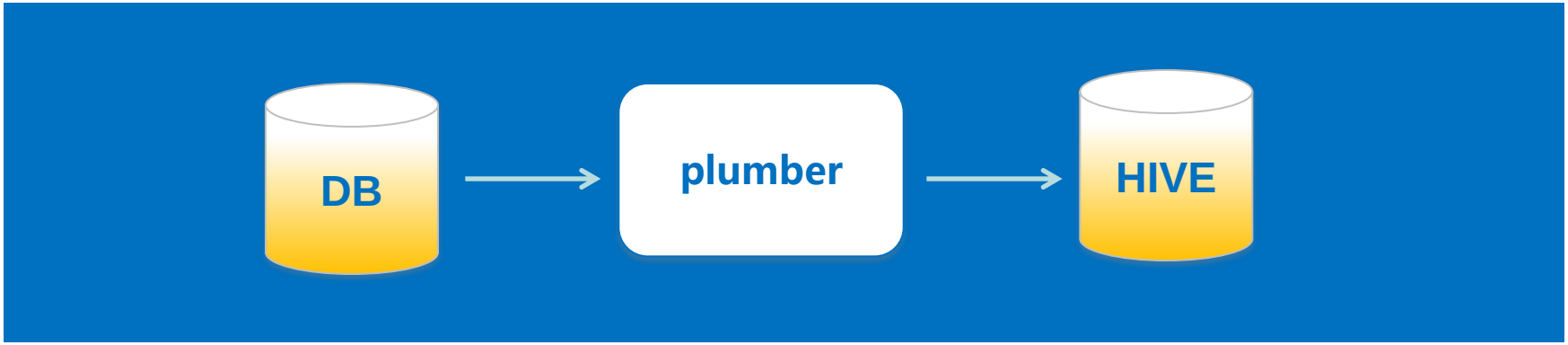
流程优化

京东技术开放日 JD.COM 京东

三次传输　三次落地　一次清洗



一次传输　无落地　实时清洗



目 录 CONTENTS

技术特点及实现方案

京东技术开放日 JD.COM 京东

- 读写分离插件化
- 多线程并行执行
- 配置化和实时统计信息
- 定制化开发全国仓库抽取

读写插件化

Reader： mysql, sqlserver, oracle, mongodb, hive, log

Writer ： hive, mysql, oracle, hbase

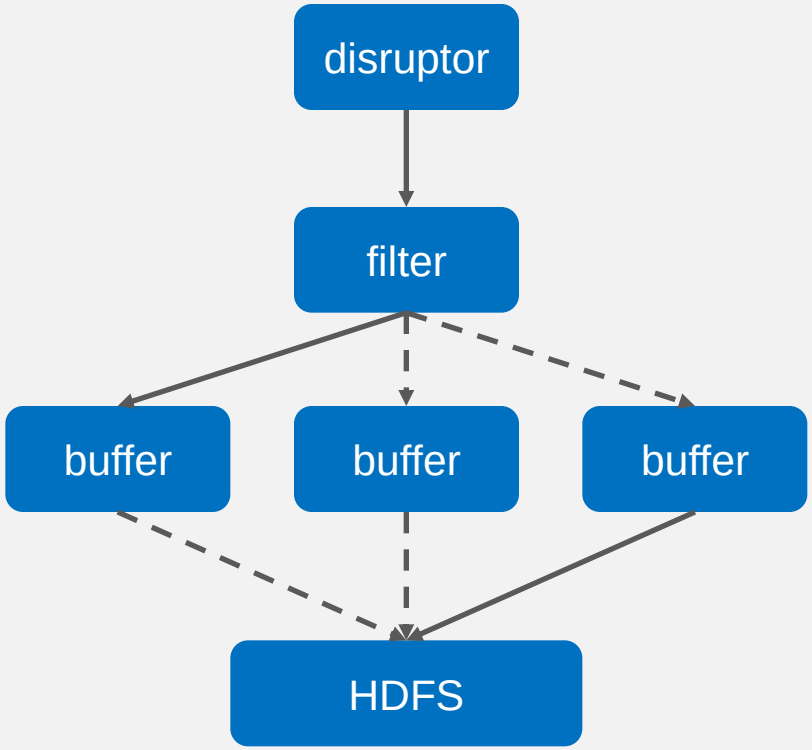
插件实现

Reader

- RDBMS – JDBC
- NOSQL – API
- LOG – http断点续传

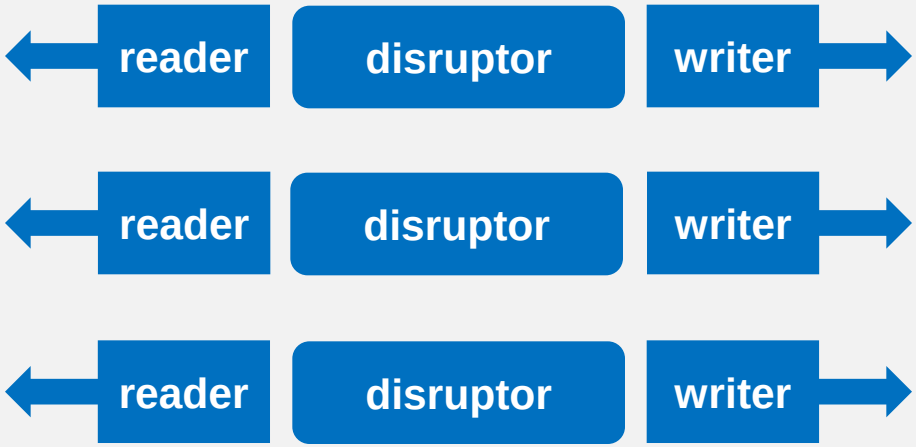
Writer

- RDBMS – JDBC
- Hive – write hdfs & add partitions



hdfs writer

多线程并行执行



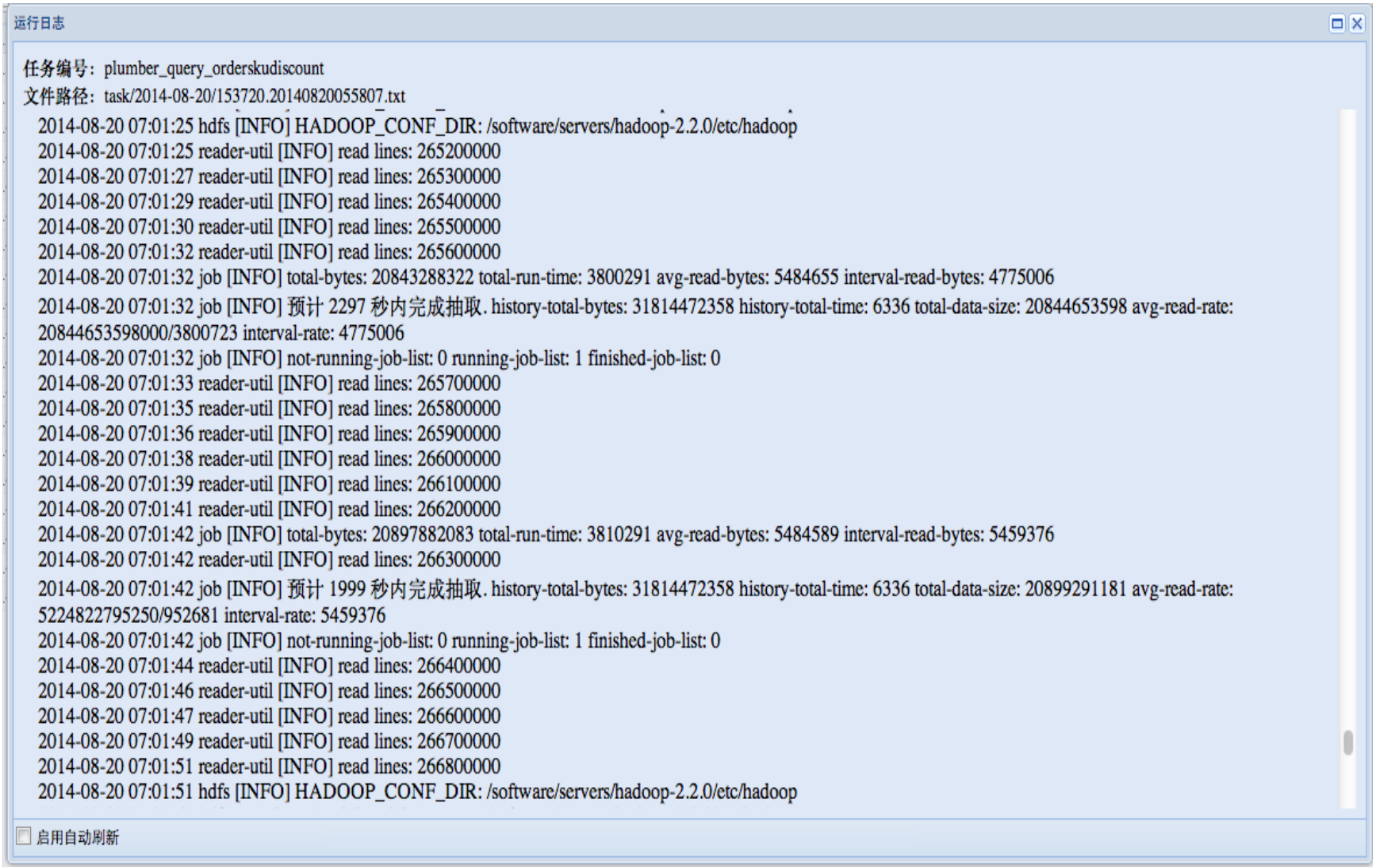
并行执行任务

- 分库分表，库名表名sql拆分

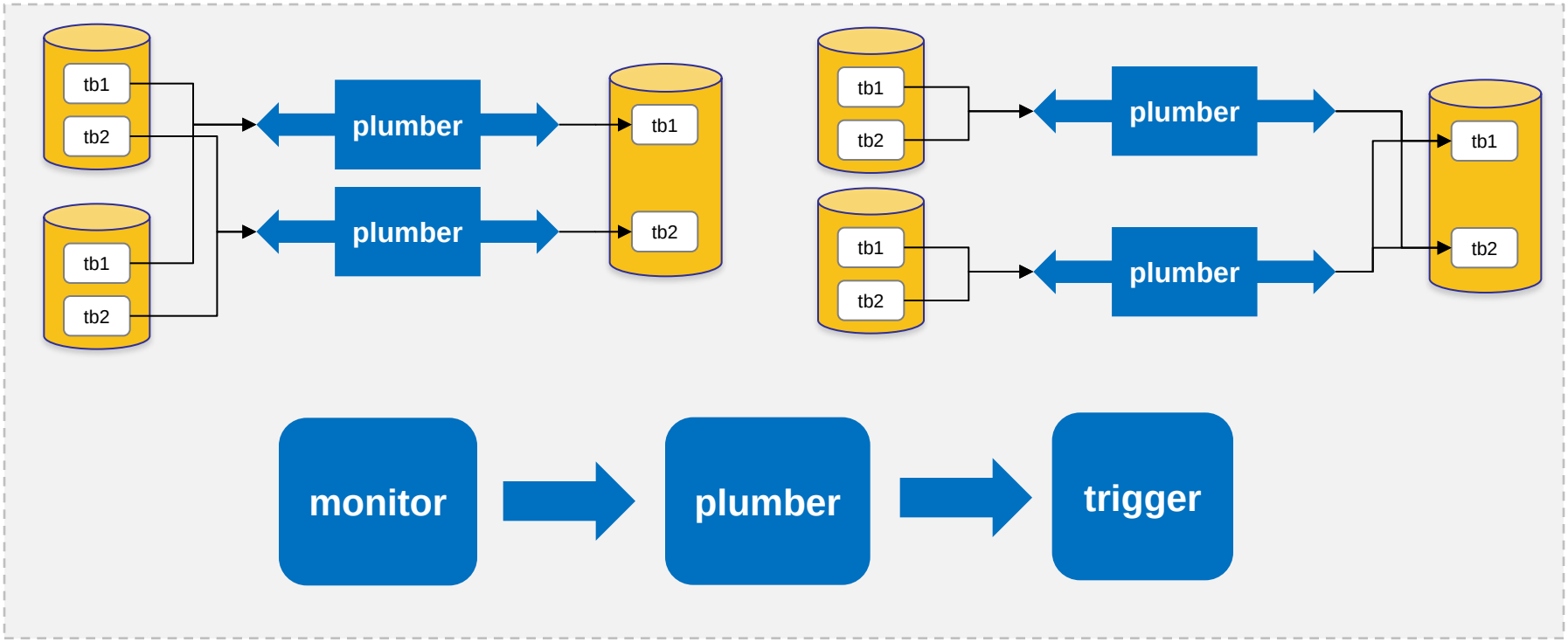
资源有效利用

- 根据系统资源增加线程数

调度模板和实时监控


运行日志
任务编号：plumber_query_orderskudiscount
文件路径：task/2014-08-20/153720.20140820055807.txt
2014-08-20 07:01:25 hdfs [INFO] HADOOP_CONF_DIR: /software/servers/hadoop-2.2.0/etc/hadoop
2014-08-20 07:01:25 reader-util [INFO] read lines: 265200000
2014-08-20 07:01:27 reader-util [INFO] read lines: 265300000
2014-08-20 07:01:29 reader-util [INFO] read lines: 265400000
2014-08-20 07:01:30 reader-util [INFO] read lines: 265500000
2014-08-20 07:01:32 reader-util [INFO] read lines: 265600000
2014-08-20 07:01:32 job [INFO] total-bytes: 20843288322 total-run-time: 3800291 avg-read-bytes: 5484655 interval-read-bytes: 4775006
2014-08-20 07:01:32 job [INFO] 预计 2297 秒内完成抽取. history-total-bytes: 31814472358 history-total-time: 6336 total-data-size: 20844653598 avg-read-rate: 20844653598000/3800723 interval-rate: 4775006
2014-08-20 07:01:32 job [INFO] not-running-job-list: 0 running-job-list: 1 finished-job-list: 0
2014-08-20 07:01:33 reader-util [INFO] read lines: 265700000
2014-08-20 07:01:35 reader-util [INFO] read lines: 265800000
2014-08-20 07:01:36 reader-util [INFO] read lines: 265900000
2014-08-20 07:01:38 reader-util [INFO] read lines: 266000000
2014-08-20 07:01:39 reader-util [INFO] read lines: 266100000
2014-08-20 07:01:41 reader-util [INFO] read lines: 266200000
2014-08-20 07:01:42 job [INFO] total-bytes: 20897882083 total-run-time: 3810291 avg-read-bytes: 5484589 interval-read-bytes: 5459376
2014-08-20 07:01:42 reader-util [INFO] read lines: 266300000
2014-08-20 07:01:42 job [INFO] 预计 1999 秒内完成抽取. history-total-bytes: 31814472358 history-total-time: 6336 total-data-size: 20899291181 avg-read-rate: 5224822795250/952681 interval-rate: 5459376
2014-08-20 07:01:42 job [INFO] not-running-job-list: 0 running-job-list: 1 finished-job-list: 0
2014-08-20 07:01:44 reader-util [INFO] read lines: 266400000
2014-08-20 07:01:46 reader-util [INFO] read lines: 266500000
2014-08-20 07:01:47 reader-util [INFO] read lines: 266600000
2014-08-20 07:01:49 reader-util [INFO] read lines: 266700000
2014-08-20 07:01:51 reader-util [INFO] read lines: 266800000
2014-08-20 07:01:51 hdfs [INFO] HADOOP_CONF_DIR: /software/servers/hadoop-2.2.0/etc/hadoop
启用自动刷新

定制化全国仓库数据抽取



- 仓库分处全国各地，网络情况不确定性大
- 各地仓库下班时间不一，可抽取时间点不一
- 个别仓库宕机不能影响第二天全国仓库报表生成时间点

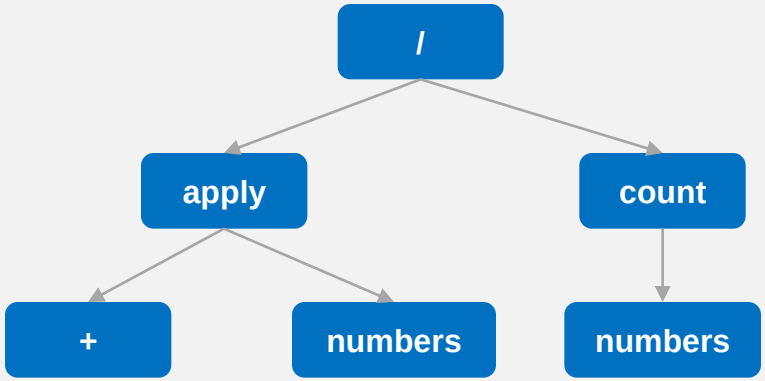
目 录 CONTENTS

为什么是clojure

Clojure

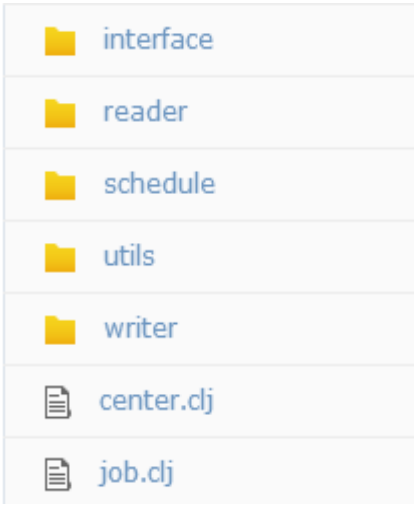
- 纯函数式
- 代码即数据
- 代码即AST
- JVM上的Lisp

```
(defn average
  [numbers]
  (/ (apply + numbers) (count numbers)))
```




```
User=> (average [ 60 80 100 400])
160
```

代码结构




```
localhost:/tmp/plumber/src/plumber/interface phoenix$wc -l *
     113 dispatch.clj
     112 dispatch_spider.clj
      25 hbase.clj
      31 hive.clj
      35 http.clj
      34 interface.clj
      34 load.clj
     111 local.clj
      35 mongo.clj
      33 rdb.clj
     563 total
```

```
localhost:/tmp/plumber/src/plumber/reader phoenix$wc -l *
      14 db_reader.clj
      29 hdfs_reader.clj
      38 hive_reader.clj
      18 http_reader.clj
      92 mongo_reader.clj
       9 mysql_reader.clj
      10 oracle_reader.clj
      56 reader.clj
      36 spider_reader.clj
      10 sqlserver_reader.clj
     312 total
```

```
localhost:/tmp/plumber/src/plumber/writer phoenix$wc -l *
     135 hbase_writer.clj
     141 hbase_writer_dispatch.clj
     156 hdfs_writer.clj
     126 local_infile.clj
     124 local_infile.clj.back
      68 localfs_writer.clj
     101 mysql_load_writer.clj
     102 rdbms_writer.clj
     147 spider_writer.clj
     108 writer.clj
    1208 total
```

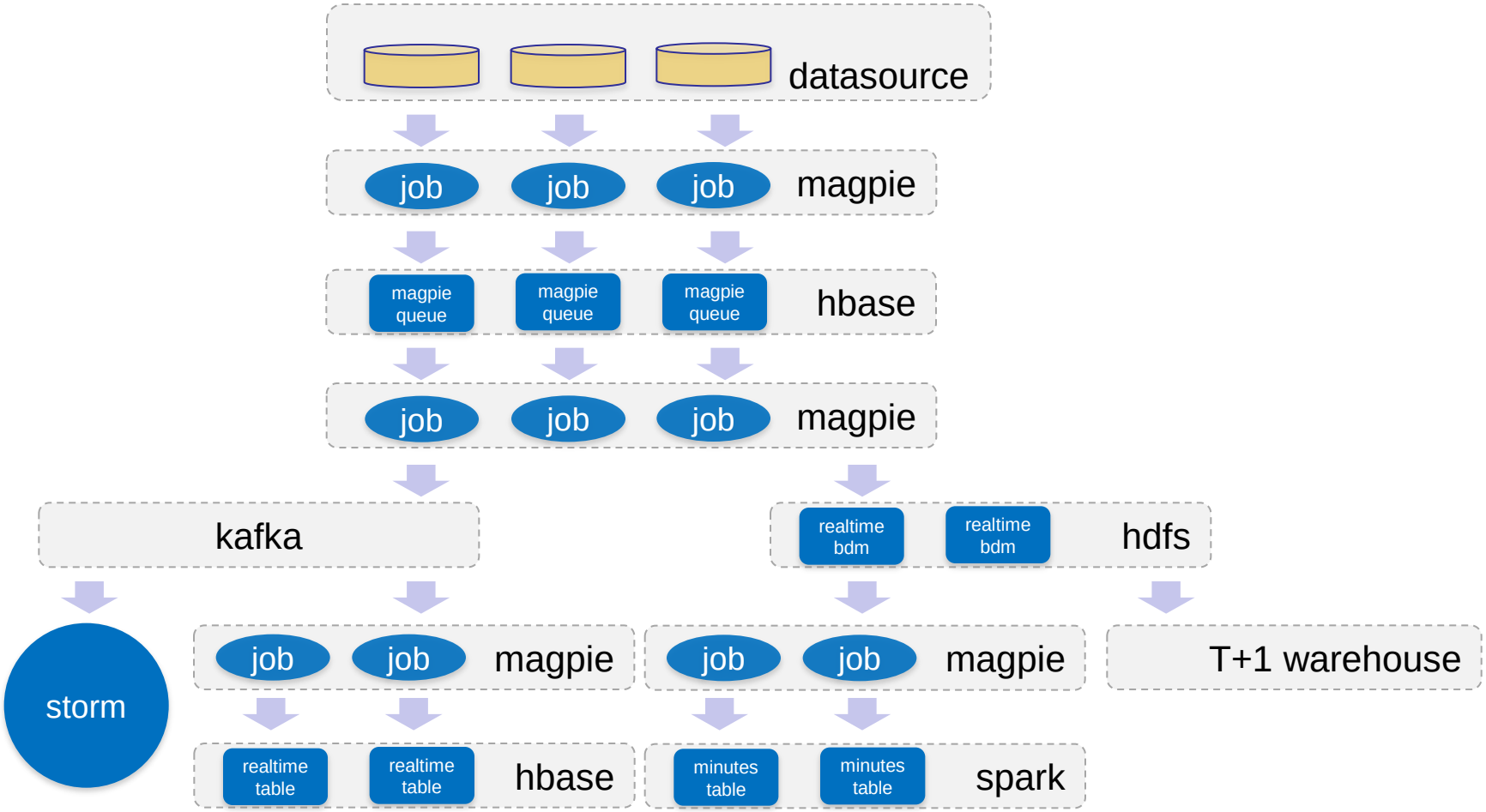
多线程编程

京东技术开放日 JD.COM 京东

```
(defn- sub-job
  "单个子任务,任务运行的最小单元,要不成功,要不失败."
  [sub-job-conf]
  (let [reader-config (atom (:reader sub-job-conf))
        writer-config (:writer sub-job-conf)
        reader-result (ref {:finished false :success false })
        writer-result (ref {:finished false :success false})
        monitor-interval 10000
        total-bytes (atom 0) ;;统计抽数字节数
        last-read-bytes (atom 0)
        avg-read-rate (atom 0) ;; 统计平均抽数速度
        interval-read-rate (atom 0);;统计interval 时间内抽数速度
        start-time (System/currentTimeMillis)
        last-static-time (atom start-time)
        static-fn (f [total-read-bytes] ;;此函数用于统计writer 的各种运行状态
                     (let [now (System/currentTimeMillis)
                           _ (reset! total-bytes total-read-bytes)
                           interval (- now @last-static-time)]
                       (if (>= interval monitor-interval)
                         (let [interval-read-bytes (- total-read-bytes @last-read-bytes)
                               total-run-time (- now start-time)
                               _ (reset! last-read-bytes total-read-bytes)
                               _ (reset! last-static-time now)
                               _ (reset! avg-read-rate (long (/ total-read-bytes (/ total-run-time 1000))))
                               _ (reset! interval-read-rate (long (/ interval-read-bytes (/ interval 1000))))]
                           (log/info "total-bytes:" @total-bytes "total-run-time:" total-run-time "avg-read-bytes:" @avg-read-rate "interval-read-bytes:" @interval-read-rate))))]
    (reify Job
      (init [this] (log/debug "init sub-job" @reader-config) )
      (start [this]
        (let [exec (java.util.concurrent.Executors/newCachedThreadPool)
              [disruptor publisher writer] (d-util/wire-up-disruptor exec writer-config static-fn)
              reader-future (future (r/reader @reader-config publisher))
              writer-future (future (writer))
              listener (f [] (do (while (not (future-done? writer-future))
                                   (try
                                     (Thread/sleep 1000)
                                     (catch Exception e (log/error e))))
                                 (d-util/shutdown-disruptor disruptor exec)
                                 (dosync (ref-set reader-result @reader-future)
                                         (ref-set writer-result (first @writer-future)))
                                 ))]
          (listener)));;监听任务是否完成,若完成,则将返回结果写入到reader-result  和 writer-result 中.

      (stop [this] (reset! reader-config (assoc @reader-config :force-stop true)))
      (finished? [this]  (:finished @writer-result))
      (success? [this] (and (:success @reader-result) (:success @writer-result)))
      (monitor [this] {:job-id (:id @reader-config) :job-start-time start-time :static-time @last-static-time :total-bytes @total-bytes :avg-read-rate @avg-read-rate :interval-read-rate @interval-read-rate})
      ;; when call finish,please check success? is true
      (result [this] (if (success? this){:reader @reader-result :writer (merge @writer-result  (monitor this))} nil))
      (info [this] sub-job-conf))))
```

Clojure in JD



magpie: realtime task scheduling system for realtime data warehouse

加入我们

京东技术开放日 JD.COM 京东

京东大数据平台

我们招聘

-分布式研发工程师

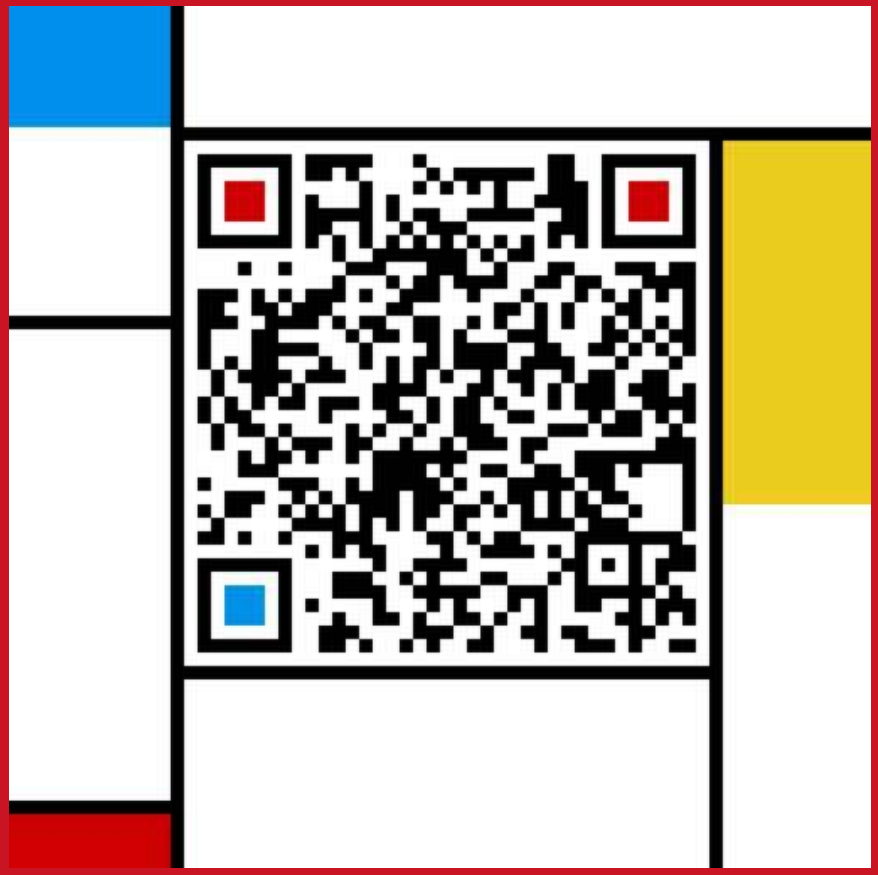
-Hadoop/Hbase/Spark/Search研发

我们欢迎所有对海量数据处理技术感兴趣的同学

简历可发送至：zhangkan@jd.com

招聘详情请关注“京东大数据”微信公众平台

谢谢！
Thank you!



北京市朝阳区北辰西路8号北辰世纪中心A座6层
6F Building A, North-Star Century Center, 8 Beichen West Street,
Chaoyang District, Beijing 100101
T. 010-5895 1234 F. 010-5895 1234
E. xingming@jd.com www.jd.com

京东技术开放日

JD.京东.COM