



Go further, faster®

# LAYOUTCOMMIT considered a noise !!

Pranoop Erasani  
pranoop@netapp.com





# Agenda

- Motivation
- Proposal
- Feedback



# Motivation

- “When to do LAYOUTCOMMIT” discussions in WG
- Workings of COMMIT
  - Servers have NVRAM’s
  - Writes are always synchronous
  - Committed to NVRAM
  - Replayed regularly to disk or replayed after a crash
  - For such servers, COMMITs are a no-op
  - Servers avoid COMMIT’s by responding to WRITE’s with FILE\_SYNC



## Motivation (contd..)

- We need something like that for LAYOUTCOMMIT
- Moreover, LAYOUTCOMMIT is being looked at as facilitating close-to-open consistency
  - when, how often can dramatically increase the frequency of operation
- Presence of LAYOUTCOMMIT allows the metadata server to be somewhat out of sync with the data server
  - But, that's not a problem with some servers(e.g.,global file systems, symmetric clusters)
- Our server keeps metadata and data sync between MDS and DS all the time
- SUMMARY: LAYOUTCOMMITs are not needed for some servers



# Prior Proposal

- Talk of adding another stable\_how4 type
  - E.g pNFS\_FILE\_SYNC
- Works for me
  - Current discussions in WG indicate that is insufficient
  - But, it's sufficient for our server



## Prior Proposal (Contd..)

- Talk of using existing stable\_how4 type
  - E.g FILE\_SYNC with WRITE
- Works for me
  - Draft needs to be updated to clarify the that FILE\_SYNC applies for MDS and DS consistency as well
- Or help deal with at the NFSv4.[1|2] server level



# Proposal

- Need a sessions flag hint

- E.g `const EXCHGID4_FLAG_PNFS_NO_LAYOUTCOMMIT = 0x00080000;`



# Comments/Feedback?

