IsolationRunner

[IsolationRunner](#) is a utility to help debug MapReduce programs.

To use the `IsolationRunner`, first set `keep.failed.task.files` to `true` (also see `keep.task.files.pattern`).

Next, go to the node on which the failed task ran and go to the `TaskTracker`'s local directory and run the `IsolationRunner`:
```
$ cd <local path>/taskTracker/${taskid}/work
$ bin/hadoop org.apache.hadoop.mapred.IsolationRunner ../job.xml
```

`IsolationRunner` will run the failed task in a single jvm, which can be in the debugger, over precisely the same input.

Note that currently IsolationRunner will only re-run map tasks.

Profiling

Profiling is a utility to get a representative (2 or 3) sample of built-in java profiler for a sample of maps and reduces.

User can specify whether the system should collect profiler information for some of the tasks in the job by setting the configuration property `mapred.task.profile`. The value can be set using the api[JobConf.setProfileEnabled(boolean)](#). If the value is set `true`, the task profiling is enabled. The profiler information is stored in the user log directory. By default, profiling is not enabled for the job.

Once user configures that profiling is needed, she/he can use the configuration property `mapred.task.profile.{maps|reduces}` to set the ranges of MapReduce tasks to profile. The value can be set using the api [JobConf.setProfileTaskRange(boolean,String)](#). By default, the specified range is `0-2`.

User can also specify the profiler configuration arguments by setting the configuration property `mapred.task.profile.params`. The value can be specified using the api[JobConf.setProfileParams(String)](#). If the string contains a `%s`, it will be replaced with the name of the profiling output file when the task runs. These parameters are passed to the task child JVM on the command line. The default value for the profiling parameters is `-agentlib:hprof=cpu=samples,heap=sites,force=n,thread=y,verbose=n,file=%s`