Implementing an NFSv4 Server Namespace
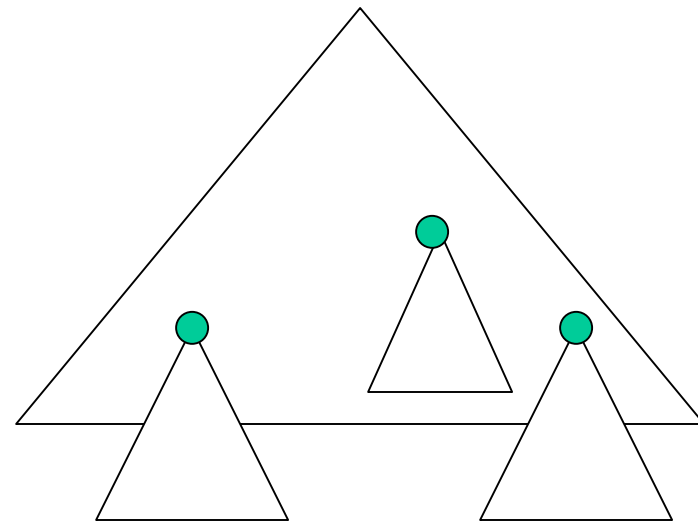
Brent Callaghan

Sun Microsystems, Inc.
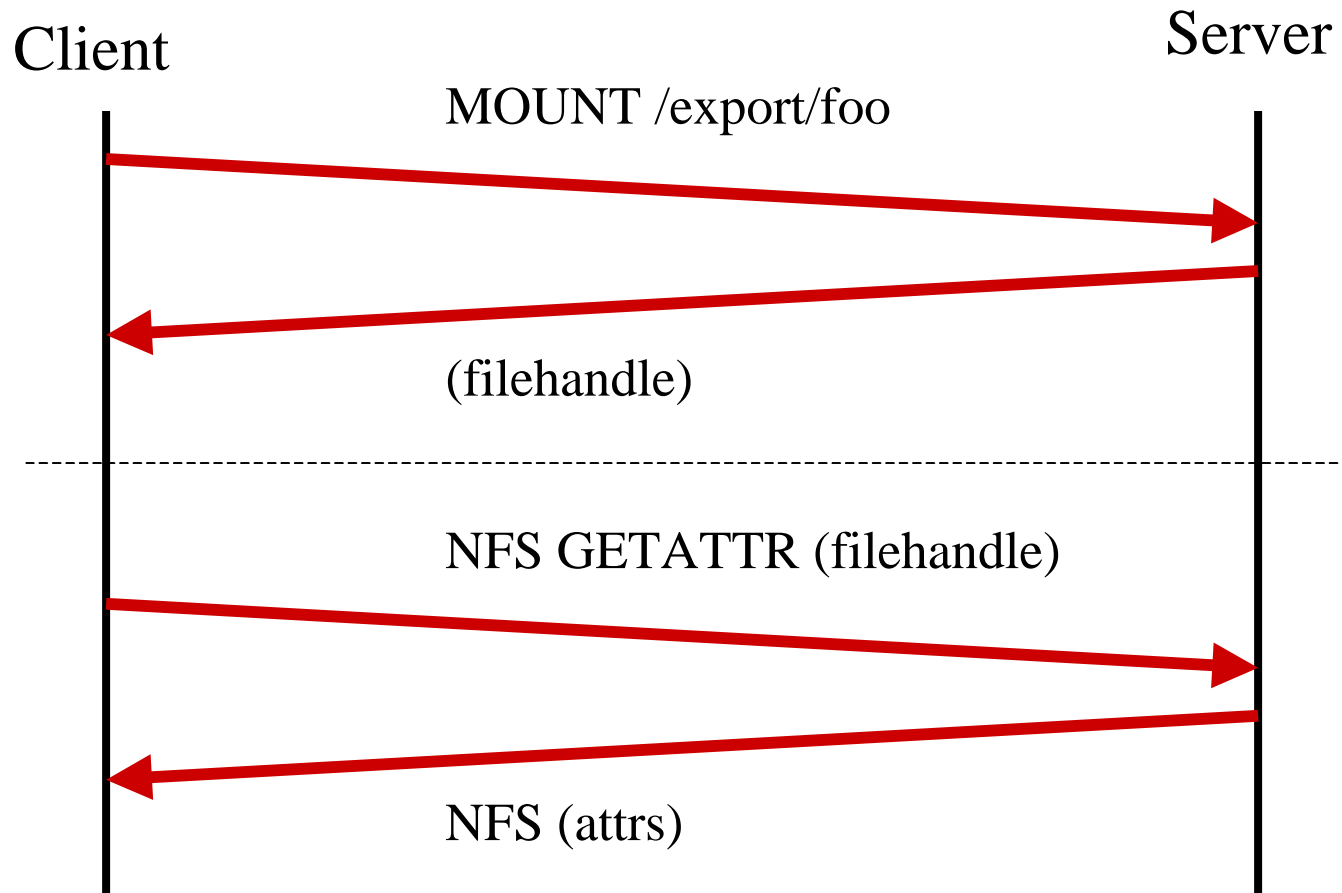
`brent@eng.sun.com`

The Mount Protocol

- For NFS access, must start with a root filehandle.

- Each NFS export has one

- Must use MOUNT protocol to get it.

Mount Protocol

Client   Server

MOUNT /export/foo

(filehandle)

NFS GETATTR (filehandle)

NFS (attrs)

Problem with Mount Protocol

- Firewalls!
  - Mount protocol has no assigned port
- NFSv4 requirement
  - One protocol, one connection
- Security
  - Mount daemon runs as superuser
- Difficult to browse exports

Mounting with NFSv4

- Need to convert a path to a filehandle
- NFS is already pretty good at that!
  - A sequence of LOOKUPs
- Start with a root filehandle: PUTROOTFH
- Must have a path to any exported directory
- Need a "pseudo" filesystem to bridge hidden parts of the server namespace.

Mounting with NFSv4

PUTROOTFH

Pseudo filesystem provides paths that bridge non-exported filesystems.
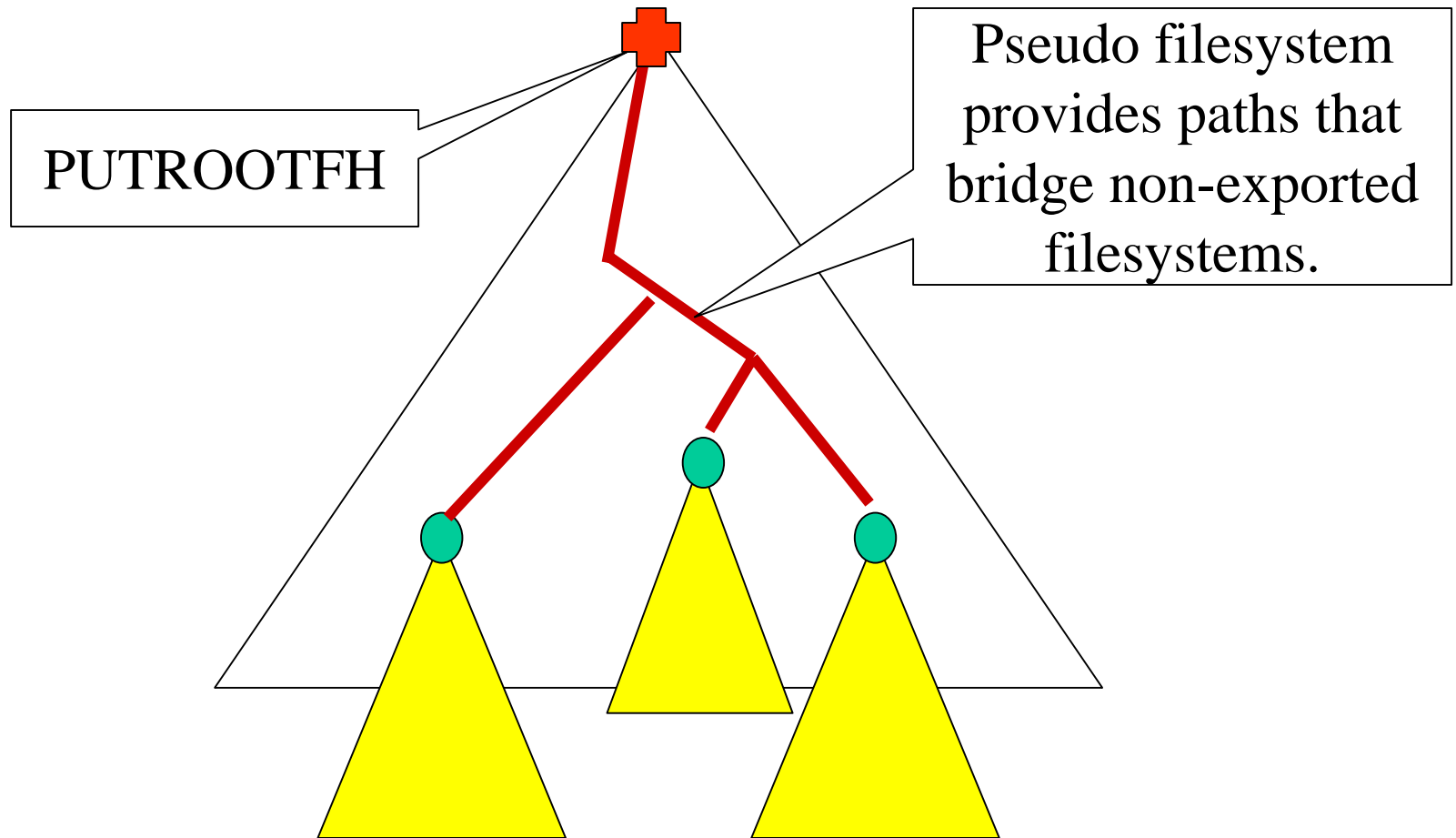
Pseudo Filesystem

Pseudo
filesystem
path
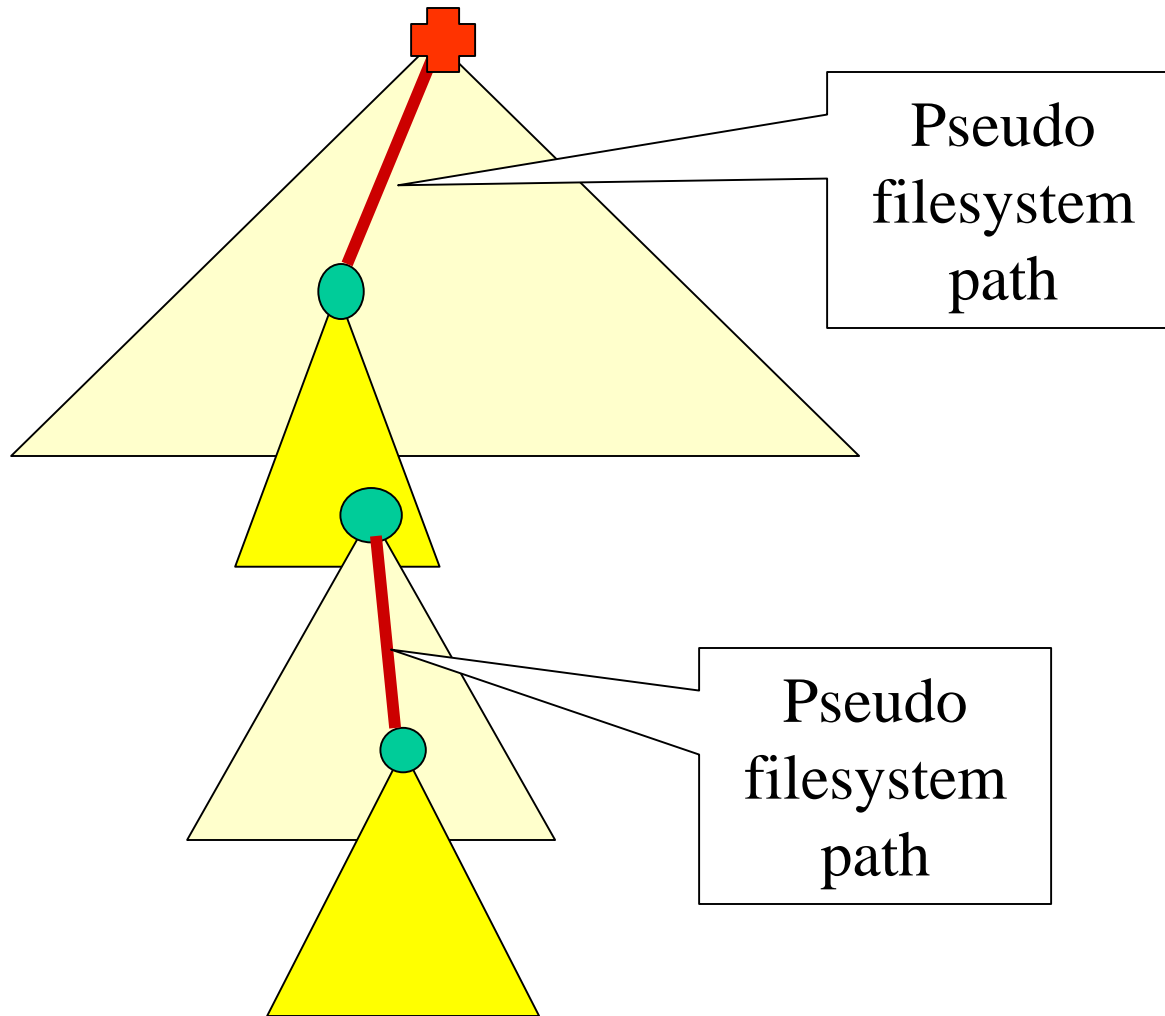
Pseudo filesystem needed wherever a path to an export crosses a non-exported path

Pseudo
filesystem
path

7

Linux Pseudo Filesystem

- An arbitrary namespace
- Need not correspond with local server path

/a

/b

/c

Server root

Pseudo root

/export

**/export/this**

/export/this/dir
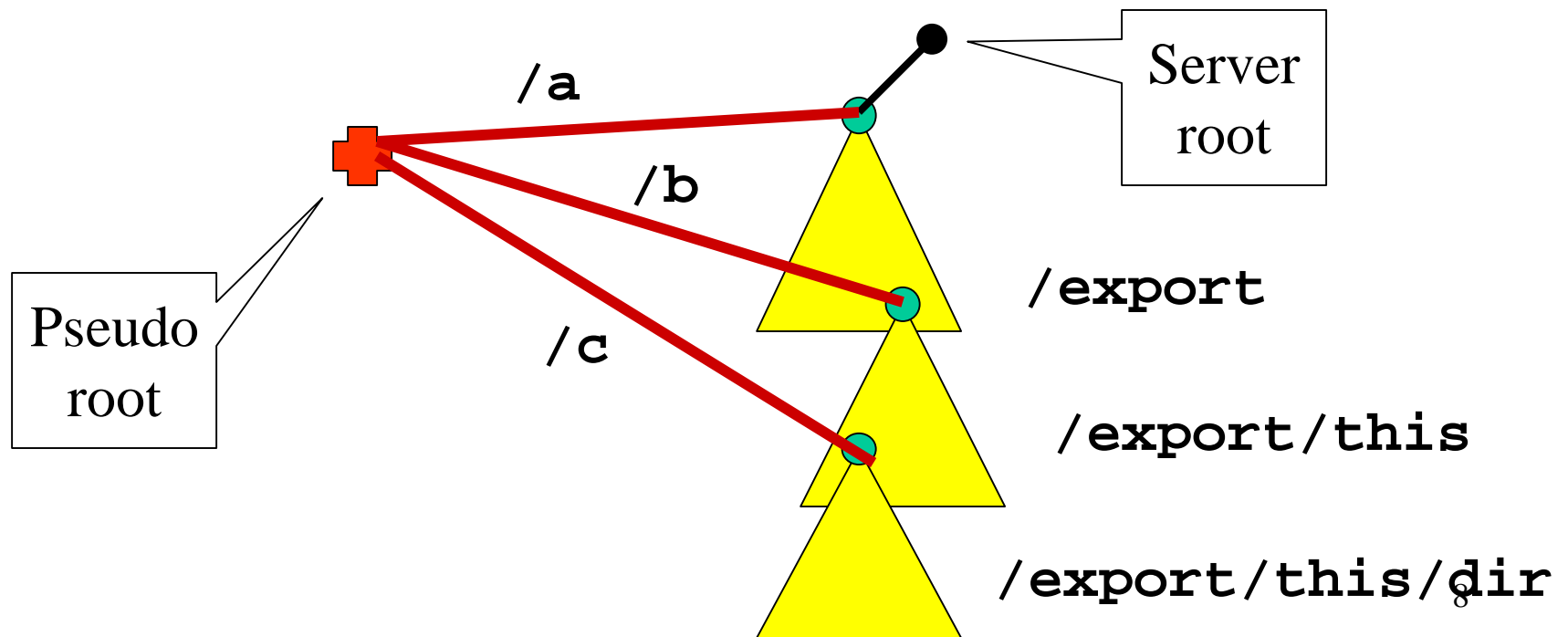
Synthetic Pseudo Filesystem

- In-memory structures within NFSv4 server

- Filehandles must be volatile

- Consistency is difficult, e.g. renames

- Significant code

In memory
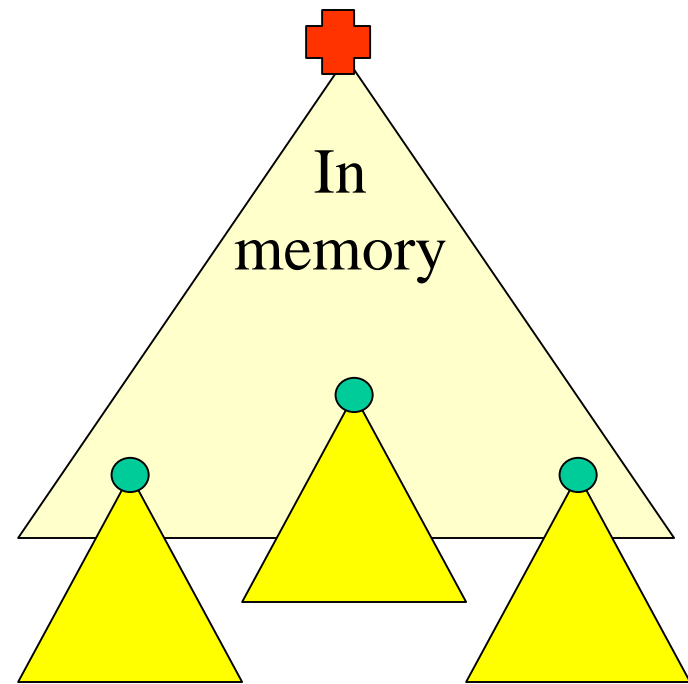
Solaris Approach: Limited View

- Provide access to unexported filesystems!
  - With limitations
- Pseudo filesystem is always consistent because there is just one filesystem.
- Filehandles can be persistent
- Easy to implement

Limited View
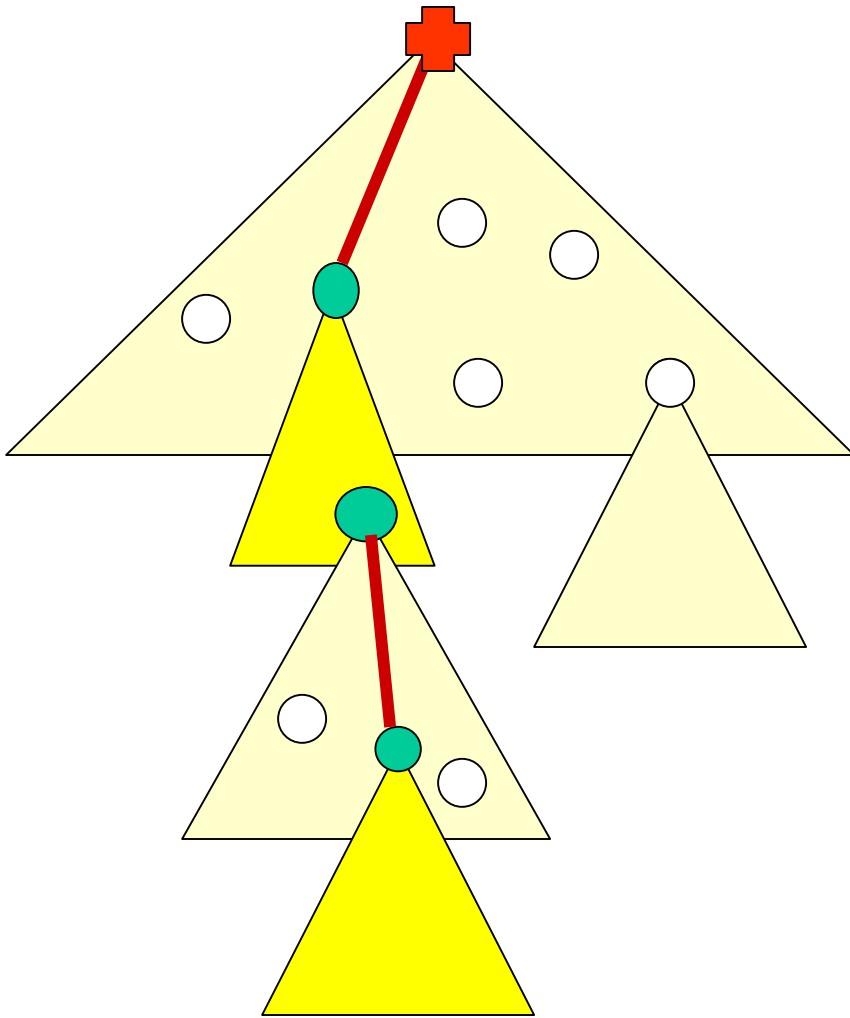
- Can see only paths to exported filesystems

- Read-only!

- LOOKUP, READDIR, GETATTR

- Restricted attributes, e.g. no space_avail

- Easy to code
  - "pseudo" exportfs bit

Implementing Limited View

Pseudo exportfs

● | fsid | fid | fid | | | |

Fidlist limits visibility

Real exportfs

● | fsid | fid | | | | |
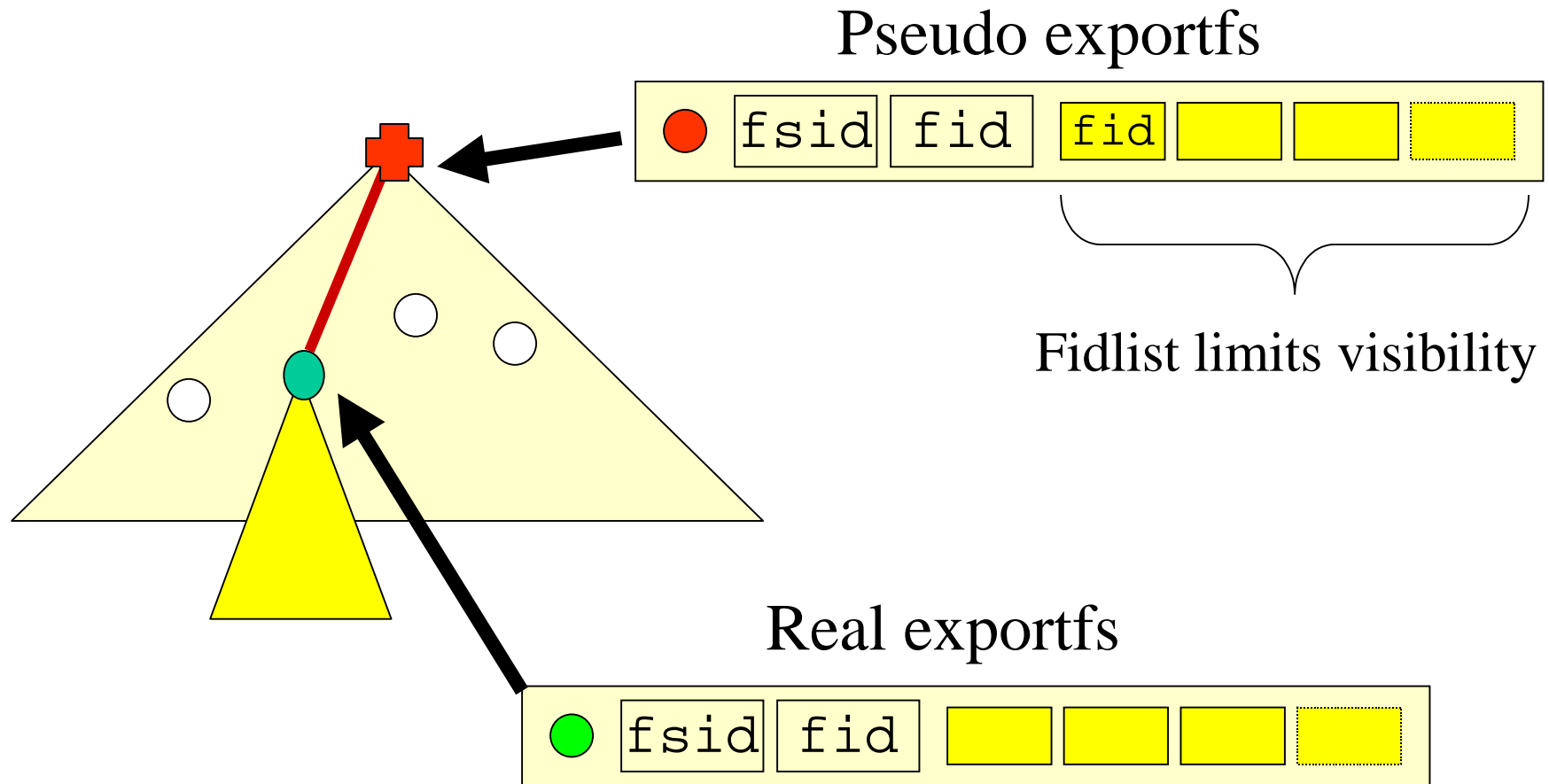
Sharing a Filesystem

- **`share /export/foo/bar`**
- No change in system administration
  - Policy: don't scare sysadmins with v4
- Pseudo-exports parent directories if necessary.
- Filehandles don't change if a parent directory is exported

Client Mirror Mounts

- NFSv4 clients cross server mountpoints
- POSIX clients must perform mirror mounts
  - To preserve POSIX compliance: fsid, fid uniqueness.
  - To avoid breaking apps like pwd, cp, df
- Mirror mounts need to happen automatically - "auto" mounts
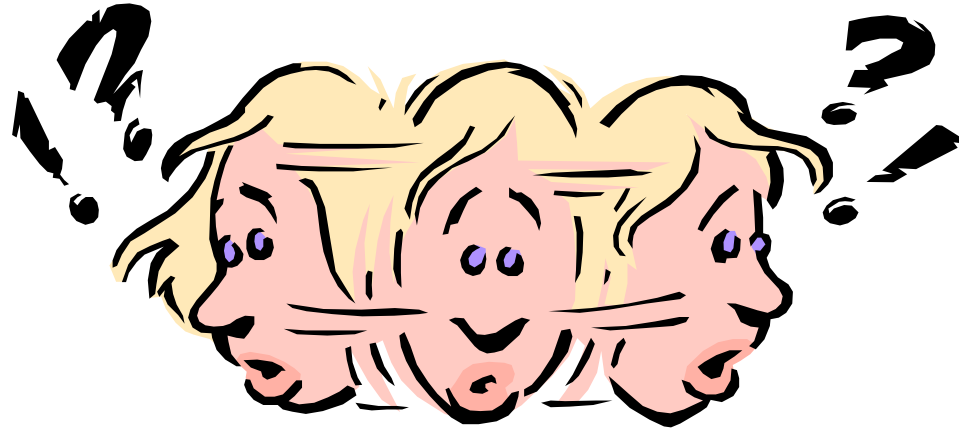
Hierarchical Exports

- Don't scare admins with v4
  - but there's a feature they might like
- Currently: one share per server mount
  - can yield dozens of identical shares on a big home directory server
- Would like to be able to have submounts inherit a parent mount

Hierarchical Exports

- `share -o `**`hier`**`,sec=krb5 /export/home`
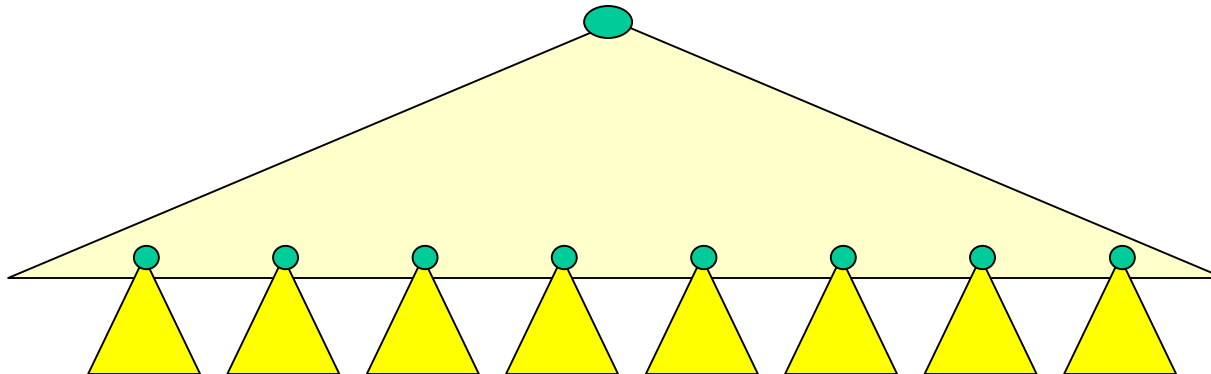
- V4 clients naturally cross mountpoints
- Need to simulate with derived shares for v2/v3 clients

Questions
&Answers