p4 add
p4 admin
p4 branch
p4 branches
p4 change
p4 changes
p4 client
p4 clients
p4 counter
p4 counters
p4 delete
p4 depot
p4 depots
p4 describe
p4 diff
p4 diff2
p4 dirs
p4 edit
p4 filelog
p4 files
p4 fix
p4 fixes
p4 flush
p4 fstat
p4 group
p4 groups
p4 have
p4 help
p4 info
p4 integrate
p4 integrated
p4 job
p4 jobs
p4 jobspec
p4 label
p4 labels
p4 labelsync
p4 lock
p4 logger
p4 obliterate
p4 opened
p4 passwd
p4 print
p4 protect
p4 rename
p4 reopen
p4 resolve
p4 resolved
p4 revert
p4 review
p4 reviews
p4 set
p4 submit
p4 sync
p4 triggers
p4 typemap
p4 unlock
p4 user
p4 users
p4 verify
p4 where

Perforce with Network Appliance Storage

Perforce User Conference 2001

Richard Geiger

PERFORCE

Introduction

- What is Network Attached storage?
- Can Perforce run with Network Attached storage?
- Why would I *want* to run Perforce with Network Attached storage?
- …and NetApp filers in particular?

Background

- This is from experience running Perforce at Network Appliance, 1997-2001
- Perforce replaced CVS; NetApp storage was a requirement
- At the time, network attached storage (NFS) was specifically discouraged by Perforce…
- …but they were willing to help us try

What is Network Attached Storage? (NAS)

- The filesystem resides on a device on a network (as opposed to one on disks attached directly to the application server host)
- Filesystem code runs on the NAS device
- NAS client hosts perform file level (as opposed to disk block level) operations
- Uses common standard networking media and protocols (e.g., TCP/UDP/IP, Ethernet, ATM, etc)
- Common NAS protocols
  - NFS (originated by Sun, Unix realm)
  - CIFS (originated by Microsoft, Windows realm)

What is a Storage Area Network? (SAN)

- Usually implemented with Fibre Channel connections from the application host to the disk array(s)
- Often though a back-end “network” of switched Fibre channel fabric
- The application host and SAN storage device exchange data in disk block units
- Filesystem code still runs in the application host
- BUT: There are no hard and fast definitions of NAS & SAN at this time, and some storage architectures can employ elements of both.

What About Performance?

- Networking technology has improved even faster than storage device capacities and processor speeds
- With Gb Ethernet, standard remote filesystem protocols can rival or exceed the performance of direct host-attached storage

What About Reliability?

- “Write caching can loose data”…
- …but modern servers use journals and nonvolatile caches.
- “Networks are subject to vagaries of access and capacity”…
- …but *we* control the network

The Common Wisdom:

Hah! NFS has a long and invidious history of (ahem) being a tad optimistic about what exactly has made it to disk just yet...Do _not_ trust NFS with database I/O is a fine maxim to work by.

- from a posting to the perforce-users mailing list, March 31, 1998

Network Appliance Filer Overview

- Special purpose: to store and serve files
- Multiprotocol: NFS, CIFS (and others)
- Capacities range from a few 100Gb to 12Tb
- Supports multiple interfaces of various kinds
- Built-in RAID reliability and performance
- Hot spares, hot pluggable drives & shelves
- Dynamically expandable filesystems
- Simple administration & OS upgrades
- Instant filesystem snapshots

A Network Appliance Filer

Perforce Server Storage

- All of the data kept by the Perforce server side
- Three categories of server files:
  - Metadata Database: `$P4ROOT/db.*`
  - The file archives: `$P4ROOT/depot/...`
  - Other: log, journal, checkpoints, license, etc.

What Can (and Should) Be Stored on a Filer?

- I believe that *all* Perforce server data can be kept on a filer, when properly configured
- Network Appliance has in fact done so, from the beginning
- We trusted RAID to protect against drive failures, and used a single filesystem for everything

`p4d` with NFS

- We only ever saw one problem, journal (non) locking, fixed in `p4d` r99.1
- We ran both under Solaris and (more recently) OSF1 (DEC, Compaq) AlphaServer

`p4d` with CIFS

- NetApp did *not* use this configuration
- Seemed to work OK, in limited experiments I've tried
- Unlike NFS, CIFS is "stateful", so special attention may be required when recovering from server reboots or network outages.

Network Configuration

- Proper network topology and configuration is essential to getting good performance
- “Network” can mean a small private dedicated network (or subnet) between a file server and application server host(s)

Back-to-back Gb Ethernet

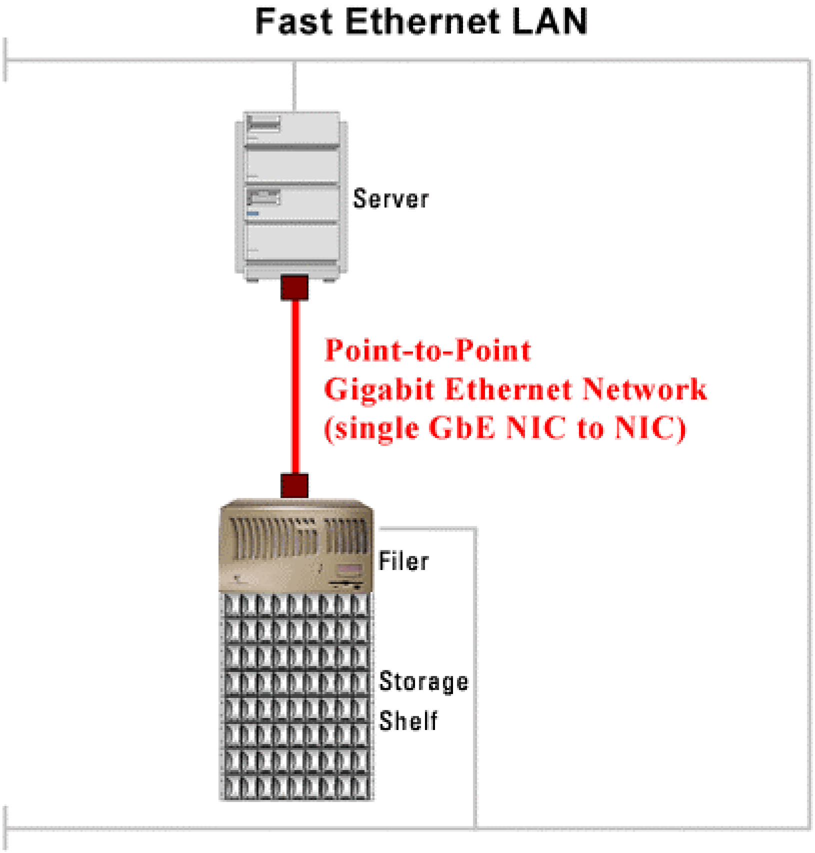

Illustration used by permission of Network Appliance, Inc.

Gb Ethernet SAN

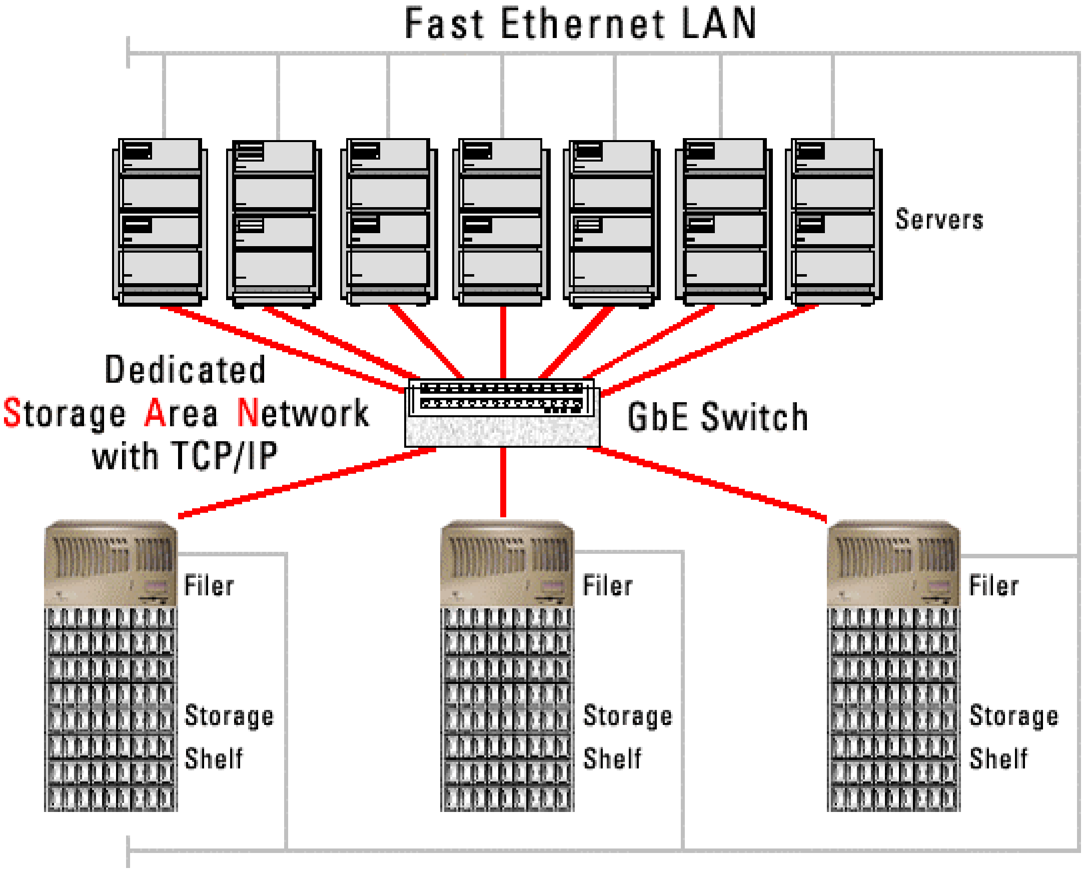

Illustration used by permission of Network Appliance, Inc.

`snap_checkpoint` ing

- Perforce checkpoints make the server unavailable while the checkpoint is written
- Large depots can make long checkpoints
- NetApp filers support filesystem snapshots
- With snap_checkpoint, the server is locked only for the few seconds required to create a NetApp snapshot
- The checkpoint is then created from the snapshot

NetApp Perforce Server Configuration

Server Host

- COMPAQ AlphaServer DS20E 500 MHz (1 cpu)
- 4 Gb memory
- Dual Gb ethernet interfaces
- OSF1 V4.0 1530

NetApp Perforce Server Configuration

Storage

- Network Appliance F760 filer
- 1 Alpha CPU
- 1 Gb memory
- 42 9Gb disks totaling 378Gb (raw)
- Dual Gb ethernet interfaces
- Two filesystems (1 for Perforce, one for defect tracking)
- 170Gb filesystem dedicated to Perforce
- NetApp Release 6.1.1R1
- NFS V3 UDP mounts

NetApp Perforce Server Configuration

Perforce Server

- Server version: P4D/OSF/2000.2/21189 (2001/03/14)
- Server license: Network Appliance 345 users
- Number of files in the depot: 877,241
- Number of client specs: 7,088
- Number of change lists: 127,230
- Number of job records: 52,418
- Metadata size: db.have: 9.2Gb; db.*: 10.2Gb
- Number Data ONTAP/Netcache source branches: 131
- Files in the main branch: 10267

Filer Tuning

- Nothing special
- UDP is fine for reliability and performance

Summary: Advantages of Perforce (server)-on-NetApp

- Of interest to large Perforce sites
- Filesystems can expand dynamically
- RAID protects against disk failures
- Snapshots allow fast-checkpointing
- Allows "hot spare" server
- SnapRestore can help with testing
- Other data management tools

Performance Measurements

- There are no established Perforce benchmarks
- So I invented one
- Measurements demonstrated performance with filer-based storage (Gb Ethernet) generally as faster or faster than local disk.
- (See the paper and references for more details)
- There's much room for better metrics and tools!

Known Issues: Server Type

- Solaris and OSF1 are known to have worked well in production environments
- ***BREAKING NEWS:*** With Solaris, be sure to place journal on local storage or turn off network locking with `–llock` mount option!

Known Issues: Other

- Report of slow database rebuilds (could be a network locking effect)
- Long-running `p4 client -d` (probably mostly a network locking effect)
- `snap_checkpoint` with NetApp R6.0 and later releases will need a new version of `p4d`

Client storage

- Just do it!
- Having Perforce clients on a network filesystem allows use of the same client workspace from many hosts
- Having Perforce clients on a multiprotocol filesystem can be interesting
  - Can use “LineEnd: share”

Further Work

- We can always use more knowledge about what platforms and configurations work well
  - In particular, it would be good to clarify the picture WRT Linux and FreeBSD server hosts
  - Information about other NAS servers would be useful.
- More and better tools for analyzing Perforce performance