



第九届中国数据库技术大会  
DATABASE TECHNOLOGY CONFERENCE CHINA 2018

# Flink 在今日头条的实践

作者：张光辉

日期：2018.05.10

邮箱：zhangguanghui@bytedance.com

DTCC  
2018

2018.05.10 – 12 北京国际会议中心



IT168.com

ChinaUnix

ITPUB

# 内容大纲

- 引入 Flink 的背景
- 构建 Flink Cluster
- 兼容 Jstorm 作业 & 迁移
- 构建流式任务管理平台
- QA



# 大纲

## ➤ 引入 Flink 的背景

- 构建 Flink Cluster
- 兼容 Jstorm 作业 & 迁移
- 构建流式任务管理平台
- QA

# 背景

应用层	广告	A/B Test	推送	数据仓库	.....
抽象 Topology	python Spout	python bolt	topology.yaml		
Jstorm 计算引擎	default	dw cluster	push cluster	ads cluster	.....

# Jstorm 集群规模

集群	25+
机器	5000+
作业	1200+

# Jstorm VS Flink

## Jstorm 的问题

- Worker 没有内存限制
- 业务团队之间没有 quota 管理，预算审核无头绪
- 重要业务采用独立集群，运维复杂
- Python topology 性能较差，推广 Java topology 比较难

## Flink on yarn 的优势

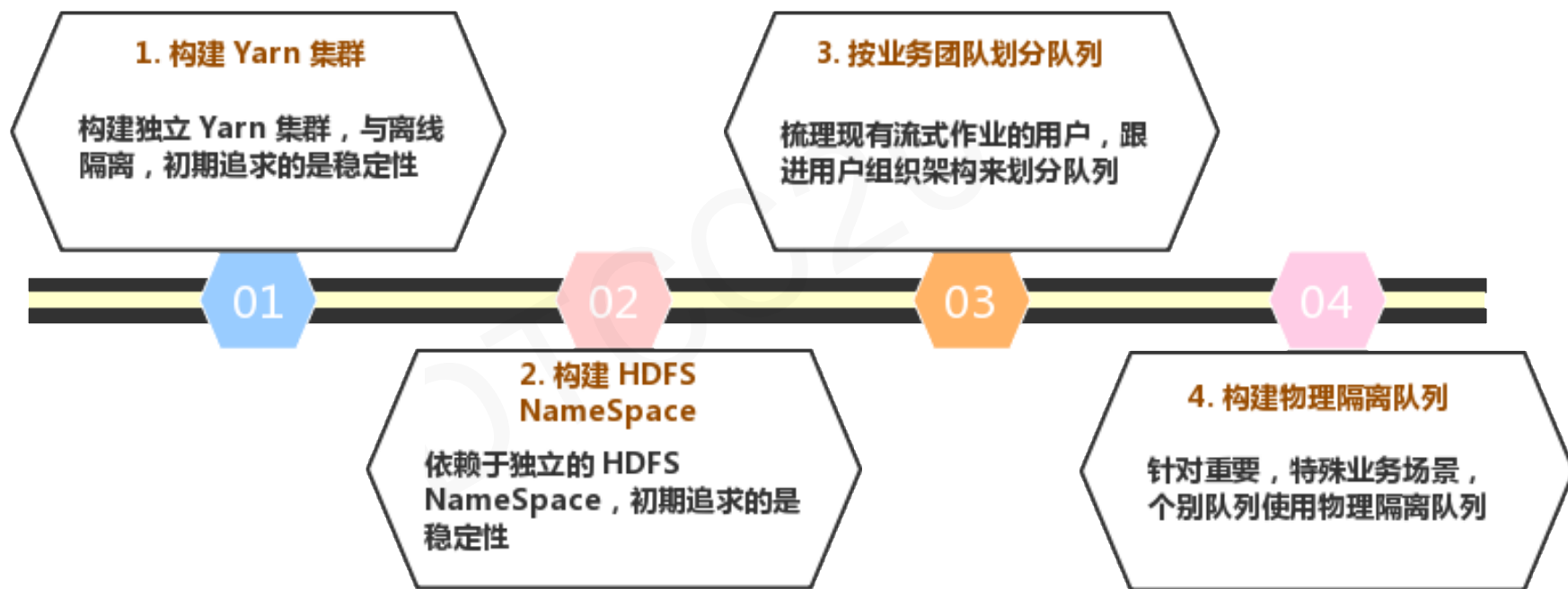
- Flink on yarn 解决 worker 粒度的物理资源限制
- Yarn 队列支持 Quota 管理
- Yarn 支持物理隔离队列
- Flink 可以兼容 storm 作业
- Flink 提供 high-level api，例如 SQL, operator
- Flink 提供窗口，checkpoint 等特性

# 大纲

- 引入 Flink 的背景
- **构建 Flink Cluster**
- 兼容 Jstorm 作业 & 迁移
- 构建流式任务管理平台
- QA

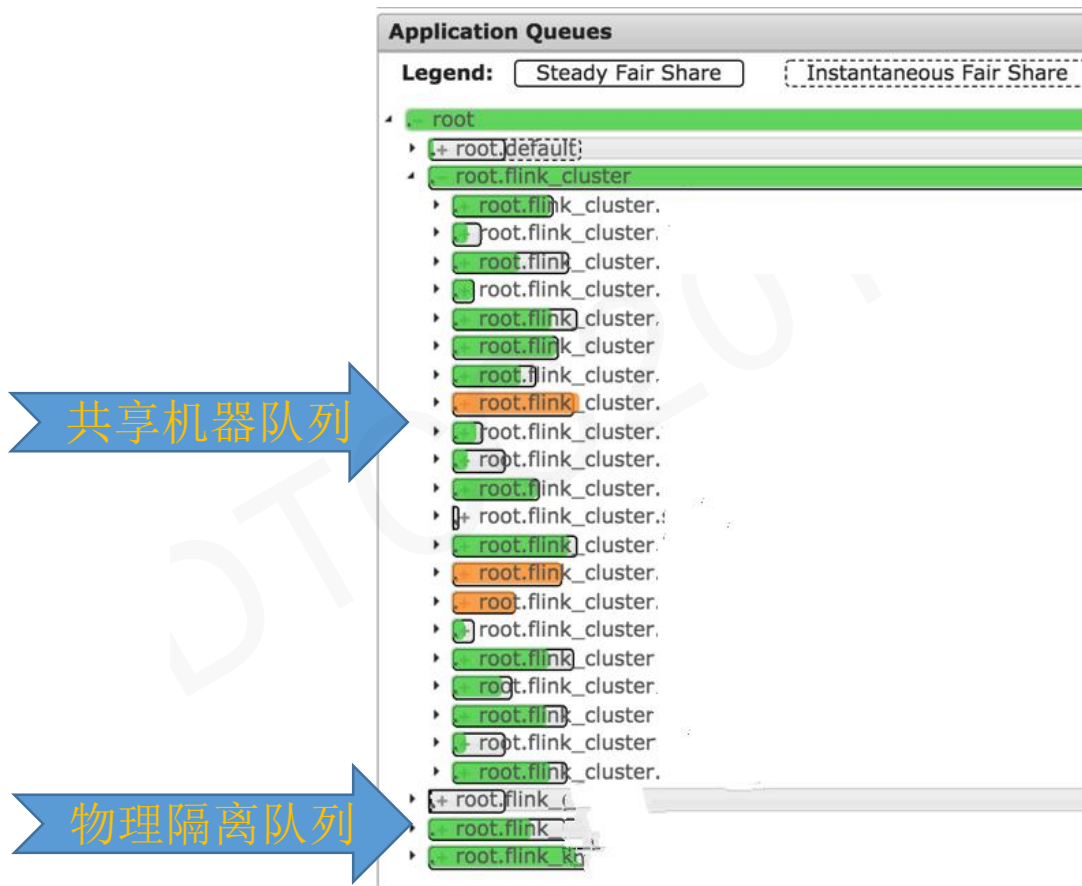


# 构建 Flink Cluster





# 构建 Flink Cluster



# 大纲

- 引入 Flink 的背景
- 构建 Flink Cluster
- 兼容 Jstorm 作业&迁移
- 构建流式任务管理平台
- QA



# 兼容 Jstorm

1. 借鉴 Flink-storm，实现了 Flink-jstorm，支持将 Jstorm topology 结构转换为 Flink job。
2. 构建脚本，能够提交 Job，停止 Job，自动注册 Metrics Dashboard，自动注册报警。
3. 构建迁移脚本，获取 Jstorm 作业资源使用情况生成 Flink 资源配置。

# 兼容Jstorm

`topology start -m jstorm.online -y topology_online.yaml`  
`topology stop -m jstorm.online -y topology_online.yaml`



`flink start -q root.flink.online -y topology_online.yaml`  
`flink stop -q root.flink.online -y topology_online.yaml`

# Jstorm->Flink 成绩表

Yarn 集群	3
机器	3000+
作业	1100+
进度	90%

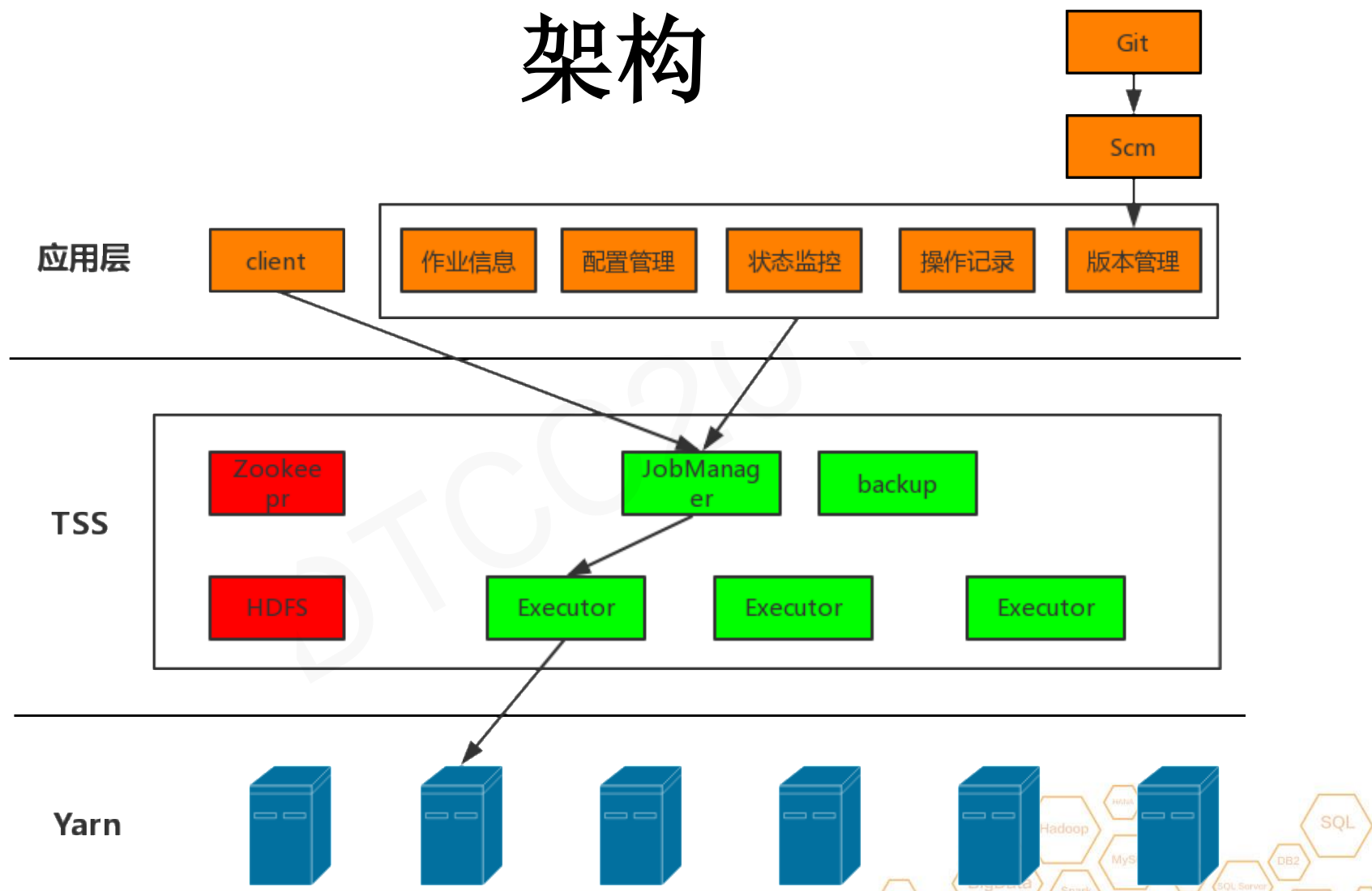
# 大纲

- 引入 Flink 的背景
- 构建 Flink Cluster
- 兼容 Jstorm 作业&迁移
- 构建流式任务管理平台
- QA

# 目标



# 架构





# 创建流式作业

创建作业模板

书写代码，完成业务逻辑

测试作业

在平台上注册作业

添加配置和参数

对代码升级版本

启动作业

查看作业运行状态

使用脚本生成作业模板，以下分别是 Python, Java

```
topology gen --name wordcount --spout="simple_spout"
--bolt="split_bolt,wordcount_bolt"
```

```
mvn archetype:generate \
-DarchetypeGroupId=com.bytedance.data \
-DarchetypeArtifactId=flink-quickstart-java-archetype \
-DarchetypeVersion=0.7.0 \
-DgroupId=com.bytedance.data \
-DartifactId=wordcount \
-Dpackage=com.bytedance.data \
-Dversion=1.0-SNAPSHOT
```

# 测试作业

创建作业模板

书写代码，完成业务逻辑

测试作业

在平台上注册作业

添加配置和参数

对代码升级版本

启动作业

查看作业运行状态

## 测试方式1

均支持本地测试，其中 Flink Java 可以使用 IDEA

## 测试方式2

使用 tss-client 提交到 stage 环境

# 注册作业

新建任务

创建作业模板

书写代码，完成业务逻辑

测试作业

在平台上注册作业

添加配置和参数

对代码升级版本

启动作业

查看作业运行状态

1. 基本信息

2. SCM仓库

3. 运行参数

4. 完成

\* 任务名:

\* 描述:

\* 负责人:

+

\* Job类型:

Flink

▼

\* 机房:

cn

▼

\* 集群:

▼

\* 队列:

▼

报警lark群:

+

下一步

DTCC  
2018

数领先机 智赢未来 (9)

IT168.com

ChinaUnix

ITPUB

SQL

DB2

SQL Server

Spark

BigData

LDAP

NewSQL

Graph

CDP

# 增加配置信息

创建作业模板

书写代码，完成业务逻辑

测试作业

在平台上注册作业

添加配置和参数

对代码升级版本

启动作业

查看作业运行状态

基本信息

配置详情

监控信息

操作历史

```
{
  "tm_num":16,
  "tm_memory_mb":10240,
  "container_vcores":8,
  "slots_per_tm":8,
  "jm_memory_mb":4096,
  "main_class":"com.bytedance.data.lis.LogIndexService",
  "custom":{}
}
```

# 更新代码版本号

创建作业模板

书写代码，完成业务逻辑

测试作业

在平台上注册作业

添加配置和参数

对代码升级版本

启动作业

查看作业运行状态

发布 inf/tss/log\_index\_service

发布方式

☒ 基于分支 ☐ 基于 Commit Hash

分支

master

\* 版本号

1.0.0.3

\* 版本描述

Merge branch 'handle\_es\_exception' into 'master'  
添加了ES异常处理；设置默认从kafka最新的offset读取数据

高级选项

取消

发布

# 提交作业

创建作业模板

书写代码，完成业务逻辑

测试作业

在平台上注册作业

添加配置和参数

对代码升级版本

启动作业

查看作业运行状态

步骤1: 确认配置信息

[编辑配置信息](#)

Id: 54(2018-04-06 00:50:48) ✓

SCM	inf/tss/helloworld 1.0.0.3
集群(cluster)	flink
队列(queue)	root.flink_cluster.online
TaskManager个数	2
Slot个数/TM	4
内存/TM(MB)	4096
CPU/TM(vCore)	2
动态参数	{}
自定义参数	{}

步骤2: 确认提交操作

取消

确认

DTCC  
2018

数领先机 智赢未来 (9)

IT168.com

ChinaUnix

ITPUB

# 查看作业状态

创建作业模板

书写代码，完成业务逻辑

测试作业

在平台上注册作业

添加配置和参数

对代码升级版本

启动作业

查看作业运行状态

1. Flink UI, 官方自带的 Web UI
2. Dashboard, 将作业的重要 metrics 汇总到一起，方便查看
3. 错误日志，将作业的错误日志收集写到 ES 上，用 kibana 查看
4. Jobtrace, 通过提取作业异常日志，匹配异常类型，提供给用户解决方案

# Flink UI

Apache Flink Dashboard

log\_index\_service  
147403504d0fc32a34581e417165eb9d

2018-04-14, 18:14:50 16h 40m 0 0 0 2 0 0 0 0 0

DashBoard dtop Cancel

Overview Timeline Exceptions Configuration

Overview  
Running Jobs  
Completed Jobs  
Task Managers  
Job Manager  
Submit new Job

Source: Custom Source -> Filter -> Map  
Parallelism: 128

HASH

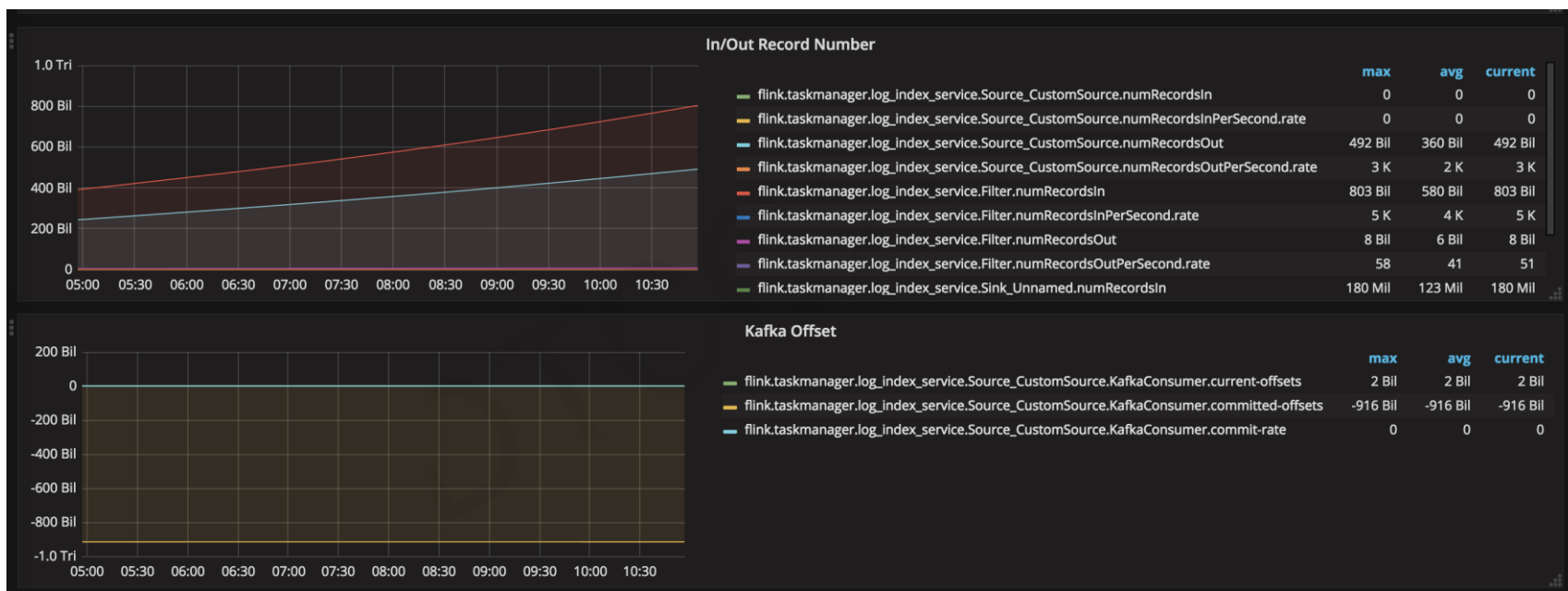
Filter -> Map -> Sink: Unnamed  
Parallelism: 128

Subtasks Subtasks by TaskManager Task Metrics Watermarks Accumulators Checkpoints Back Pressure

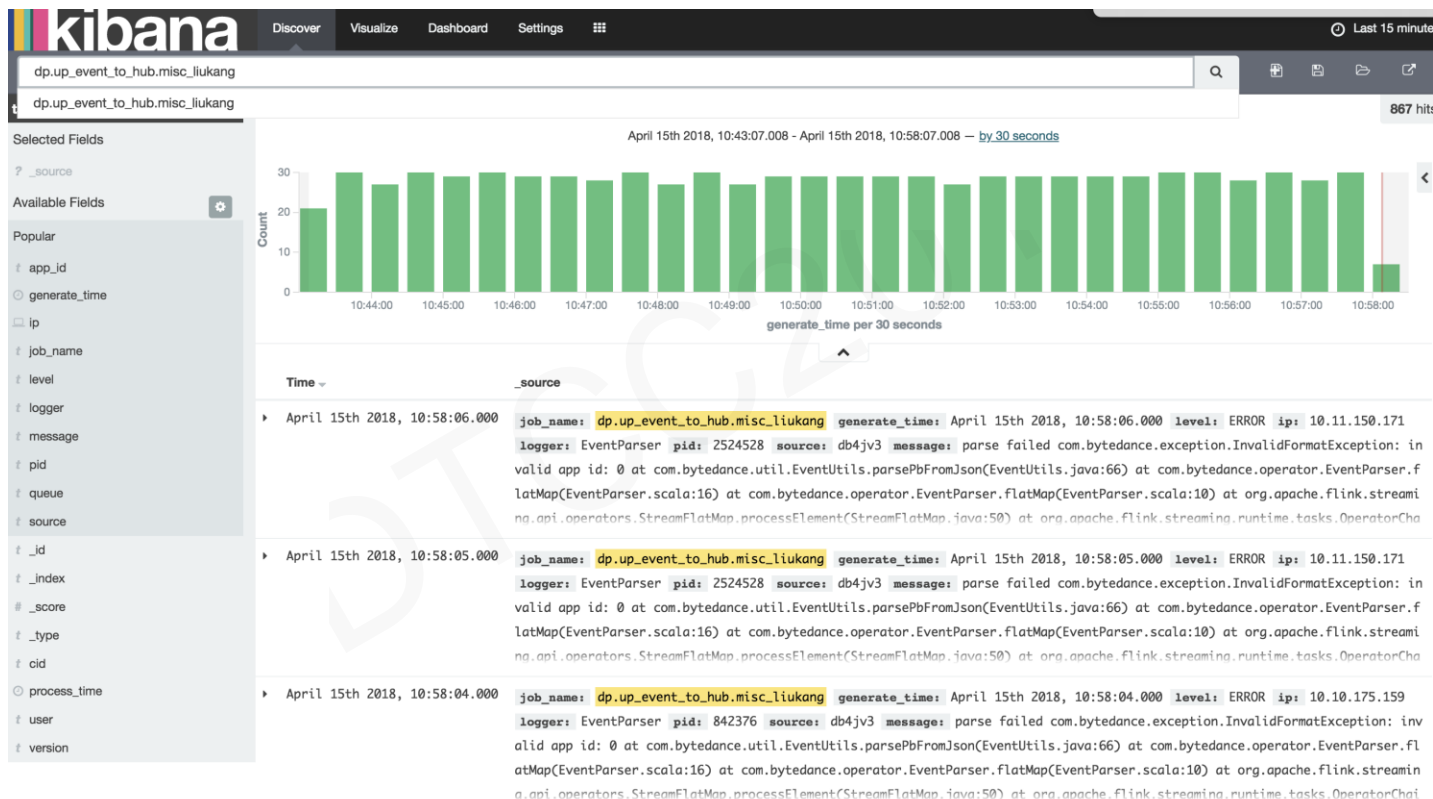
Start Time	End Time	Duration	Name	Bytes received	Records received	Bytes sent	Records sent	Parallelism	Tasks	Status
2018-04-14, 18:14:50	2018-04-15, 10:55:27	16h 40m	Source: Custom Source -> Filter -> Map	0 B	0	164 GB	290,695,206	128	0 0 128 0 0	RUNNING
2018-04-14, 18:14:50	2018-04-15, 10:55:27	16h 40m	Filter -> Map -> Sink: Unnamed	164 GB	290,694,991	0 B	0	128	0 0 128 0 0	RUNNING



# DashBoard



# Kibana



# JobTrace

将每个 Flink Job 的 ERROR 日志收集汇总，根据 Error 日志来诊断作业问题，将已知问题的解决方案反馈给用户，下面是常见异常日志。

- java.lang.OutOfMemoryError
- Insufficient number of network buffers
- ImportError: No module named
- is running beyond physical memory limits
- diagnostics=Container released on a .\*lost.\* node
- ERROR not support cluster name
- Unable to allocate further port in port range
- java.lang.IndexOutOfBoundsException
- No such file or directory



# 操作历史

基本信息

配置详情

监控信息

操作历史

- zhangguanghui 停止任务 (2018-04-14 17:18:25)
- zhangguanghui 启动任务 (2018-04-14 01:09:43)
- zhangguanghui 创建任务 (2018-04-14 01:09:14)



# 致谢

1. 感谢今日头条 Data-inf 和 Data-platform 伙伴们，以上的分享内容，是我们一起共同努力的成果。
2. 感谢 DTCC 的组委会成员，使我能够站在这里分享。

# THANKS







讲师申请

联系电话（微信号）：18612470168

关注“ITPUB”更多  
技术干货等你来拿~

与百度外卖、京东、魅族等先后合作系列分享活动



## 让学习更简单

微学堂是以ChinaUnix、ITPUB所组建的微信群为载体，定期邀请嘉宾对热点话题、技术难题、新产品发布等进行移动端的在线直播活动。

截至目前，累计举办活动期数60+，参与人次40000+。

## ITPUB学院

ITPUB学院是盛拓传媒IT168企业事业部（ITPUB）旗下  
企业级在线学习咨询平台  
历经18年技术社区平台发展  
汇聚5000万技术用户  
紧随企业一线IT技术需求  
打造全方式技术培训与技术咨询服务  
提供包括企业应用方案培训咨询（包括企业内训）  
个人实战技能培训（包括认证培训）  
在内的全方位IT技术培训咨询服务

ITPUB学院讲师均来自于企业  
一些工程师、架构师、技术经理和CTO  
大会演讲专家1800+  
社区版主和博客专家500+

## 培训特色

无限次免费播放  
随时随地在线观看  
碎片化时间集中学习  
聚焦知识点详细解读  
讲师在线答疑  
强大的技术人脉圈

## 八大课程体系

基础架构设计与建设  
大数据平台  
应用架构设计与开发  
系统运维与数据库  
传统企业数字化转型  
人工智能  
区块链  
移动开发与SEO



## 联系我们

联系人：黄老师  
电话：010-59127187  
邮箱：edu@itpub.net  
网址：edu.itpub.net  
培训微信号：18500940168