

pNFS block disk protection

IETF 81 NFSv4 WG Meeting, July 27, 2011

Sorin Faibish – EMC (sfaibish@emc.com)

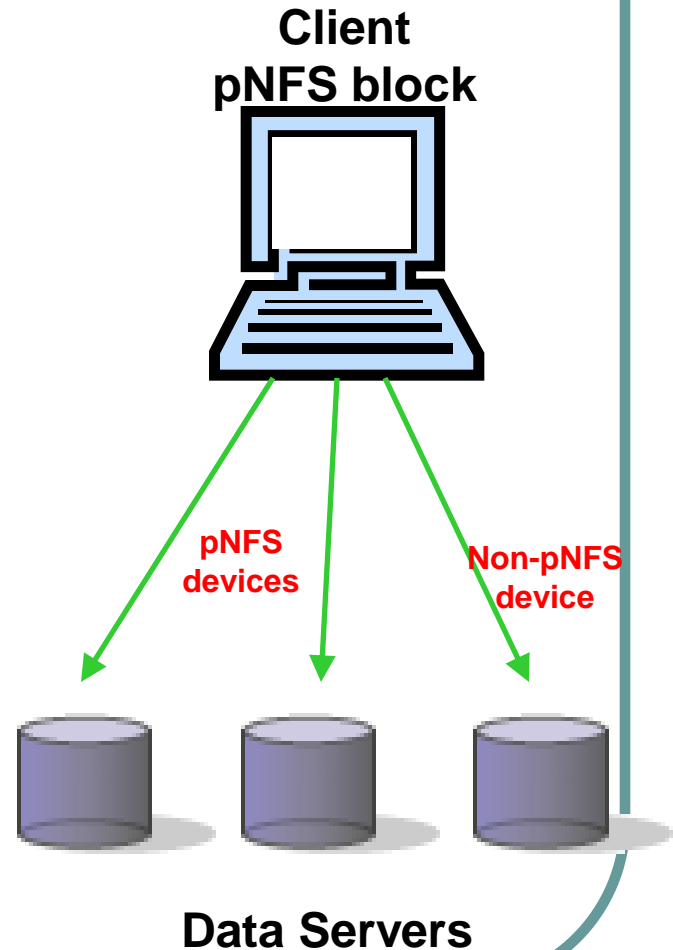
David Black - EMC

Jason Glasgow – Google

Problem Statement

- Client OS cannot identify pNFS devices from non-pNFS devices
- pNFS devices are discovered after mount time via GETDEVICEINFO
- During boot kernel/apps may write to pNFS devices and destroy pNFS FS
- There is no protocol way (5663) to report non-pNFS devices used in layouts
- Problem observed when complex volumes support was implemented

Will be a separate I-D



Usecases

- Allow protection of pNFS file system from non-pNFS access so that Clients OS MAY protect pNFS devices at boot
- Allow pNFS clients to identify unsigned volumes that are included in the layout and prevent data corruption
- Allow non-pNFS and pNFS disk partition to coexist (now must be separated)

Solution

- Use GPT with special GUID to identify pNFS devices:

http://en.wikipedia.org/wiki/GUID_Partition_Table

- pNFS Block Storage partitions are identified in the GPT with special GUID
e5b72a69-23e5-4b4d-b176-16532674fc34.
- NFS clients do not issue block I/O operations for non-pNFS access to any storage identified as pNFS Block Storage by that GUID.

Client Behavior

- Client OS will be responsible to prevent non-pNFS access to the pNFS signed devices
- Client OS will prevent pNFS access to devices without pNFS specific GPT.
- pNFS clients report an error on unsigned pNFS devices included in layouts (see Permission Access draft)
- Fallback to MDS.

Goal

- Allow to become WG draft or informational I-D?
- Addition to 5663, errata?