现在监控的近300个参数中，涵盖hadoop 2.x版本中的大部分，下面给出参数的类别，和使用的意义。可参考http://hadoop.apache.org/docs/current/hadoop-project-dist/hadoop-common/Metrics.html

##### Ganglia系统参数

##### StartupProgress参数

1.显示了NameNode启动过程的具体统计参数，比如LoadingFsImage, LoadingEdits,SavingCheckpoint,和 SafeMode

2. 意义：能够让用户了解NameNode启动过程的细节和可能发生的错误。

3. V0.1版本忽略，由于异常探测主要针对正在运行（NameNode已经运行）的系统而言，所以此版本暂时忽略，以后可能会加入。

##### FSNamesystem参数

1. 显示了整个文件系统块的操作（删除，复制），锁，文件系统容量等系列参数

2.反应了datanode和文件系统的健康程度和活动状态

##### Namenode参数

1. 显示了整个namenode的工作状态，比如和外界请求的交互，一系列的报告和命令操作

##### JVM参数

1. 包括整套JVM参数的全部信息，如线程运行的情况，垃圾回收的时间，日志erro的个数等

##### RetryCache 参数

用于监控NameNode的失败情况

##### RPC参数

显示了RPC用于认证的信息，已经队列排队时长的信息

##### RPC detailed参数

更加详细的展现了RPC操作的种种细节

##### Queue 参数

一系列的表现系统排队情况的参数，对于评价系统运行任务情况有很大的意义

##### UgiMetrics参数

关于用户和用户组访问hadoop系统的信息，可以用于作为用户行为异常探测

V0.1版本忽略

综上，V0.1制定了如下几个类别的数据，分别进行异常探测，而非将200+的参数一起使用。

1. 系统类别:ganglia系统参数
2. Namenode类别:包括namenode和retrycache参数
3. RPC类别：包括RPC和RPC detailed
4. 文件系统类别：FSNamesystem
5. Queue参数类别：

V0.1会过滤部分状态变量，比如dfs.FSNamesystem.LastWrittenTransactionId。