NFSv4.1 Architecture and Tradeoffs in Windows Server 2012

Roopesh Battepati
Principal Development Lead

Mike Johnson
Senior Developer

Microsoft Corporation

Agenda

- What's new in NFS in Windows Server 2012?
- High Level Windows NFS Server Architecture
- RFC Tradeoffs in Windows NFS Server
- User Account Mapping
- Pseudo File System
- Volume Mount Points
- Supported File Systems
- Diagnostics
- Q & A

What's new in NFS in Windows Server 2012?

- New RPC/XDR Implementation
  - Winsock Kernel Based
  - Auto-tuned
- Flexible User Account Mapping Support
  - passwd/group file mapping
- New Powershell Support for User Identity Store Management:
  - Active Directory
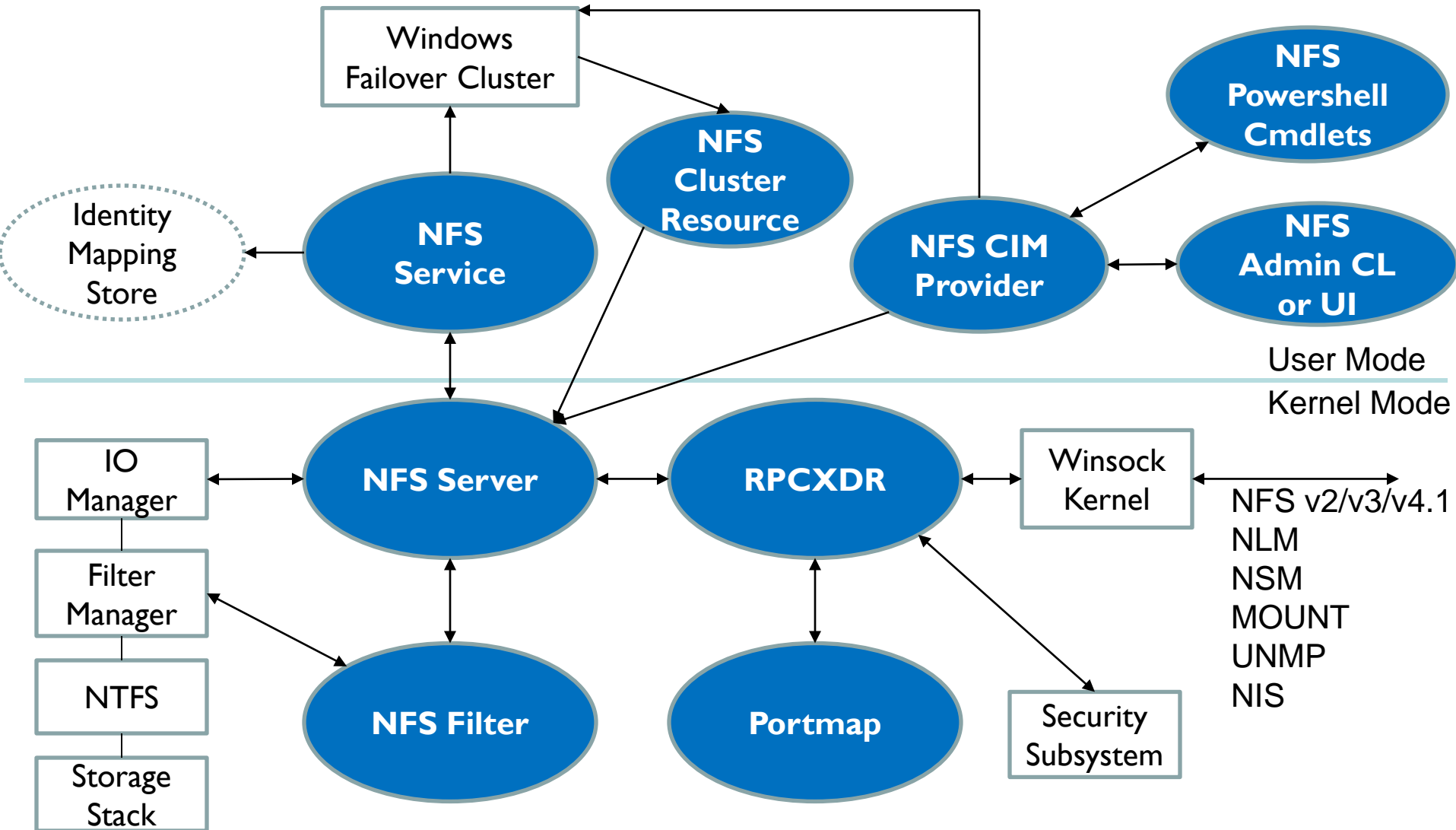 - ADLDS

What's new in NFS in Windows Server 2012?

- NFS v4.1 Server
  - Compliant with Mandatory Aspects of RFC 5661
  - Supports Volume Mount Points
  - No Support for:
    - ACL's
    - Delegations
    - Migration & Replication
    - pNFS
    - RDMA
    - Other Optional Aspects of RFC 5661

What's new in NFS in Windows Server 2012?

- ❑ Continuous Availability
  - ❑ New NFS Server Architecture for seamless integration with Windows Failover Clustering
- ❑ RPCSEC_GSS
  - ❑ Added Kerberos Privacy (Krb5p) support to existing krb5 and krb5i
- ❑ Central NFS Team Blog http://aka.ms/nfs

Agenda

- What's new in NFS in Windows Server 2012?
- **High Level Windows NFS Server Architecture**
- RFC Tradeoffs in Windows NFS Server
- User Account Mapping
- Pseudo File System
- Volume Mount Points
- Supported File Systems
- Diagnostics
- Q & A

Windows NFS Server Architecture

New RPC-XDR Transport Layer

Enhancements	Benefits
• Powered by new gen networking stack • TDI to WinSock (WSK) Infrastructure • Both Server for NFS & Client for NFS	• Better scalability • Better support (TDI is deprecated) • Receive Side Scaling (RSS)
• RPC Port Multiplexer • NFS, NLM, NSM, MOUNT all traffic on 2049	• Ease of deployment • Firewall friendliness - less ports to manage
• Auto tuned caches & thread pools • Dynamic resource management	• Out of the box performance & scale • Less tuning tasks • Remove guess work for admins
• Dynamic endpoints with per interface registration • PNP notification for network interfaces	• Faster failovers
• New Kerberos Privacy Implementation • Authentication options (KRB5, KRB5i & KRB5p)	• Secure deployments

Agenda

- What's new in NFS in Windows Server 2012?
- High Level Windows NFS Server Architecture
- **RFC Tradeoffs in Windows NFS Server**
- User Account Mapping
- Pseudo File System
- Volume Mount Points
- Supported File Systems
- Diagnostics
- Q & A

RFC tradeoffs in Windows NFS Server

- Post-op attributes are not atomic
  - RFC requires post-op attributes to be atomic
  - Windows has no support
  - NFS Server approximates…
- Open Upgrade/Downgrade
  - Windows doesn't have such support
  - NFS tries to open a new handle before closing the old handle
  - Handle that NFS Server uses might have more permission than needed

RFC tradeoffs in Windows NFS Server

- ❑ Lock Upgrade/Downgrade
  - ❑ Optional RFC feature – no support in Windows
- ❑ Partial Unlock
  - ❑ Windows Byte Range Lock doesn't support partial unlock
  - ❑ We approximate by releasing the current lock and relocking the remaining range
  - ❑ Serialized within NFS 4.1 access, but runs the risk with other local app or SMB access.

RFC tradeoffs in Windows NFS Server

- Anonymous stateid
  - NFS Server always opens a file system handle for a request using anonymous stateid
  - Performance degradation expected when NFS Clients send too many requests using anonymous stateids

RFC tradeoffs in Windows NFS Server

- Lease Management
  - Expired leases are revoked lazily
  - Conflicting requests revoke expired stateids
  - Idle sessions revoked by session inactivity timer

- Delete/Rename on Windows
  - Name doesn't go away from namespace until all handles are closed
  - Different from *NIX where name disappears from namespace immediately following the operation regardless of open handles
  - Rename on directories allowed only when there are no outstanding opens on all files within the directory tree

RFC tradeoffs in Windows NFS Server

- File Handle Size
  - Fixed 32 byte handles for NFS v2/v3
  - Variable size for v4.1
  - Some NFS Clients have difficulty in handling file handles greater than 56 bytes
- Volume Mount Points
  - RFC and NFS Client implementations assume that a directory can have *exactly* one parent
  - A volume can be mounted on more than one location in Windows

Agenda

- What's new in NFS in Windows Server 2012?
- High Level Windows NFS Server Architecture
- RFC Tradeoffs in Windows NFS Server
- **<u>User Account Mapping</u>**
- Pseudo File System
- Volume Mount Points
- Supported File Systems
- Diagnostics
- Q & A

User Account Mapping – What's New?

- Passwd and Group File Mapping
  - Unix style passwd and group file
    - %SystemRoot%\system32\drivers\etc\passwd
    - %SystemRoot%\system32\drivers\etc\group
- Comprehensive Powershell and UI support for identity mapping store setup and configuration
  - New Server Manager UI
  - 40+ NFS Powershell cmdlets

User Account Mapping- PowerShell Cmdlets

- ❑ Configure Identity Mapping Store
    - ❑ Set-NfsMappingStore
    - ❑ Get-NfsMappingStore
    - ❑ Install-NfsMappingStore
    - ❑ Test-NfsMappingStore
- ❑ Provision User Accounts
    - ❑ Set-NfsMappedIdentity
    - ❑ Get-NfsMappedIdentity
    - ❑ New-NfsMappedIdentity
    - ❑ Remove-NfsMappedIdentity
- ❑ Check/Validate User Accounts
    - ❑ Resolve-NfsMappedIdentity
    - ❑ Test-NfsMappedIdentity

User Account Mapping - Demo

- Scenario - Migrate Account Mapping Information from File Mapping to ADLDS Mapping Store
  - Existing Passwd and Group files
  - Install ADLDS Role and Provision ADLDS Instance
  - Migrate Mapping Information from Passwd and Group Files using Get-NfsMappedIdentity and New-NfsMappedIdentity PowerShell cmdlets
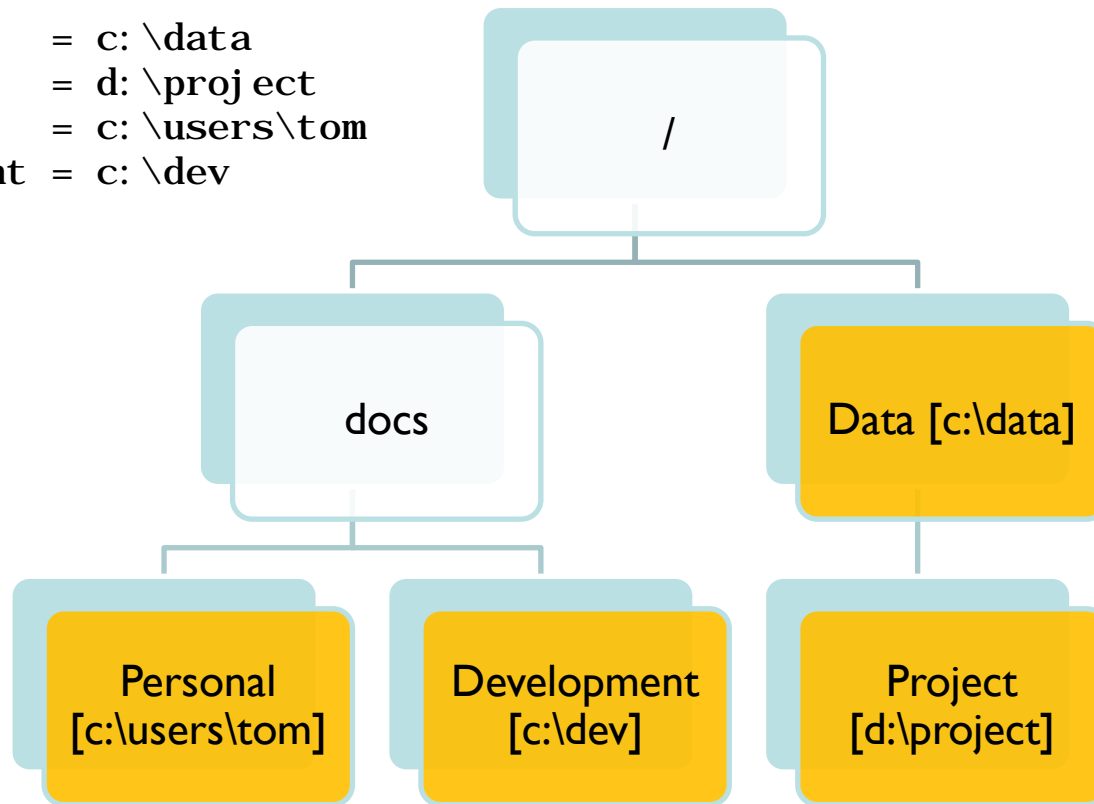
Agenda

- What's new in NFS in Windows Server 2012?
- High Level Windows NFS Server Architecture
- RFC Tradeoffs in Windows NFS Server
- User Account Mapping
- **Pseudo File System**
- Volume Mount Points
- Supported File Systems
- Diagnostics
- Q & A

Pseudo File System

- A separation between virtual and physical namespace

- Read-only FS from NFS Client's point of view

- Real FS can be attached at any layer

- Admin provides an alias for exports. NFS Server infers PFS from it

- Back compatible with NFS v2 and v3

- Volume snapshots are a special case for the PseudoFS

Pseudo File System - Example

```
/data              = c:\data
/data/project      = d:\project
/docs/personal     = c:\users\tom
/docs/development  = c:\dev
```

```
                        /
            ┌───────────┴───────────┐
          docs                 Data [c:\data]
     ┌──────┴──────┐                │
  Personal    Development        Project
[c:\users\tom] [c:\dev]        [d:\project]
```

Pseudo File System - Snapshots

- VSS snapshots available via NFS Server
  - Each VSS snapshot is exposed as a new FS
  - NFS Server automatically discovers newly available VSS snapshots
- Namespace is constructed via predefined rules
  - In NFS v2/v3, snapshots are presented as hidden directories at the root of the NFS mount point
  - In NFS v4.1, snapshots are presented at the terminal node inside PFS and all the Volume Mount Point transition points

Pseudo File System - Demo

❑ Demo

Agenda

- What's new in NFS in Windows Server 2012?
- High Level Windows NFS Server Architecture
- RFC Tradeoffs in Windows NFS Server
- User Account Mapping
- Pseudo File System
- **Volume Mount Points**
- Supported File Systems
- Diagnostics
- Q & A

Volume Mount Points

- Special Windows FS object (directory) which is used to provide entry point to other volumes
- Windows NFS v2/v3 implementation don't support VMPs – no crossing FS boundary!
- Windows NFS v4.1 implementation supports VMP's
  - Can export more than 26 volumes!

Volume Mount Points - Challenges

- A volume can be mounted in multiple place inside PFS
- Problem with LOOKUPP
  - A volume can be mounted at multiple places which means that there is no reliable way to find the parent
  - Limited FH size – unless we can encode all the ancestor FS id for an object, we cannot go across the VMP upward
  - An object in the namespace could have multiple different FH's depending on path taken to the object

Agenda

- What's new in NFS in Windows Server 2012?
- High Level Windows NFS Server Architecture
- RFC Tradeoffs in Windows NFS Server
- User Account Mapping
- Pseudo File System
- Volume Mount Points
- **Supported File Systems**
- Diagnostics
- Q & A

Supported File Systems

- NTFS and NTFS-like file systems
    - FILE_UNICODE_ON_DISK
    - FILE_PERSISTENT_ACLS
    - FILE_SUPPORTS_OPEN_BY_FILE_ID
    - FILE_SUPPORTS_REPARSE_POINTS
    - FILE_LINKS_INFORMATION
- No Support for ReFS
    - 128 bit FileId in ReFs
    - NFS RFC's allow max fileid size of 64 bits
- No Support for CsvFs/FAT/FAT32/CDFS

Agenda

- What's new in NFS in Windows Server 2012?
- High Level Windows NFS Server Architecture
- RFC Tradeoffs in Windows NFS Server
- User Account Mapping
- Pseudo File System
- Volume Mount Points
- Supported File Systems
- **<u>Diagnostics</u>**
- Q & A

- Event Channels (Eventvwr path ➔ Applications & Service/Microsoft/Windows/ServicesForNFS-Server)
  - Admin
  - IdentityMapping
  - Operational
  - Analytic
- End to End Tracing Support
  - Activity ID
  - Enables Analytics/Performance Tracing across the kernel
  - Analyze with Netmon or custom tools

Diagnostics - Performance Counters

- Server for NFS-NFSv4 Statistics
- Server for NFS-Netgroup
- Server for NFS-User Mapping
- Server for NFS-NFSv4 Read Write Statistics
- Server for NFS-NFSv4 Request/Response Sizes
- Server for NFS-NFSv4 Throughput
- Server for NFS- NFSv4 Operation Statistics
- Server for NFS - Session and Connection Statistics

Diagnostics - Demo

- NFS Server with exports

- Start ETW tracing

- Mount share from NFS Client and enumerate files and perform file copy

- Stop ETW trace and analyze events with NetMon

Agenda

- What's new in NFS in Windows Server 2012?
- High Level Windows NFS Server Architecture
- RFC Tradeoffs in Windows NFS Server
- User Account Mapping
- Pseudo File System
- Volume Mount Points
- Supported File Systems
- Diagnostics
- **Q & A**

THANK YOU!