SDC
STORAGE DEVELOPER CONFERENCE
SNIA ■ SANTA CLARA, 2014

Choices for implementing SMB 3 on non Windows Servers

Dilip Naik
HvNAS Pty Ltd

Australians good at NAS protocols!

Focus & contents of this talk

- Why SMB 3? How SMB 3?
- Implementing an SMB 3 Server on Linux/UNIX (non Windows)
  - Not focusing on implementing an SMB 3 client
  - Some of the solutions discussed apply to CIFS and also other non SMB 3 protocols
  - Breadth first, not depth first discussion

Why SMB 3 – Speed, Reliability, Cost

- Microsoft investments are in NAS & not DAS
  - SMB 3 Multi Channel provides speed & failover, NAS only
  - SMB Direct/RDMA is NAS only
  - SMB 3 Persistent Handles is NAS only
  - SMB 3 File/Directory leases (allows client to cache) NAS, not DAS
- SMB 3 NAS is cheaper, faster, more reliable than DAS!
  - Multiple NICs, Persistent Handles vs Multi Path
  - NAS management versus SAN management

Why SMB 3? Product Perspective

- Required for Hyper-V & SQL workloads on NAS
  - SQL system d/b MUST be on SMB 3 NAS
- Increasingly important in Hyper-V
  - Hyper-V 2012 R2 live migration uses SMB 3
  - Windows Server 2012 R2 VDI dedupe works ONLY when VDI files on NAS (not DAS)
- Windows "internals" tuned to leverage SMB 3 features – e.g. CopyFileEx API
- Quick way for storage startups to monetize Windows/Hyper-V after starting with VMware

SMB 3 and Storage Spaces

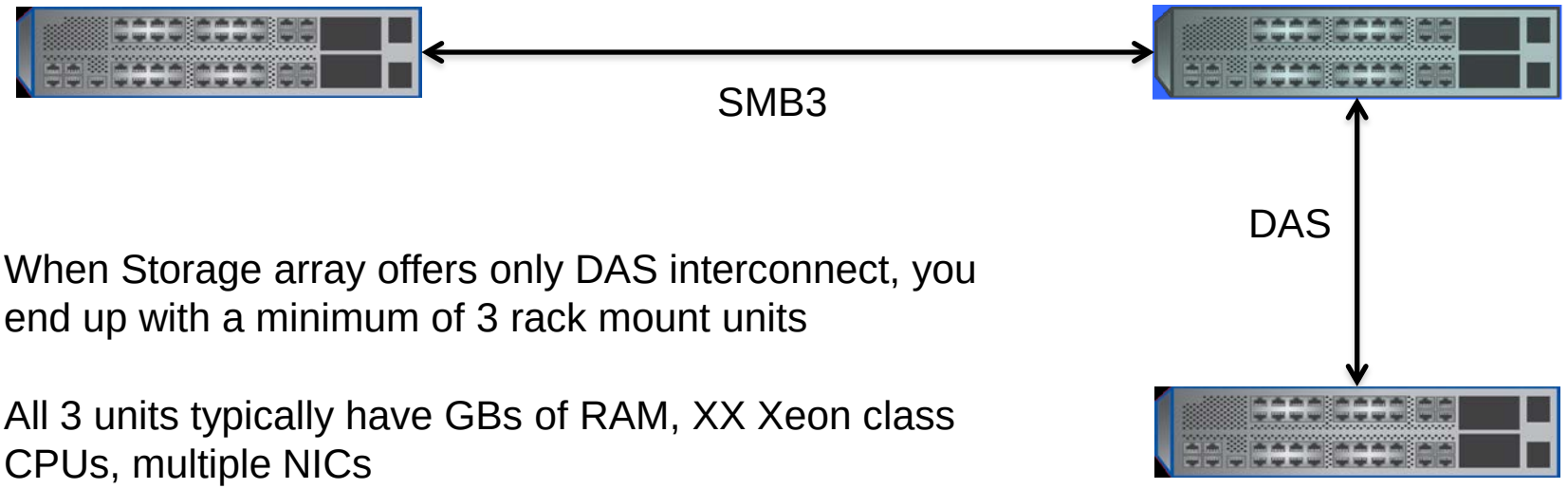
Yes, Microsoft invested in Storage Spaces & its great features such as Tiering!

Spaces is even more valuable when used as DAS attached to NAS rather than directly to compute node!

Why SMB 3 – CapEx & OpEx – DAS – 1 of 2

Hyper-V host aka Compute Node

Windows Server 2012 NAS



When Storage array offers only DAS interconnect, you end up with a minimum of 3 rack mount units

All 3 units typically have GBs of RAM, XX Xeon class CPUs, multiple NICs

Compute Node is hosting 50+ VMs in this example

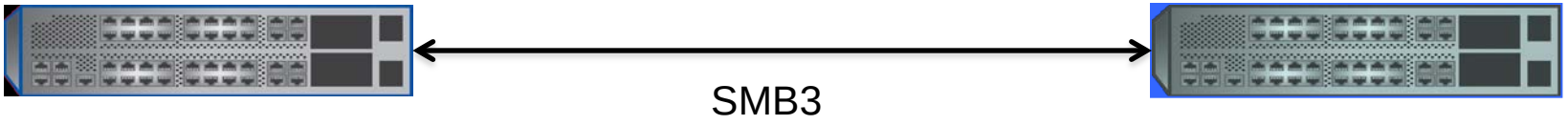
NAS is hosting files for 50+ VMs & driving storage array

Storage Array offering only DAS interconnect

Why SMB 3 – CapEx & OpEx – NAS – 2 of 2

Hyper-V host aka Compute Node

Storage Array with
Native SMB 3 stack



When Storage array offers a Native SMB 3 interface, you end up with a minimum of just 2 rack mount units !

Both CapEx and OpEx savings!

Why SMB 3 – backup appliances

Primary Storage

Backup Appliance

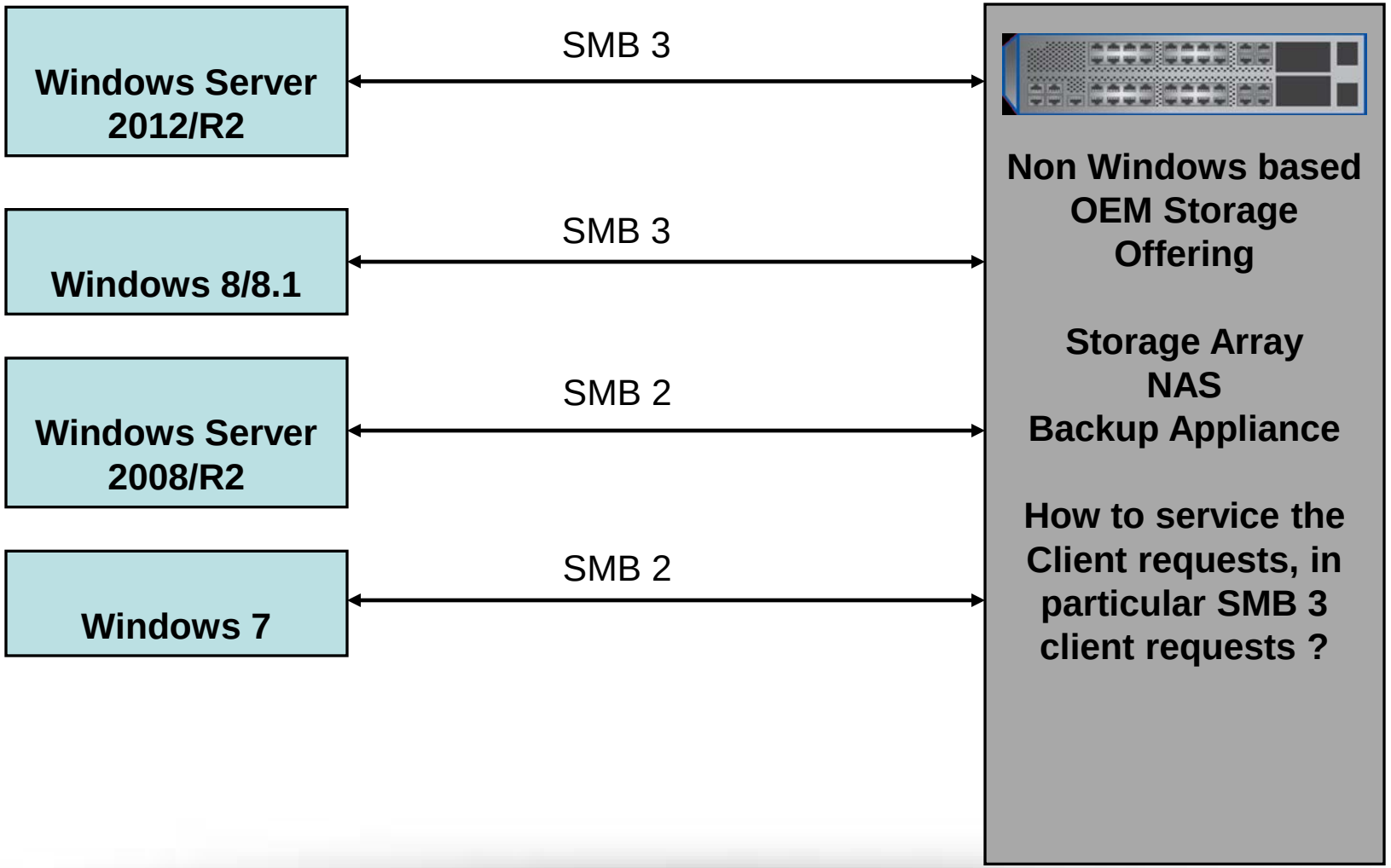
SMB3,
SMB 2,
CIFS

CIFS effectively 60kb, 1 I/O in flight
SMB 2.0 1 MB, 1 I/O in flight
SMB 2.1 1 MB, N I/Os in flight
SMB 3 1 MB, N I/Os in flight per TCP Channel
SMB Direct hardware assisted I/O

If Primary Storage has 2 NICs & Backup Appliance has 2 NICs, that is 4 NOT 2 TCP Channels

One way to gain market share in backup appliance market is implement SMB 3!

Problem Statement



Does your SMB 3 stack support SMB 2?

- SMB 3 = SMB 2.2
  - SMB 3 was SMB 2.2 as per Microsoft
  - Even today, SMB 3 is defined in specification "MS-SMB2.pdf"
  - SMB 2 defines 1 Read, 1 Write command
  - "SMB 3" defines none – uses SMB 2's Read, Write Command
  - Incremental cost of supporting SMB 2 low

SMB 3 server implementation choices

- 1. Deploy a Linux/UNIX native SMB 3 server
  - Multiple choices here
- 2. Deploy bits on the SMB 3 client
  - Multiple choices
- 3. Install bits on on Linux/UNIX storage
- 4. Man in middle device

Choice 1A Develop your own Linux/UNIX SMB 3 stack

- Pros
  - Control on architecture, choices made
  - Protocol Specs, support from Microsoft
  - SNIA talks, Plugfests
- Cons
  - Time to market, Resource costs
- Some companies on this path

Choice 1B Start with Likewise

- Last publicly available stack is probably SMB 2.0+ but not quite SMB 2.1
- Add SMB 3 functionality
- Find/train developers to be familiar with code
- Still some time to market considerations & cost
- Any companies on this path?

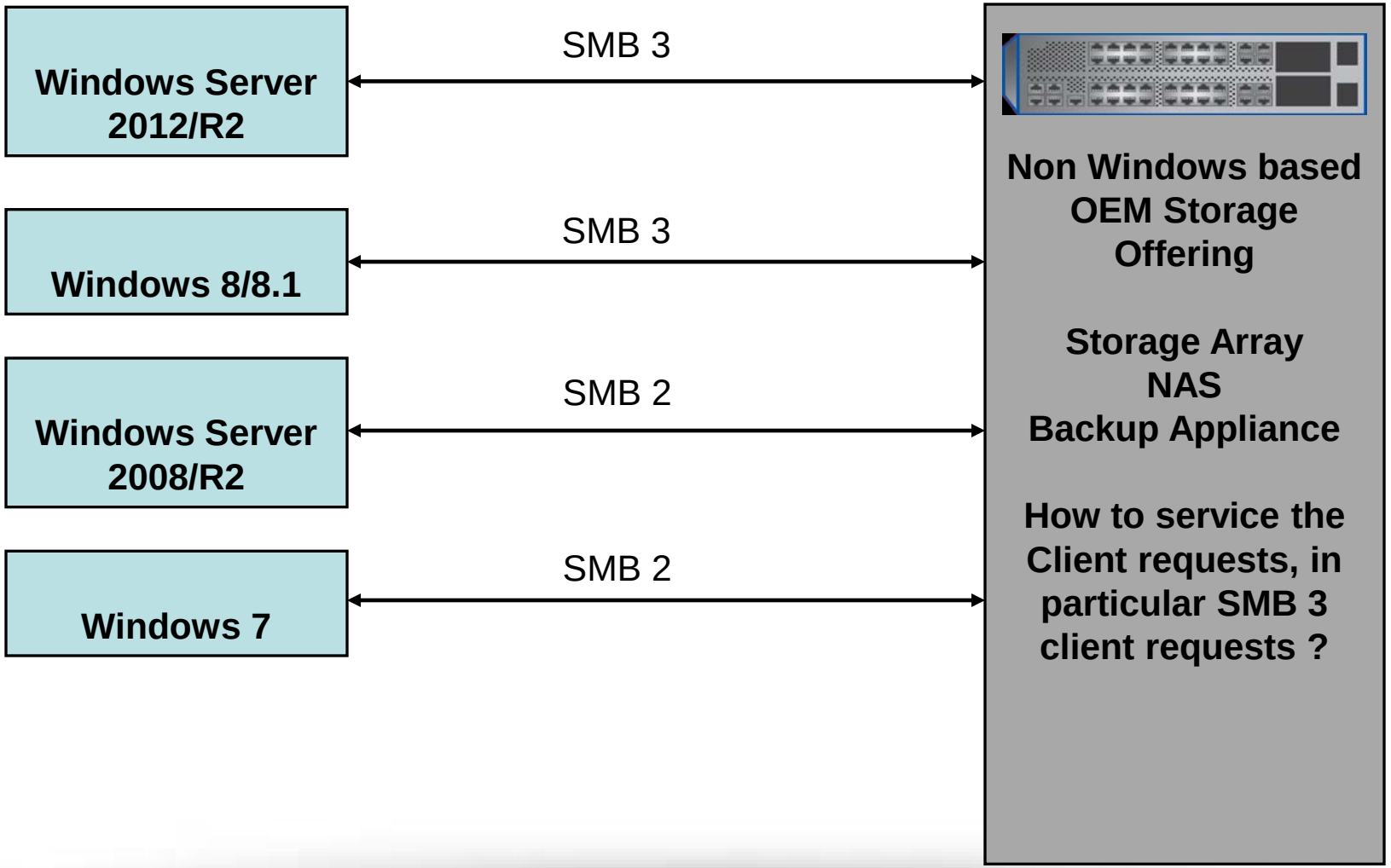
Choice 1C Deploy Samba 4

- Pros
  - Widely deployed code base
  - Well understood best fit scenarios
  - Expertise available
- Cons
  - Cons of GPL apply
  - Some SMB 3 features still in development
- Some companies may adopt this path after Samba has complete SMB 3 implementation

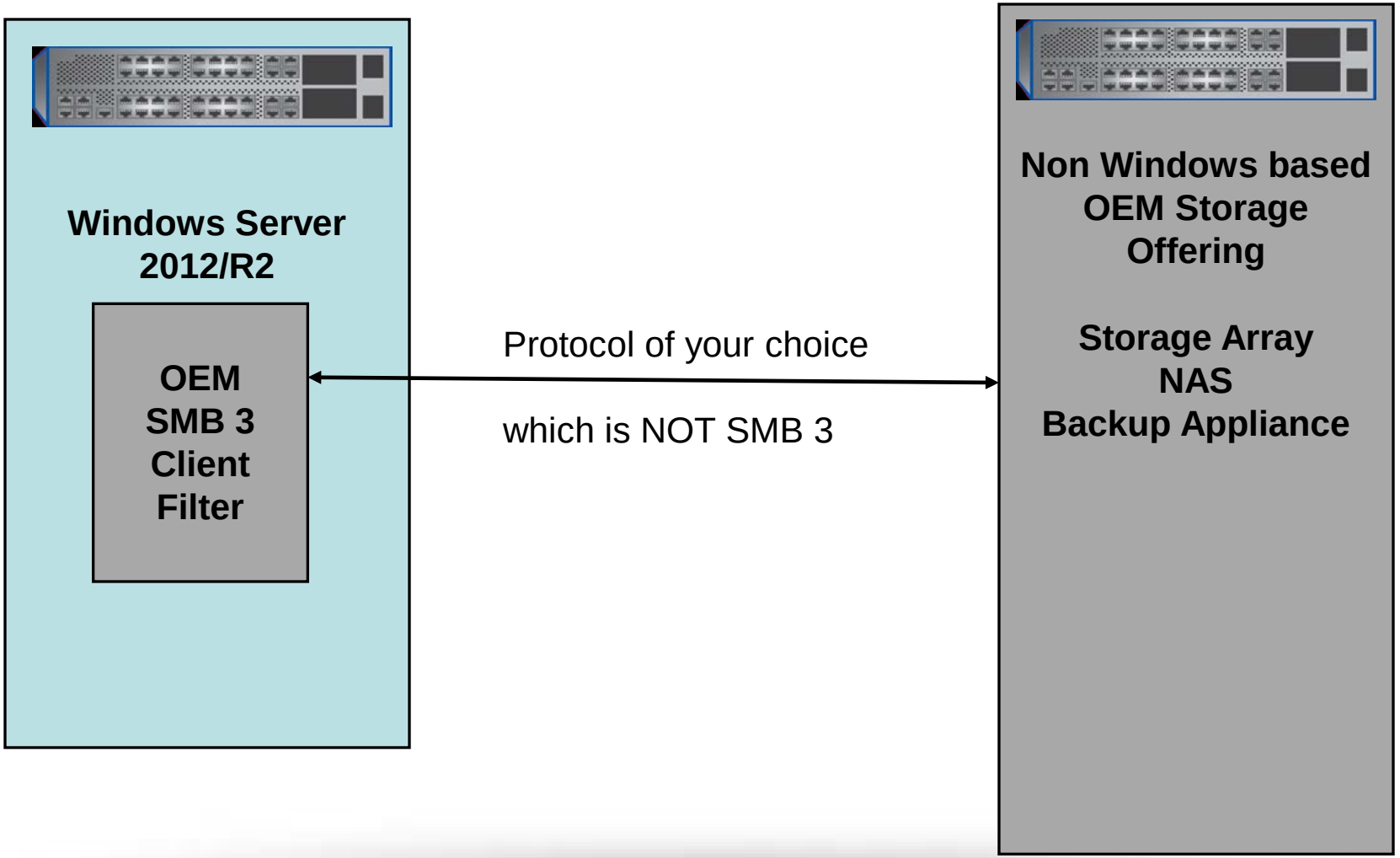
Choice 1D License & ship HvNAS

- non GPL source implements ALL optional features (Multi Channel, Encryption, Dir Leases, Persistent Handles)
  - Both data access plane and control plane
  - 99+% pass rate on MS Protocol Test Suite
- Pros
  - Time to market, non GPL so link into file sys (no IPC & buffer copy overheads)
- Cons
  - Proprietary

Problem Statement



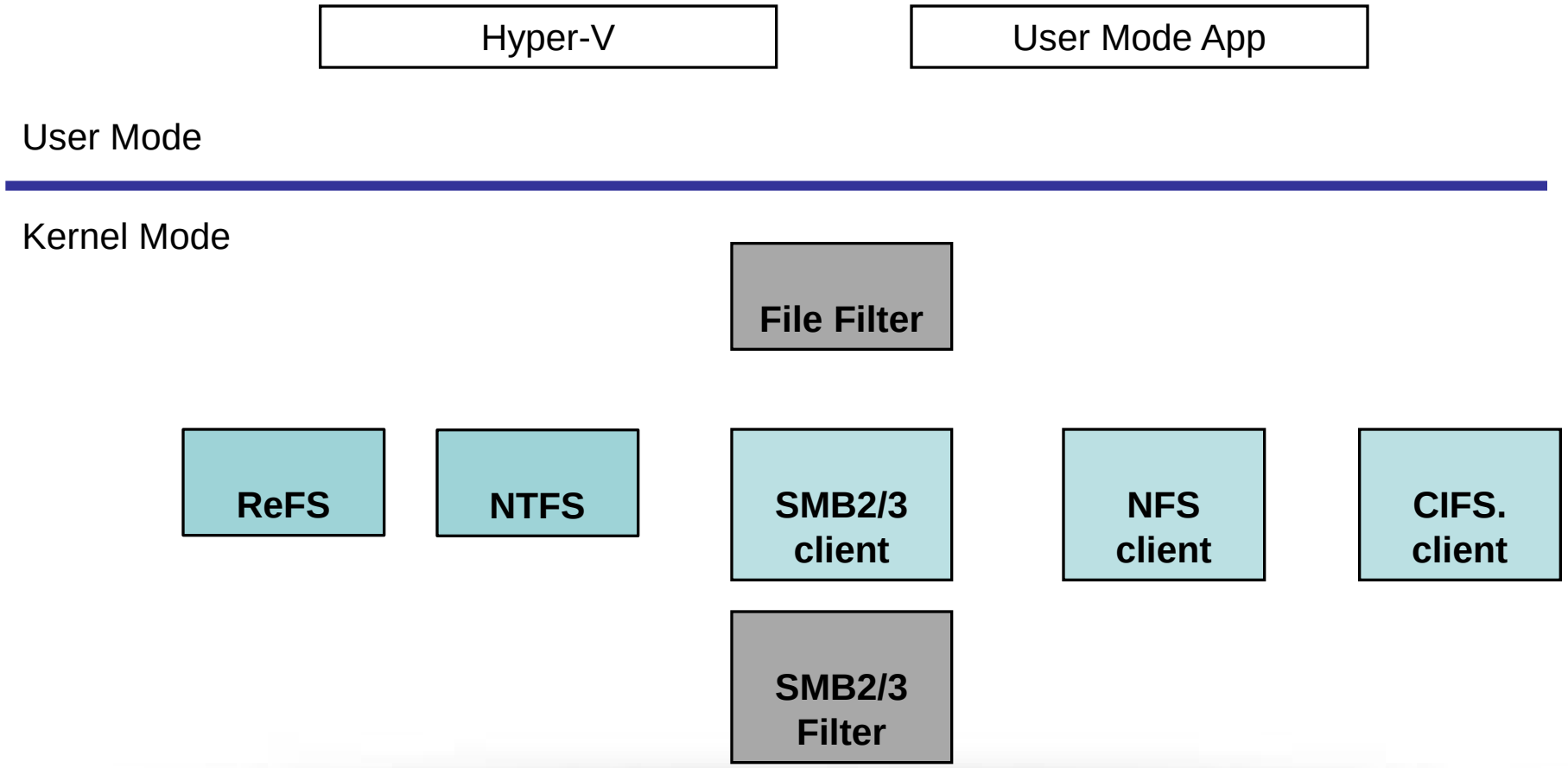
Choice 2 Deploy bits on SMB 3 client



Deploy bits on SMB 3 client

- Pros
  - Wide range of choices, complete control
  - Time to market - Needed pieces of technology may be available off the shelf
- Cons
  - Install non trivial filter on every SMB 3 client
  - What about SMB 3++?

2 choices in SMB 3 client filter



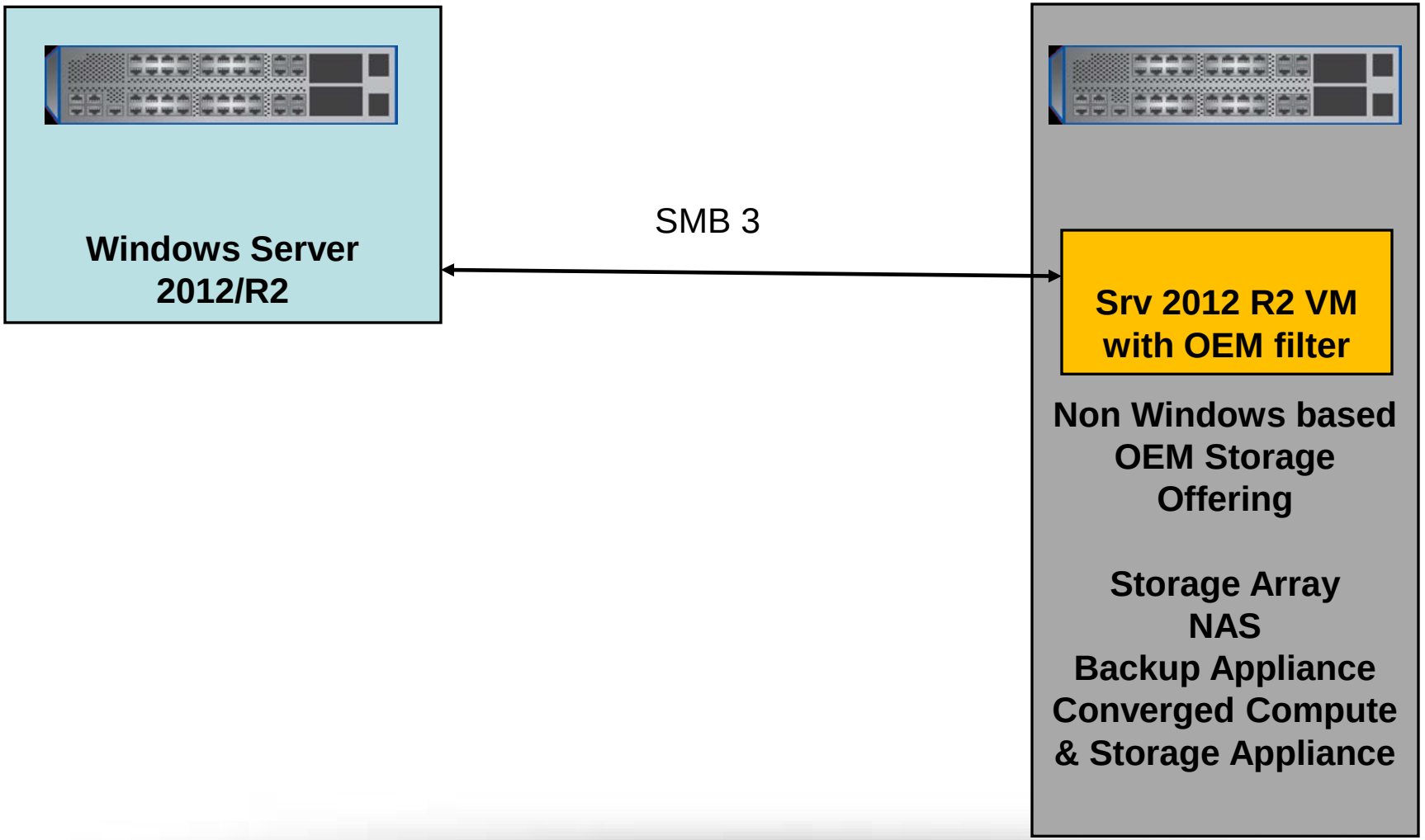
2A File Filter in SMB 3 client

- Pros
  - Extremely well documented i/f
  - Lots of sample code & tests
  - Developer expertise available
- Cons
  - Catch I/O before it is SMB 3- loose all Client side SMB 3 intelligence
  - App specific tuning e.g. every time Hyper-V changes…

2B SMB 3 Filter in SMB 3 client

- Pros
  - Leverage intelligence in SMB 3 client
  - Convert to protocol of choice
- Cons
  - Non trivial
  - Partially documented i/f, few samples
  - Developer expertise relatively scarce

Choice 3 License the SMB 3 stack from Microsoft !



Server 2012 R2 VM in OEM Storage

- Pros
  - Perfect SMB 3 & RPC implementation!
  - Smooth upgrade path for SMB 3.0, 3.1++
  - Convert SMB 3 to your choice of protocol (NFS, ATA etc) using a file mini filter driver – available from HvNAS or write your own http://winntfs.com/2014/05/12/protocol-converter-between-cifs-smb2-smb3-and-nfs/
    - Effectively man in middle approach
- Cons
  - Added cost of Server 2012 license

Summary

- SMB 3 is important!
- Do implement it, choice of how is yours
- Dilip@HvNAS.com

References

- Microsoft SMB 3 talks at SNIA
  - Too many to list here!
  - http://blogs.technet.com/b/josebda/ - LOTS & I mean LOTS of SMB3/Hyper-V blogs including http://blogs.technet.com/b/josebda/archive/2014/03/30/updated-links-on-windows-server-2012-r2-file-server-and-smb-3-0.aspx
- www.HvNAS.com
- www.winntfs.com