## Hadoop

1.只删除或修改属于自己的目录的文件或目录（内部约束），虽然在系统上做了限制，但有些文件目录是作为公共目录存在的，误删可能影响其他用户。（系统管理员删除文件时亦应先和文件所有者确认后才可删除）

2.不要在商用环境用rm/rmr操作，因为是没有必要的，除非特殊情况，如手动清理Trash。

如果确有需要删除的文件，建议删除时有其他人检查命令，防止误删。

3.GateWay上新建用户时，需明确用户的组别，现在的HDFS根据业务分类，通过组来实现权限隔离。检查确保同组用户有rx权限，其他用户在HDFS中只有x权限（同步数据需要），在linux系统中，无任何权限。

4.Gate新建的用户不可与集群的hadoop启动用户同名。

5.建立用户时需要确保hive的配置文件中指向的元数据库是正确的，且需要进入对指定数据库赋予建表的权限。

6.非系统管理员禁止登陆集群主机，尤其是namenode所在主机，防止造成不可恢复的数据损坏。

7.由于停电或其他异常导致的主机重启后，需检查HDFS中的数据完整性(hadoop fsck /)

8.系统升级或修改配置后需要重启集群，需要征得系统所有用户的同意，重启后，需要向所有用户确认系统是否可用。修改前的文件需要备份，一旦出现用户反映不可用，考虑回滚配置。

## Hive

1.用户自己不要修改配置文件，里面有资源分配和安全控制的限制，是为了所有用户无冲突的利用平台所做的必要限制。

2.建表时指明存放路径，重要数据用外部表存放，临时数据用内部表存放，存放的目录尽量置于自己的业务目录下。

3.商用环境运行的脚本需在测试环境通过，不要为图方便，直接在商用环境测试。

4.每设置一个客户端，都要确保其Trash可用，防止数据误删后无法恢复。

5.Hive进行查询时，需加上limit条件。