kylin资讯频道排行迁移方案说明书

1. 组件描述：

资讯频道排行由kylin提供，在bi系统中展示电视猫和微鲸的资讯频道播放排行。

1. 迁移步骤：
2. 前置条件：

基础组建kylin；

logstash

bi报表系统

1. 迁移步骤：

因为仅需要保留半天的信息，所以只需要在新机房中重新部署启动即可

* 通过logstash做kafka topc的切分
* kylin上cube的构建
* 在报表系统中配置报表

1. 检查清单：

|  |  |  |
| --- | --- | --- |
| 阶段 | 检测点 | 检测方法 |
| Kafka的topic分流 | 查看打到日志文件正确性 | Json格式化查看是否过滤成功 |
| 查看kafka里日志正确性 | Json格式化查看是否过滤成功 |
| Cube构建 | 查看构建是否成功 | 观察kylin的monitor |
| 查看数据正确性 | 编写测试sql在kylin的insight里查询 |
| 报表系统配置 | 生成报表 | 报表展示正确 |

1. kafka topic切分条件
2. 电视猫资讯频道

[logType] == "play" and [contentType] == "hot" and [event] == "startplay"

1. 微鲸资讯频道

[logType] == "play" and [contentType] == "hot"

1. cube构件维度值和度量值详情

可以参见kylin系统内配置http://115.231.96.78:7070/kylin/models

具体如下：

1. 电视猫资讯频道

**模型**

"dimensions": [

{

"table": "DEFAULT.MEDUSA\_HOT\_PLAY",

"columns": [

"PATHSUB",

"LOGTIMESTAMP",

"DATETIME",

"VIDEONAME",

"USERID",

"PATHMAIN",

"PATHSPECIAL",

"ACCOUNTID",

"VIDEOSID",

"CONTENTTYPE",

"DAY"

]

}

],

"metrics": [

"VIDEOSID",

"ACCOUNTID",

"USERID",

"VIDEONAME"

],

"filter\_condition": "",

"partition\_desc": {

"partition\_date\_column": "DEFAULT.MEDUSA\_HOT\_PLAY.DAY",

"partition\_time\_column": null,

"partition\_date\_start": 0,

"partition\_date\_format": "yyyy-MM-dd",

"partition\_time\_format": "HH:mm:ss",

"partition\_type": "APPEND",

"partition\_condition\_builder": "org.apache.kylin.metadata.model.PartitionDesc$DefaultPartitionConditionBuilder"

},

Cube

{

"uuid": "f48ac86b-1e34-413d-839c-1862091402d6",

"last\_modified": 1477035568045,

"version": "1.5.4",

"name": "medusa\_hot\_play\_cube",

"model\_name": "medusa\_hot\_play\_model",

"description": "",

"null\_string": null,

"dimensions": [

{

"name": "DEFAULT.MEDUSA\_HOT\_PLAY.DATETIME",

"table": "DEFAULT.MEDUSA\_HOT\_PLAY",

"column": "DATETIME",

"derived": null

},

{

"name": "DEFAULT.MEDUSA\_HOT\_PLAY.VIDEONAME",

"table": "DEFAULT.MEDUSA\_HOT\_PLAY",

"column": "VIDEONAME",

"derived": null

},

{

"name": "DEFAULT.MEDUSA\_HOT\_PLAY.CONTENTTYPE",

"table": "DEFAULT.MEDUSA\_HOT\_PLAY",

"column": "CONTENTTYPE",

"derived": null

},

{

"name": "DEFAULT.MEDUSA\_HOT\_PLAY.USERID",

"table": "DEFAULT.MEDUSA\_HOT\_PLAY",

"column": "USERID",

"derived": null

},

{

"name": "DEFAULT.MEDUSA\_HOT\_PLAY.PATHMAIN",

"table": "DEFAULT.MEDUSA\_HOT\_PLAY",

"column": "PATHMAIN",

"derived": null

},

{

"name": "DEFAULT.MEDUSA\_HOT\_PLAY.VIDEOSID",

"table": "DEFAULT.MEDUSA\_HOT\_PLAY",

"column": "VIDEOSID",

"derived": null

},

{

"name": "DEFAULT.MEDUSA\_HOT\_PLAY.PATHSPECIAL",

"table": "DEFAULT.MEDUSA\_HOT\_PLAY",

"column": "PATHSPECIAL",

"derived": null

}

],

"measures": [

{

"name": "\_COUNT\_",

"function": {

"expression": "COUNT",

"parameter": {

"type": "constant",

"value": "1",

"next\_parameter": null

},

"returntype": "bigint"

},

"dependent\_measure\_ref": null

},

{

"name": "COUNT\_DISTINCT\_USER",

"function": {

"expression": "COUNT\_DISTINCT",

"parameter": {

"type": "column",

"value": "USERID",

"next\_parameter": null

},

"returntype": "bitmap"

},

"dependent\_measure\_ref": null

},

{

"name": "VIDEONAME",

"function": {

"expression": "RAW",

"parameter": {

"type": "column",

"value": "VIDEONAME",

"next\_parameter": null

},

"returntype": "raw"

},

"dependent\_measure\_ref": null

}

],

"dictionaries": [],

"rowkey": {

"rowkey\_columns": [

{

"column": "DATETIME",

"encoding": "dict",

"isShardBy": false

},

{

"column": "VIDEONAME",

"encoding": "dict",

"isShardBy": false

},

{

"column": "CONTENTTYPE",

"encoding": "dict",

"isShardBy": false

},

{

"column": "USERID",

"encoding": "dict",

"isShardBy": false

},

{

"column": "PATHMAIN",

"encoding": "dict",

"isShardBy": false

},

{

"column": "VIDEOSID",

"encoding": "dict",

"isShardBy": false

},

{

"column": "PATHSPECIAL",

"encoding": "dict",

"isShardBy": false

}

]

},

"hbase\_mapping": {

"column\_family": [

{

"name": "F1",

"columns": [

{

"qualifier": "M",

"measure\_refs": [

"\_COUNT\_",

"VIDEONAME"

]

}

]

},

{

"name": "F2",

"columns": [

{

"qualifier": "M",

"measure\_refs": [

"COUNT\_DISTINCT\_USER"

]

}

]

}

]

},

"aggregation\_groups": [

{

"includes": [

"DATETIME",

"VIDEONAME",

"CONTENTTYPE",

"USERID",

"PATHMAIN",

"VIDEOSID",

"PATHSPECIAL"

],

"select\_rule": {

"hierarchy\_dims": [],

"mandatory\_dims": [],

"joint\_dims": []

}

}

],

"signature": "3S3ZO/99k8ZqEp35gf4TtA==",

"notify\_list": [],

"status\_need\_notify": [

"ERROR",

"DISCARDED",

"SUCCEED"

],

"partition\_date\_start": 0,

"partition\_date\_end": 3153600000000,

"auto\_merge\_time\_ranges": [

604800000,

2419200000

],

"retention\_range": 43200000,

"engine\_type": 2,

"storage\_type": 2,

"override\_kylin\_properties": {}

}

1. 微鲸资讯频道

模型

"dimensions": [

{

"table": "DEFAULT.HOT\_PLAY",

"columns": [

"DATETIME",

"ACCOUNTID",

"DATEMINUTE",

"USERID",

"VIDEOSID",

"DATEHOUR",

"LOGTIMESTAMP",

"PATH",

"CONTENTTYPE"

]

}

],

"metrics": [

"ACCOUNTID",

"USERID",

"VIDEOSID"

],

"filter\_condition": "",

"partition\_desc": {

"partition\_date\_column": "DEFAULT.HOT\_PLAY.DATETIME",

"partition\_time\_column": null,

"partition\_date\_start": 0,

"partition\_date\_format": "yyyy-MM-dd HH:mm:ss",

"partition\_time\_format": "HH:mm:ss",

"partition\_type": "APPEND",

"partition\_condition\_builder": "org.apache.kylin.metadata.model.PartitionDesc$DefaultPartitionConditionBuilder"

},

Cube

{

"uuid": "90d8b28c-c852-43d4-937b-2bfe71dc13bf",

"last\_modified": 1476755453834,

"version": "1.5.4",

"name": "hot\_play\_cube",

"model\_name": "hot\_play\_model\_2",

"description": "",

"null\_string": null,

"dimensions": [

{

"name": "DEFAULT.HOT\_PLAY.PATH",

"table": "DEFAULT.HOT\_PLAY",

"column": "PATH",

"derived": null

},

{

"name": "DEFAULT.HOT\_PLAY.CONTENTTYPE",

"table": "DEFAULT.HOT\_PLAY",

"column": "CONTENTTYPE",

"derived": null

},

{

"name": "DEFAULT.HOT\_PLAY.DATEMINUTE",

"table": "DEFAULT.HOT\_PLAY",

"column": "DATEMINUTE",

"derived": null

},

{

"name": "DEFAULT.HOT\_PLAY.USERID",

"table": "DEFAULT.HOT\_PLAY",

"column": "USERID",

"derived": null

},

{

"name": "DEFAULT.HOT\_PLAY.VIDEOSID",

"table": "DEFAULT.HOT\_PLAY",

"column": "VIDEOSID",

"derived": null

}

],

"measures": [

{

"name": "\_COUNT\_",

"function": {

"expression": "COUNT",

"parameter": {

"type": "constant",

"value": "1",

"next\_parameter": null

},

"returntype": "bigint"

},

"dependent\_measure\_ref": null

},

{

"name": "COUNT\_DISTINCT\_USER",

"function": {

"expression": "COUNT\_DISTINCT",

"parameter": {

"type": "column",

"value": "USERID",

"next\_parameter": null

},

"returntype": "bitmap"

},

"dependent\_measure\_ref": null

}

],

"dictionaries": [],

"rowkey": {

"rowkey\_columns": [

{

"column": "PATH",

"encoding": "dict",

"isShardBy": false

},

{

"column": "CONTENTTYPE",

"encoding": "dict",

"isShardBy": false

},

{

"column": "DATEMINUTE",

"encoding": "dict",

"isShardBy": false

},

{

"column": "USERID",

"encoding": "dict",

"isShardBy": false

},

{

"column": "VIDEOSID",

"encoding": "dict",

"isShardBy": false

}

]

},

"hbase\_mapping": {

"column\_family": [

{

"name": "F1",

"columns": [

{

"qualifier": "M",

"measure\_refs": [

"\_COUNT\_"

]

}

]

},

{

"name": "F2",

"columns": [

{

"qualifier": "M",

"measure\_refs": [

"COUNT\_DISTINCT\_USER"

]

}

]

}

]

},

"aggregation\_groups": [

{

"includes": [

"PATH",

"CONTENTTYPE",

"DATEMINUTE",

"USERID",

"VIDEOSID"

],

"select\_rule": {

"hierarchy\_dims": [],

"mandatory\_dims": [],

"joint\_dims": []

}

}

],

"signature": "ygAaacrVlWXT8rnq1RBn4w==",

"notify\_list": [

"li.tuo@whaley.cn"

],

"status\_need\_notify": [

"ERROR",

"DISCARDED",

"SUCCEED"

],

"partition\_date\_start": 0,

"partition\_date\_end": 3153600000000,

"auto\_merge\_time\_ranges": [

604800000,

2419200000

],

"retention\_range": 43200000,

"engine\_type": 2,

"storage\_type": 2,

"override\_kylin\_properties": {}

}