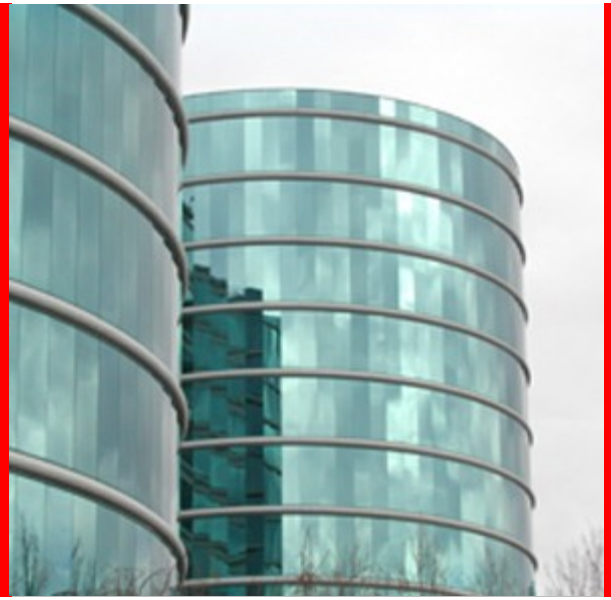

ORACLE®



Sun

ORACLE

ORACLE THIRTY YEARS 30

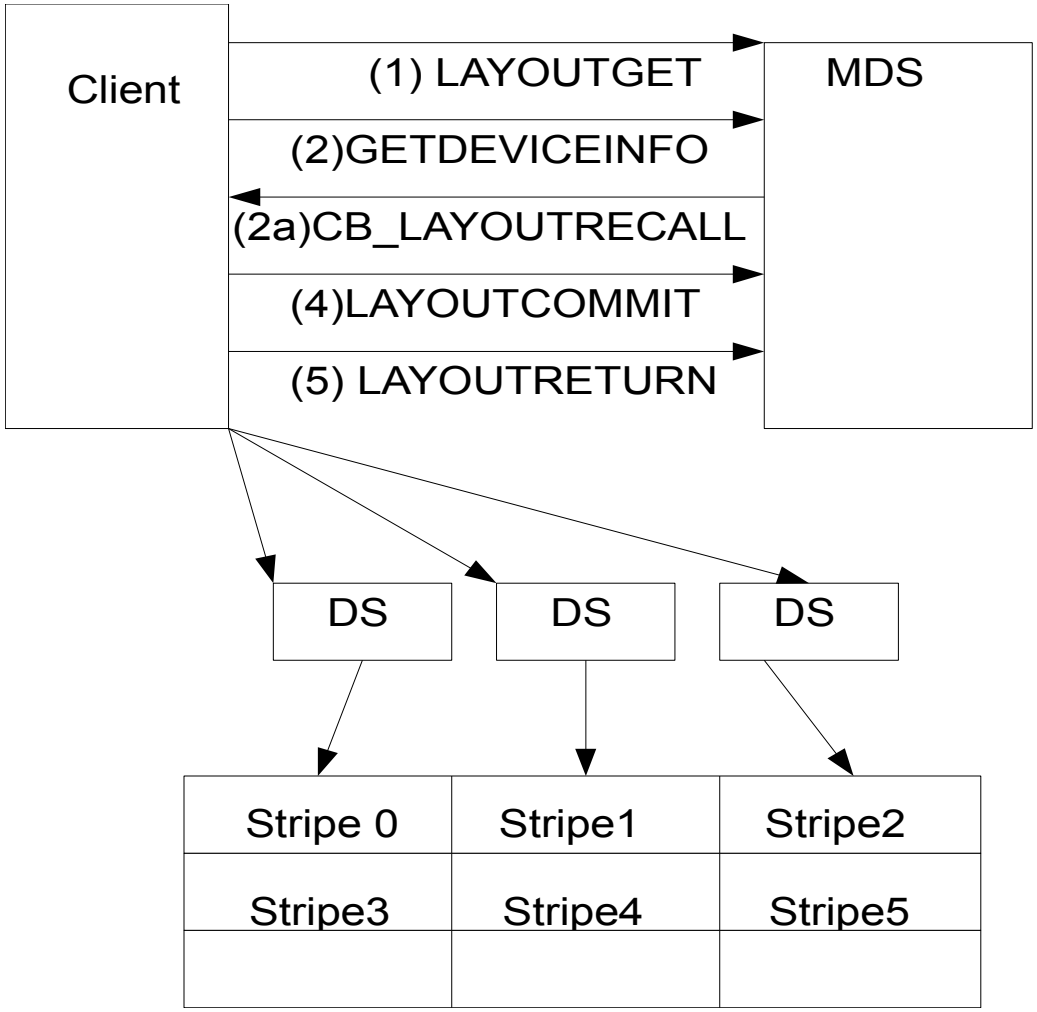
NFS V4.1 Client Layout Segments

Karen Rochford

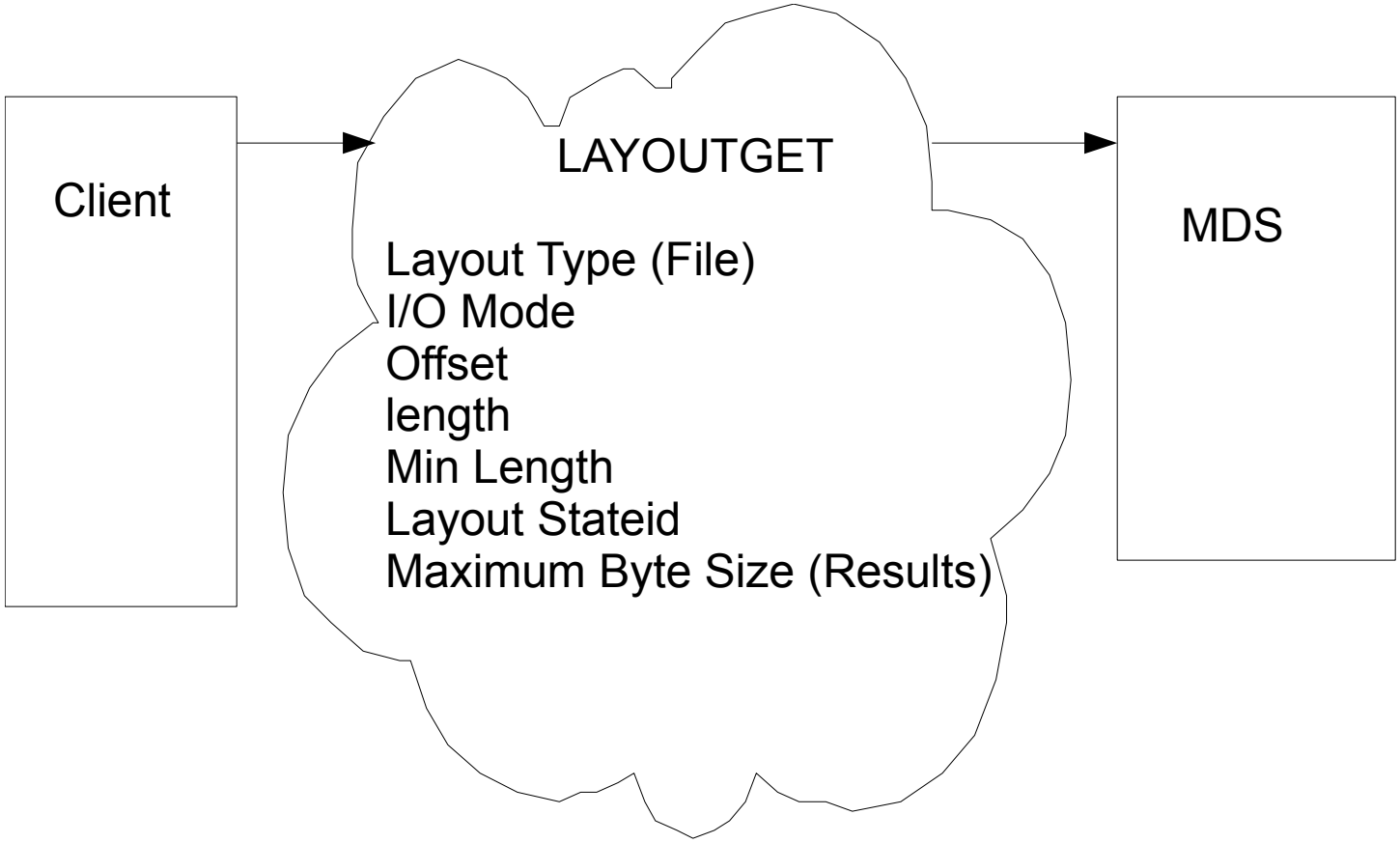
Client Layout Segments
What Am I Talking About?

- File Type Layouts Only
- Layout Segment, what are they?
- LAYOUTGET considerations
- I/O Considerations

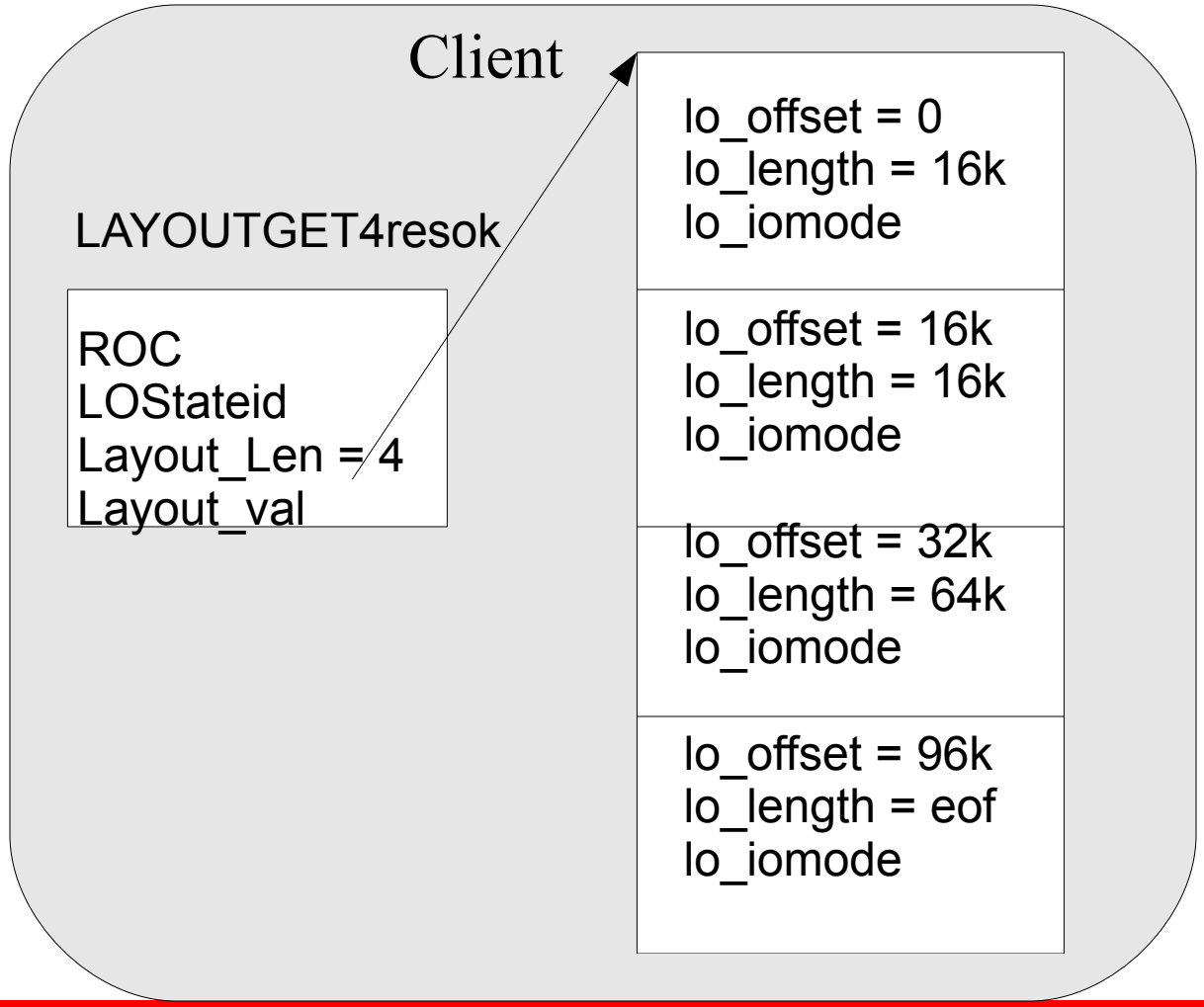
File Layouts - Review



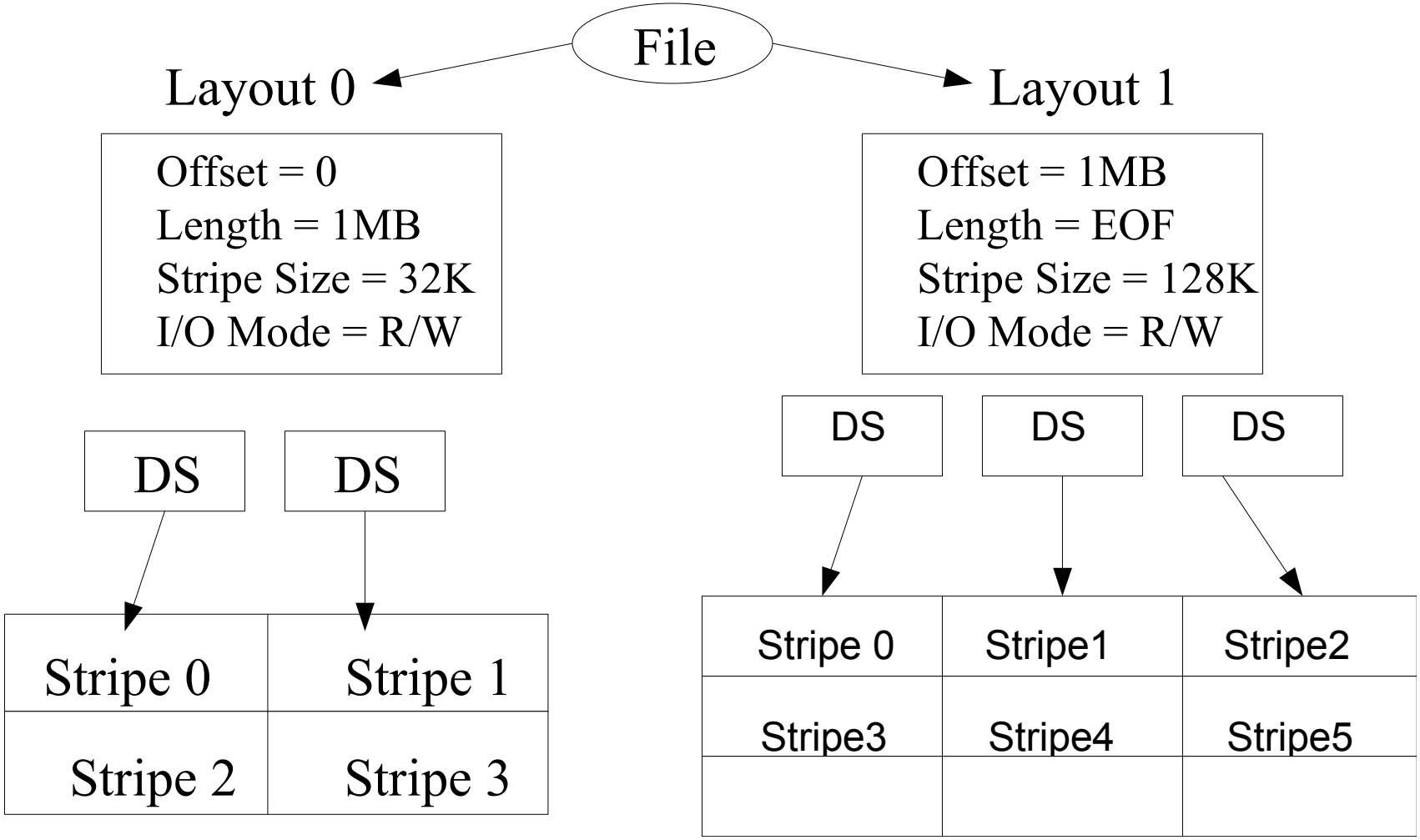
LAYOUTGET Arguments



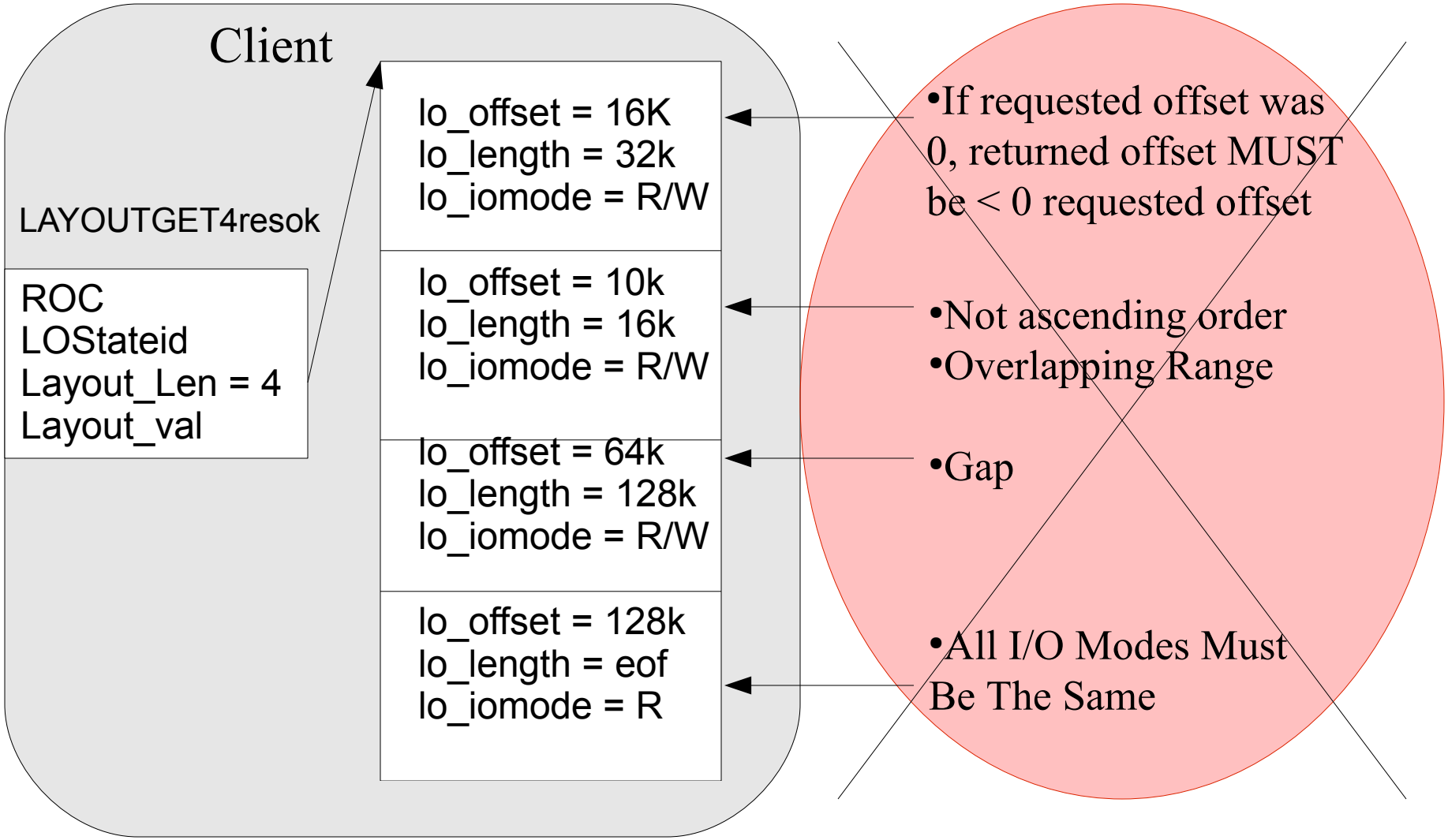
LAYOUTGET Results
Example 4 Layout Segments



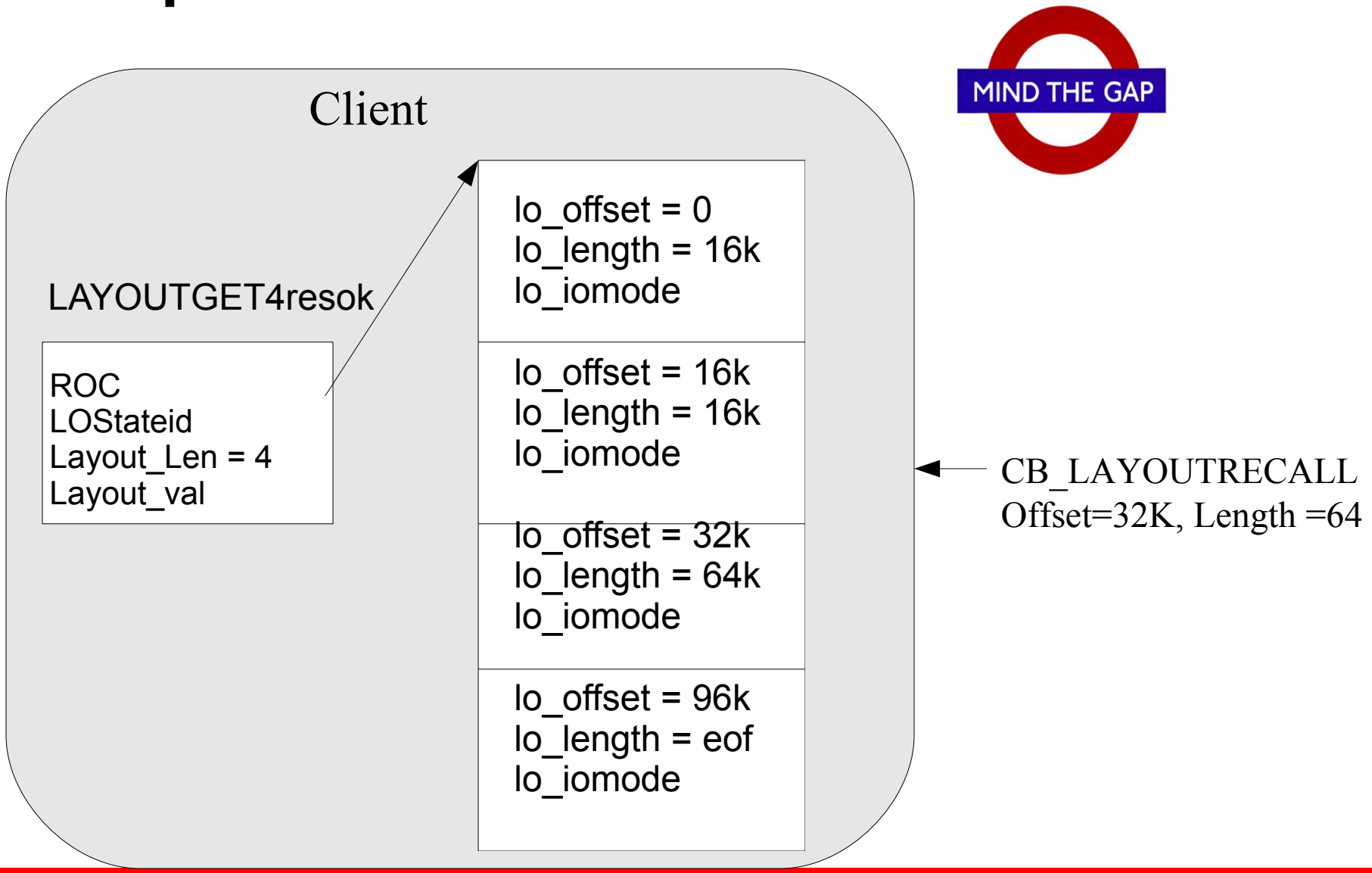
Multiple Layouts, why?



Invalid LAYOUTGET Results

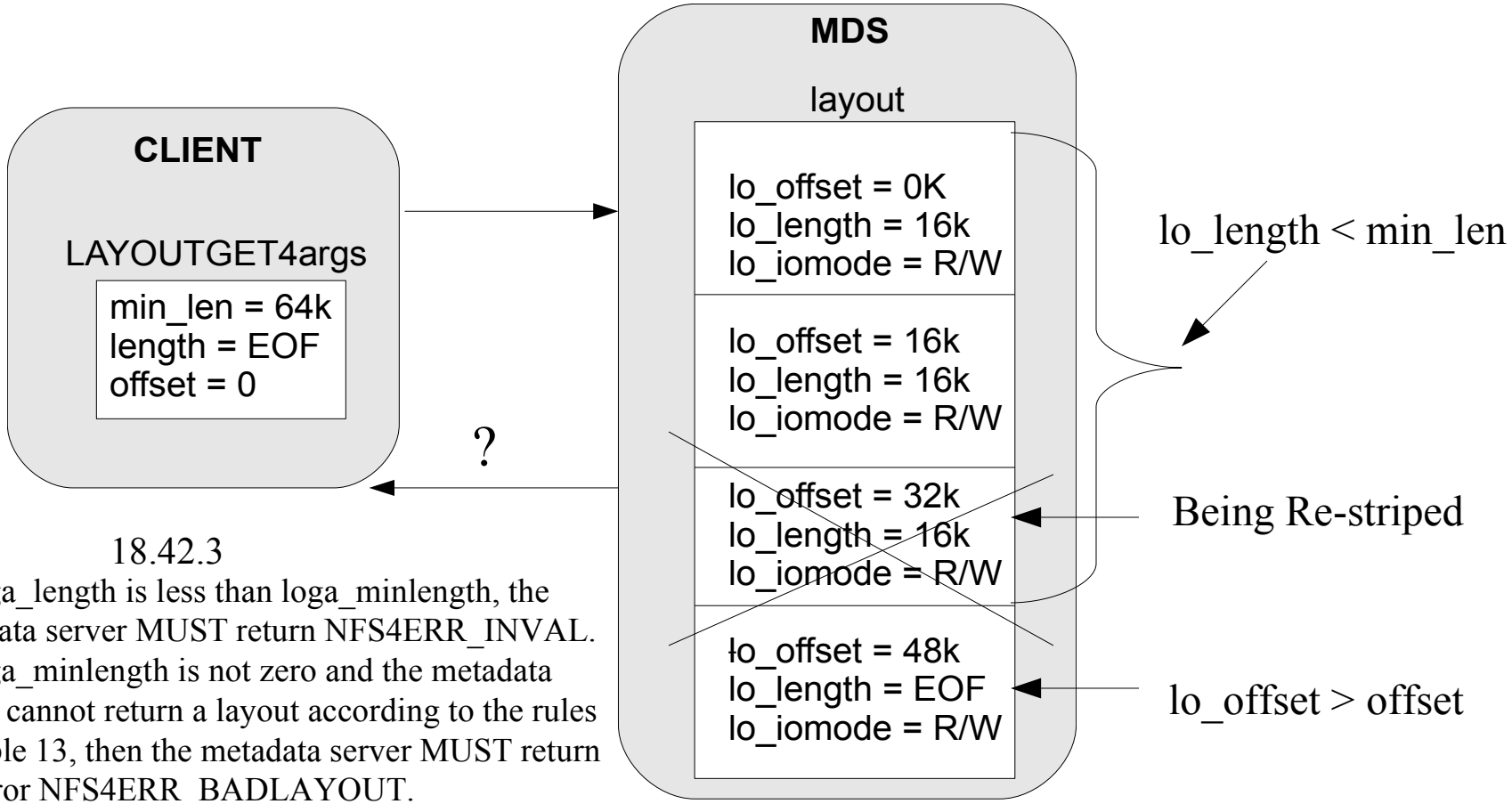


Gaps?


MIND THE GAP
Client
LAYOUTGET4resok
ROC
LOStateid
Layout_Len = 4
Layout_val
lo_offset = 0
lo_length = 16k
lo_iomode
lo_offset = 16k
lo_length = 16k
lo_iomode
lo_offset = 32k
lo_length = 64k
lo_iomode
lo_offset = 96k
lo_length = eof
lo_iomode
CB_LAYOUTRECALL
Offset=32K, Length =64

LAYOUTGET Results – Only 1 Status

NFS4ERR_INVAL Or NFS4ERR_BADLAYOUT?

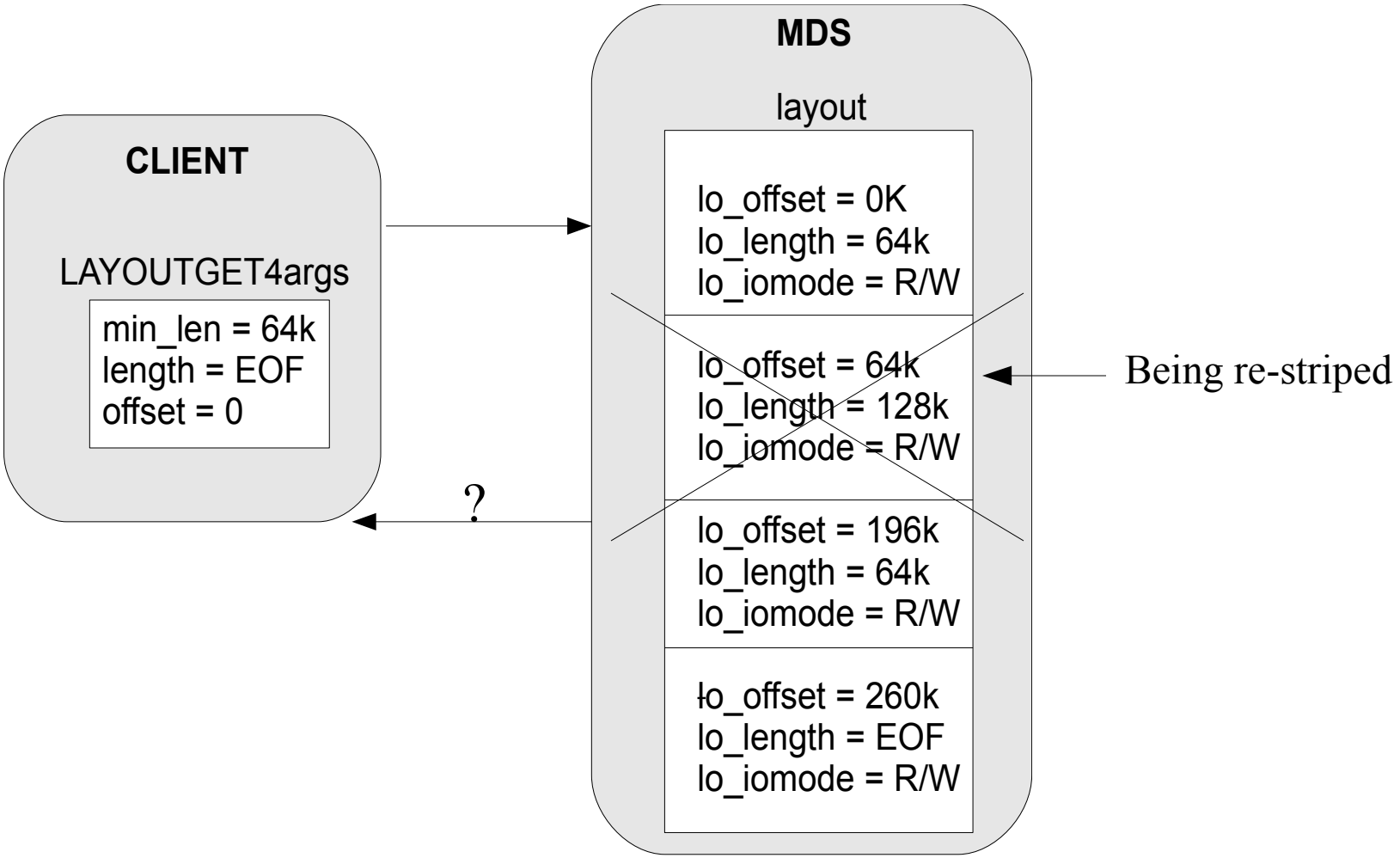


18.42.3

- If loga_length is less than loga_minlength, the metadata server MUST return NFS4ERR_INVAL.
- If loga_minlength is not zero and the metadata server cannot return a layout according to the rules in Table 13, then the metadata server MUST return the error NFS4ERR_BADLAYOUT.

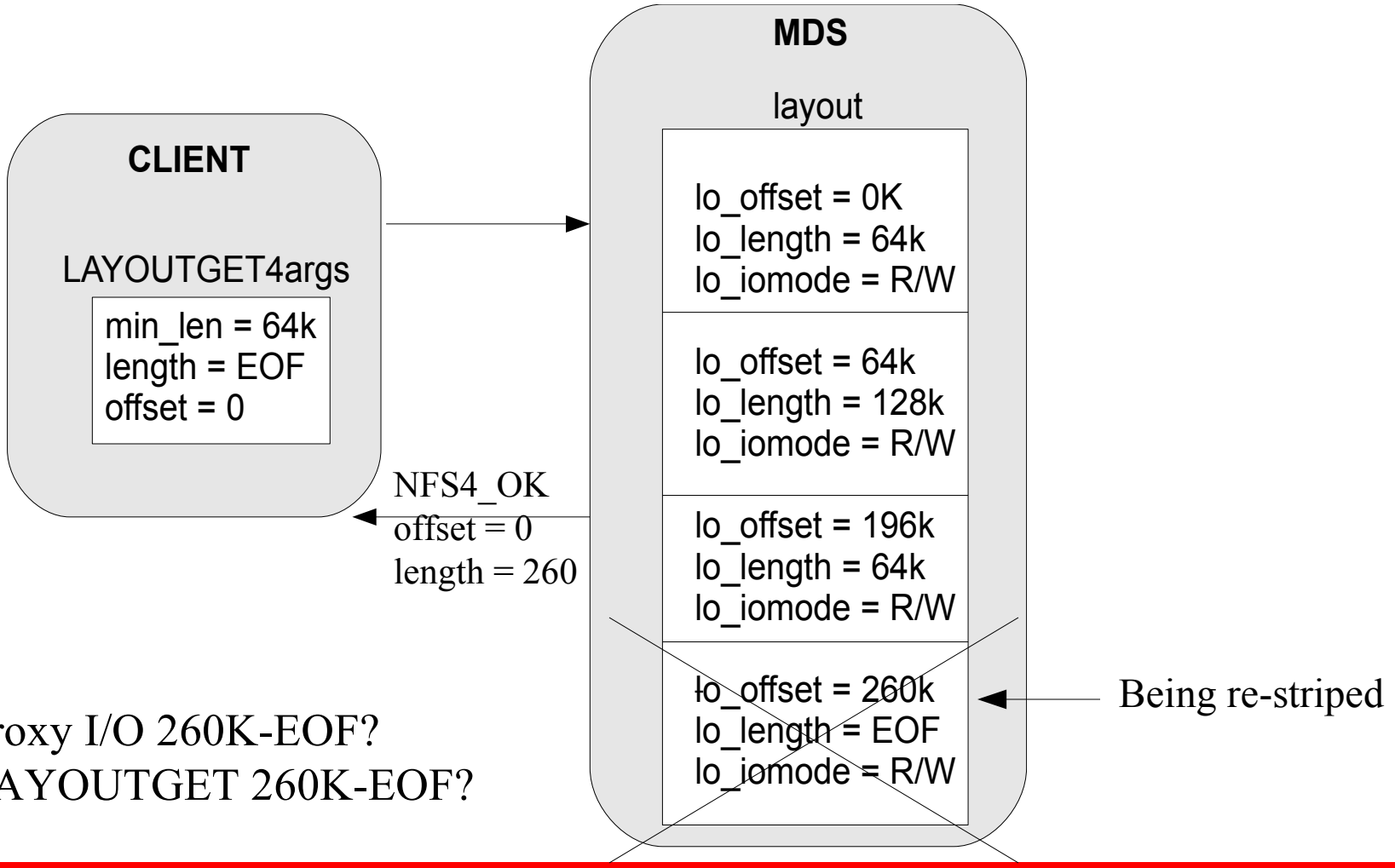
LAYOUTGET Results – Only 1 Status

NFS4ERR_TRYLATER Or NFS4_OK?

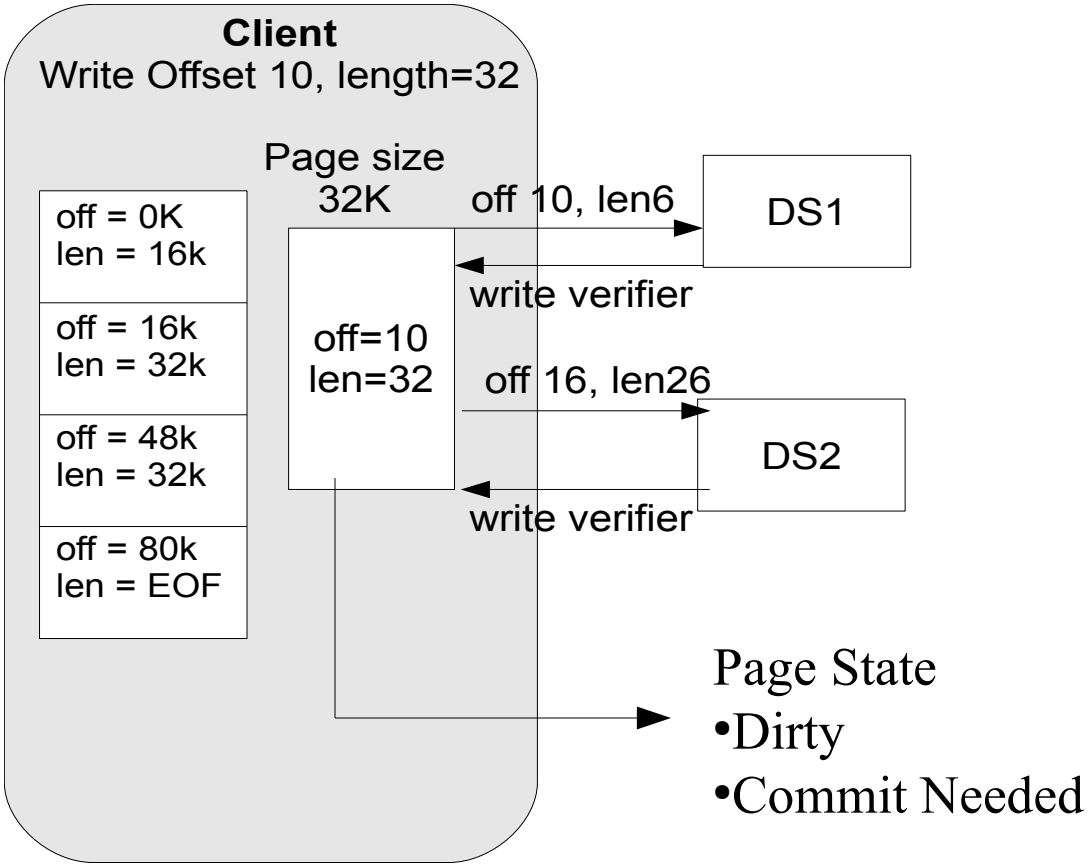


LAYOUTGET Results – Only 1 Status

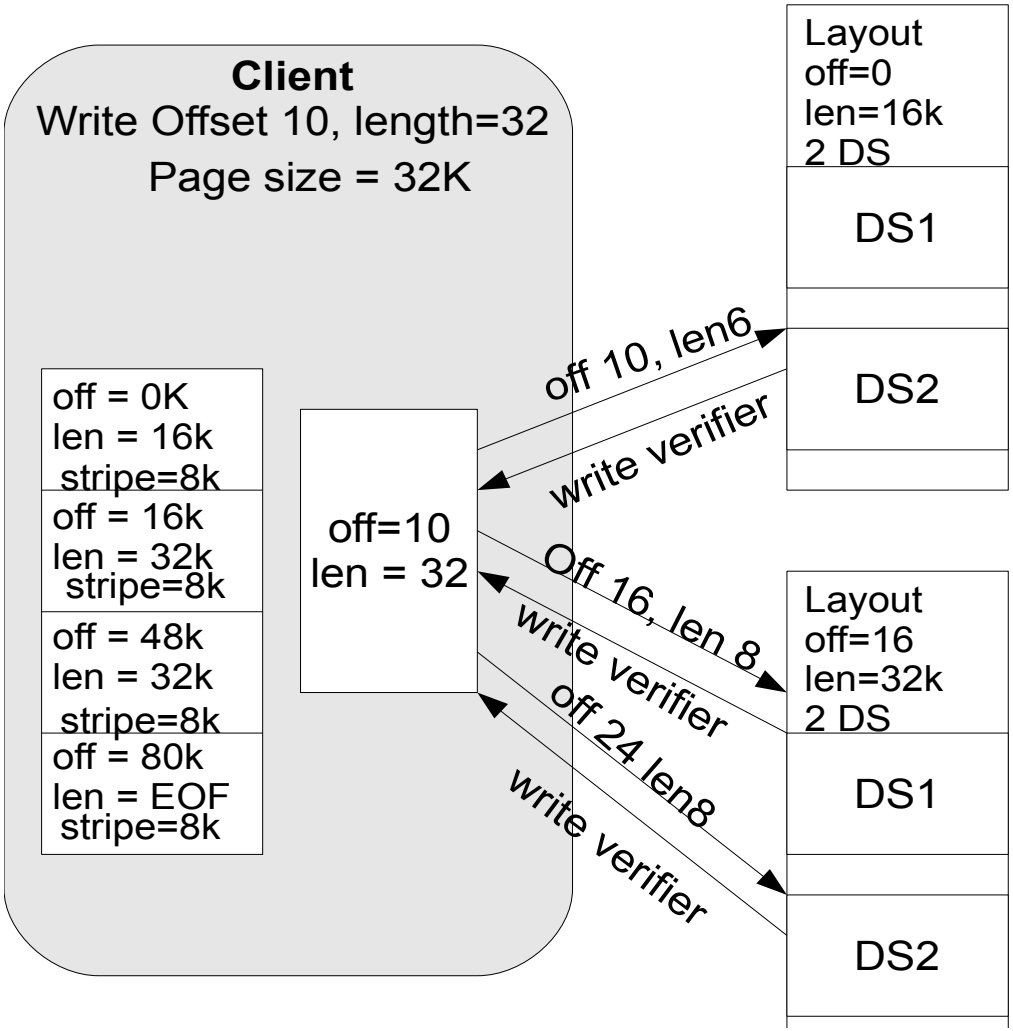
NFS4_OK, but wait, what about the missing piece?



Page size considerations



Same Issue With Stripe Size

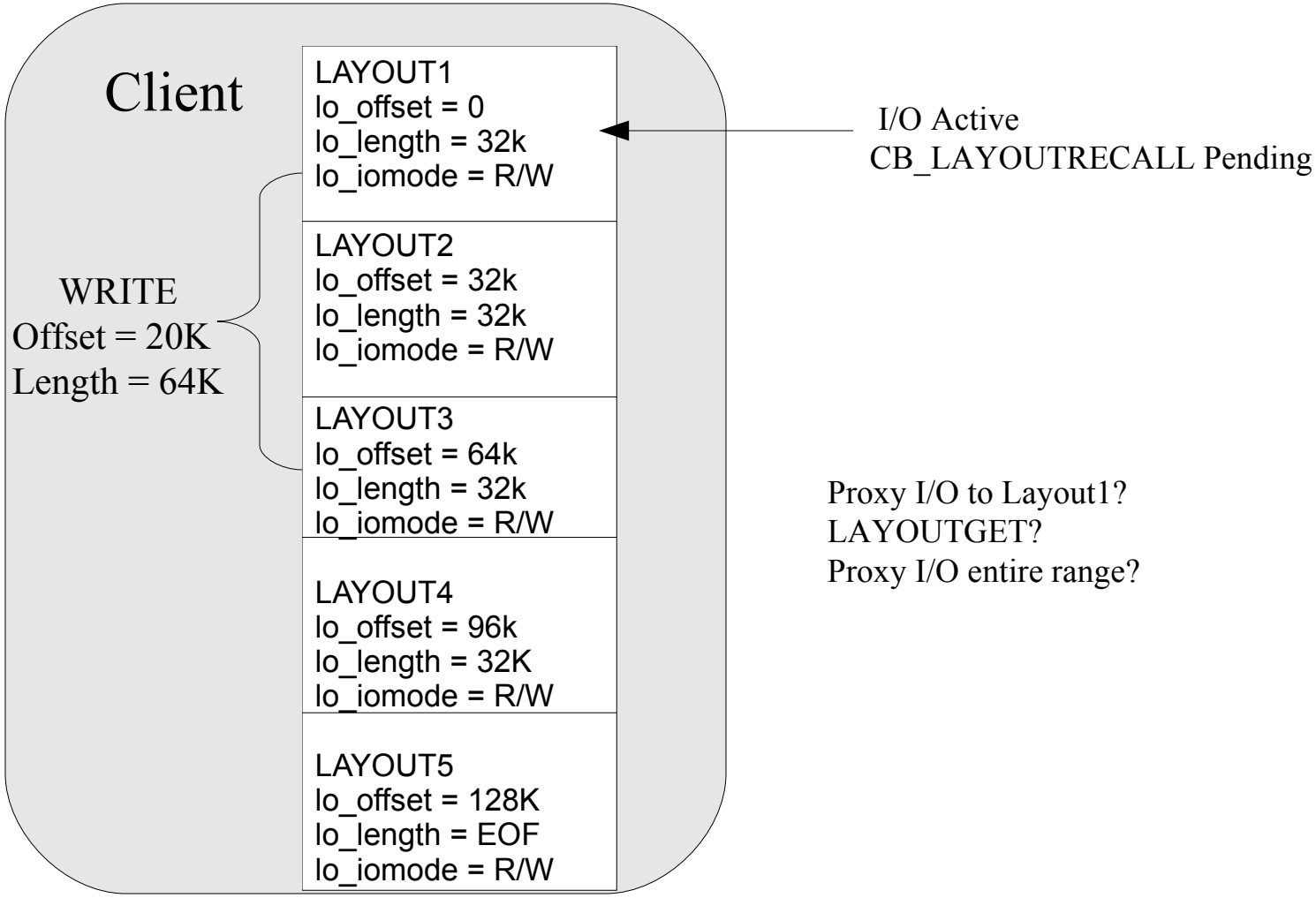


Multiple Layouts
What Does This Mean To The Client

- Needs To Maintain Layout Information Per Segment
  - Offset
  - Length
  - IOMode
  - Filehandles
  - Deviceid
  - Active/Waiting I/O
  - Layout Return Active/Waiting
  - Layout Get Active/Waiting
  - Layout Commit Active/Waiting
  - Layout Recall Active/Waiting
  - Commit Active/Waiting
- I/O which covers multiple layouts can get interesting.

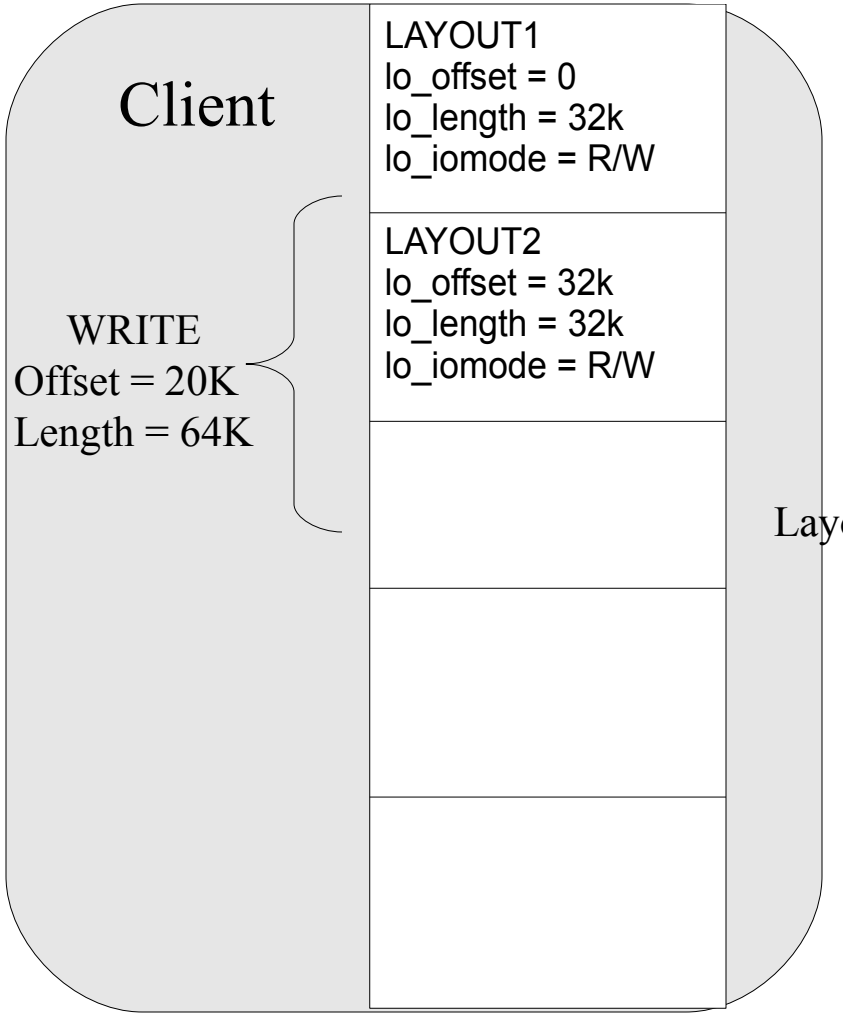
I/O Covering Multiple Segments

CB_LAYOUTRECALL Pending



I/O Covering Multiple Segments

Layout Missing



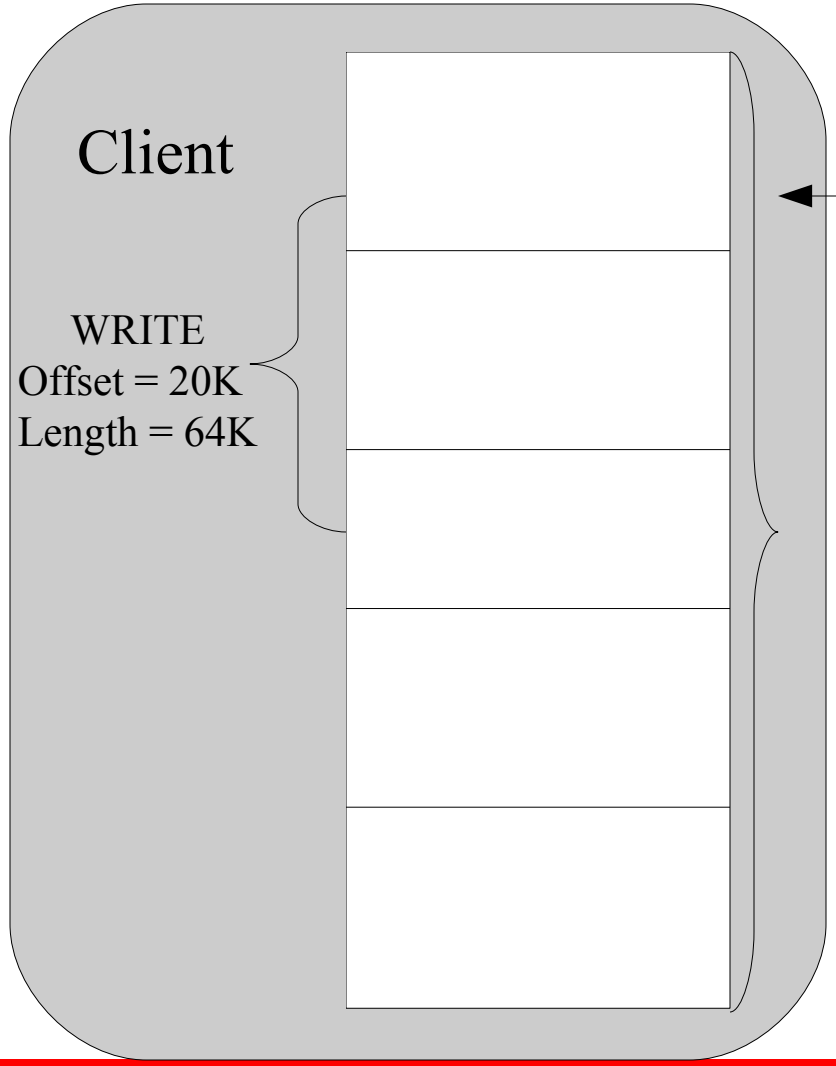
Layout Missing

LAYOUTGET?
- What range?
  - 64-EOF?
  - 64-84?
  - 0 – EOF?

Proxy I/O
- 20-84K?
- 64-84K?

I/O Covering Multiple Segments

LAYOUTGET Active



LAYOUTGET In Progress For 0 – EOF

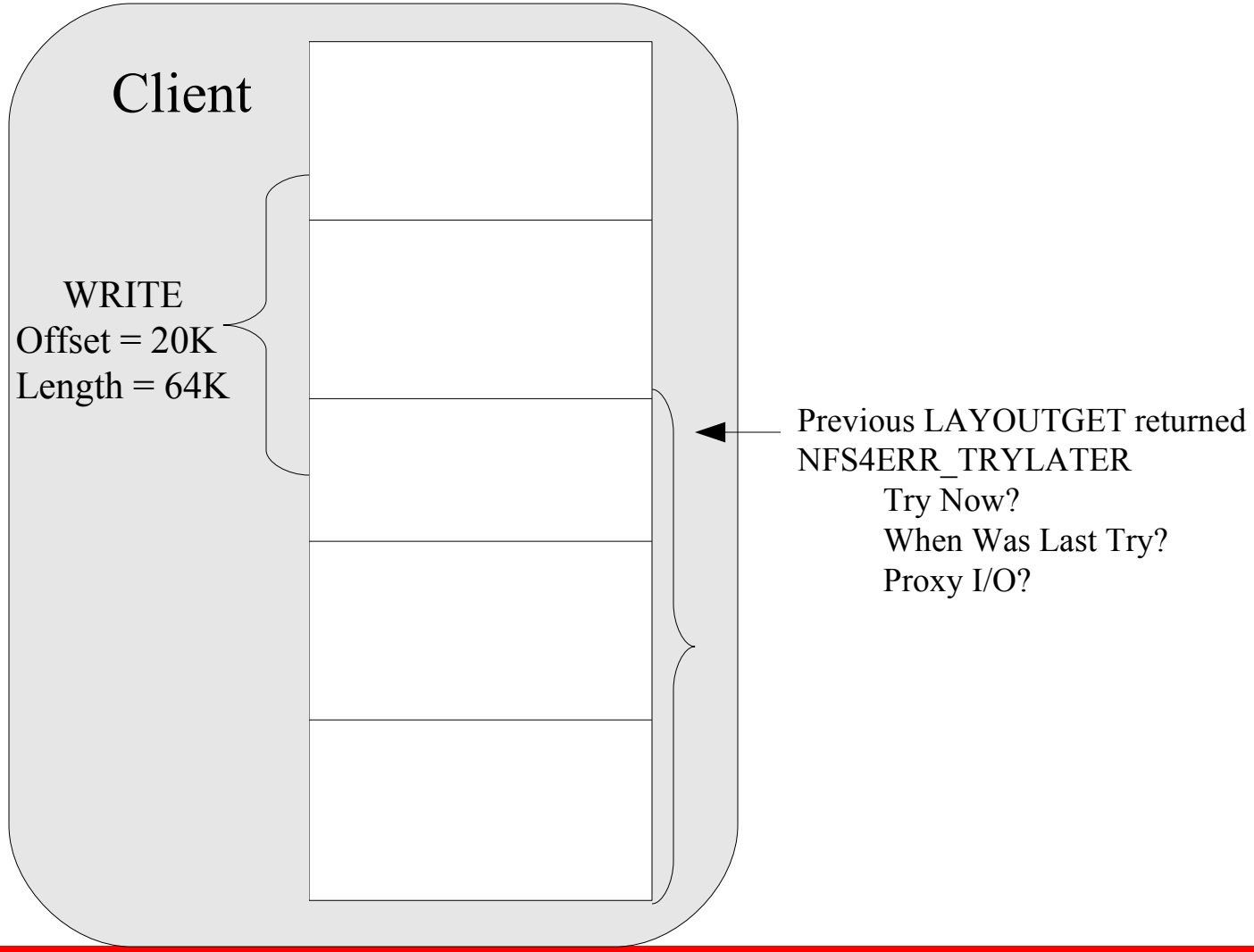
Wait?

What if Results Don't Cover range?

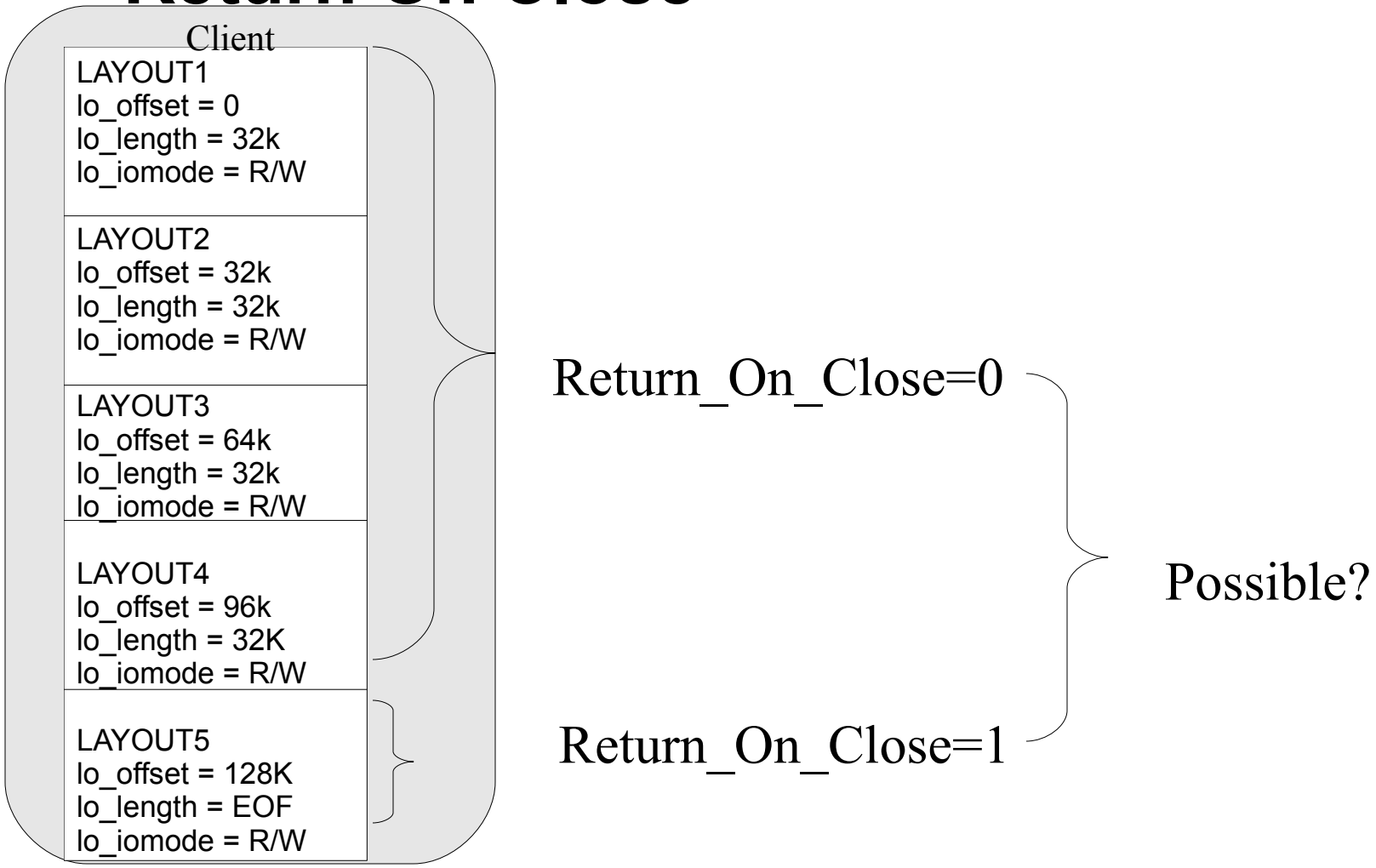
Proxy I/O Now?

I/O Covering Multiple Segments

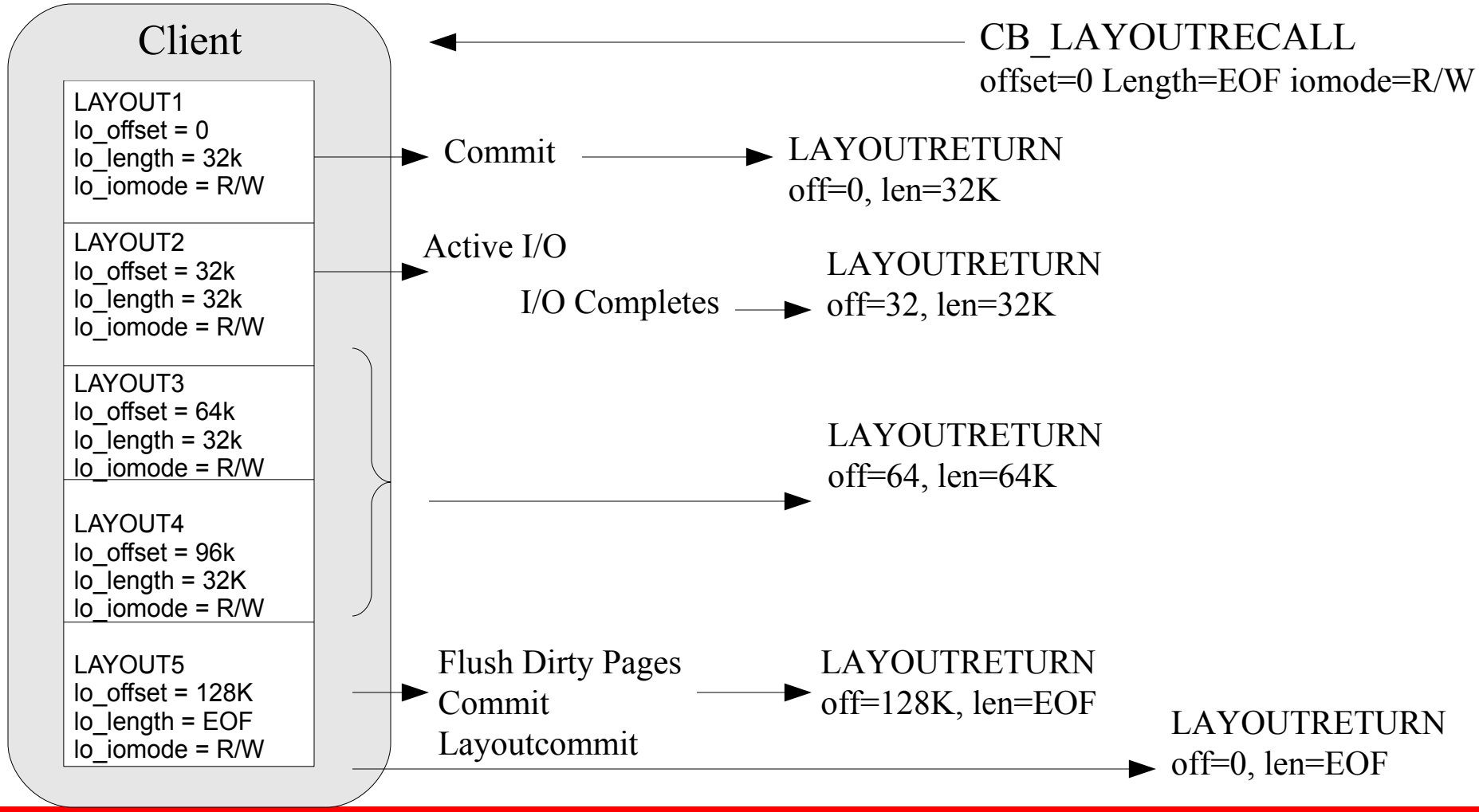
NFS4ERR_TRYLATER



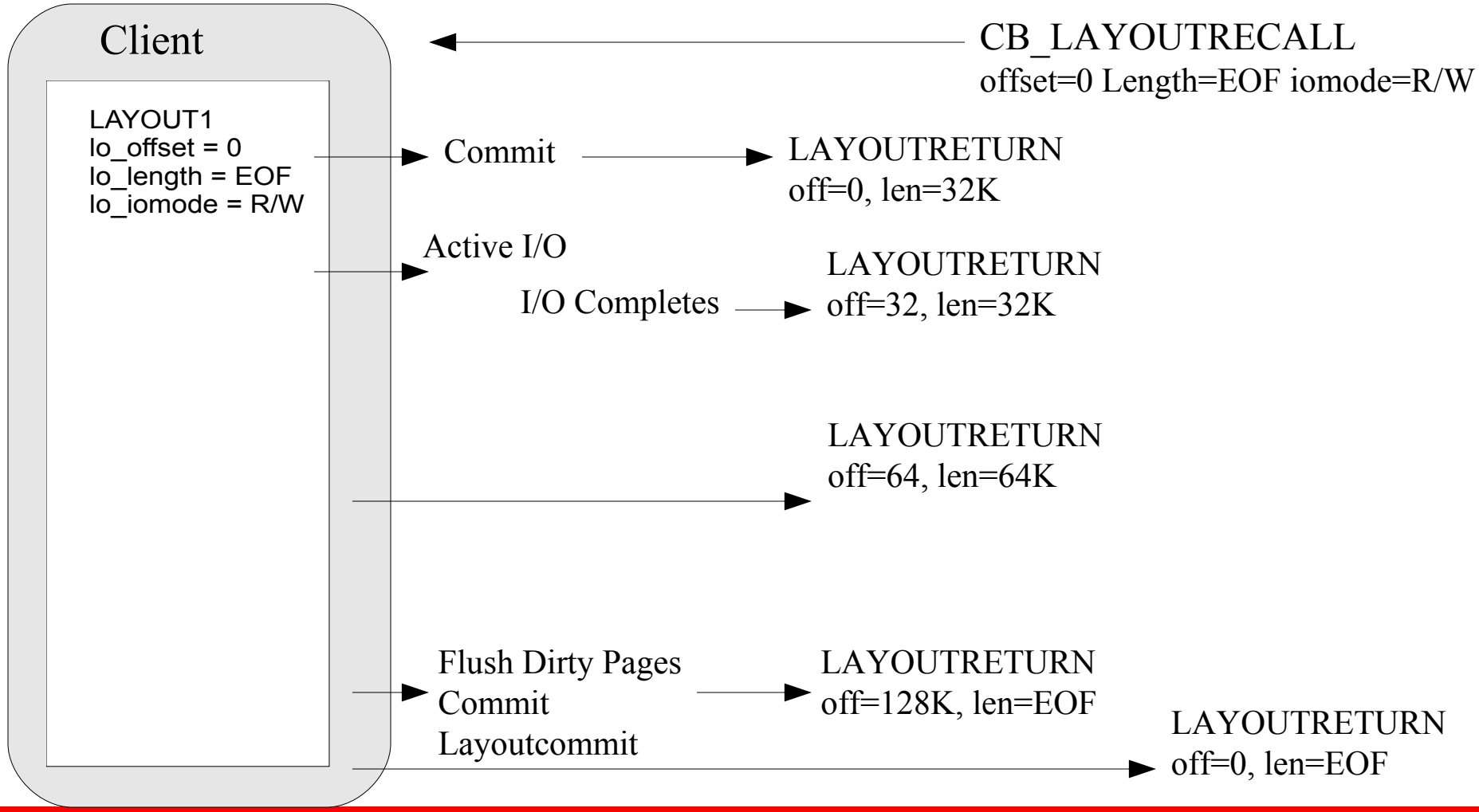
What about LAYOUTRETURN? Return On Close



CB_LAYOUTRECALL LAYOUTRETURN
More interesting



CB_LAYOUTRECALL LAYOUTRETURN Even with 1 Layout



Conclusion

- Be aware of what LAYOUTGET results can be based on arguments.
- Watch out for page boundary issues.
- Lots of considerations regarding I/O, LAYOUTRETURN, since each layout segment may have different states.
- Want to avoid it (for now?) and only deal with files that have a single layout?
  - LAYOUTGET offset = 0, length = EOF
  - LAYOUTRETURN any layout with segments
  - Must do Proxy I/O then for files with multiple segments


ORACLE®