Database Performance on NAS:
A Tutorial

Darrell Suggs

Overview

- **Database on NFS – Today's Reality**

- **Retrospective**

- **High Level Performance Analysis**

- **Real State-of-the-Union**

- **Keys to Future Success**

- **Brave New Worlds**

- **Simple Data Points**

- **Q&A**

Network Appliance - Darrell Suggs

NetApp®

Database on NFS

- **Today's Reality**

  - **This is not fiction!**

  - **Real customers**

  - **Real deployments**

  - **Real business critical applications**

  - **Large number of deployments & growing**

- **Database on NFS IS a Datacenter Solution**

Network Appliance - Darrell Suggs

NetApp®

Database on NFS

- **Consider Oracle**

  - **Dominant presence in business databases**

  - **Beats the NFS/Oracle drum often and loud**

  - **Specifically,**
    - **Oracle + NFS + Linux + RAC**

- **Customer visits Oracle World**

  - **Returns asking for NFS/Oracle presentations**

- **Many Oracle + Vendor + NetApp examples**

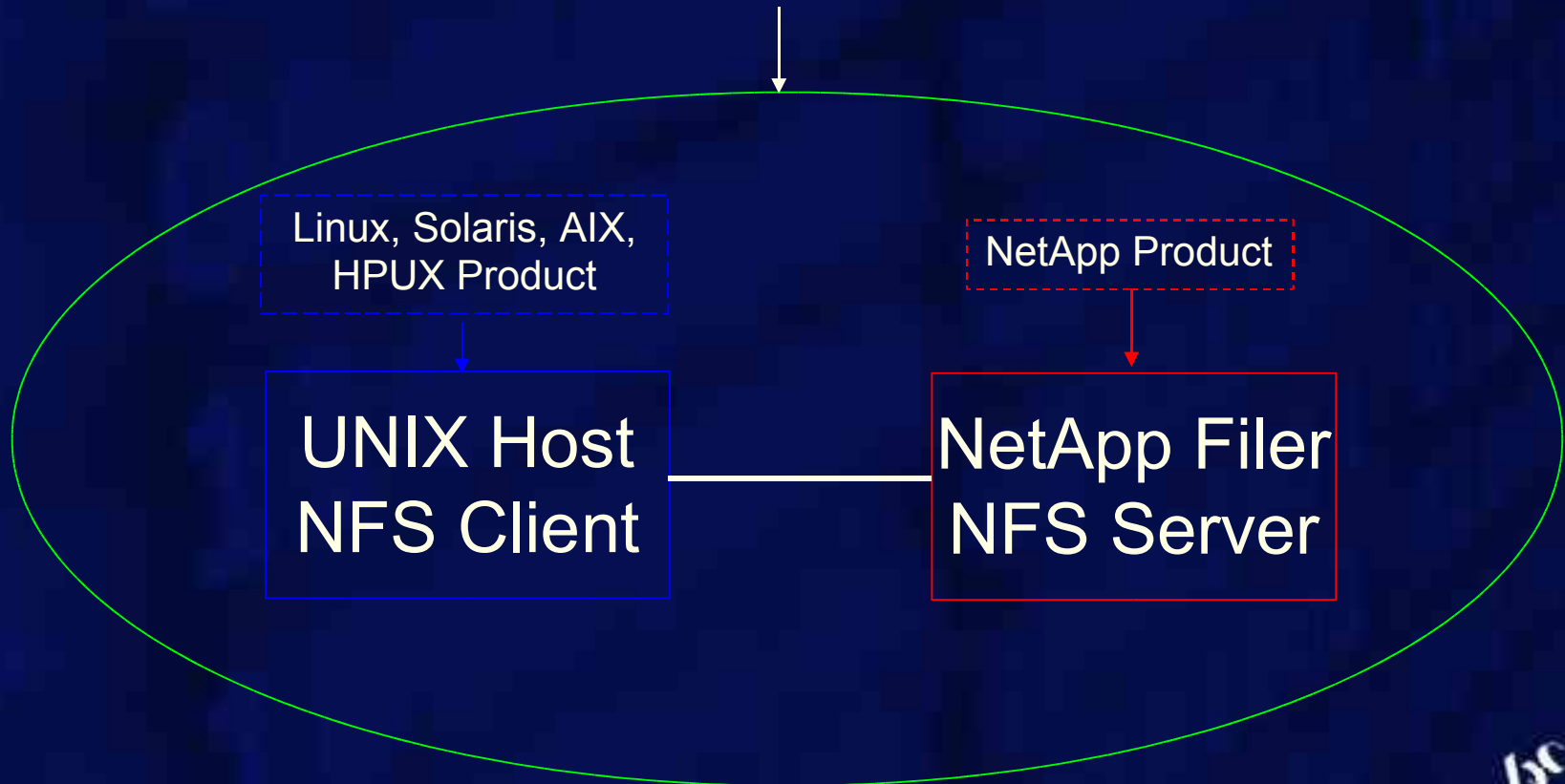 - **Worked with many of you on specific deployments**

Network Appliance - Darrell Suggs

NetApp

Database on NFS

- **Why the NFS + DB combination?**

- **NFS is simply compelling**

  - **More manageable**
    - **Expansion, movement, and backup of files**
  - **Cost effective**
    - **People, Tools, and Infrastructure**
  - **Sharing paradigm**
    - **Makes deployment and maintenance easier**
  - **Also fits new server technologies well**
    - **E.g. Blades + NFS are a heaven-made-match**

Network Appliance - Darrell Suggs

NetApp®

Retrospective

- ## Why we care

**What the Customer Purchases and Deploys
An NFS Solution**

Linux, Solaris, AIX, HPUX Product

NetApp Product

UNIX Host NFS Client

NetApp Filer NFS Server

Network Appliance - Darrell Suggs

Retrospective

- **Beepy (NetApp) - NAS 2003**

  - **Database workload is challenging, details matter**

  - **Networking is Cool**

  - **Database + NFS → Good match**

  - **//http://www.nasconf.com/pres03/beepyondatabase.pdf**

- **Colaco/McDougall (Sun) - NAS 2003**

  - **Database on NFS is compelling**

  - **http://www.nasconf.com/pres03/colaco.pdf**

Network Appliance - Darrell Suggs

NetApp

Retrospective

- **Suggs/Daniel (NetApp) – Connectathon 2003**
  - **NFS Clients have issues**
  - **Database on NFS is compelling**
  - **http://www.connectathon.org/talks04/suggs.pdf**
- **Colaco (Sun) / Suggs (NetApp) – Joint Whitepaper**

  - **Oracle on Solaris/NFS/Filer – Best Practice**
    - **http://www.sun.com/bigadmin/content/nas/sun_netapps_rd**
    - **or**
    - **http://www.netapp.com/tech_library/ftp/3322.pdf**
- **Other best practices guides coming**

Network Appliance - Darrell Suggs

NetApp

High Level Performance Analysis

- **Performance Dimensions of Interest**

    - **Ethernet vs Fibre Channel Throughput**
        - **Ethernet = 1 Gb → 10 Gb**
        - **FC = 2 Gb -> 4 Gb**

    - **NIC vs HBA**
        - **Typical NIC ~15,000 IOPs**
        - **Typical HBA ~25,000 IOPs**

    - **CPU cost of NFS vs Local FS**
        - **Consider a well behaved NFS client**
        - **Host CPU cost of NFS ~= 2X to 4X Local FS**

    - **Latency**
        - **Latency differences are typically measured in uSec**
        - **Average storage latencies are measured in mSec**

Network Appliance - Darrell Suggs

NetApp®

High Level Performance Analysis

- **Summary of Performance Differences**

    - **Ethernet vs Fibre Channel Throughput**

        - **Wire speed differences easily leveled out (multiple wires)**

    - **NIC vs HBA**

        - **HBA IOPs are higher, but are rarely an issue**
        - **Multiple NICs are common**

    - **CPU cost of NFS vs Local FS**

        - **I/O cycles are only a percentage of CPU load**
        - **Database servers operate with plenty of headroom**

    - **Latency**

        - **Latency is a storage subsystem issue, not a protocol issue**

Network Appliance - Darrell Suggs

NetApp

High Level Performance Analysis
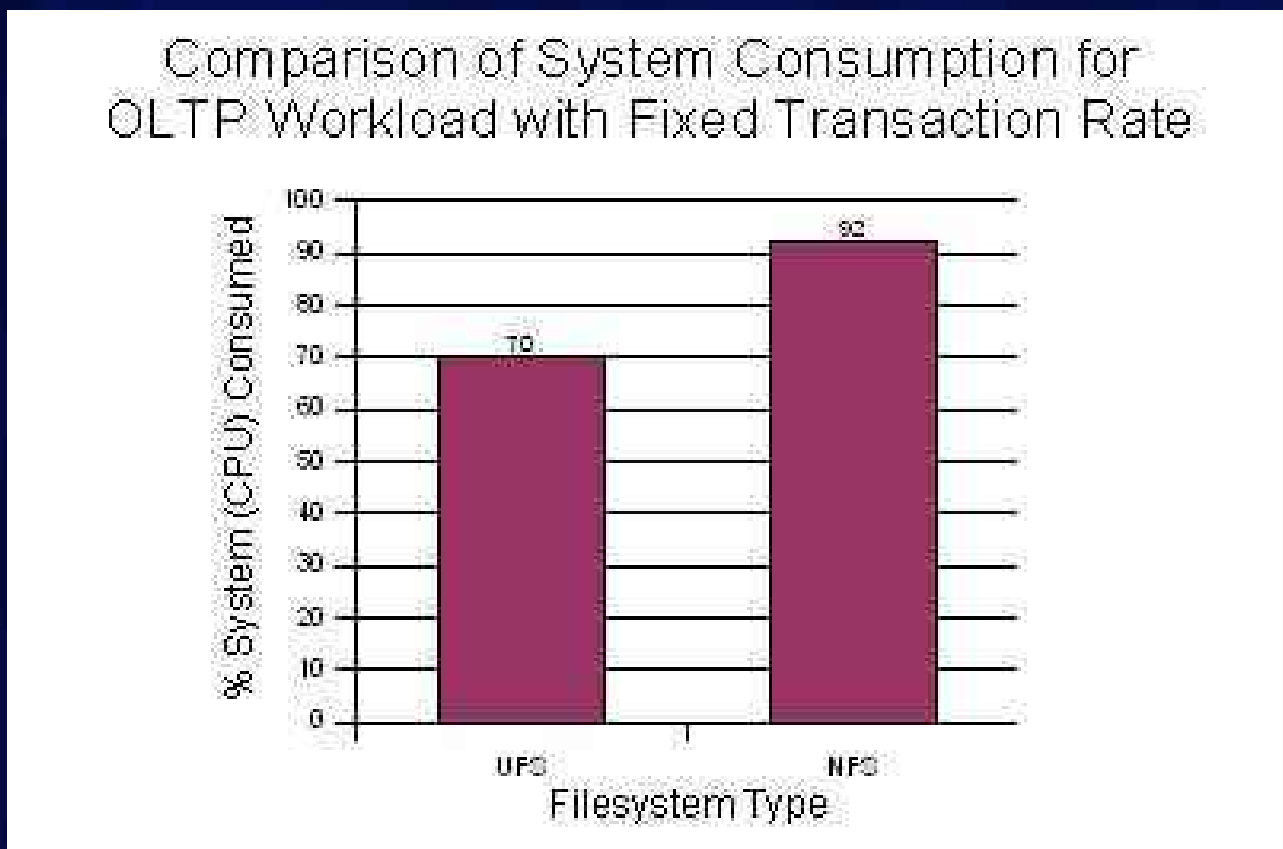
- **NFS vs iSCSI vs FCP**

    - **iSCSI is blocks on Ethernet**

        - **Cost advantages of Ethernet**
        - **Performance characteristics of Ethernet**
        - **Host CPU similar to NFS for Software iSCSI**
        - **Host CPU similar to FCP for Hardware iSCSI**

    - **Continuum**

        - **DB on NFS: Good Cost, Files Manageability**
        - **DB on iSCSI: Good Cost, Blocks Manageability**
        - **DB on FCP: Poor Cost, Blocks Manageability**

Network Appliance - Darrell Suggs

NetApp

State-of-the-Union

- **There exist two scenarios today:**

  - **Deployments Done Right**
    - **E.g. the Sun/NetApp Best Practice paper**
  - **Deployments Done Wrong**
    - **Real customer mistakes**

- **Our challenge (NAS industry)**

  - **Increase the first and decrease the second**

Network Appliance - Darrell Suggs

NetApp®

State-of-the-Union

- **The good deployment story**

Network Appliance - Darrell Suggs

NetApp®

State-of-the-Union

- ## The bad deployment story

Database Server	SAN/FC/SCSI	Cheesy Local Storage
Local FS		

Performance
-Some metric with Speed X
-    *TPM or User Latency or Wall clock time*

Customer Consumes the "NFS Kool-Aid"

Database Server	GigE/IP	Beefy NFS Server
NFS		

Performance
-*Speed is now 2 to 10 times WORSE !!!*

-*Steve and I usually get a phone call…*

State-of-the-Union

- **What causes the bad story?**

  - **Local FS is well tuned environment**

  - **NFS**

    - **Old versions, no patches**
      - **It works, what's the issue?**
    - **Default mount options**
      - **Didn't need to specify mount options for Local FS**
    - **No baseline testing for performance/network**
      - **The mount worked, must be all set!**
    - **Perception**
      - **Of course NFS is a poor performer, everyone knows that.**

Network Appliance - Darrell Suggs

NetApp

State-of-the-Union

- **Why does the bad story happen? Is NFS a bad solution?**

- **Absolutely not!**

  - **NFS began with a specific mission**
    - **Semi-wide area sharing**
    - **Home directories and shared data**
  - **Note: problems are NOT with NFS protocol**

- **The Mission has changed!**

  - **Narrow sharing environment**
  - **Typically dedicated (often p2p) networks**
  - **Data sharing → High-speed I/O Interconnect**
  - **Mission evolved to Mission Critical Workloads**

- **Actually, NFS has done ok**

  - **Credit a strong protocol design**
  - **Credit decent engineering on the implementations**

Network Appliance - Darrell Suggs

NetApp®

Keys to Future Success

- **Components of Successful Deployments**

    – **Strong NFS Server (of course)**

      • **Rarely an issue**

    – **Strong NFS Client (less obvious)**

      • **Multiple aspects of a good client**

    – **Strong Best Practices guides (just beginning)**

      • **Imperative for customer success**

Network Appliance - Darrell Suggs

NetApp

Keys to Future Success

- **Remaining Challenges**

    - **Perception. This is still a big challenge**
        - **Many sales/support folks still say NO.**
    - **Out-of-box success.**
        - **Local FS just works well with no tuning.**
        - **NFS/Database deployments often falter.**
    - **Guidelines for deployment.**
        - **Clear deployment best practices are imperative.**

- **Reflect on positive**

    - **Much progress made in the last couple of years.**

    - **v4 provides opportunity, perception, and vehicle for improvements.**

Network Appliance - Darrell Suggs

NetApp

Keys to Future Success

- **Client Challenges**

  - **NetApp will deliver a strong server☺**

  - **Clients must deliver**
    - **Strong Out-of-box behavior**
    - **Flexible semantics (cache/lock) via mount options**
    - **Strong wire efficiency (app request == wire transfer)**
    - **Scalability (single mount point, multiple wires, etc)**
    - **Reasonable per IO CPU cost vs Local FS (e.g. 2X)**
    - **Guidelines for good deployments**

  - **v4 success helps**
    - **Perception + Reality of improvements**

Network Appliance - Darrell Suggs

NetApp®

Brave New Worlds

- **NFS v4**

  - **Demonstrates continued growth of NFS**

- **Blades**

  - **Most natural deployment environment**
    - **iSCSI boot each blade over Ethernet**
    - **NFS environment for data sharing**
    - **Uses same shared Gigabit Ethernet infrastructure for both**

- **Oracle RAC**

  - **Demands NFS for manageability**

  - **Demands Local FS semantics (e.g. no caching) PLUS Shared paradigm**

- **Hardware vendors and application vendors**

  - **Herding customers toward NFS**

  - **NFS simply needs to take advantage of the opportunity**

Network Appliance - Darrell Suggs

NetApp

NFS CPU Cost Comparison

- **Consider two Unix's**

  – **4K Read I/O's, 16 concurrent threads**

  – **Compare Local FS vs NFS, CPU Cost per I/O**

  – **Normalized**

CPU Cost of I/O

Network Appliance - Darrell Suggs
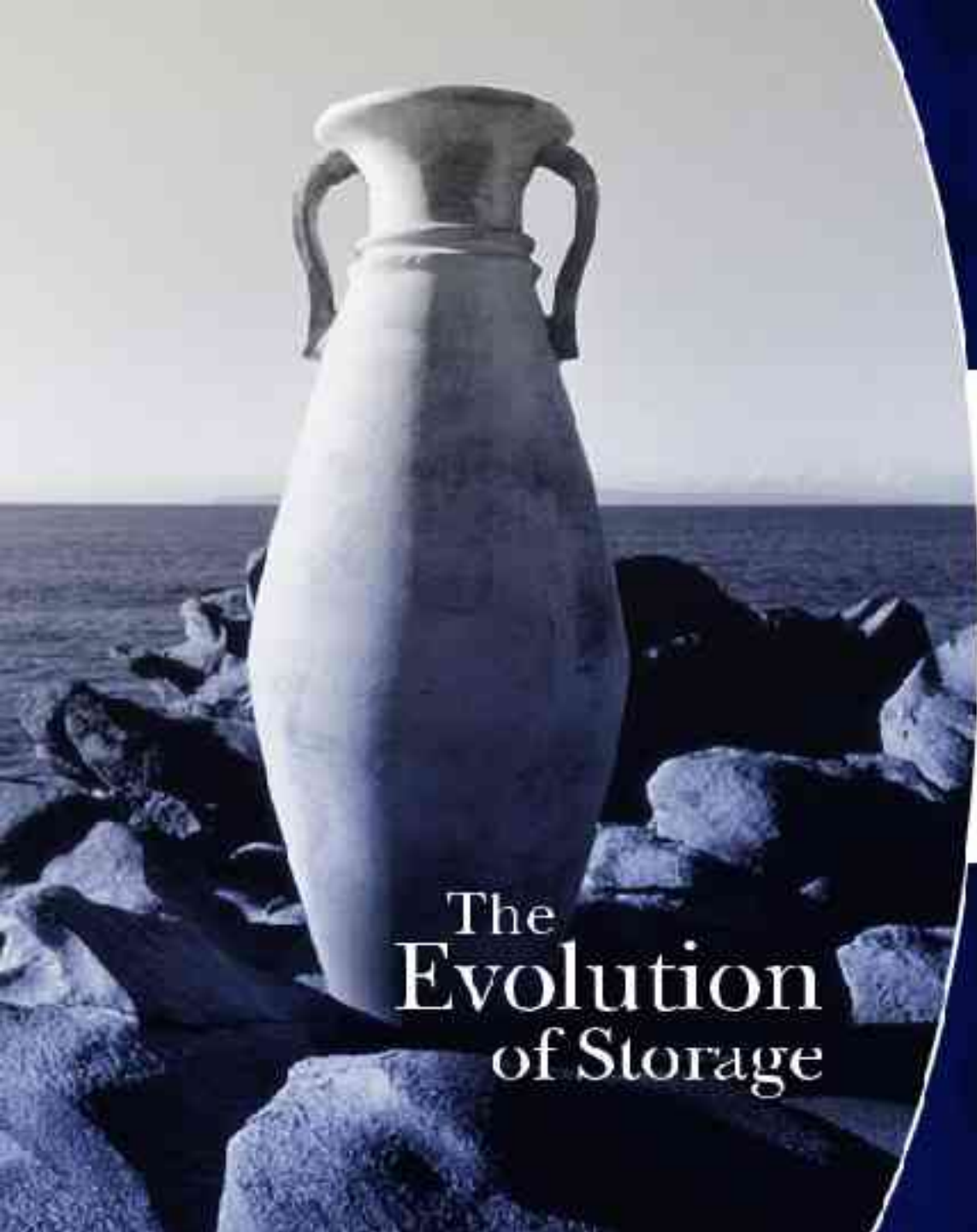
NetApp®
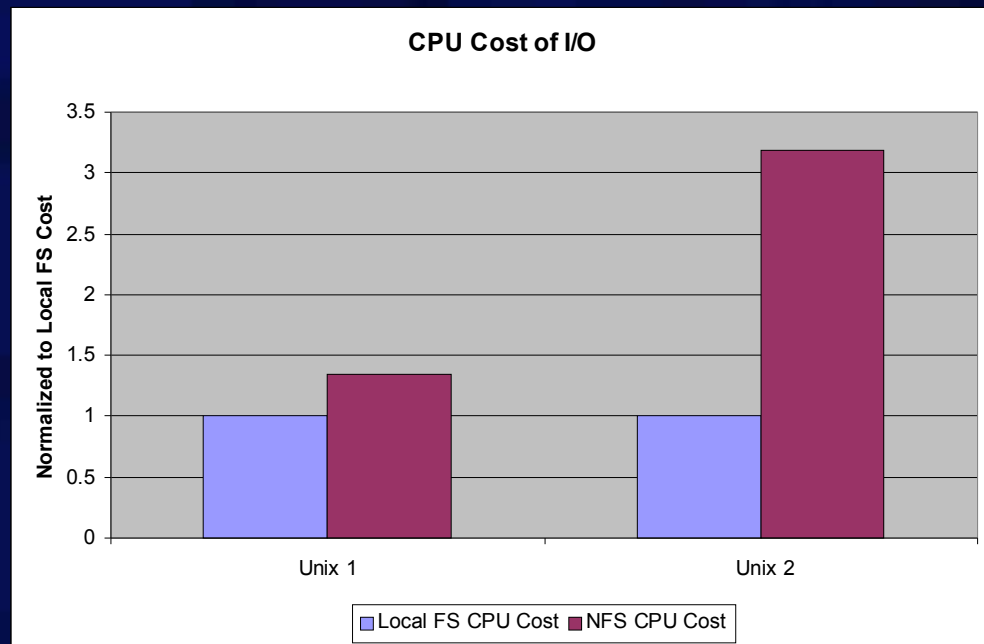
Summary

- **Database on NFS is a reality in today's market**

- **NFS is NOT a poor performing paradigm**

- **Good deployments require effort**

- **Success in DB deployments is good for all**

- **Ongoing improvements are important**

- **Best Practice guides are imperative**

- **The future of Database on NFS is bright.**

Network Appliance - Darrell Suggs

NetApp®

The
Evolution
of Storage

NetApp®