
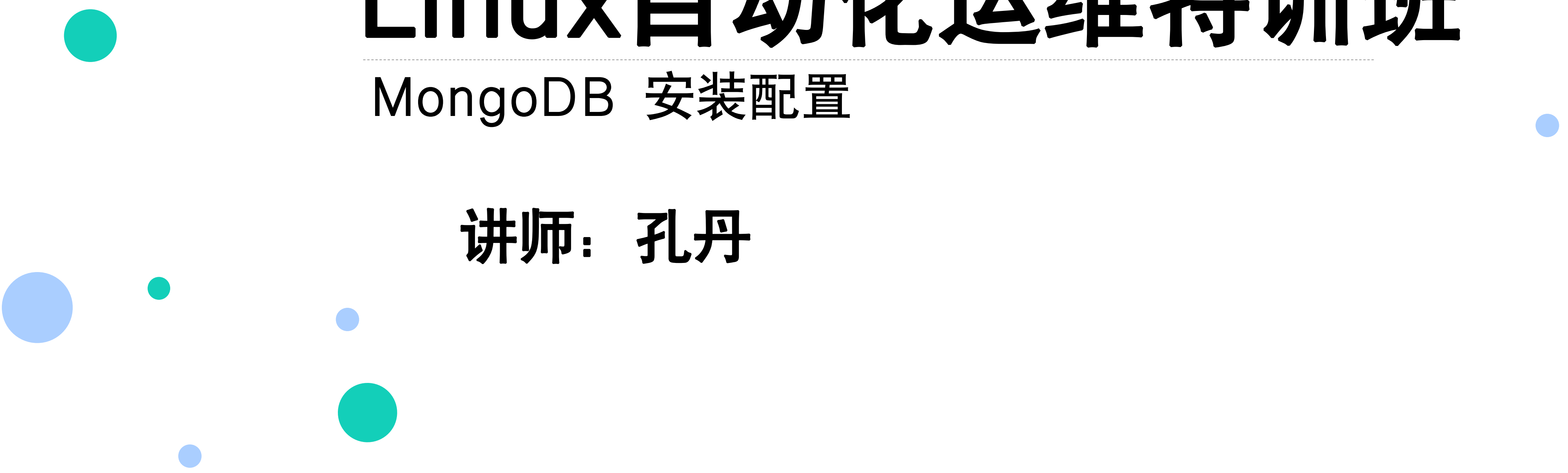




Linux自动化运维特训班

企业级NFS实战

讲师：孔丹



大纲

- NFS简介
- NFS原理
- NFS配置实战
- NFS优化

NFS简介

- NFS是Network File System的缩写，中文意思是网络文件系统。它的主要功能是通过网络（一般是局域网）让不同的主机系统之间可以共享文件或目录。NFS客户端（一般为应用服务器，例如web）可以通过挂载（mount）的方式将NFS服务器端共享的数据目录挂载到NFS客户端本地系统中（就是某一个挂载点下）。从客户端本地看，NFS服务器端共享的目录就好像是客户端自己的磁盘分区或者目录一样，而实际上却是远端的NFS服务器的目录。
- NFS网络文件系统很像Windows系统的网络共享，安全功能，网络驱动器影射，这也和Linux系统里的samba服务类似。只不过一般情况下，Windows网络共享服务或samba服务用于办公局域网共享，而互联网中小型网站集群架构后端常用NFS进行数据共享。

NFS简介

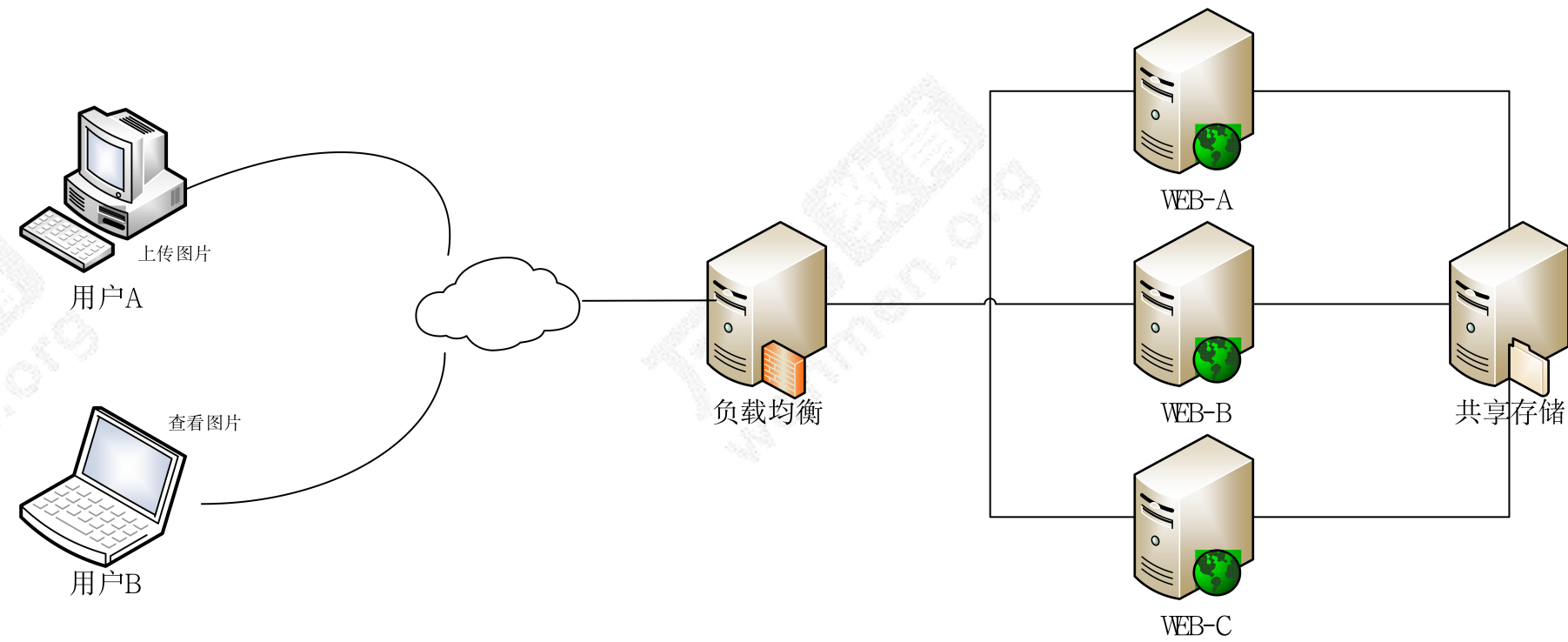
□ NFS在企业中的应用场景

在企业集群架构的工作场景中，NFS网络文件系统一般被用来存储共享视频，图片，附件等静态资源文件，通常网站用户上传的文件都会放到NFS共享里，例如：BBS产品的图片，附件，头像（注意网站BBS程序不要放NFS共享里），然后前端所有的节点访问这些静态资源时都会读取NFS存储上的资源。NFS是当前互联网系统架构中最常用的数据存储服务之一，前面说过，中小型网站公司应用频率更高，大公司或门户除了使用NFS外，还可能会使用更为复杂的分布式文件系统，比如Moosefs（mfs），GlusterFS，FastDFS，ceph等

在企业生产集群架构中，NFS作为所有前端Web服务的共享存储，存储的内容一般包括网站用户上传的图片，附件，头像等，注意，网站的程序代码不要放NFS共享里，因为网站程序是开发运维人员统一发布的，不存在发布延迟问题，直接批量发布到Web节点提供访问比共享到NFS里访问效率更高。

NFS简介

□ 企业生产集群为什么需要共享存储角色



例如：A用户上传图片到WEB-A服务器，然后让B用户访问这张图片，结果B用户访问的请求分发到了WEB-B，因为WEB-B上没有这张图片，这就导致它无法看到A用户上传的图片，如果此时有一个共享存储，A用户上传图片的请求无论是分发到WEB-A还是WEB-B上，最终都会存储到共享存储上，而在B用户访问图片时，无论请求分发到WEB-A还是WEB-B上，最终也都会去共享存储上找，这样就可以访问到需要的资源了。这个共享存储的位置可以通过开源软件和商业硬件实现，互联网中小型集群架构会用普通PC服务器配置NFS网络文件系统实现。

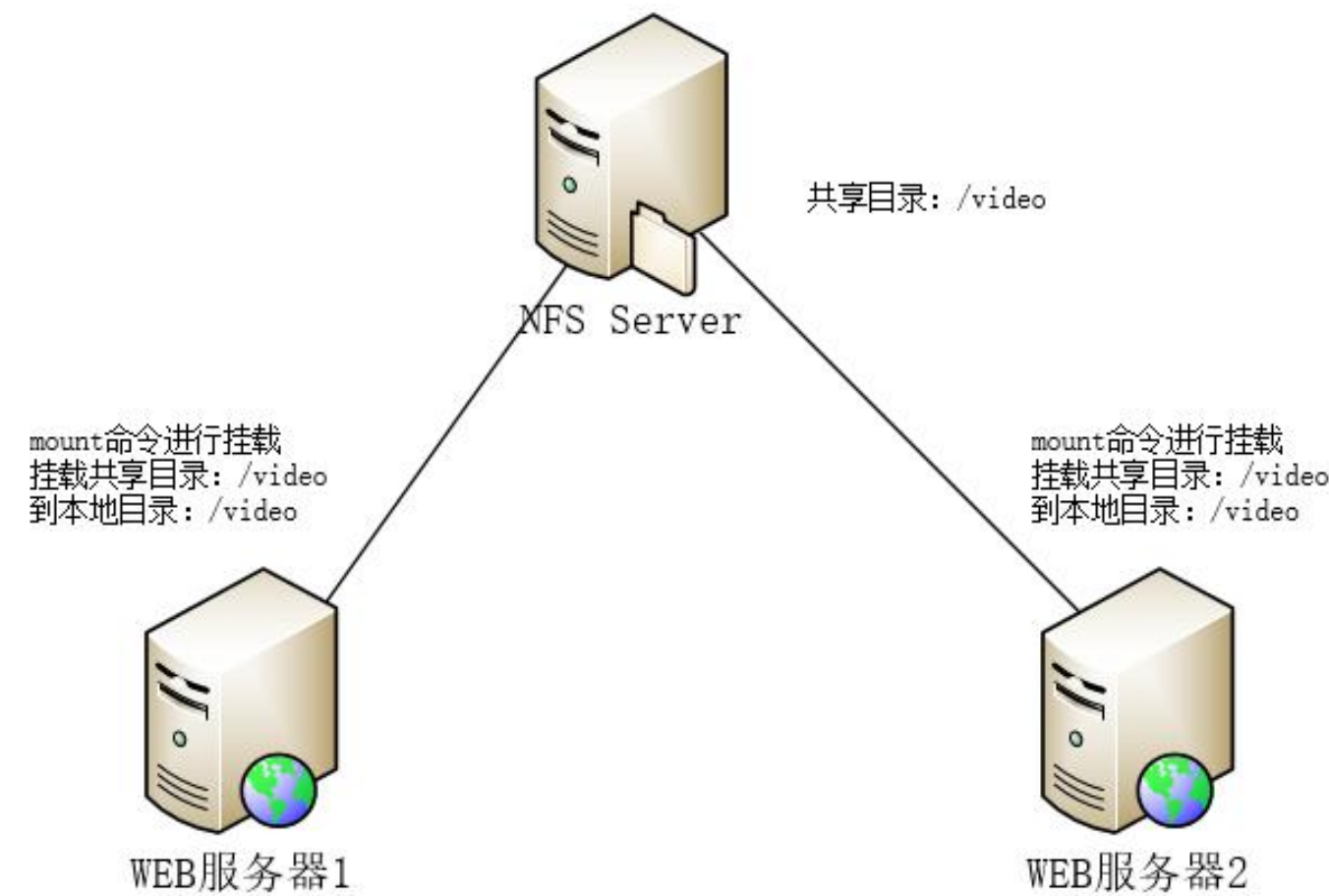
NFS简介

□ 企业生产集群为什么需要共享存储角色

中小型互联网企业一般不会买硬件存储，因为太贵，大公司如果业务发展很快的话，可能会临时买硬件存储顶一下网站压力，当网站并发继续加大时，硬件存储的扩展相对就会很费劲，且价格成几何级数增加。例如：淘宝网就曾替换掉了很多硬件设备，比如，用lvs+haproxy替换了netscaler负载均衡设备，用FastDFS，TFS配合PC服务器替换了netapp，emc等商业存储设备，去IOE正在成为互联网公司的主流。

NFS系统原理

□ NFS系统挂载结构图解与介绍



NFS服务器端/video共享目录挂载到了两台NFS客户端上。在客户端查看时，NFS服务器端的/video目录就相当于客户端本地的磁盘分区或目录，几乎感觉不到使用上的区别，根据NFS服务端授予的NFS共享权限以及共享目录的本地系统权限，只要在指定的NFS客户端操作挂载/video的目录，就可以将数据轻松地存取到NFS服务器端上的/video目录中了。

NFS系统原理

□ NFS系统挂载结构图解与介绍

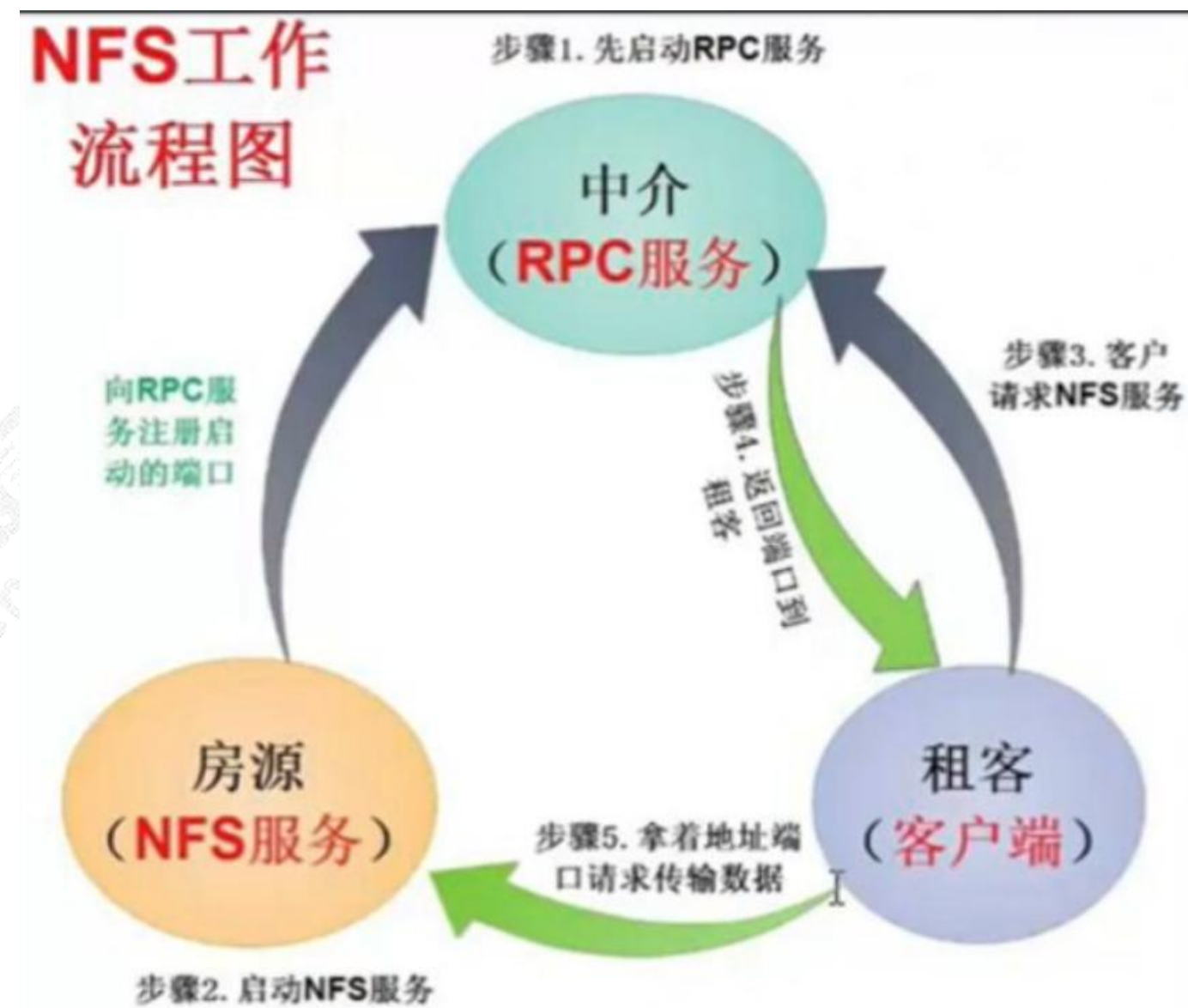
NFS系统是通过网络来进行数据传输的（所以叫做网络文件系统）因此，NFS会使用一些端口来传输数据，那么，NFS到底使用哪些端口来进行数据传输呢？

NFS在传输数据时使用的端口会随机选择。既然如此，NFS客户端是怎么知道NFS服务端使用的哪个端口呢？

答案就是通过RPC（中文意思远程过程调用，英文Remote Procedure Call简称RPC）协议/服务来实现，这个RPC服务的应用在门户级的网站有很多，例如：百度等。

NFS系统原理

□ 什么是RPC (Remote Procedure Call)



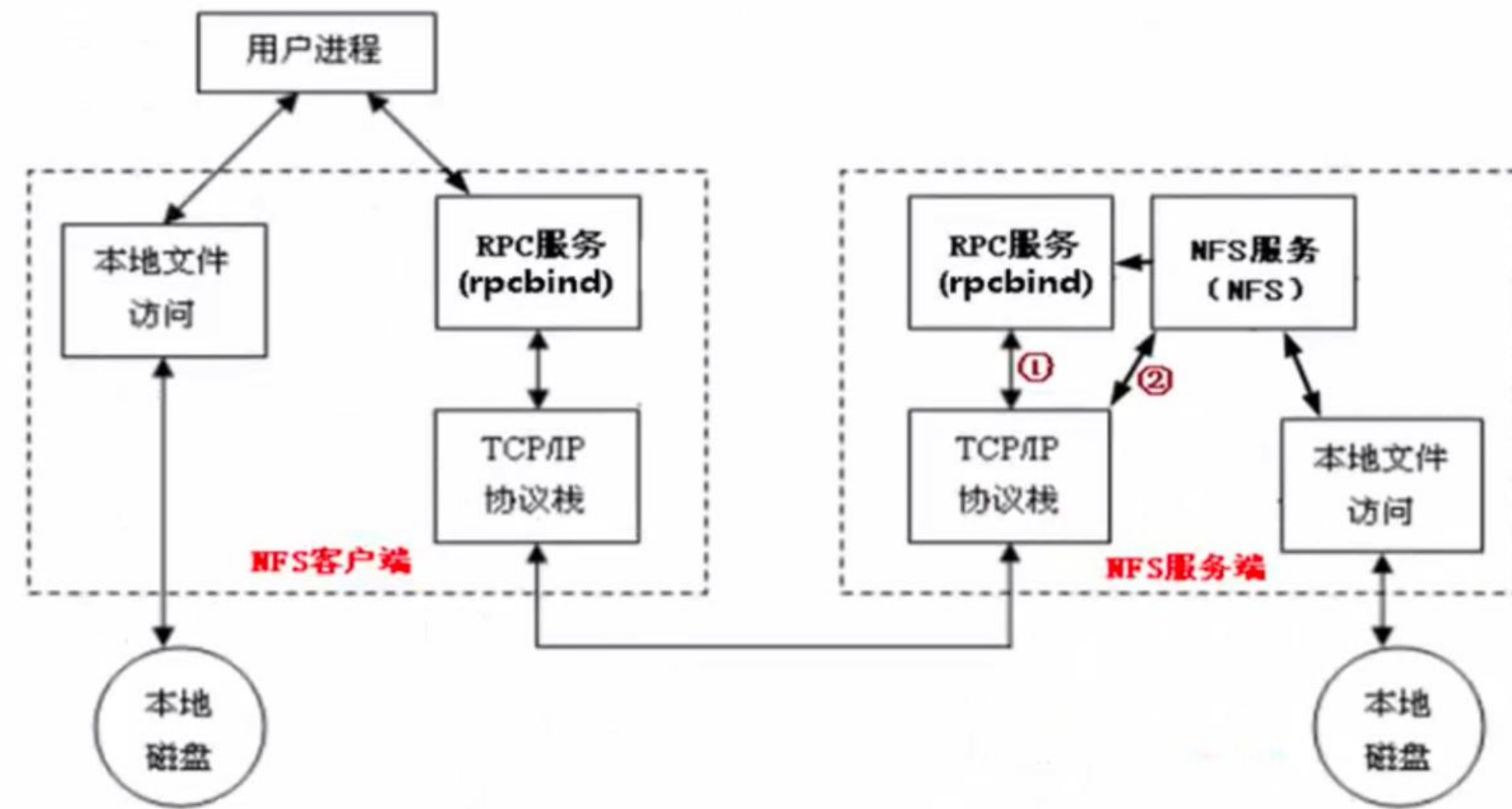
那么RPC服务又是如何知道每个NFS的端口呢？

这是因为，当NFS服务端启动服务时会随机取用若干端口，并主动向RPC服务注册取用的相关端口及功能信息，如此一来，RPC服务就知道NFS每个端口对应的NFS功能了，然后RPC服务使用固定的111端口来监听NFS客户端提交的请求，并将正确的NFS端口信息回复给请求的NFS客户端，这样一来，NFS客户端就可以与NFS服务端进行数据传输了。

在启动NFS SERVER之前，首先要启动RPC服务，，否则NFS SERVER就无法向RPC服务注册了

NFS系统原理

□ NFS的工作流程原理



当访问程序通过NFS客户端向NFS服务端存取文件时，其请求数据流程大致如下：

- 1) 首先用户访问网站程序，由程序在NFS客户端上发出存取NFS文件的请求，这时NFS客户端（即执行程序的服务端）的RPC服务（rpcbind服务）就会通过网络向NFS服务器端的RPC服务（rpcbind服务）的111端口发出NFS文件存取功能的询问请求。
- 2) NFS服务端的RPC服务（rpcbind服务）找到对应的已注册的NFS端口后，通知NFS客户端的RPC服务（rpcbind服务）
- 3) 此时NFS客户端获取到正确的端口，并与NFS daemon联机存取数据
- 4) NFS客户端把数据存取成功后，返回给前端访问程序，告知给用户存取结果，作为网站用户，就完成了一次存取操作。

NFS配置实战

- 实例一：共享/data 目录给192.168.150.0/24整个网段可读可写。
- NFS服务器部署角色

服务器系统	角色	IP
CentOS release 7.x	NFS服务器端(NFS-Sever)	192.168.150.11
CentOS release 7.x	NFS客户端1(NFS-Client1)	192.168.150.12
CentOS release 7.x	NFS客户端2(NFS-Client2)	192.168.150.13

NFS配置实战

□ 实例一：共享/data 目录给192.168.150.0/24整个网段可读可写。

□ 1、安装软件

```
yum -y install nfs-utils rpcbind
```

□ 2、编写nfs配置文

nfs配置文件默认存在/etc/exports

```
vim /etc/exports
```

```
#share /data by kongd for share at 20191020
```

```
/data 192.168.150.0/24(rw,sync)
```

/etc/exports文件说明：

第一部分： /data --指定共享目录信息

第二部分： 192.168.150.0/24 --指定了一个网段信息，表示允许指定的网段主机挂载到我本地的共享目录上

第三部分： (rw,sync) --表示定义共享参数信息，

rw 表示读写，对共享目录设置的权限

sync 同步，数据会先写入到NFS服务器内存中，会立刻同步到磁盘里面==直接存储硬盘中

NFS配置实战

□ 实例一：共享/data 目录给192.168.150.0/24整个网段可读可写。

□ 3、创建共享目录，进行权限设定

```
mkdir /data -p
```

```
chown -R nfsnobody.nfsnobody /data
```

说明：NFS共享目录管理用户为nfsnobody，此用户不用创建，安装nfs软件时会自动创建

□ 4、启动服务

首先，启动rpc服务；然后启动nfs服务

```
systemctl enable rpcbind
```

```
systemctl start rpcbind
```

```
systemctl enable nfs-server
```

```
systemctl start nfs-server
```

rpcbind服务启动信息查看：rpcinfo -p localhost

NFS配置实战

□ 实例一：共享/data 目录给192.168.150.0/24整个网段可读可写。

□ NFS客户端配置

1、客户端安装nfs-utils软件

```
yum -y install nfs-utils
```

2、检查远端showmount

```
showmount -e 192.168.150.11
```

3、客户端挂载

```
mount -t nfs 192.168.150.11:/data /mnt
```

通过/etc/fstab开机自动挂载

NFS配置实战

□ NFS服务相关进程信息

- rpcbind rpc启动进程 主进程
- rpc state 检查数据存储的一致性
- rpc.rquotad 磁盘配额
- rpc.mountd 权限管理验证
- nfsd NFS主进程
- rpc.idmapd 用户压缩映射

□ 指定 NFS客户端地址的配置详细说明

客户端地址	具体地址	说 明
授权单一客户端访问NFS	10.0.0.30	一般情况，生产环境中此配置不多
授权整个网段可访问NFS	10.0.0.0/24	其中的24等同于255.255.255.0，指定网段为生产环境中最常见的配置。配置简单，维护方便
授权整个网段可访问NFS	10.0.0.*	指定网段的另外写法（不推荐使用）
授权某个域名客户端访问	nfs.wanmenedu.com	此方法生产环境中一般情况不常用
授权整个域名客户端访问	*.wanmenedu.com	此方法生产环境中一般情况不常用

NFS配置实战

□ 常见案例

常用格式说明	要共享的目录客户端IP地址或IP段(参1,参2,)
配置例一	/data 10.0.0.0/24(rw,sync)说明：允许客户端读写，并且数据同步写入到服务器端的磁盘里注意：24和"("之间不能有空格
配置例二	/data 10.0.0.0/24(rw,sync/all_squash,anonuid=2000,anongid=2000) 说明：允许客户端读写，并且数据同步写到服务器端的磁盘里，并且指定客户端的用户UID和GID，早期生产环境的一种配置，适合多客户端共享一个NFS服务单目录，如果所有服务器的nfsnobody账户UID都是65534,则本例没什么必要了.早期centos5.5的系统默认情况下nfsnobody的UID不一定是65534,此时如果这些服务器共享一个NFS目录，就会出现访问权限问题.
配置例三	/home/tom 10.0.0.0/24(ro)说明：只读共享用途：例如在生产环境中，开发人员有查看生产服务器日志的需求，但又不希望给开发生产服务器的权限，那么就可以给开发提供从某个测试服务器NFS客户端上查看某个生产服务器的日志目录（NFS共享）的权限，当然这不是唯一的方法，例如可以把程序记录的日志发送到测试服务器供开发查看或者通过收集日志等其它方式展现

NFS配置实战

□ NFS配置权限设置常用参数说明

参数	说明
rw	Read-write,表示可读写权限
ro	Read-only, 表示只读权限
sync	(同步, 实时) 请求或吸入数据时, 数据同步写入到NFS Server的硬盘后才返回
async	(异步) 写入时数据会先写到内存缓冲区, 只到硬盘有空档才会写入磁盘, 这样可以提升写入速率! 风险为若服务器挂掉或不正常关机, 会损失缓冲区中未写入磁盘的数据
no_root_squash	访问NFS Server共享目录的用户如果是root, 它对该共享目录具有root权限。
root_squash	如果访问目录的是root, 则它的权限将被压缩成匿名用户。
all_squash	不管访问共享目录的用户身份如何, 它的权限都被压缩成匿名用户。
anonuid	指定共享文件夹里文件所有者的uid号: 例如: (rw,squash,anonuid=12306,anongid=12306)
anongid	指定共享文件夹里文件所有者的gid号: 例如: (rw,squash,anonuid=12306,anongid=12306)

NFS配置实战

□ 相关命令

exportfs命令

如果我们在启动了NFS之后又修改了/etc/exports，是不是还要重新启动nfs呢？这个时候我们就可以用exportfs命令来使改动立刻生效，该命令格式如下：

格式：exportfs [-aruv]

- a 全部挂载或卸载 /etc/exports中的内容

- r 重新读取/etc/exports 中的信息，并同步更新/etc/exports、
/var/lib/nfs/xtab

- u 卸载单一目录（和-a一起使用为卸载所有/etc/exports文件中的目录）

- v 在export的时候，将详细的信息输出到屏幕上。

具体例子：

```
# exportfs -au 卸载所有共享目录
```

```
# exportfs -ra 重新共享所有目录并输出详细信息
```

rpcinfo命令

利用rpcinfo -p 可以查看出RPC开启的端口所提供的程序有哪些

其中nfs 开启的是2049，portmapper(rpcbind) 开启的是111，其余则是rpc开启的

NFS优化

□ 有关系统安全挂载参数选项

在企业工作场景，一般来说，NFS服务器共享的只是普通静态数据（图片，附件，视频），不需要执行suid，exec等权限，挂载的这个文件系统只能作为数据存取之用，无法执行程序，对于客户端来讲增加了安全性，例如：很多木马篡改站点文件都是由上传入口上传的程序到存储目录，然后执行的。

```
mount -t nfs -o nosuid,noexec,nodev,rw 192.168.150.11:/data /mnt
```

□ mount挂载性能优化参数选项

1) 禁止更新目录及文件时间戳挂载，命令如下：

```
mount -t nfs -o noatime,nodiratime 192.168.150.11:/data /mnt
```

2)安全加优化的挂载方式如下：

```
mount -t nfs -o  
nosuid,noexec,nodev,noatime,nodiratime,intr,rsiz=131072,wsiz=131072 192.168.150.11:/data /mnt
```

3) 默认的挂载方式如下：

```
mount -t nfs 192.168.150.11:/data /mnt
```

NFS优化

□ NFS内核优化建议

[x] /proc/sys/net/core/rmem_default:该文件指定了接收套接字缓冲区大小的默认值（以字节为单位），默认设置：124928 建议：8388608

[x] /proc/sys/net/core/rmem_max: 该文件指定了接收套接字缓冲区大小的最大值（以字节为单位） 建议：16777216

[x] /proc/sys/net/core/wmem_default:该文件指定了发送套接字缓冲区大小的默认值（以字节为单位），默认设置：124928 建议：8388608

[x] /proc/sys/net/core/wmem_max:该文件指定了发送套接字缓冲区大小的最大值（以字节为单位）。默认设置：124928. 建议：16777216

```
cat >>/etc/sysctl.conf<<EOF
net.core.wmem_default=8388608
net.core.rmem_default=8388608
net.core.rmem_max=16777216
net.core.wmem_max=16777216
EOF
```


总结

- NFS简介
- NFS原理
- NFS配置实战
- NFS优化

作业

- 架设一台NFS服务器，并按以下要求配置：
- 1、开放/nfs/shared目录，供所有用户查询资料。
- 2、开放/nfs/upload目录，供服务器本网段内的数据上传目录，并将所有用户及所属目录的用户组映射为nfs-upload，其UID和GID均为210。
- 3、将/home/tom目录仅共享给192.168.150.12这台主机，并且只有用户tom可以完全访问该目录。



谢谢观看

更多好课，请关注[万门大学APP](#)

