

# 阿里云 文件存储NAS

## 常见问题

文档版本：20190612

## 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

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。	 <b>禁止：</b> 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。	 <b>警告：</b> 重启操作将导致业务中断，恢复业务所需时间约10分钟。
	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	 <b>说明：</b> 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
<b>粗体</b>	表示按键、菜单、页面名称等UI元素。	单击 <b>确定</b> 。
<code>courier</code> 字体	命令。	执行 <code>cd /d C:/windows</code> 命令，进入Windows系统文件夹。
<code>##</code>	表示参数、变量。	<code>bae log list --instanceid</code> <code>Instance_ID</code>
<code>[ ]</code> 或者 <code>[a b]</code>	表示可选项，至多选择一个。	<code>ipconfig [-all -t]</code>
<code>{ }</code> 或者 <code>{a b}</code>	表示必选项，至多选择一个。	<code>swich {stand   slave}</code>

# 目录

---

法律声明.....	I
通用约定.....	I
1 NFS 客户端已知问题.....	1
2 规避 NFS 连接卡住的风险.....	3
3 一般性问题.....	6
3.1 NAS 地域与支持的存储类型和协议类型.....	6
3.2 创建经典网络挂载点时为什么需要RAM授权? .....	9
3.3 文件存储怎么收费? .....	9
3.4 文件存储NAS是什么、适合什么应用场景? .....	9
3.5 文件存储NAS都有哪些相关的概念? .....	10
3.6 每个账户可以创建多少个文件系统、文件系统有什么限制? .....	10
3.7 文件存储NAS支持哪些访问协议? .....	10
3.8 文件存储的性能怎样? .....	10
3.9 挂载点是什么概念、有什么作用? .....	10
3.10 权限组是什么概念、有什么作用? .....	11
3.11 如何避免NFS 4.0监听端口被误认为木马? .....	11
3.12 如何避免多进程/多客户端并发写同一日志文件可能出现的异常? .....	12
3.13 为什么 Windows server SMB 协议 I/O 性能上不去? .....	14
3.14 如何提升IIS访问NAS性能.....	15
3.15 文件存储NAS是否支持inotify? .....	15
3.16 linux上NFS性能只有几MB速度.....	16
3.17 不同共享客户端间的数据不同步.....	16
4 SMB FAQ.....	18
4.1 SMB 基本操作常见问题.....	18
4.2 SMB 挂载常见问题.....	18
4.3 SMB 性能问题.....	21
4.4 SMB 挂载失败的原因分析.....	22

# 1 NFS 客户端已知问题

NFS 客户端为内核的一部分，由于部分内核存在一些缺陷，会影响 NFS 的正常使用。为了获得更好的 NFS 稳定性体验，请使用阿里云推荐的内核版本。

## 推荐 Linux 系统版本

推荐使用阿里云官方内核镜像，选择经过阿里云严格测试的内核版本，确保稳定性，如下所示。

- CentOS 6.9 及以上
- Redhat 6.9 及以上
- Ubuntu 14.04 / 16.04 / 18.04
- Debian 8 及以上
- SUSE 11 及以上
- OpenSUSE 42.3 及以上
- AliyunLinux 17 及以上

## 网络抖动导致文件系统无响应

当系统内核版本为以下两种情况时，NFS 客户端遇到网络抖动、服务端故障切换等情况，可能会触发 TCP 连接异常，导致 NFS 连接中断。



### 说明：

您可以通过 `uname-a` 命令查看内核版本。

- 内核版本低于 2.6.32-696.16.1（不包括 2.6.32-696.16.1 版本）。
- 内核版本高于 2.6.32-696.16.1（包括 2.6.32-696.16.1 版本），但是 NFS 挂载没有使用 `noresvport` 选项。

当发生连接异常后，需要在 ECS 实例上重新执行挂载操作。挂载操作请参见在 [Linux 系统中挂载 NFS 文件系统](#)，使用推荐的参数进行挂载。

使用 `noresvport` 选项进行挂载后，请执行 `ss -ant | grep 2049` 命令检查端口号，如下图框选部分所示，端口号大于 1024，说明挂载生效。反之，请先卸载所有已挂载的文件系统（执行 `mount | grep nas` 命令确保已不存在挂载），再进行重新挂载。

```
[root@ ~]# ss -ant | grep 2049
ESTAB      0      0      192.168.0.163:30024      192.168.0.168:2049
```



### 注意：

请避开使用 2.6.32-696 ~ 2.6.32-696.10.1 版本的内核，该版本发生文件系统无响应的概率高于其他内核版本，详情请参见[RHEL6.9: NFSv4 TCP transport stuck in FIN\\_WAIT\\_2 forever](#)。

#### 内核缺陷导致文件系统无响应

当系统的内核版本为以下几个版本时，NFS 服务端故障转移，可能造成 NFS 客户端的打开/读/写操作出现死锁情况，从而导致文件系统持续无响应。

- Redhat 6/CentOS 6 2.6.32-696.3.1.el6。
- Redhat 7/CentOS 7 3.10.0-229.11.1.el7 之前的所有内核版本。
- Ubuntu 15.10 Linux 4.2.0-18-generic。

当操作无响应时，请重启 ECS 实例，详情请参见[RHEL7: NFSv4 client loops with WRITE / NFS4ERR\\_STALE\\_STATEID - if NFS server restarts multiple times within the grace period](#)。

#### 不支持 chown 命令和系统调用

系统的内核版本为 2.6.32 时，不支持 NFS 客户端执行 chown 命令和系统调用。

#### ls 操作无法终止

当系统的内核版本为 2.6.32-696.1.1.el6 及之前版本时，在系统中执行 ls 操作的同时还在进行添加、删除文件/子目录操作，将导致 ls 操作永远无法终止。

请升级内核版本，避免此问题。

## 2 规避 NFS 连接卡住的风险

在网络抖动或者服务端升级时，可能发生 TCP 连接异常，在某些情况下会造成 NFS 文件系统卡住，需要在卡住之后，手动重新挂载才能恢复。

### 规避方法

为了避免以上风险，让异常的 TCP 连接自动恢复，请确保挂载 NFS 文件系统的 ECS 实例同时满足以下两个条件：

1. 拥有 2.6.32-696.16.1 及以上版本的 Linux 内核。
2. 使用 `noresvport` 参数挂载 NFS 文件系统并且确认生效。

### 确认内核版本

您可以在 ECS 实例上执行 `uname -a` 命令查看当前内核版本。

请务必确保内核版本高于 2.6.32-696.16.1，否则必须尽快升级内核。继续使用低版本内核会对您的业务带来巨大的风险，在升级过程中可能会造成 NFS 文件系统卡住，在卡住之后，必须手动重新挂载才能恢复业务。如果在卡住之后没有及时手动重新挂载，会造成无限长时间的业务停滞。

要让异常的 TCP 连接自动恢复，规避 NFS 文件系统卡住的风险，请在谨慎评估应用与内核的兼容性之后，逐台升级内核。假如评估发现无法升级内核，请尽快通过钉钉用户群[联系我们](#)。



说明：

NFS 客户端是 Linux 内核的一部分，请避免使用有缺陷的内核版本，详细信息请参考 [NFS 客户端已知问题](#)。

内核升级后，请重启 ECS 实例使新内核生效。

### 使用 noresvport 参数

确认 ECS 实例内核版本符合要求后，请执行 `ss -nt | grep 2049 | awk '{print $4}' | awk -F ':' '{print $2}'` 命令检查挂载 NFS 的端口号是否大于 1024，确认 `noresvport` 参数是否已经生效。

- 如果没有输出结果，或者所有输出结果都大于 1024，则说明升级过程不会对 TCP 连接造成影响。

- 如果有任何结果小于1024，请按照下面推荐的方法，将NFS文件系统全部卸载，再重新挂载。

此处挂载点地址以 `foo.aliyuncs.com` 为例，挂载的本地目录路径以 `/mnt` 为例。

1. 执行以下命令记录已经挂载的所有 NFS 文件系统，找到一个挂载的本地目录（例如 `/mnt`）。

```
mount | grep aliyun > /tmp/mount_aliyun && cat /tmp/mount_aliyun
```

2. 停止对 `/mnt` 中的任何文件进行操作的所有应用。
3. 卸载文件系统。

```
sudo umount /mnt
```

如果卸载失败，请确认所有相关应用已停止后，执行 `sudo umount -lf /mnt`。

4. 检查是否所有的 NFS 文件系统卸载完成。

```
mount | grep aliyun
```

- 如果返回为空，则表示所有的 NFS 文件系统卸载完成。
- 如果返回不为空，请重复 [步骤 ii](#) 和 [步骤 iii](#)，直到所有的 NFS 文件系统卸载完成。

5. 挂载文件系统。

- 如果希望使用 NFSv4 挂载，请执行以下命令。

```
sudo mount -t nfs -o vers=4.0,noresvport foo.aliyuncs.com:/ /mnt || sudo mount -t nfs4 -o noresvport foo.aliyuncs.com:/ /mnt
```

- 如果希望使用 NFSv3 挂载，请执行以下命令。

```
sudo mount -t nfs -o vers=3,noexec,proto=tcp,noresvport foo.aliyuncs.com:/ /mnt
```

6. 执行以下命令查找其他需要挂载的文件系统，并执行 [步骤 v](#) 完成挂载。

```
cat /tmp/mount_aliyun
```

7. 当所有 NFS 文件系统重新挂载成功后，请再次执行检查端口号的命令，确保所有端口号大于 1024。

```
ss -nt | grep 2049 | awk '{print $4}' | awk -F ':' '{print $2}'
```

如果端口号仍然有小于 1024 的情况，请执行 `uname -a` 确认内核版本高于

2.6.32-696.16.1。如果按照上述步骤升级内核后，显示的内核版本仍然低于要求，请确认内核升级后是否已重启 ECS 实例。

## 升级注意事项

完成所有上述操作后，升级过程中的 TCP 连接异常可以全部自动恢复。虽然无须人工干预，但仍然建议您在升级期间关注业务运行状况。

升级期间我们的研发团队会密切关注 TCP 连接恢复情况，如果需要您手动重新挂载恢复业务，我们会在第一时间联系您，请保持阿里云账号的手机号码畅通。

## 3 一般性问题

### 3.1 NAS 地域与支持的存储类型和协议类型

NAS 所在地域不同，支持功能会有所差异。

地域是指物理的数据中心，资源创建后地域不可更改。

NAS 所在地域与支持的存储类型、协议类型以及可用区对应情况如下。

· 中国大陆：

地域名称	所在城市	支持的存储类型、协议类型以及可用区
华北 1	青岛	容量型：NFS、SMB 可用区：C
华北 2	北京	- 性能型：NFS、SMB 可用区：C、E - 极速型 可用区：A、B、C、D、E、F
华北 3	张家口	- 性能型：NFS、SMB 可用区：B - 极速型 可用区：A、B
华北 5	呼和浩特	- 容量型：NFS、SMB 可用区：A - 性能型：NFS、SMB 可用区：A

地域名称	所在城市	支持的存储类型、协议类型以及可用区
华东 1	杭州	<ul style="list-style-type: none"> <li>- 容量型：NFS、SMB</li> <li>可用区：B、G</li> <li>- 性能型：NFS、SMB</li> <li>可用区：G、F</li> <li>- 极速型</li> <li>可用区：B、E、F、H、I</li> </ul>
华东 2	上海	<ul style="list-style-type: none"> <li>- 容量型：NFS、SMB</li> <li>可用区：B</li> <li>- 性能型：NFS</li> <li>可用区：B</li> <li>- 极速型</li> <li>可用区：A、B、C、D、E、F</li> </ul>
华南 1	深圳	<ul style="list-style-type: none"> <li>- 容量型：NFS、SMB</li> <li>可用区：A、D</li> <li>- 性能型：NFS、SMB</li> <li>可用区：A、B、D</li> <li>- 极速型</li> <li>可用区：A、B、C、D、E</li> </ul>

· 其他国家和地区：

地域名称	所在城市	支持的存储类型、协议类型以及可用区
香港	香港	<ul style="list-style-type: none"> <li>- 容量型：NFS、SMB</li> <li>可用区：B</li> <li>- 性能型：NFS、SMB</li> <li>可用区：B</li> </ul>
亚太东南 1	新加坡	<ul style="list-style-type: none"> <li>- 容量型：NFS、SMB</li> <li>可用区：A</li> <li>- 性能型：NFS、SMB</li> <li>可用区：A</li> </ul>

地域名称	所在城市	支持的存储类型、协议类型以及可用区
亚太东南 2	悉尼	- 容量型：NFS、SMB 可用区：A - 性能型：NFS、SMB 可用区：A
亚太东南 3	吉隆坡	容量型：NFS、SMB 可用区：A
亚太东南 5	雅加达	容量型：NFS、SMB 可用区：A
亚太东北 1	东京	容量型：NFS、SMB 可用区：A、B
亚太南部 1	孟买	- 容量型：NFS、SMB 可用区：A - 性能型：NFS、SMB 可用区：A
欧洲中部 1	法兰克福	- 容量型：NFS、SMB 可用区：A - 性能型：NFS、SMB 可用区：A
英国	伦敦	- 容量型：NFS、SMB 可用区：A - 性能型：NFS、SMB 可用区：A
美国西部 1	硅谷	- 容量型：NFS、SMB 可用区：B - 性能型：NFS、SMB 可用区：B

地域名称	所在城市	支持的存储类型、协议类型以及可用区
美国东部 1	弗吉尼亚	<ul style="list-style-type: none"> <li>- 容量型：NFS、SMB</li> <li>可用区：A</li> <li>- 性能型：NFS、SMB</li> <li>可用区：A</li> </ul>

### 3.2 创建经典网络挂载点时为什么需要RAM授权？

与专有网络环境（VPC）不同的是，经典网络环境缺少了网络层面的隔离，为了确保您的 NAS 文件系统数据安全，NAS 需要对挂载和访问经典网络挂载点的 ECS 实例进行验证，只允许属于您自己的 ECS 实例挂载和访问文件系统，即文件系统实例的主账号与 ECS 实例的主账号相同。因此为了进行上述的验证，NAS 需要您通过 RAM 授权阿里云获取您的 ECS 实例列表。



说明：

- 通过 RAM 授权后，NAS 系统仅有权限调用您的 DescribeInstances 接口，无法调用其他任何接口；NAS 系统通过 DescribeInstances 接口获取的 ECS 实例列表不会做任何形式的记录，仅用于权限验证。
- 通过 RAM 授权后，请不要随意删除或编辑 RAM 中的 AliyunNASDefaultRole 角色，否则可能遇到无法挂载或文件系统操作异常。

### 3.3 文件存储怎么收费？

文件存储 NAS 的基础收费模式为按实际使用量付费，同时提供预付费存储包。

具体收费细则请参阅[计量项和计费说明](#)。

### 3.4 文件存储NAS是什么、适合什么应用场景？

文件存储是阿里云推出的全新存储产品服务，是提供标准的 NAS 文件存储接口，无限容量，单一命名空间，共享，安全，高可用，高可靠，高性能的分布式文件存储服务。传统的应用程序及工作负载不做任何修改即可访问文件存储，支持 ECS 实例、E-HPC 和容器服务等多个计算节点同时读写文件系统。

阿里云 ECS 云服务器使用文件存储的典型应用场景如下：

- 使用 SLB + 多台 ECS（如 web 服务器）部署业务，多台 ECS 需要访问同一个存储空间，以便多台 ECS 能共享数据。

- 日志共享：多台 ECS 应用，需要将日志写到同一个存储空间，以方便做集中的日志数据处理与分析。
- 企业办公文件共享：企业有公共的文件需要共享给多组业务使用，需要集中的共享存储来存放数据。

### 3.5 文件存储NAS都有哪些相关的概念？

文件存储提供无限扩展的文件系统来给云服务器存取数据，主要包括以下概念：

- 文件系统：文件系统是文件存储的实例，将文件系统挂载（mount）到 ECS 云服务器、E-HPC 或容器服务后，可以像使用本地文件系统一样使用文件存储。
- 挂载点：挂载点是计算节点访问文件存储的入口，定义了什么类型网络的计算节点、采用怎样的权限来访问文件存储。
- 权限组：定义访问文件存储的权限信息，包括授权 IP 地址、读写权限、用户权限等。

### 3.6 每个账户可以创建多少个文件系统、文件系统有什么限制？

每个账户最多可以创建 10 个文件系统，每个文件系统容量上限为 1PB，每个文件系统最多可以支持 10 亿个文件。

### 3.7 文件存储NAS支持哪些访问协议？

文件存储 NAS 支持以下协议：

1. NFS v3.0和 NFS v4.0。
2. SMB 2.1 及以上，对应支持 Windows 7/Windows Server 2008 R2 及以上的各 Windows 版本，不支持 Windows Vista/ Windows Server 2008 及以下的各 Windows 版本。

### 3.8 文件存储的性能怎样？

单个文件系统的吞吐性能上限（峰值）与文件系统的当前使用容量线性相关，即存储量越大，吞吐性能上限（峰值）越高。

具体的容量与性能关系请参见 [定价页面](#) 中“吞吐能力与存储空间的关系”部分。

### 3.9 挂载点是什么概念、有什么作用？

挂载点是计算节点（ECS 实例、E-HPC 或容器服务）访问文件存储的入口。挂载点定义了什么类型网络的计算节点、采用怎样的权限来访问文件存储。同一个挂载点可以被多个计算节点同时挂载，共享访问。

### 3.10 权限组是什么概念、有什么作用？

权限组定义了访问文件存储的权限信息，包括授权 IP 地址、读写权限、用户权限等。

### 3.11 如何避免NFS 4.0监听端口被误认为木马？

#### 问题现象

使用NFS 4.0挂载NAS文件系统后，有一个0.0.0.0的随机端口被监听，并且无法通过netstat定位该监听端口的进程。

由于端口不固定，并且无法确定监听的程序，很容易被误判为受到木马攻击。

```
[root@iz8v-111111111111:~]# netstat -tanpl | grep tcp | g
tcp        0      0 0.0.0.0:111          0.0.0.0:*
tcp        0      0 0.0.0.0:22         0.0.0.0:*
[root@iz8v-111111111111:~]# mount -t nfs -o vers=4.0 9156
[root@iz8v-111111111111:~]# netstat -tanpl | grep tcp | g
tcp        0      0 0.0.0.0:38862       0.0.0.0:*
tcp        0      0 0.0.0.0:111        0.0.0.0:*
tcp        0      0 0.0.0.0:22         0.0.0.0:*
tcp6       0      0 :::41841           :::*
```

#### 问题原因

此随机端口是NFS 4.0为了支持callback而监听的。因为内核参数fs.nfs.nfs\_callback\_tcpport默认是0，所以NFS 4.0 client会随机挑选一个端口进行监听，而随机端口本身并不会带来安全风险。

不过为了更方便地管理端口，用户可以选择使用以下方法固定该callback端口。

#### 解决方案

在挂载文件系统之前，用户可以通过配置参数fs.nfs.nfs\_callback\_tcpport到一个非零的确定值，以固定该端口。

```
sudo sysctl fs.nfs.nfs_callback_tcpport=<port>
```

在以下示例中，在用户将fs.nfs.nfs\_callback\_tcpport手动配置到端口45450，并使用NFS 4.0挂载文件系统之后，netstat显示被监听的端口就是手动配置的45450。

(请注意以下实例中使用的是root用户，所以不需要使用sudo执行sysctl命令。)

```
[root@iz8c1m146t32a9cux1t0z /]# netstat -tanpl | grep tcp | gr
tcp        0      0 0.0.0.0:111          0.0.0.0:*
tcp        0      0 0.0.0.0:22         0.0.0.0:*
[root@iz8c1m146t32a9cux1t0z /]# sysctl fs.nfs.nfs_callback_tcp
fs.nfs.nfs_callback_tcpport = 45450
[root@iz8c1m146t32a9cux1t0z /]# mount -t nfs -o vers=4.0 9150
[root@iz8c1m146t32a9cux1t0z /]# netstat -tanpl | grep tcp | gr
tcp        0      0 0.0.0.0:45450       0.0.0.0:*
tcp        0      0 0.0.0.0:111        0.0.0.0:*
tcp        0      0 0.0.0.0:22        0.0.0.0:*
tcp6       0      0 :::45450           :::*
```

### 3.12 如何避免多进程/多客户端并发写同一日志文件可能出现的异常?

#### 问题现象

文件存储 NAS 为多客户端提供了统一名字空间的文件共享读写能力，但在多进程/多客户端并发写同一个文件的场景中（典型的例如并发写同一个日志文件），各进程分别维护了独立的文件描述符及写入位置等上下文信息，而 NFS 协议本身并没有提供 Atomic Append 语义的支持，因此可能会出现写覆盖、交叉、串行等异常现象，推荐的解决方案有两种：

- （推荐）不同进程/客户端写入同一文件系统的不同文件中，后续分析处理时再进行归并，这个方案能够很好地解决并发写入导致的问题，同时无需使用文件锁，不会对性能造成影响。
- 对于并发追加写同一个文件（如日志）的场景，可以使用文件锁 + seek 机制来保证写入的原子性和一致性。但是文件锁 + seek 是一个比较耗时的操作，可能会对性能产生显著的影响。下面将对这种方式进行一个简单的介绍，以供参考。

#### flock + seek 使用方法

由于 NFS 协议本身没有提供对 Atomic Append 语义的支持，因此当并发写入同一文件末尾（如日志）时，很可能会出现相互覆盖的情况。在 Linux 中，通过使用 flock + seek 的方式，可以在 NFS 文件系统上做到模拟 Atomic Append，对并发追加写入同一文件提供保护和支持。

使用方式如下：

1. 调用 `fd = open(filename, O_WRONLY | O_APPEND | O_DIRECT)` 以追加写的方式打开文件，并且指定 `O_DIRECT`（直写，不通过 Page Cache），获得文件描述符 `fd`。
2. 调用 `flock(fd, LOCK_EX|LOCK_NB)` 尝试获取文件锁，如果获取失败（如锁已被占用）则会返回错误，此时可以继续重试或进行错误处理。
3. 文件锁获取成功后，调用 `lseek(fd, 0, SEEK_END)` 将 `fd` 当前的写入偏移定位到文件末尾。

4. 执行正常的 write 操作，此时写入位置应该是文件的末尾，并且由于有文件锁的保护，不会出现并发写入相互覆盖的问题。
5. 写操作执行完成后，调用 flock(fd, LOCK\_UN) 释放文件锁。

下面是一个简单的 C 语言示例程序，仅供参考：

```
#define _GNU_SOURCE
#include<stdlib.h>
#include<stdio.h>
#include<fcntl.h>
#include<string.h>
#include<unistd.h>
#include<sys/file.h>
#include<time.h>

const char *OUTPUT_FILE = "/mnt/blog";
int WRITE_COUNT = 50000;

int do_lock(int fd)
{
    int ret = -1;
    while (1)
    {
        ret = flock(fd, LOCK_EX | LOCK_NB);
        if (ret == 0)
        {
            break;
        }
        usleep((rand() % 10) * 1000);
    }
    return ret;
}

int do_unlock(int fd)
{
    return flock(fd, LOCK_UN);
}

int main()
{
    int fd = open(OUTPUT_FILE, O_WRONLY | O_APPEND | O_DIRECT);
    if (fd < 0)
    {
        printf("Error Open\n");
        exit(-1);
    }
    for (int i = 0; i < WRITE_COUNT; ++i)
    {
        char *buf = "one line\n";

        /* Lock file */
        int ret = do_lock(fd);
        if (ret != 0)
        {
            printf("Lock Error\n");
            exit(-1);
        }

        /* Seek to the end */
        ret = lseek(fd, 0, SEEK_END);
        if (ret < 0)
```

```
    {
        printf("Seek Error\n");
        exit(-1);
    }

    /* Write to file */
    int n = write(fd, buf, strlen(buf));
    if (n <= 0)
    {
        printf("Write Error\n");
        exit(-1);
    }

    /* Unlock file */
    ret = do_unlock(fd);
    if (ret != 0)
    {
        printf("UnLock Error\n");
        exit(-1);
    }
}
return 0;
}
```

更详细的 flock() 使用方式可以参考以下链接：<http://www.hackinglinuxexposed.com/articles/20030616.html>。



#### 说明:

在NAS文件系统中使用 flock() 需要您的Linux内核版本在2.6.12及以上，如果您的内核版本较低，请使用 fcntl() 调用。

## 3.13 为什么 Windows server SMB 协议 I/O 性能上不去？

### 现象

Windows SMB 客户端默认不打开 `large mtu` 选项，因此影响 I/O 性能提升。

### 解决方案

您可以去修改注册表项来打开 `large mtu` 特性。

注册表路径：`HKLM\System\CurrentControlSet\Services\LanmanWorkstation\Parameters`。

在该路径下，增加DWORD类型的键，命名为`DisableLargeMtu`，设置其值为0，重新启动后才能生效。

## 3.14 如何提升IIS访问NAS性能

### 问题现象

对于 IIS 使用 NAS share 的方式，访问一个文件时，IIS 后台会有很多次访问 NAS 操作。不同于访问本地文件系统，访问 NAS 每次至少要有一次网络交互，因此虽然每次访问的时间不长，但是多次的叠加可能会造成客户端总时间比较长。

### 解决方案

改进方式请参见[SMB2 Client Redirector Caches Explained](#)。

您可以将文章中提到的三个注册表项都调大，例如调为600或以上。

注册表项所在路径为 `HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\services\LanmanWorkstation\Parameters`。

注册表名称分别为：

- FileInfoCacheLifetime
- FileNotFoundCacheLifetime
- DirectoryCacheLifetime



说明：

- 三个注册表项都不存在：
  1. 先确认使用的是 SMB 而不是 NFS。
  2. 然后确认客户使用的 windows 版本支持这三个注册表项。如果 windows 版本支持而注册表项不存在，手动创建一下。详情请参见[Performance tuning for file servers](#)
- 建议把 IIS 频繁访问的 js/css 等网页程序相关的内容放在本地。

## 3.15 文件存储NAS是否支持inotify?

使用 inotifywait 配合 rsync 是一种常见的准实时数据备份/同步方案，但在NAS文件系统上 inotifywait 无法正常工作，这是由于 inotify 本身的实现机制导致的。

### inotifywait 原理简介

inotifywait 是Linux 内核子模块 inotify 的用户态接口实现，inotify 实现在 VFS 层，当文件操作到达 VFS 层时，inotify 模块会将操作类型（创建/删除/属性改变等）和操作对象（文件名）反馈给用户态，用户态的 inotifywait 即可将本次操作信息输出给用户。

## NAS 上使用 inotifywait 存在的问题

由于 inotify 是在 kernel 的 VFS 层实现的，因此在 NFS 文件系统中，远程客户端对 NFS 文件系统的操作无法被本地 kernel 所感知，inotify 也就无法感知远程客户端的文件修改操作。

因此，在 NAS 上使用 inotifywait 会出现以下现象：

- 在客户端 A 和 B 同时挂载一个 NAS 文件系统，在客户端 A 启动 inotifywait 监听挂载目录。
- 在客户端 A 上操作挂载目录中的文件，可以被 inotifywait 感知。
- 在客户端 B 上操作挂载目录中的文件，inotifywait 无法感知任何文件操作。

## 替代方案

一个可行的替代方案是使用 [FAM](#)，FAM 是一个用来监听文件或目录的库，全部在用户态实现，原理是在后台运行一个 daemon，定时扫描目录，获取文件变化情况。

但是使用 FAM 存在以下几个缺陷：

- 需要自己写程序调用 FAM 接口实现功能。
- 对于文件数目很多的场景，使用该方案性能会较差，可能消耗大量资源，无法做到很好的实时性。

## 3.16 linux上NFS性能只有几MB速度

Linux nfs客户端对于同时发起的NFS请求数量进行了控制，若该参数配置较小会导致IO性能较差，请查看该参数：`cat /proc/sys/sunrpc/tcp_slot_table_entries`

默认编译的内核该参数最大值为256，可适当提高该参数的值来取得较好的性能，请以root身份执行以下命令：

```
echo "options sunrpc tcp_slot_table_entries=128" >> /etc/modprobe.d/sunrpc.conf
echo "options sunrpc tcp_max_slot_table_entries=128" >> /etc/modprobe.d/sunrpc.conf
sysctl -w sunrpc.tcp_slot_table_entries=128
```

修改完成后，您需要重新挂载文件系统或重启机器。

## 3.17 不同共享客户端间的数据不同步

### 问题现象

当您用台ECS挂载同一个NFS文件系统，在ECS-A上append写文件，在ECS-B用tail -f观察文件内容的变化。在ECS-A写完之后，在ECS-B看到文件内容变化会有10-30秒的延时。然而相同的场景下，如果直接在ECS-B上打开文件（比如vi）却是立即可以看到更新的内容的。

## 原因

该现象与mount的选项以及tail -f实现相关。

用户使用的mount命令为：`mount -t nfs4 /mnt/`

对于在ECS-B上以这一方式NFS mount的文件系统，默认情况下kernel对文件和目录的属性维护了一份metadata缓存，文件和目录属性（包括许可权、大小、和时间戳记）缓存的目的是减少NFSPROC\_GETATTR 远程过程调用（RPC）的需求。

tail -f的实现是sleep+fstat来观察文件属性（主要是文件大小）的变化，然后读入文件并输出。可见，tail -f是否能实时输出文件内容主要取决于fstat的结果，由于metadata cache的存在，fstat轮询到的并不是实时的文件属性。因此，即使在NFS服务器端文件已经更新了，但tail -f却没法知道文件已经改动了，于是输出就会出现延时。

## 解决方法

使用mount的noac选项可以disable文件和目录属性的缓存。

```
mount -t nfs4 -o noac /mnt/
```

## 4 SMB FAQ

---

### 4.1 SMB 基本操作常见问题

为什么用 net use 会看到 mount 点处于已断开状态

如果15分钟内在文件系统上没有操作，连接会自动断开。只要有操作，连接自动重连。

CIFS/SMB文件系统的容量和性能如何？

现在 SMB 文件系统是部署在容量型 NAS 集群上的，所以单文件系统的容量上限和带宽上限与容量型 NAS 是一样。其他属性，比如对单一命名空间，VPC、经典网络的支持等，都和 NFS 是一样的。

详情请参见[文件 NAS](#)。

SMB 文件系统支持的协议和操作系统

SMB 文件系统支持的协议和 ECS OS的详情请参见[SMB 协议限制说明](#)。

SMB 文件系统不支持的特性请参见[SMB 文件系统不支持特性](#)。

访问 SMB 文件系统的一些限制

同 NFS 文件系统一样，不能跨 region 访问，不能直接通过公网访问，需要专线连到 VPC。

如果需要在有文件系统挂载点的VPC之外访问，请参见：

- [本地/DC VPN网络访问阿里云文件存储](#)
- [本地/DC NAT网络访问阿里云文件存储](#)
- [跨 VPC 挂载阿里云文件存储 NAS](#)
- [跨账户挂载阿里云文件存储NAS](#)

### 4.2 SMB 挂载常见问题

Windows 2016 以后的版本不支持挂载 NAS SMB

- 问题现象

命令及错误提示：

```
C:\Users\Administrator>net use z: \\xxxxx-xxxx.xxxxx.nas.aliyuncs.com\myshare
System error 1272 has occurred.
```

You can't access this shared folder because your organization's security policies block unauthenticated guest access. These policies help protect your PC from unsafe or malicious devices on the network.

· 解决方案:

造成以上问题的原因是 windows server 的这个版本的安全特性默认不支持guest 用户访问远程共享目录。

解决方法如下:

- 修改注册表:

```
[HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\LanmanWorkstation\Parameters]
"AllowInsecureGuestAuth"=dword:0
```

修改为:

```
[HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\LanmanWorkstation\Parameters]
"AllowInsecureGuestAuth"=dword:1
```

- 切换到 Powershell 键入以下命令:

```
New-ItemProperty -Path $registryPath -Name $name -Value $value -PropertyType DWORD -Force
```

具体解决方案请参见[Guest access in SMB2 disabled by default in Windows 10, Windows Server 2016 version 1709, and Windows Server 2019](#)。

## Linux如何挂载SMB

NAS 目前并不支持 Linux 挂载 SMB。

如果要挂载, 一般 linux 系统挂载命令如下:

```
mount -t cifs [MOUNT POINT] [LOCAL DIRECTORY] -o username=guest,vers=3.0
```



说明:

Linux 版本不同挂载命令会有所区别。

较新的 Linux 系统使用 `-o username=guest` 选项需要输入密码, 可以改用以下命令:

```
mount -t cifs [MOUNT POINT] [LOCAL DIRECTORY] -o guest,vers=3.0
```

## SMB 挂载有时候连接不上

问题现象:

有时候用户会混用 NFS 和 SMB 文件系统，导致第一次通过 net use 挂载 NFS 文件系统连接失败后，挂载正确的 SMB 文件系统也会出现问题。

解决方案：

检查确保挂载正确的文件系统后，暂时停止挂载，5分钟后再次挂载。如果还失败，请发工单。

Administrator 能看见挂载的 SMB 目录，其他用户看不到

造成以上原因是由于 windows 下的用户账户隔离机制引起的。

要实现多用户的共享，需要创建一个目录链接。比如 C 盘下创建一个 myshare，命令如下：

```
mklink /D C:\myshare \\xxxxxxx-xxxx.cn-beijing.nas.aliyuncs.com\myshare\
```

Windows Server 2016 的 IIS 无法加载 SMB volume 文件的问题

如果遇到 Windows Server 2016 的 IIS 无法加载 SMB 文件卷文件的问题，请参见云栖文章：[安装和配置 AD 域](#)。

IIS 服务在阿里云 NAS 上的最佳实践

如果遇到 Windows 2016 挂载 SMB 失败，HTTP错误500.19，错误码0x8007003a。请参见云栖文章：[IIS 服务在阿里云 NAS 上的最佳实践](#)。

是否支持 NFS 和SMB 同时挂载一个文件系统

不能以 NFS 和 SMB 同时挂载同一个文件系统。

我们建议不要使用 Linux 作为客户端访问 SMB，因为存在一些操作上的问题。比如支持的字符集、文件名的长度（Windows 支持255宽字符，Linux 支持255 UTF8 字节）等等。

但用户如果确实需要的话，可以在支持 SMB2 及以上的 kernel 上挂载。

挂载命令：`mount -t cifs -o vers=2.0 \\挂载点\myshare /mnt` 或者 `mount -t cifs -o vers=2.0 //挂载点/myshare /mnt`。如果弹出需要密码，直接回车就可以。

确认 kernel 是否支持 CIFS 挂载：在/boot下，检查CONFIG\_CIFS的配置，y 或 m 表示支持。

如图所示：

```
$grep -i cifs /boot/config-2.6.18-274.alios5.1  
CONFIG_CIFS=m
```



说明：

- 执行以上命令前需要安装 `cifs-utils`。以 CentOS 系统为例，安装命令如下：

```
yum install samba-client samba-common cifs-utils
```

- 如果遇到一些 Linux 系统对 `cifs` 支持方面的问题，建议升级 `kernel` 到 3.10.0-514 及以上版本。

## 4.3 SMB 性能问题

### 开始 IO 操作前的延迟问题

当通过挂载点地址直接访问 SMB 服务器，在开始 IO 操作之前会有几分钟的等待时间。

应该通过什么操作减少访问时的等待时间？

### 解决方案

产生等待时间主要是 NFS Client 或 WebClient 造成的。

- 是否安装 NFS Client，如果用不到 NFS 服务，请删除。
- 查看注册表配置项。注册表路径：`HKEY_LOCAL_MACHINE->System->CurrentControlSet->Control->NetworkProvider->Order->ProviderOrder`。

比如，注册值为 `LanmanWorkstation,RDPNP,Nfsnp`，请删除 `Nfsnp`，然后重启 ECS。

- 如果存在 `webclient`，也会影响到文件浏览器登录 `smb` 服务，请删除。



说明：

客户端初次连接 SMB 服务器较慢时，可以 `ping` 挂载地址查看能否 `ping` 通，或者查看延时是否正。

- 无法 `ping` 通，请检查您的网络设置。
- 如果延时较长，请 `ping` 挂载 `ip`。如果 `ping ip` 比 `ping dns` 延时小很多，可能是 DNS 问题，请检查您的配置。

### 关于性能问题的解决步骤

1. 查看修改 `ProviderOrder` 的注册表值。初次访问等待时间久时，请先检查该值。
2. 使用 `fio` 进行性能测试，查看有无异常。

```
fio.exe --name=./iotest1 --direct=1 --rwmixread=0 --rw=write --bs=4K --numjobs=1 --thread --iodepth=128 --runtime=300 --group_reporting --size=5G --verify=md5 --randrepeat=0 --norandommap --refill_bufers --filename=\\<mount point dns>\myshare\testfio1
```

```
fio.exe --name=./iotest1 --direct=1 --rwmixread=0 --rw=write --bs=4K --numjobs=1 --thread --iodepth=128 --runtime=300 --group_reporting
```

```
--size=5G --verify=md5 --randrepeat=0 --norandommap --refill_buffers --filename=\\<mount point dns>\myshare\testfio1
```

- 查看程序，尽量以大数据块进行 IO 读写。如果数据块较小，会消耗更多的网络资源。如果不能修改数据块大小，可以使用BufferedOutputStream。

## 4.4 SMB 挂载失败的原因分析

在Windows 系统中挂载 SMB 文件系统，可能发生系统错误导致挂载失败。本文主要介绍挂载失败的主要原因及解决方法。

### 系统错误 53

#### 错误描述

找不到网络路径。

#### 主要原因

网络未连通。

#### 解决方法

请根据如下步骤验证网络连通性。

- 使用ping 命令检查挂载点地址是否可连通，延时是否正常。

```
ping <挂载点地址>
```

- 若网络 ping 通，则执行[步骤 2](#)。
- 若网络 ping 不通，请从以下方面排查。
  - 确认挂载命令正确，无多余或缺少 /、\、空格及myshare等内容。

正确挂载 SMB 文件系统的命令格式：

```
net use <挂载目标盘符> \\<挂载点地址>\myshare
```

样例：

```
net use z: \\xxxx.cn-hangzhou.nas.aliyuncs.com\myshare
```

- 确认文件系统类型为 SMB。



- 确认挂载点地址填写正确。
- 确认客户端的 ECS 与挂载点在同一个 VPC 中。
- 确认跨 VPC 或通过 VPN 连入的客户端，网络配置正确。

2. 使用 telnet命令检查 SMB 服务是否可用。

```
telnet <挂载点地址> 445
```

### 系统错误 58

#### 错误描述

指定的服务器无法运行请求的操作。

#### 主要原因

客户端 SMB 协议版本支持不兼容。

#### 解决方法

请确认 Windows 系统版本为Windows 2008 R2 及以上版本（不包括 Windows 2008）。

### 系统错误 64

#### 错误描述

指定的网络名不可用。

#### 主要原因

- NAS 权限组未允许目标 ECS 访问。
- 服务欠费。
- 选择经典网络进行挂载时，ECS 和 NAS 不属于同一阿里云 UID。
- 文件系统类型不是 SMB。

#### 解决方法

无权访问 NAS 文件系统资源，请从以下方面进行排查。

1. 确认文件系统挂载点权限组已包含该机器的内网 IP/VPC IP。
2. 确认阿里云 UID 未欠费。
3. 确认经典网络挂载时，ECS 和 NAS 属于同一个阿里云UID。
4. 确认文件系统类型为 SMB。



### 系统错误 67

#### 错误描述

找不到网络名。

#### 主要原因

关键的网络服务未启动。

解决方法

启动如下服务，具体操作可参见[挂载SMB文件系统](#)。

1. 启用 Workstation 服务。
2. 启用 TCP/IP NetBIOS Helper 服务。

#### 系统错误 85

错误描述

本地设备名已在使用中。

主要原因

目标盘符已被占用。

解决方法

请更换目标盘符重新挂载文件系统。

#### 系统错误 1272

错误描述

系统提示：“不能访问此共享文件夹，因为您组织的安全策略阻止未经身份验证的来宾访问。这些策略可帮助保护您的电脑免受网络上不安全设备或恶意设备的威胁。”

主要原因

Windows 系统因安全策略阻挡了以来宾访问权限（Guest Auth）访问SMB文件系统。

解决方法

若您的系统为 Windows Server 2016 之后版本（不包括 WindowsServer 2016），请修改以下注册表项允许来宾访问权限（Guest Auth）。

```
[HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\LanmanWorkstation\Parameters]
"AllowInsecureGuestAuth"=dword:1
```

具体解决方案请参考[Guest access in SMB2 disabled by default in Windows](#)。