NFSv4 Delegation in Solaris 10

- Bill.Baker@sun.com

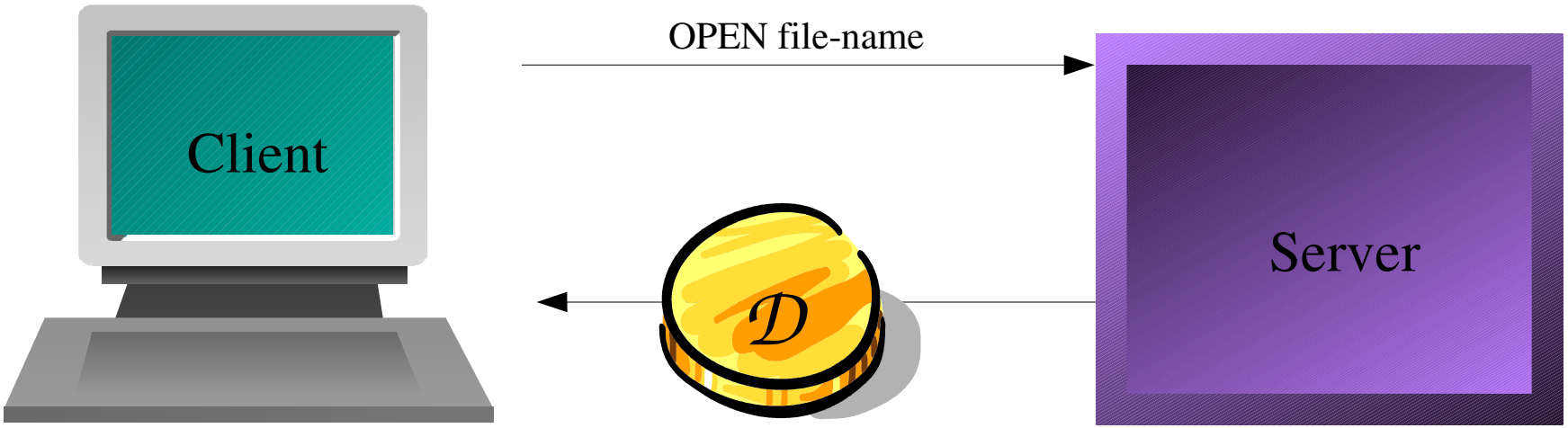
Sun microsystems

NFSv4 Delegation in Solaris 10

The technologies behind the scenes

NFSv4 Delegation

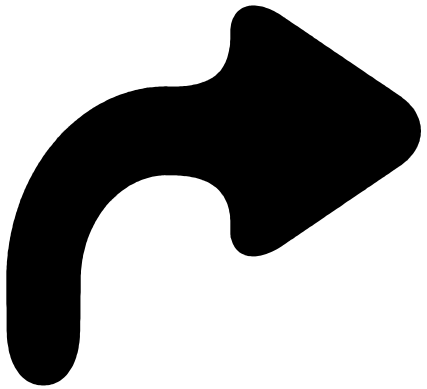
What is a delegation?



A guarantee by the server that no conflicting access to a file will occur without notification.

NFSv4 Delegation

Solaris Implementation

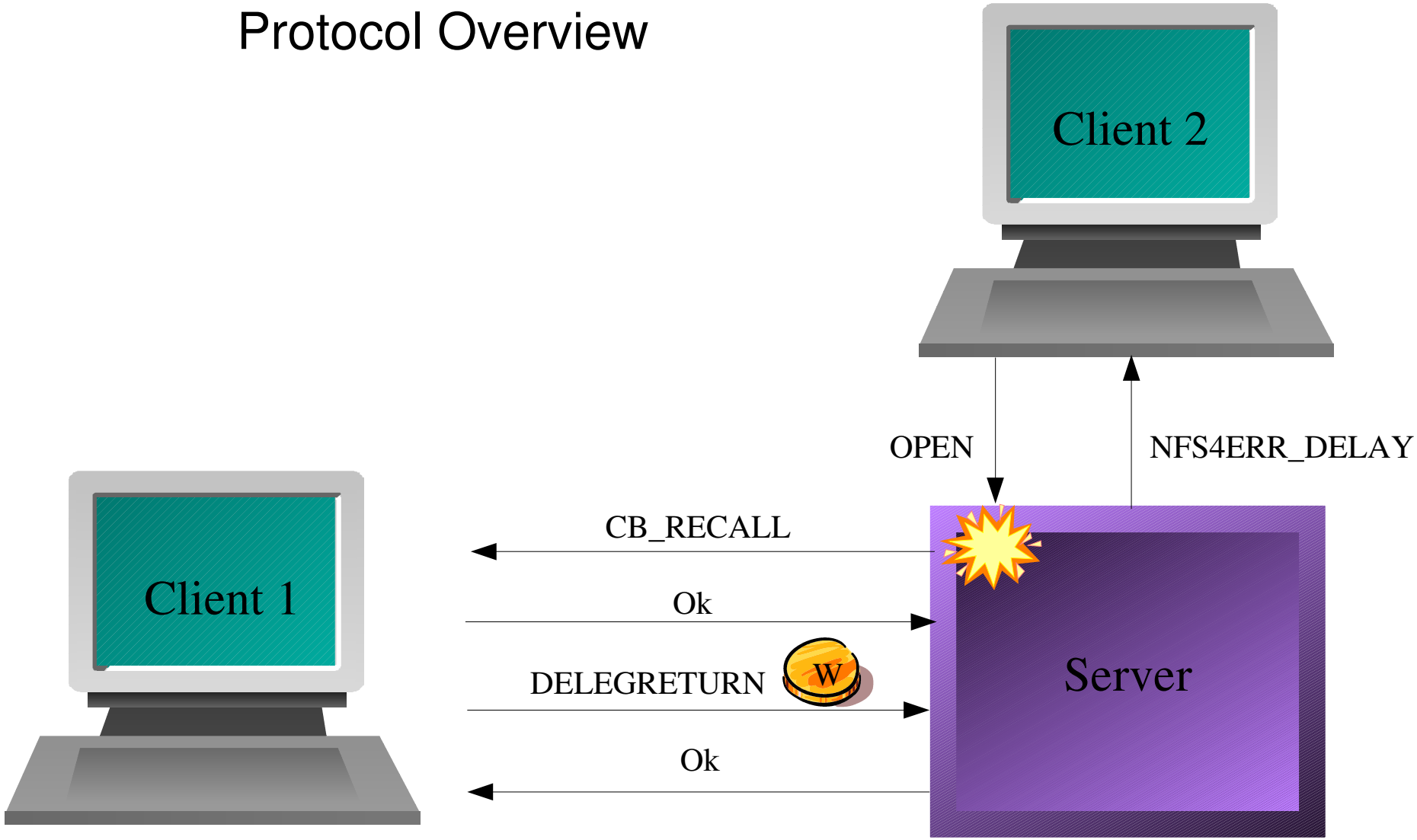


- Enabled by default
- Server offers both read & write
- Integrated with local access
- Integrated with v2 & v3 access
- Disabled via /etc/default/nfs

NFS_SERVER_DELEGATION=off

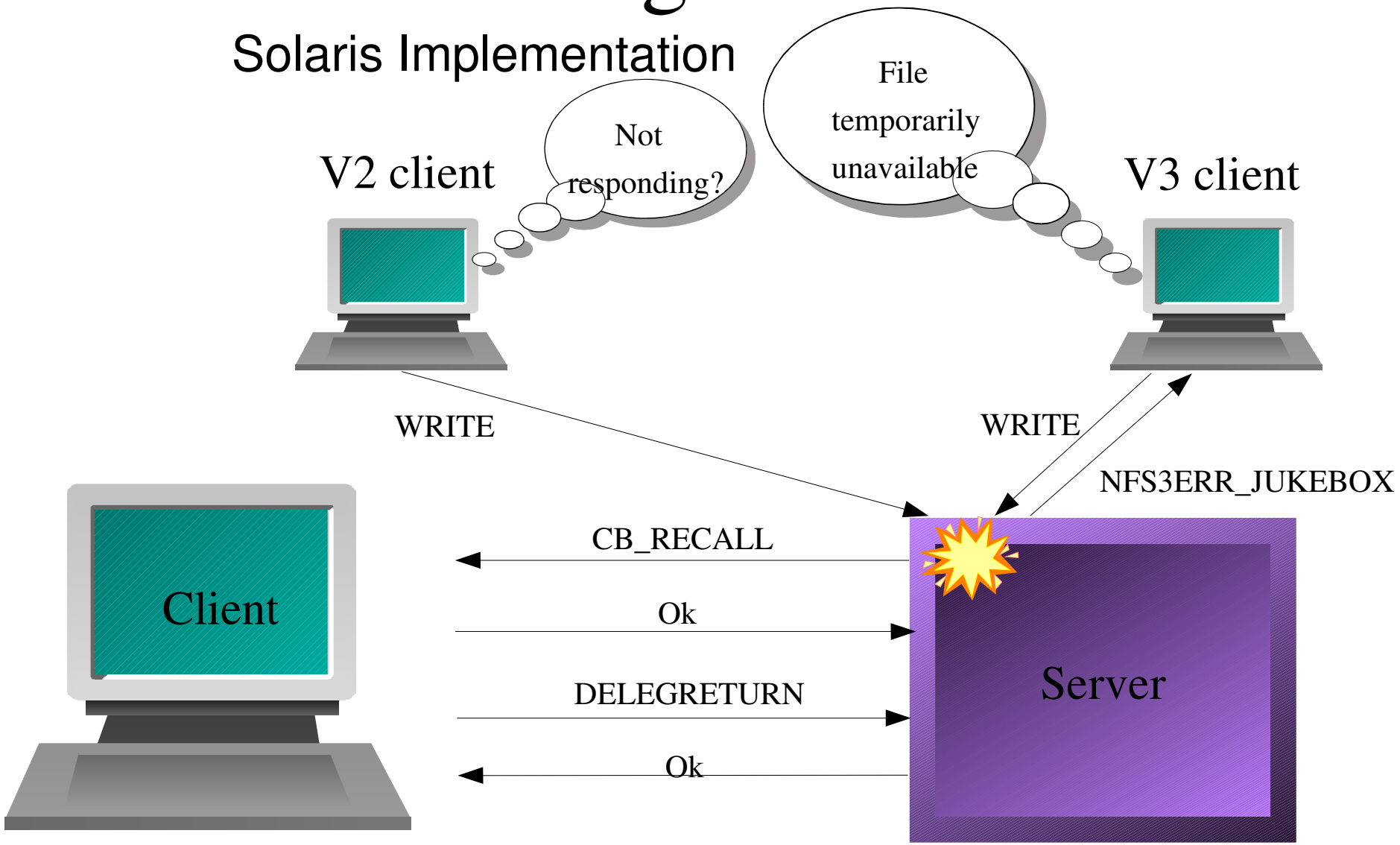
NFSv4 Delegation

Protocol Overview



NFSv4 Delegation

Solaris Implementation

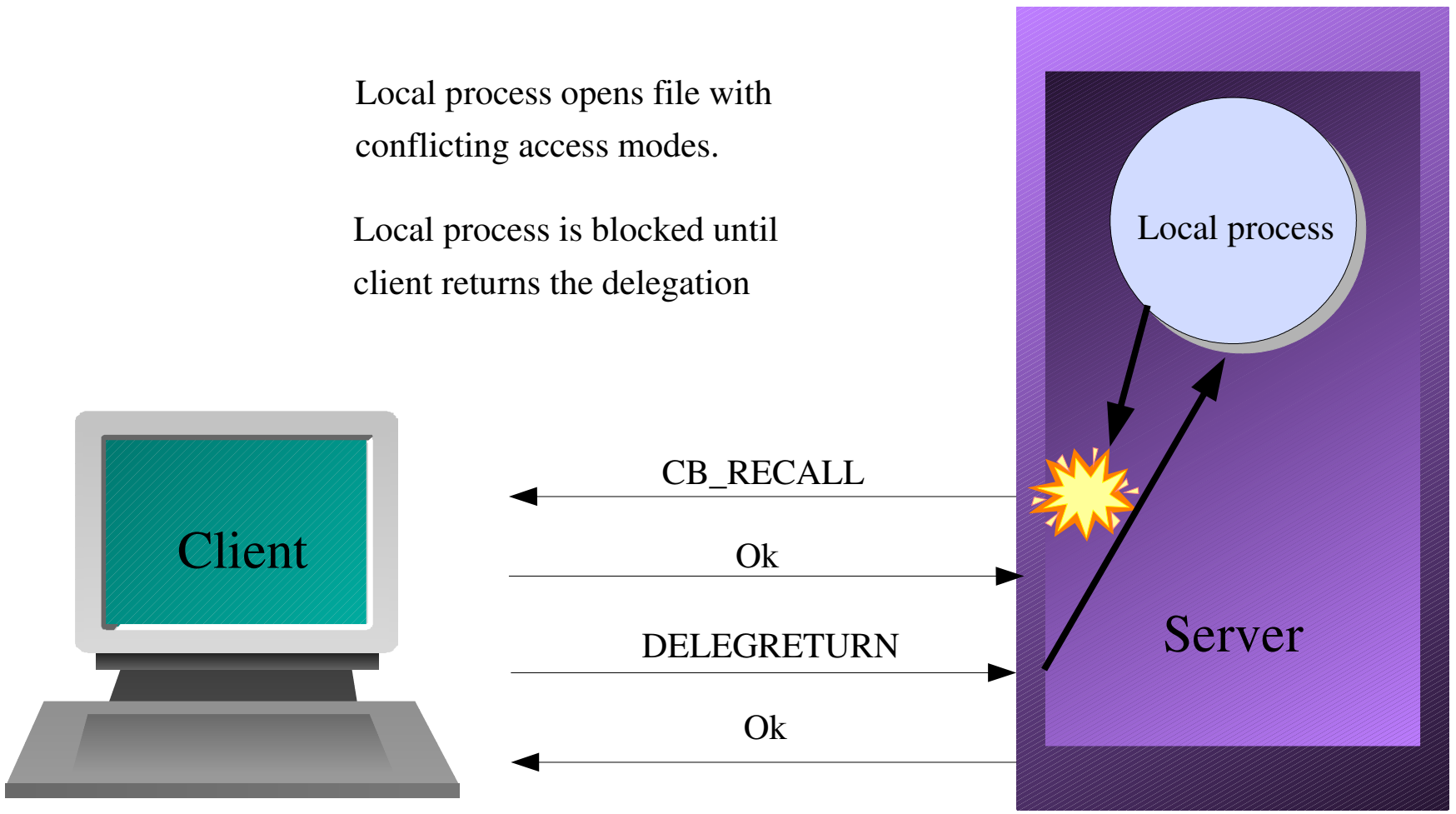


NFSv4 Delegation

Solaris Implementation

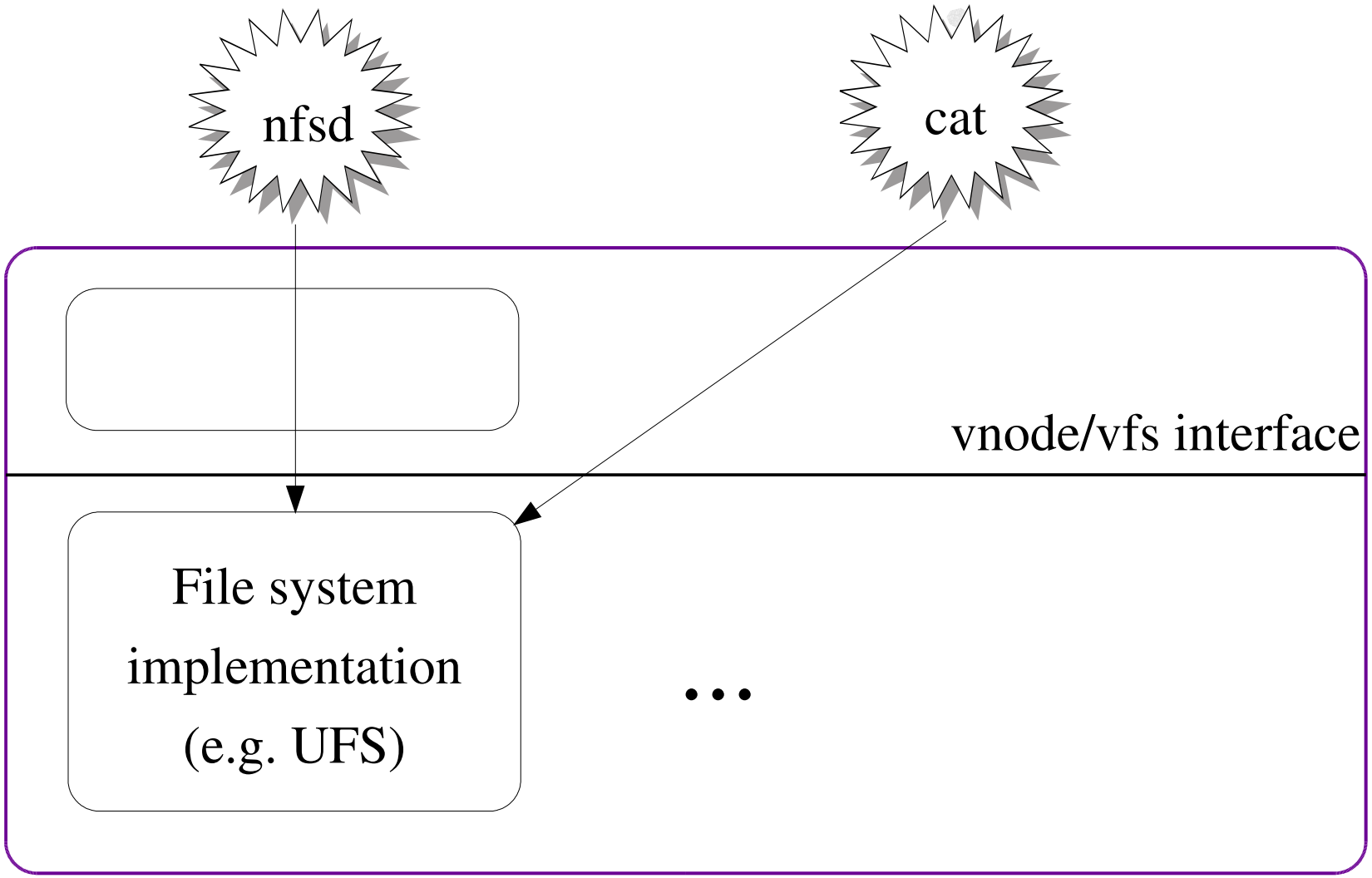
Local process opens file with conflicting access modes.

Local process is blocked until client returns the delegation



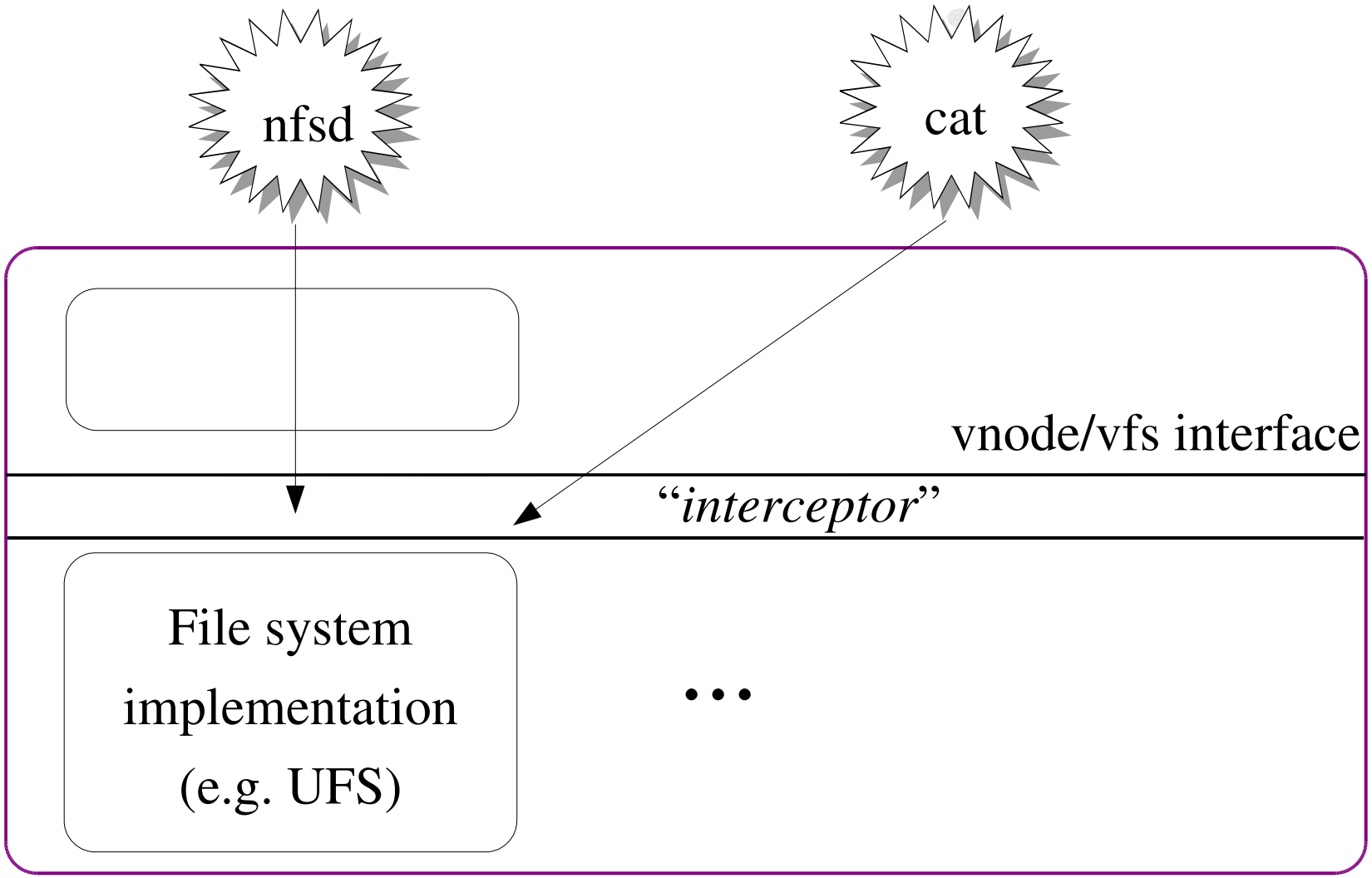
NFSv4 Delegation

Integration with Local Access



NFSv4 Delegation

Integration with Local Access - File Event Monitors



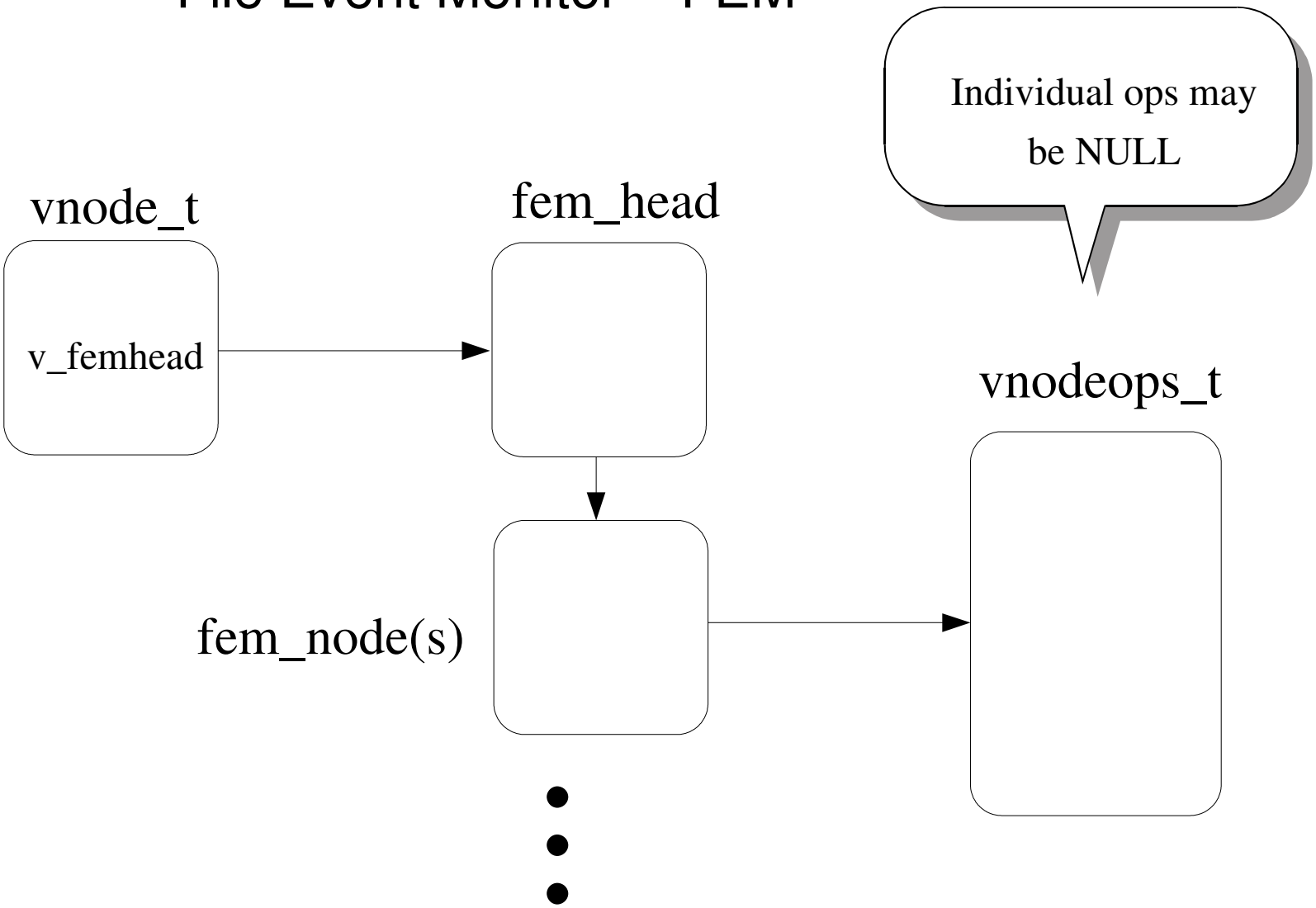
NFSv4 Delegation

File Event Monitors – FEM

- Enable integration with local access
- Minimize changes to individual file systems
- Very low overhead on non-monitored files

NFSv4 Delegation

File Event Monitor – FEM



NFSv4 Delegation

File Event Monitor – FEM

The names have been changed to ...

```
fs_operation_def_t rfs4_deleg_templ [] =
     {{ VOPNAME_OPEN, deleg_open4read } ... };

server_init()
{
     fem_create("name", rfs4_deleg_templ, dops);
}
grant_delegation()
{
     fem_install(vp, dops, opaque, ...);
}
```

NFSv4 Delegation

File Event Monitor – FEM

```
deleg_open4read(femarg_t *fa, int mode, cred_t *cr)
{
     if (mode & (FWRITE|FTRUNC) {
          initiate_recall();
          wait4delegreturn();
     }
     return (vnext_open(fa, mode, cr));
}
```

Monitors may:

- Do extra processing
- Return an error
- Pass thru to the next level
- Both pre and post-op processing

NFSv4 Delegation

File Event Monitor – FEM

What are the User Level APIs?

"*We don't need no stinking APIs!*"

But, seriously, folks ...

NFSv4 Delegation

File Event Monitor – FEM

"*We don't need no stinking APIs!*"

A Tactical solution to an interesting problem

More to follow, watch this space

Possibilities:

- DMAPI
- FAM/fileschanged
- OpenSolaris?

NFSv4 Delegation

FEM – Lessons Learned

- Ill Defined vnode OPS are "Bad"
  - Open w/o close? How rude!
- Sufficiency is difficult to prove
  - Vnode interception interesting, but ...
- Deadlocks are hard to avoid
  - Monitor programming guidelines
  - Sample Monitors?

NFSv4 Delegation

Sun microsystems