

Hadoop On Demand: 配置指南

目录

1	1. 简介.....	2
2	2. 段.....	2
3	3. HOD配置项.....	2
3.1	3.1 一般的配置项.....	2
3.2	3.2 hod的配置项.....	3
3.3	3.3 resouce_manager的配置项.....	3
3.4	3.4 ringmaster的配置项.....	4
3.5	3.5 gridservice-hdfs的配置项.....	4
3.6	3.6 gridservice-mapred的配置项.....	4
3.7	3.7 hodring的配置项.....	5

1. 1. 简介

这个文档讲述了一些最重要和常用的Hadoop On Demand (HOD) 的配置项。这些配置项可通过两种方式指定: INI风格的配置文件, 通过`--section.option[=value]`格式指定的HOD shell的命令行选项。如果两个地方都指定了同一个选项, 命令行中的值覆盖配置文件中的值。

你可以通过以下命令获得所有配置项的简要描述:

```
$ hod --verbose-help
```

2. 2. 段

HOD配置文件分成以下几个配置段:

- `hod`: HOD客户端的配置项
- `resource_manager`: 指定要使用的资源管理器的配置项, 以及使用该资源管理器时需要的一些其他参数。
- `ringmaster`: RingMaster进程的配置项
- `hodring`: HodRing进程的配置项
- `gridservice-mapred`: Map/Reduce守护进程的配置项
- `gridservice-hdfs`: HDFS守护进程的配置项

3. 3. HOD配置项

接下来的一节会先描述大部分HOD配置段中通用的一些配置项, 再描述各配置段特有的配置项。

3.1. 3.1 一般的配置项

某些配置项会在HOD配置中的多个段定义。在一个段中定义的配置项, 会被该段所适用的所有进程使用。这些配置项意义相同, 但在不同的段中可以有不同的取值。

- `temp-dir`: HOD进程使用的临时目录。请确保运行hod的用户有权限在这个指定的目录下创建子目录。如果想在每次分配的时候都使用不同的临时目录, 可以使用环境变量, 资源管理器会让这些环境变量对HOD进程可用。例如, 在Torque设置的时候, 使`--ringmaster.temp-dir=/tmp/hod-temp-dir.$PBS_JOBID`会让ringmaster在每一次申请时使用不同的临时目录; Troque会在ringmaster启动前展开这个环境变量

量。

- debug: 数值类型, 取值范围是1-4。4会产生最多的log信息。
- log-dir: 日志文件的存放目录。缺省值是<install-location>/logs/。temp-dir变量的限制和注意事项在这里同样使用。
- xrs-port-range: 端口范围, 会在这之中挑选一个可用端口用于运行XML-RPC服务。
- http-port-range: 端口范围, 会在这之中挑选一个可用端口用于运行HTTP服务。
- java-home: 给Hadoop使用的Java的位置。
- syslog-address: syslog守护进程要绑定的地址。格式为host:port。如果配置了这个选项, HOD日志信息会被记录到这个位置的syslog。

3.2. 3.2 hod的配置项

- cluster: 集群的描述性名称。对于Torque, 这个值被指定为集群中所有节点的'Node property'。HOD使用这个值计算可用节点的个数。
- client-params: 逗号分割的hadoop配置参数列表, 其中的每一项都是一个key-value对。在提交节点上会据此产生一个hadoop-site.xml, 用于运行Map/Reduce作业。
- job-feasibility-attr: 正则表达式, 用于指定是否和如何检查作业的可行性 — 资源管理器限制或调度限制。目前是通过torque作业的'comment'属性实现的, 缺省情况下没有开启这个功能。设置了这个配置项后, HOD会使用它来确定哪些种类的限制是启用的, 以及请求超出限制或者累积超出限制时是回收机群还是留在排队状态。torque comment属性可以被某个外部机制周期性地更新。比如, comment属性被hod/support目录下的[checklimits.sh](#)更新, 这样设置job-feasibility-attr的值等于TORQUE_USER_LIMITS_COMMENT_FIELD, "User-limits exceeded. Requested:([0-9]*) Used:([0-9]*) MaxLimit:([0-9]*)"会使HOD产生相应的行为。

3.3. 3.3 resouce_manager的配置项

- queue: 资源管理器中配置的队列名, 作业会被提交到这里。
- batch-home: 个安装目录, 其下的'bin'中有资源管理器的可执行文件。
- env-vars: 逗号分隔的key-value对的列表, 形式是key=value, 它会被传递给运行在计算节点的作业。例如, 如果python没有安装在常规位置, 用户可以通过设置环境变量'HOD_PYTHON_HOME'指定python可执行文件的路径。之后, 在计算节点运行的HOD的进程就可以使用这个变量了。

3.4. 3.4 ringmaster的配置项

- `work-dirs`: 这是一个由逗号分隔的路径列表，这些路径将作为HOD产生和传递给Hadoop，用于存放DFS和Map/Reduce数据的目录的根目录。例如，这是DFS数据块存放的路径。一般情况下，有多少块磁盘就指定多少路径，以确保所有的磁盘都被利用到。`temp-dir`变量的限制和注意事项在这儿同样适用。
- `max-master-failures`: hadoop主守护进程启动前可以失败的次数，超出这个次数后，HOD会让这次集群分配失败。在HOD集群中，有时候由于某些问题，比如机器没安装java，没有安装Hadoop，或者Hadoop版本错误等，会存在一个或几个“坏”节点。当这个配置项被设为正整数时，只有当hadoop master (JobTracker或者NameNode) 在上述的坏节点上，由于上面提到的种种原因启动失败的次数超过设定的值时，RingMaster才会把错误返回给客户端。如果尝试启动的次数没有超过设定值，当下一个HodRing请求运行一个命令时，同一个hadoop master会指定给这个HodRing。这样，即使集群中存在一些坏的节点，HOD也会尽全力使这次分配成功。

3.5. 3.5 gridservice-hdfs的配置项

- `external`: 如果被置为false，HOD必须在通过`allocate`命令分配的节点上自己创建HDFS集群。注意，在这种情况下，如果集群被回收，HDFS集群会停止，所有数据会丢失。如果被置为true，它会尝试链接外部的已配置的HDFS系统。通常，因为在作业运行之前作业的输入需要被放置在HDFS上，并且作业的输出需要持久保留，在生产环境中一个内部的HDFS集群意义不大。
- `host`: 外部配置好的NameNode的主机名。
- `fs_port`: NameNode RPC服务绑定的端口。
- `info_port`: NameNode web UI服务绑定的端口。
- `pkgs`: 安装目录，其下有bin/hadoop可执行文件。可用来使用集群上预先安装的Hadoop版本。
- `server-params`: 一个逗号分割的hadoop配置参数列表，每一项为key-value对形式。这些将用于产生被NameNode和DataNode使用到的hadoop-site.xml文件。
- `final-server-params`: 除会被标记为final外和上面相同。

3.6. 3.6 gridservice-mapred的配置项

- `external`: 如果被置为false，HOD必须在通过`allocate`命令分配的节点上自己创建Map/Reduce集群。如果被置为true，它会尝试链接外部的已配置的Map/Reduce系统。

- `host`: 外部配置好的JobTracker的主机名。
- `tracker_port`: JobTracker RPC服务绑定的端口。
- `info_port`: JobTracker web UI服务绑定的端口。
- `pkgs`: 安装目录, 其下有bin/hadoop可执行文件。
- `server-params`: 一个逗号分割的hadoop配置参数列表, 每一项为key-value对形式。这些将用于产生被JobTracker和TaskTracker使用到的hadoop-site.xml文件。
- `final-server-params`: 除会被标记为final外和上面相同。

3.7. 3.7 hodring的配置项

- `mapred-system-dir-root`: DFS上的目录, HOD会在这个目录下创建子目录并把全路径作为参数'`mapred.system.dir`'的值传递给Hadoop守护进程。全路径的格式为`value-of-this-option/userid/mapredsystem/cluster-id`。注意, 如果HDFS启用了权限, 这里指定的路径下应允许所有用户创建子目录。设置此配置项的值为`/user`会使HOD使用用户的home目录来产生`mapred.system.dir`的值。
- `log-destination-uri`: 一个URL, 能反映一个外部的静态的DFS或者集群节点上的本地文件系统上的路径。当集群被回收时, HOD会把Hadoop日志上传到这个路径。要指定DFS路径, 使用'`hdfs://path`'格式。要指定一个集群节点上的本地文件系统路径, 使用'`file://path`'格式。当HOD回收集群时, 作为HOD的清除过程的一部分, hadoop日志会被删除。要做到持久储这些日志, 你可以使用这个配置项。路径的格式会是`values-of-this-option/userid/hod-logs/cluster-id`。注意, 应该保证所有的用户能在这里指定的目录下创建子目录。把这个值设为`hdfs://user`会使这些日志被转移到用户在DFS上的home目录下。
- `pkgs`: 安装目录, 其下有bin/hadoop可执行文件。如果给`log-destination-uri`指定了一个HDFS URL, HOD上传日志时会用到这个配置项。注意, 当用户使用了和外部静态HDFS不同版本的tarball时, 这个配置项会派上用场。