阿里云
对象存储OSS

运维指南

产品版本：V1.1.0

文档版本：20180416

为了无法计算的价值 | 阿里云

法律声明

阿里云提醒您在阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。
2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。
3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。
4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的”现状“、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。
5. 阿里云网站上所有内容，包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含”阿里云”、Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。
6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。

通用约定

表 -1: 格式约定

格式	说明	样例
	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	**禁止**： 重置操作将丢失用户配置数据。
	该类警示信息可能导致系统重大变更甚至故障，或者导致人身伤害等结果。	**警告**： 重启操作将导致业务中断，恢复业务所需时间约10分钟。
	用于警示信息、补充说明等，是用户必须了解的内容。	**说明**： 导出的数据中包含敏感信息，请妥善保存。
	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	**说明**： 您也可以通过按**Ctrl** + **A**选中全部文件。
>	多级菜单递进。	**设置** > **网络** > **设置网络类型**
粗体	表示按键、菜单、页面名称等UI元素。	单击 **确定**。
courier字体	命令。	执行 cd /d C:/windows 命令，进入Windows系统文件夹。
斜体	表示参数、变量。	bae log list --instanceid *Instance_ID*
[]或者[a|b]	表示可选项，至多选择一个。	ipconfig *[-all|-t]*
{}或者{a|b}	表示必选项，至多选择一个。	swich *{stand | slave}*

目录

法律声明.....I
通用约定.....I
1 工具的使用.....1
1.1 tsar 工具常用命令.....1
2 例行维护.....2
2.1 目的.....2
2.2 如何查看集群飞天 Service 是否正常.....2
2.3 前端机运维.....2
2.3.1 如何重启 OSS 前端机服务.....2
2.3.2 LVS操作说明.....3
2.3.3 检查集群 KV 服务.....3
3 安全维护.....5
3.1 建立故障响应机制.....5
3.1.1 明确各类型故障责任人.....5
3.1.2 建立备货机制.....5
3.2 日志审计.....5
4 主要监控项说明及告警处理.....8
4.1 OSS 5xx 监控说明.....8
4.1.1 获取 5xx 的 access_log 的日志信息.....8
4.1.2 根据 5XX access_log 的 request trace 到 fastcgi 日志信息.....8
4.2 关于 OSS 前端机内核参数设置.....8
4.3 配置 Tsar 采集.....9
5 附录：KV 模块返回 OSS 模块码对照表.....10

1 工具的使用

1.1 tsar 工具常用命令

- 查看 tsar 命令详细帮助

  命令：`tsar –help`

- 查看 tsar 最近两天每一分钟的 nginx 运行数据

  命令：`tsar –n 2 –i 1 –nginx`

  -n 表示获取两天的数据。 -i 表示每一分钟输出一条结果记录。

- 查看 tsar 最近两天每一分钟的负载情况运行数据

  命令：`tsar --load -n 2 -i 1`

2 例行维护

2.1 目的

OSS 交付后会执行各类运维活动以保障其稳定性、可用性以及基础软件的可用性。运维活动包括但不限于如下内容。

2.2 如何查看集群飞天 Service 是否正常

查看集群启动了那些 Service，以及 Service 的具体信息。

图 2-1: 查看飞天 Service 是否正常

```
[admin@[redacted]_AG ~]
$/apsara/deploy/rpc_wrapper/rpc.sh al
WorkItemName    | NuwaAddress
sys/oss         | nuwa://[redacted]:10240/sys/oss/ServiceMaster
test/KMAnalyse  | nuwa://[redacted]:10240/test/KMAnalyse/ServiceMaster
sys/ossquota    | nuwa://[redacted]:10240/sys/ossquota/ServiceMaster
sys/pangu       | nuwa://[redacted]:10240/sys/pangu/ServiceMaster
sys/kvengine    | nuwa://[redacted]:10240/sys/kvengine/ServiceMaster
```

2.3 前端机运维

2.3.1 如何重启 OSS 前端机服务

禁止擅自重启服务，如有需要请联系技术支持。

2.3.2 LVS操作说明

图 2-2: Member实时数据

在 VIP 上的七层健康检查状态

- OSS 健康检查方式：HTTP

  检测URL：/systemoperation/checkossstatus

- OCM 健康检查方式：HTTP

  检测 URL：/systemoperation/checkocmstatus

- IMG 健康检查方式：HTTP

  检测 URL：/systemoperation/checkossstatus

2.3.3 检查集群 KV 服务

检查 KV 服务是否正常，其实就是检查 KV APP 是否存在异常 Partition。

检查命令

```
/apsara/deploy/ku
 --interactive=false
--command=checkready
--appname=file_data
/apsara/deploy/ku
--interactive=false
--command=checkready
```

```
--appname=file_meta
```

结果分析

图 2-3: 结果分析

```
[admin@[redacted] AG ~]
$/apsara/deploy/ku --interactive=false --command=checkready --appname=file_data
ERROR Partition Count: 0

KV engine status: YES
[admin@[redacted] AG ~]
$/apsara/deploy/ku --interactive=false --command=checkready  --appname=file_meta
ERROR Partition Count: 0

KV engine status: YES
```

没有异常Partition

3 安全维护

3.1 建立故障响应机制

3.1.1 明确各类型故障责任人

当 OSS 的监控出现报警时由驻场同学根据报警提示去排查问题。若驻场同学解决不了问题时，需要联系阿里云专有云服务团队。在生产当中遇到 OSS 产品相关的问题时需要联系阿里云专有云服务团队。

3.1.2 建立备货机制

在专有云 OSS 中，为了保证向用户提供高可靠的云存储服务，OSS 存储数据占用的磁盘空间不能超过 OSS 总容量的 70%。当超过 OSS 总存储容量的 70% 时，需要补充磁盘空间，使得 OSS 总的可利用空间大于 50%。

3.2 日志审计

在专有云中，当有用户访问 OSS 存储的数据时，会将用户的访问行为等信息记录在 access_log 中，每天都会产生一个新的 access_log.xxx（xxx 表示当天的日期）。若用户请求遇到问题时，可以通过记录的日志去排查。

日志格式

在nginx.conf中，日志格式如下：

```
log_format  main  '$remote_addr $remote_port $remote_user [$time_local] "$request" '
                  '$status $body_bytes_sent $request_time_msec "$http_referer" '
                  '"$http_user_agent" "$host" "$upstream_http_x_oss_request_id" '
                  '"$upstream_http_x_oss_bucket_log_enable" "$upstream_http_x_oss_requester"
'
                  '"$upstream_http_x_oss_operation" "$upstream_http_x_oss_bucket_name" '
                  '"$upstream_http_x_oss_object_name" $upstream_http_x_oss_object_size '
                  '$upstream_http_x_oss_turn_around_time "$upstream_http_x_oss_error_code" '
                  '$request_length "$upstream_http_x_oss_bucket_owner" '
                  '$upstream_http_x_oss_quota_delta_size "$upstream_http_x_oss_sync_request"
'
                  '"$upstream_http_x_oss_location" "$upstream_http_x_oss_bucket_version" '
                  '"$vpcid" "$vpcaddr" "$upstream_http_x_oss_bucket_storage_type" $scheme "$
upstream_http_x_oss_process
```

```
_type"'
```

日志格式内容说明

表 3-1: 日志格式内容说明

格式	说明
$remote_addr	客户端地址
$remote_port	客户端地址端口
$remote_user	记录客户端用户名称
$time_local	通用日志格式下的本地时间
$request	记录请求的 URL 和 HTTP 协议
$status	记录请求状态
$body_bytes_sent	发送给客户端的字节数，不包括响应头的大小
$request_time_msec	请求处理时间，单位是毫秒
$http_referer	记录从哪个页面链接访问过来的
$http_user_agent	记录客户端浏览器相关信息
$host	主机名
$upstream_http_x_oss_request_id	请求 ID
$upstream_http_x_oss_bucket_log_enable	后端机 bucket 日志记录
$upstream_http_x_oss_requester	对 OSS 进行访问的用户
$upstream_http_x_oss_operation	对 OSS 进行的请求类型
$upstream_http_x_oss_bucket_name	请求的 bucket 名称
$upstream_http_x_oss_object_name	请求的 object 名称
$upstream_http_x_oss_object_size	请求的 object 的大小
$upstream_http_x_oss_turn_around_time	请求 OSS 时的处理时间

格式	说明
$upstream_http_x_oss_error_code	OSS 请求错误代码
$request_length	请求的长度（包括请求行，请求头和请求正文）
$upstream_http_x_oss_bucket_owner	所请求 bucket 的所有者
$upstream_http_x_oss_quota_delta_size	进行计量收费数据的大小
$upstream_http_x_oss_sync_request	同步请求到 OSS
$upstream_http_x_oss_location	oss_server 所在的可用区
$upstream_http_x_oss_bucket_version	bucket 的版本
$vpcid	VPC 的 ID
$vpcaddr	VPC 的地址
$upstream_http_x_oss_bucket_storage_type	OSS的bucket存储类型
$scheme	是http还是https请求
$upstream_http_x_oss_process	用于视频截图，方便后续计量计费时使用

日志列表

日志名称	路径及文件名	日志说明
access_log	/apsara/apache/logs/access_log.xxx	访问日志
oss_server.LOG	/apsarapangu/disk1/logs/oss_server.LOG	fastcgi日志

4 主要监控项说明及告警处理

4.1 OSS 5xx 监控说明

4.1.1 获取 5xx 的 access_log 的日志信息

命令

从 access log 最末尾 10000 行获取最近的错误日志，如果第一遍抓取不到，可以增大过滤的日志行数。

```
tail -n 10000 /apsara/apache/logs/access_log.20150713 |grep "HTTP/1.1\" 5"
```

结果分析

图 4-1: 结果分析

4.1.2 根据 5XX access_log 的 request trace 到 fastcgi 日志信息

命令

```
grep $request_id /apsarapangu/disk1/logs/oss_server.LOG
```

如果在当前fastcgi日志中获取不到该access log对应的fastcgi日志信息，需要扩大过滤范围，最后根据fastcgi 具体日志信息，做出相应的处理策略。

4.2 关于 OSS 前端机内核参数设置

OSS 前端机需要重新配置 TCP/IP 协议栈内核参数。

需要配置的参数

修改 */etc/sysctl.conf*

```
sudo vi /etc/sysctl.conf
```

对比和添加其他参数

缺少的参数需要添加，参数 key 需要一致，value 不同的要修改。

```
vm.max_map_count = 8388608
net.ipv4.tcp_rmem = 4096 87380 4194304
net.ipv4.tcp_wmem = 4096 16384 4194304
net.core.wmem_default = 8388608
net.core.rmem_default = 8388608
net.core.rmem_max = 16777216
net.core.wmem_max = 16777216
net.core.netdev_max_backlog = 204800
net.core.somaxconn = 204800
net.ipv4.tcp_max_orphans = 3276800
net.ipv4.tcp_max_syn_backlog = 204800
net.ipv4.tcp_tw_recycle = 0
net.ipv4.tcp_tw_reuse = 1
net.ipv4.tcp_fin_timeout = 15
net.ipv4.ip_local_port_range = 1024 65000
net.ipv4.tcp_syncookies = 0
```

修改完毕后重新加载内存参数

```
sudo /sbin/sysctl -p
```

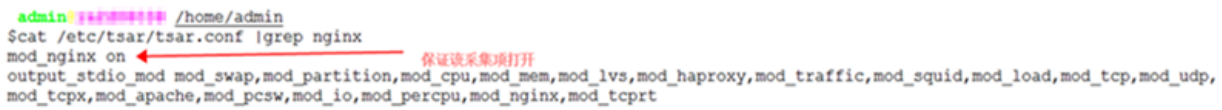
4.3 配置 Tsar 采集

命令

```
cat /etc/tsar/tsar.conf |grep nginx
```

结果分析

图 4-2: 结果分析



5 附录：KV 模块返回 OSS 模块码对照表

KVENGINE_TIMEOUT	-1
KVENGINE_INVALID_PARAMETER	-2
KVENGINE_SERVICE_UNAVAILABLE	-3
KVENGINE_APPLICATION_NOT_FOUND	-4
KVENGINE_KEY_OUT_OF_RANGE	-5
KVENGINE_CLIENT_BUSY	-6
KVENGINE_SERVER_BUSY	-7
KVENGINE_SERVER_EMPTY	-8
KVENGINE_OPERATION_FAIL	-10
KVENGINE_NETWORK_ERROR	-11
KVENGINE_CELL_NOT_EXIST	-21
KVENGINE_ATOMIC_WRITE_FAIL	-22
KVENGINE_UNKNOWN_ATOMIC_WRITE_TYPE	-23
KVENGINE_SCANNER_REACH_END	-24
KVENGINE_INVALID_METADATA	-25
KVENGINE_PARTITION_NOT_FOUND	-200
KVENGINE_OK_HAS_MORE_ENTRIES	-210
KVENGINE_ERROR	-999