VMware & Storage

Presented By:

- Marc Staimer, President & CDS
- Dragon Slayer Consulting
- marcstaimer@mac.com
- 503-579-3763

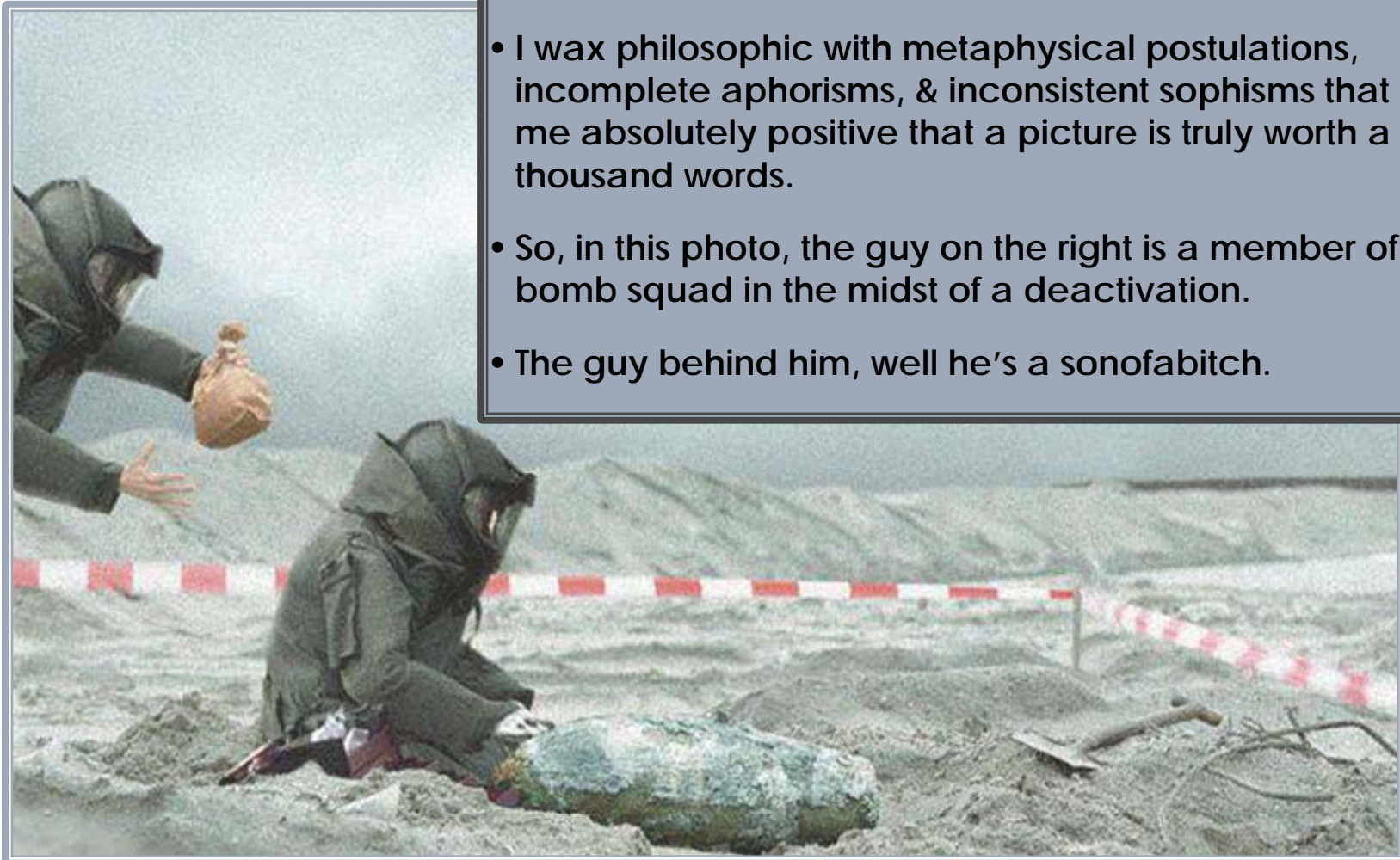# Dragon Slayer Consulting Intro

- **Marc Staimer - President & CDS**
  - 12+ years
    - Storage, SANS, SW, Networks, Servers
    - Consults vendors (> 100)
    - Consults end users (> 400)
    - Analysis at trade shows
    - Publishes consistently with Tech Target
    - Periodically published for trade magazines
 - 30+ years industry experience

marcstaimer@mac.com
503-579-3763

Picture is Worth a 1000 Words

- In the course of my consultancy practice I'm often asked "How do I define a sonofabitch"?

- I wax philosophic with metaphysical postulations, incomplete aphorisms, & inconsistent sophisms that make me absolutely positive that a picture is truly worth a thousand words.

- So, in this photo, the guy on the right is a member of a bomb squad in the midst of a deactivation.

- The guy behind him, well he's a sonofabitch.

Agenda
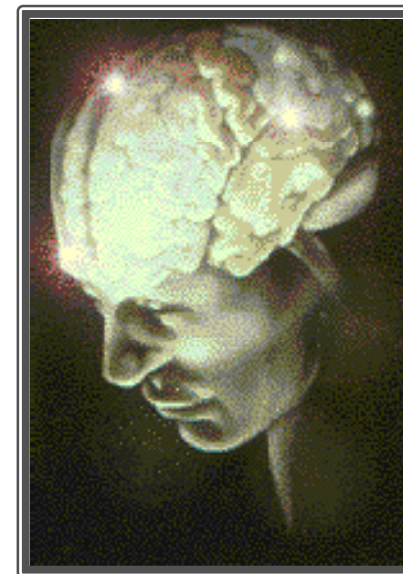
- **Real World Level Setting**

- **How Server Virtualization Works with Storage**

- **Server Virtualization Storage "Gotcha" Issues**

- **Server Virtualization Data Protection Issues**
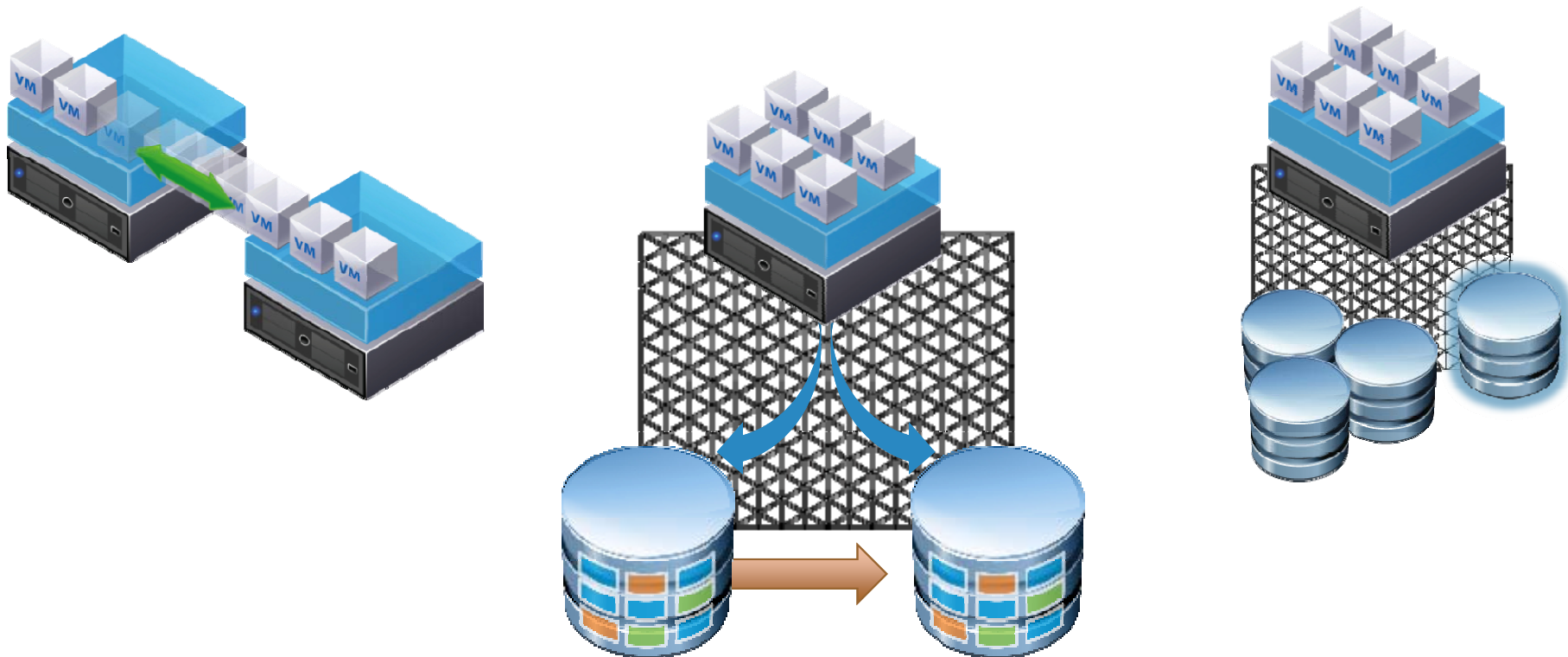
- **How to resolve**

Why Server Virtualization?

- Greater Application Availability
- Faster/simpler server management
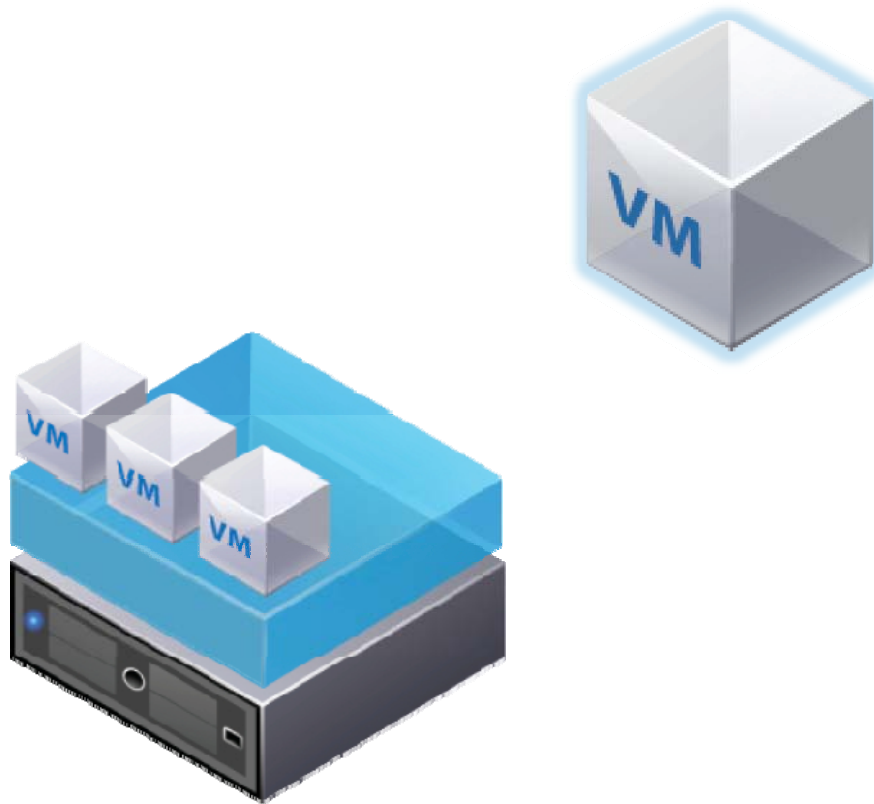- Faster/simpler DR & Bus Continuity
- Server consolidation

Greater Application Availability VMware

- **Reduces or eliminates scheduled downtime**
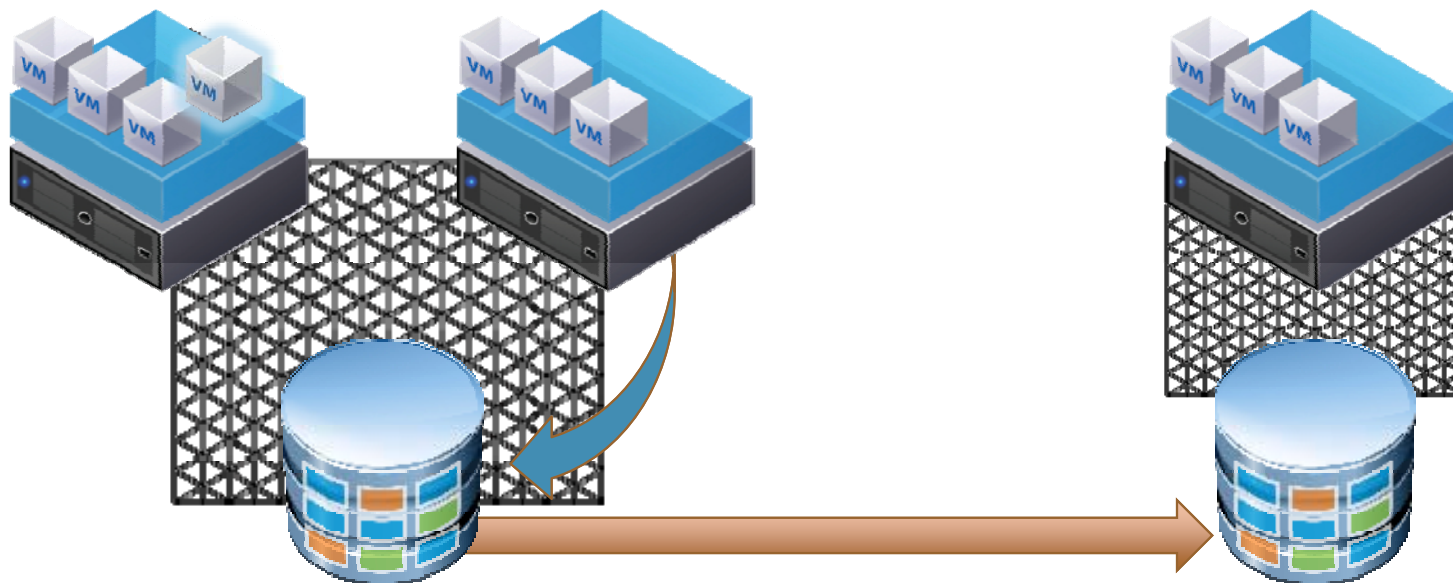 - Vmotion, Storage Vmotion, Thin Provisioning

Faster Simpler Server Management

- **Easier to implement, provision, operate, manage**

Faster Simpler DR & Bus Continuity

- **Better able to handle unscheduled downtime**
  - Local or Remote (SRM)

Server Consolidation

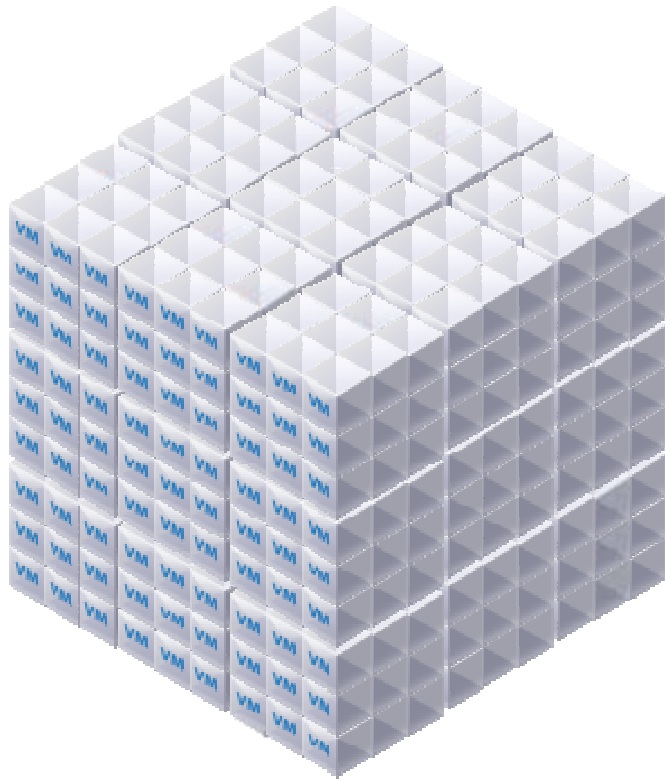- **Business case to justify the above**
  - Less
    - Servers
    - NICs and/or HBAs
    - Cables
    - Switch ports
    - Switches
    - Rack space
    - Floor space
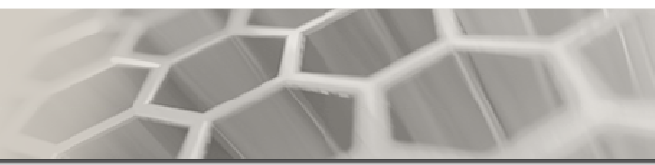   - Power
    - Cooling
   - Less CapEx & OpEx

vSphere 4.1 = "Cloud Scale"

- 32 hosts / cluster
- 320 VMs / host
- 3,000 VMs / cluster (2x)
- 1,000 hosts / vCenter (>3x)
- 10,000 VMs / vCenter (>3x)
- 15,000 register VMs / vCenter (>3x)

99% of VMware's 170K Customers
Can Run Their Entire Datacenter in a Single VMware Cluster

Seriously Enhanced Scalability in Other Ways Too

	vSphere 4	vSphere 4.1	Ratio
Concurrent VI Clients	**30**	**120**	**4x**
Hosts per DC	100	500	**5x**
VMs per DC	**2,500**	**5,000**	**2x**
Linked Mode	10,000	30,000	**3x**
Concurrent vMotions per host	**2 @ up to 2.67GBps**	**8 @ 8GBps**	**~ 4x**
Hosted footprint & mem consumption	NA	40% less	3x > IOPS
Transparent memory compression	**NA**	**> 15% perf increase**	**>1.15x**
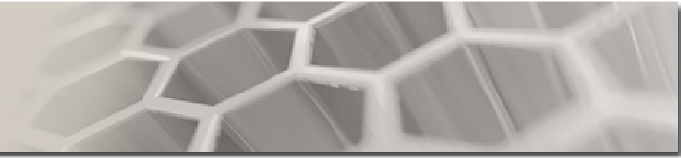
The Secret of Life...



There Really is No Such Thing As a Free Lunch!

Significant VMware vStorage Improvements

- **VMware vSphere really enhanced vStorage**
  - Virtual Disk Thin Provisioning
  - Improved iSCSI Software Initiator Efficiency
  - vCenter Server Storage Enhancements
  - Enhanced Storage Vmotion
  - Dynamic Expansion of VMFS Volumes
  - vStorage APIs for Multipathing

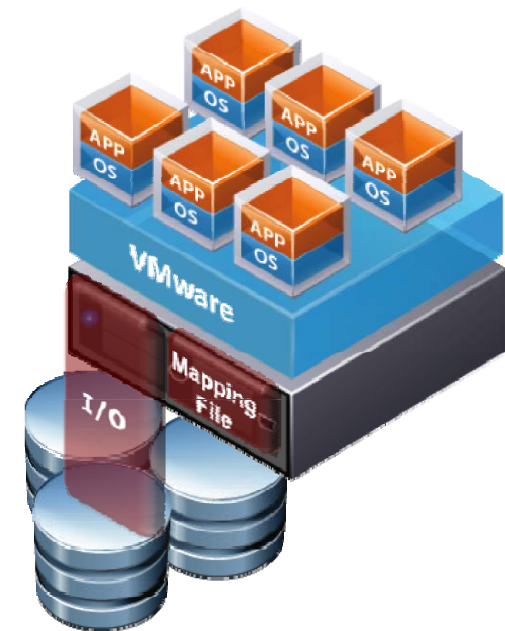# VMware Storage System Options

- **Raw device mapping – SAN**

- **VMFS shared block – DAS or SAN**

- **Shared NFS storage – NAS**
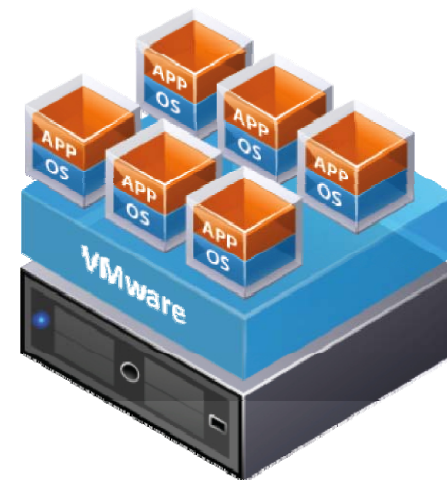
Raw Device Mapping (RDM)

- **Guest VMs access storage directly over iSCSI or FC**
  - VMs can even boot from raw devices
- **Great**
  - Per-server queues for performance
  - Easier measurement
  - Methodology for clustering
    - Virtual-to-virtual or virtual-to-physical
- **But…**
  - Tricky VMotion and DRS
  - No storage Vmotion
  - No SRM
  - A lot more manually intensive mgmt overhead
  - Limited to 256 LUNs per data center

VMFS *(Virtual Machine File System)*
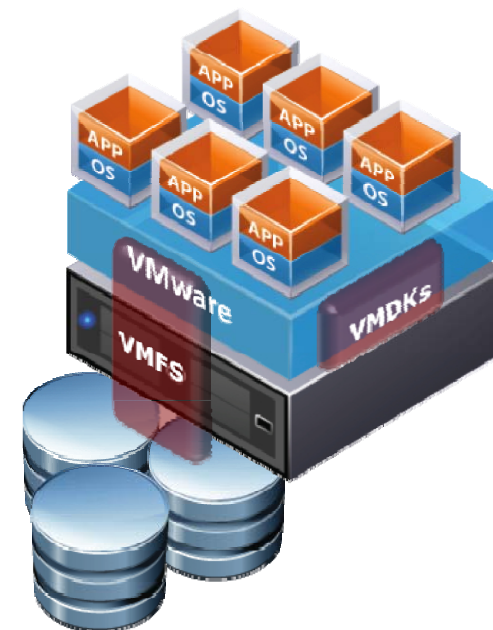## Shared Block Storage

- ## Shared storage - common/ workstation approach
  - Stores VMDK image in VMFS datastores
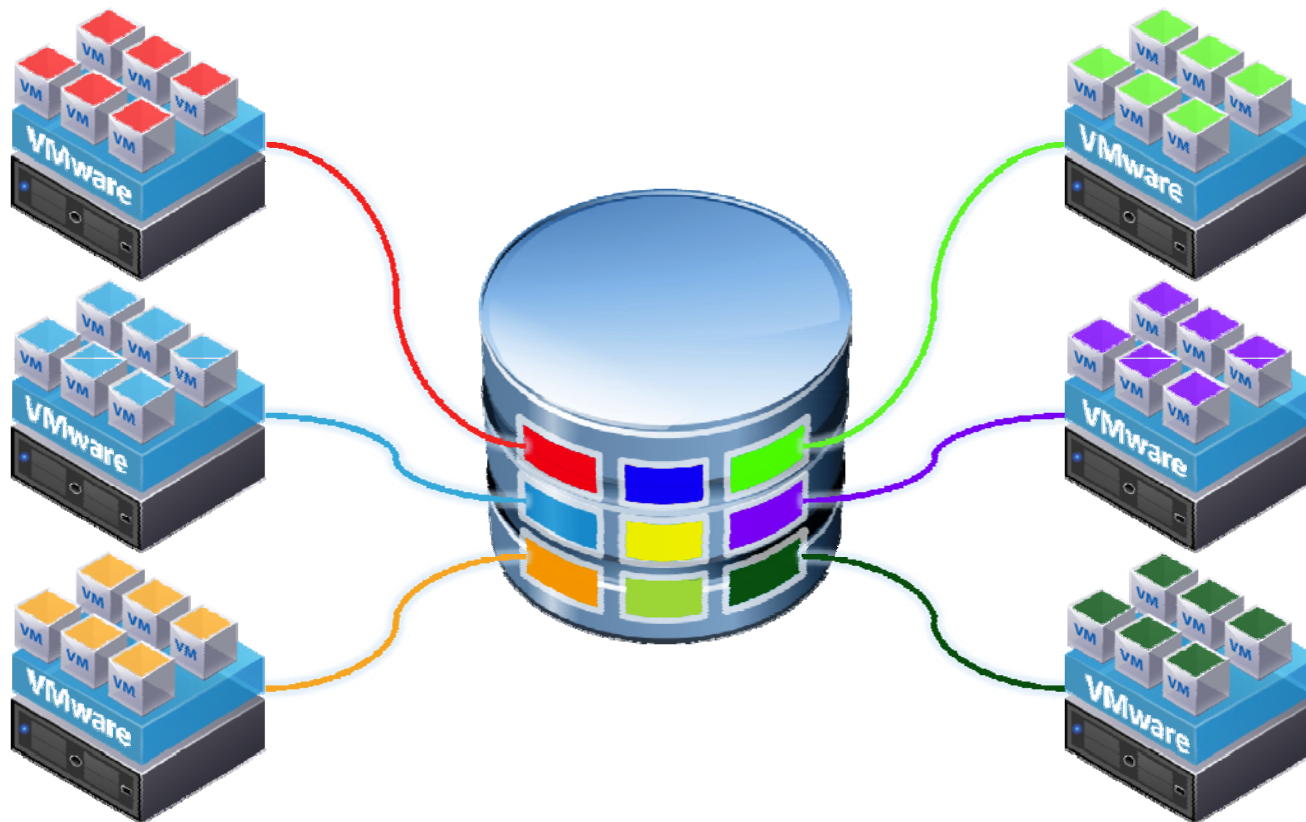 - DAS or FC/iSCSI SAN
- ## Dominant VMware Storage
  - VMFS is VMware's own FS
    - Develops 1st on VMFS
  - Traditional, familiar, mature, robust
  - VMware admin independence
  - Prime features (Storage VMotion, etc)
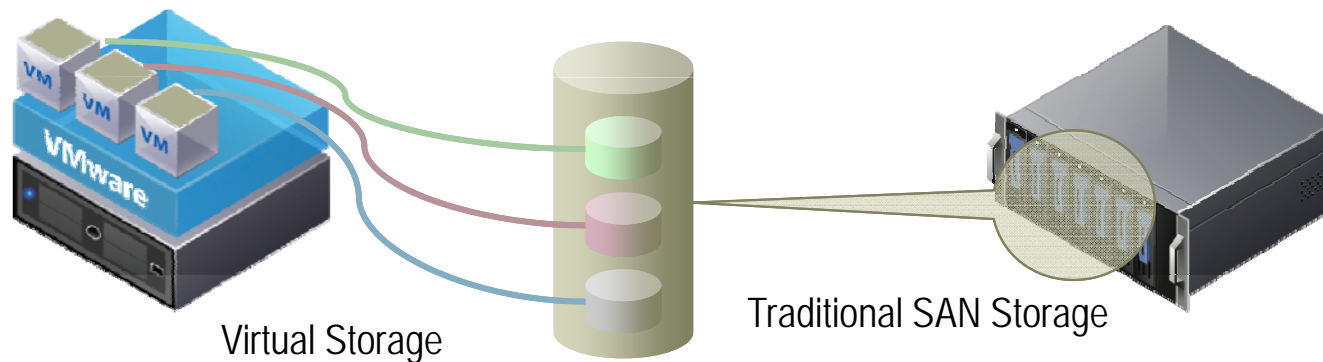 - Multipathing, load balancing, failover

How Server Virtualization Works w/SAN Storage

- **Storage LUNs are**
  - Virtualized and shared by Virtual Server

LUN Oversubscription

- **Storage systems can't distinguish between VMs**
  - Each VM contests for same HDDs
    - Contention seriously decreases performance
    - There is no VM QoS today

Virtual Storage

Traditional SAN Storage

Consequences – Especially With SATA HDDs

- **Slower SATA drives don't handle contention well**
  - Nominal buffers or ques = higher response times

FC/SAS

Que depth of 0 to 32
Usually 0

SATA

7,200 RPM
5,400 RPM

Que depth of 256 to 512
15,000 RPM
10,000 RPM
7,200 RPM

Ends up consuming
the controller's que
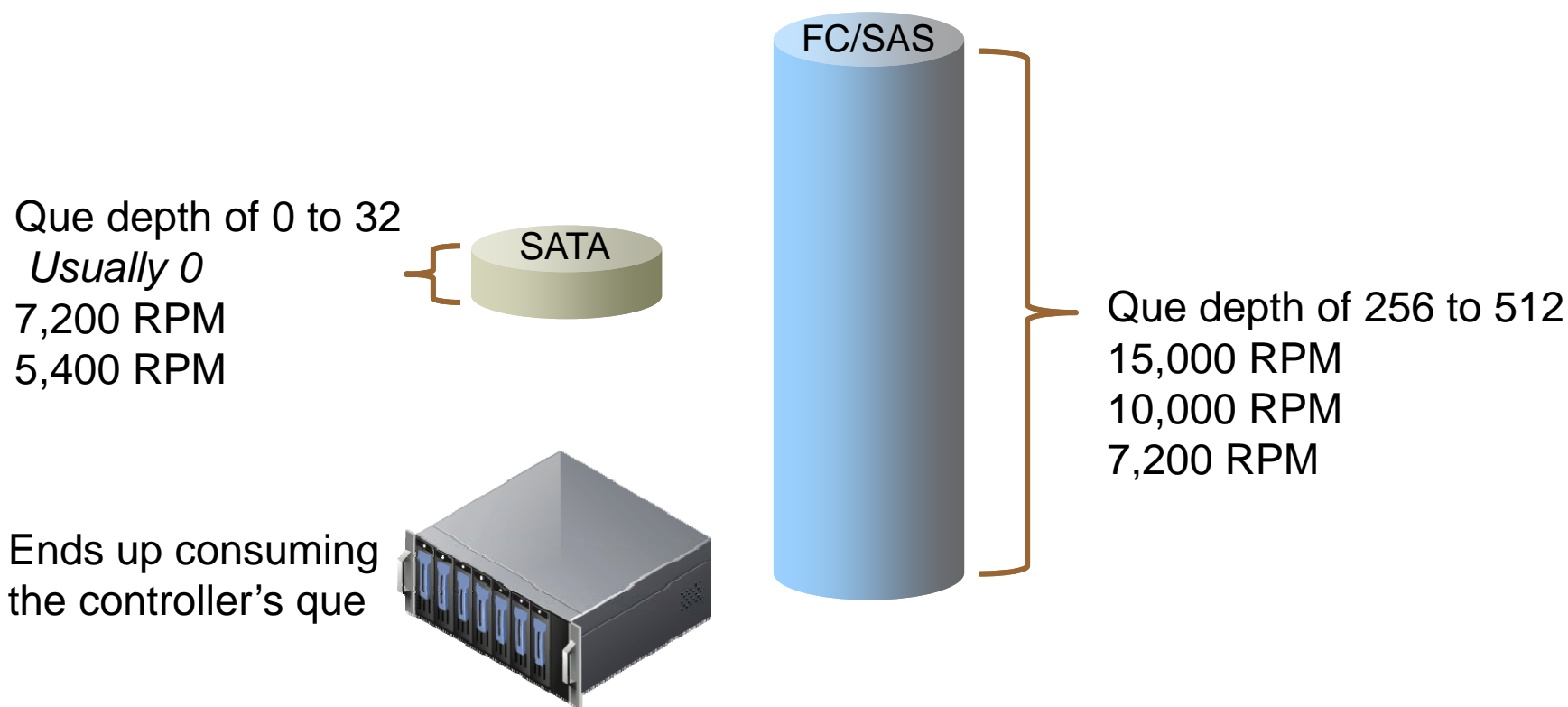
vSphere 4.1 Storage IO Control (a.k.a. QoS)

- **Solves some IO contention issues**
  - User definable QoS prioritization
  - Enables/disabled per specific datastore
  - Tied to datastore latency thresholds
    - Exceeds threshold, then congestion is assumed
      - At that point VMs are prioritized by policy
  - Admins can also define amt of IOPS reachable / VM

Other Problematic VMware VMFS Storage Issues

- **Storage Provisioning**
- **Too much oversubscription**
- **Painful scaling**
- **Data Protection**

Storage Provisioning

- **VM & application disruptive on storage side**

- **Time consuming**

- **Manually intensive**

- **The complete opposite of VM provisioning**

  - Requires knowledge, expertise, & knowhow

    - RAID group setup, DP, snapshots, replication, etc.

Key Issue: Performance Troubleshooting

- **VM App performance declines noticeably**
  - When moving from physical to virtual servers
  - Often causing fruitless VM migrations
  - Lots of admin frustration looking for root cause
  - Even more troubling when problem is intermittent
    - Significant potential for VM SCSI timeouts
    - Mission critical VMs require own unique LUNs

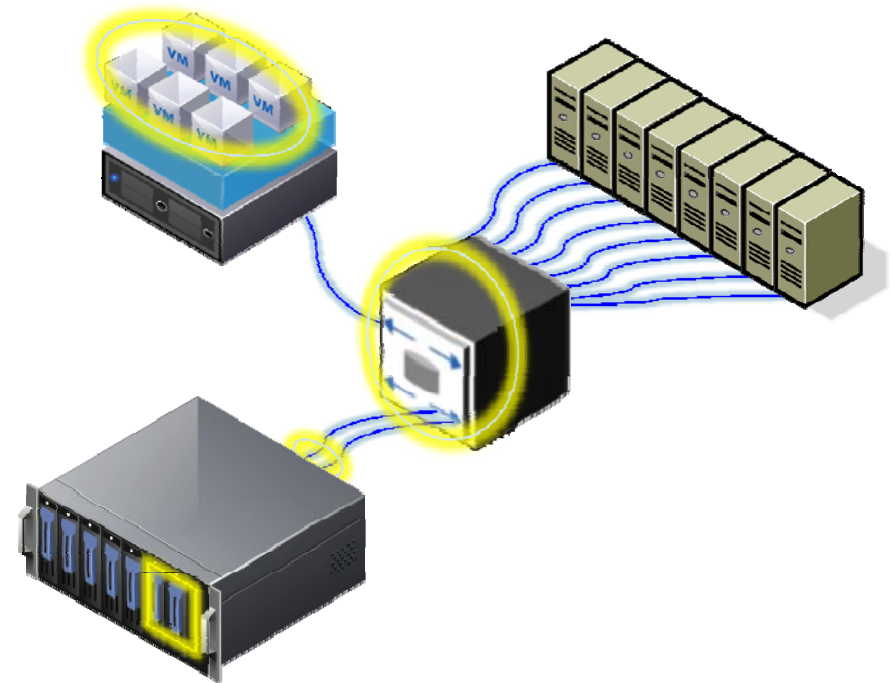# Often Comes From Too Much Oversubscription

- ## Too much within the:
  - Hypervisor
  - LUN
  - HDDs
  - SAN fabric
  - Target Storage ports
- ## Creates positive loop
  - Problems feed on themselves
- ## Finding root cause can be a hair pulling experience
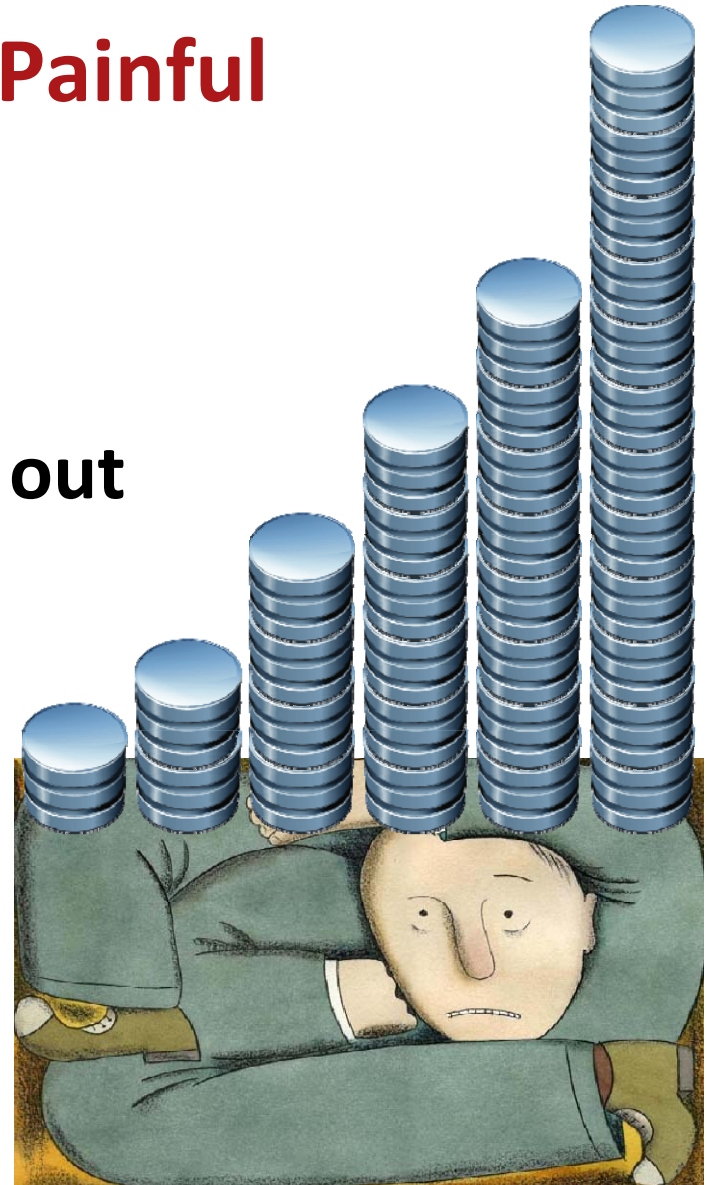 - Especially intermittent – excessive time consumption
    - Must capture/trap problem while it's occurring
      - Which means it has to occur again

Storage System Scaling Still Painful

- **Manually Intensive planning**
- **Storage typically scales up, not out**
  - Rip-out-and-replace
  - Ongoing data migration
    - App disruptive

A Major P.I.T.R. for Large Environments

- **Specifically when VM # exceeds double digits**
  - Time consuming pedantic manually intensive tasks
    - Setting LUN assignments & RAID levels
    - LUN management
  - Most SAN storage LUN expansion is manually intensive

# Ways to Avoid These VMFS Storage Issues

- **Storage that**
    - Is as easy to provision, change, expand as VMs
    - Makes troubleshooting easy
    - Requires minimal to no expertise
        - To implement, operate, and manage
    - Helps avoid LUN oversubscription issues
    - There are 2 types
        - Very simple iSCSI storage
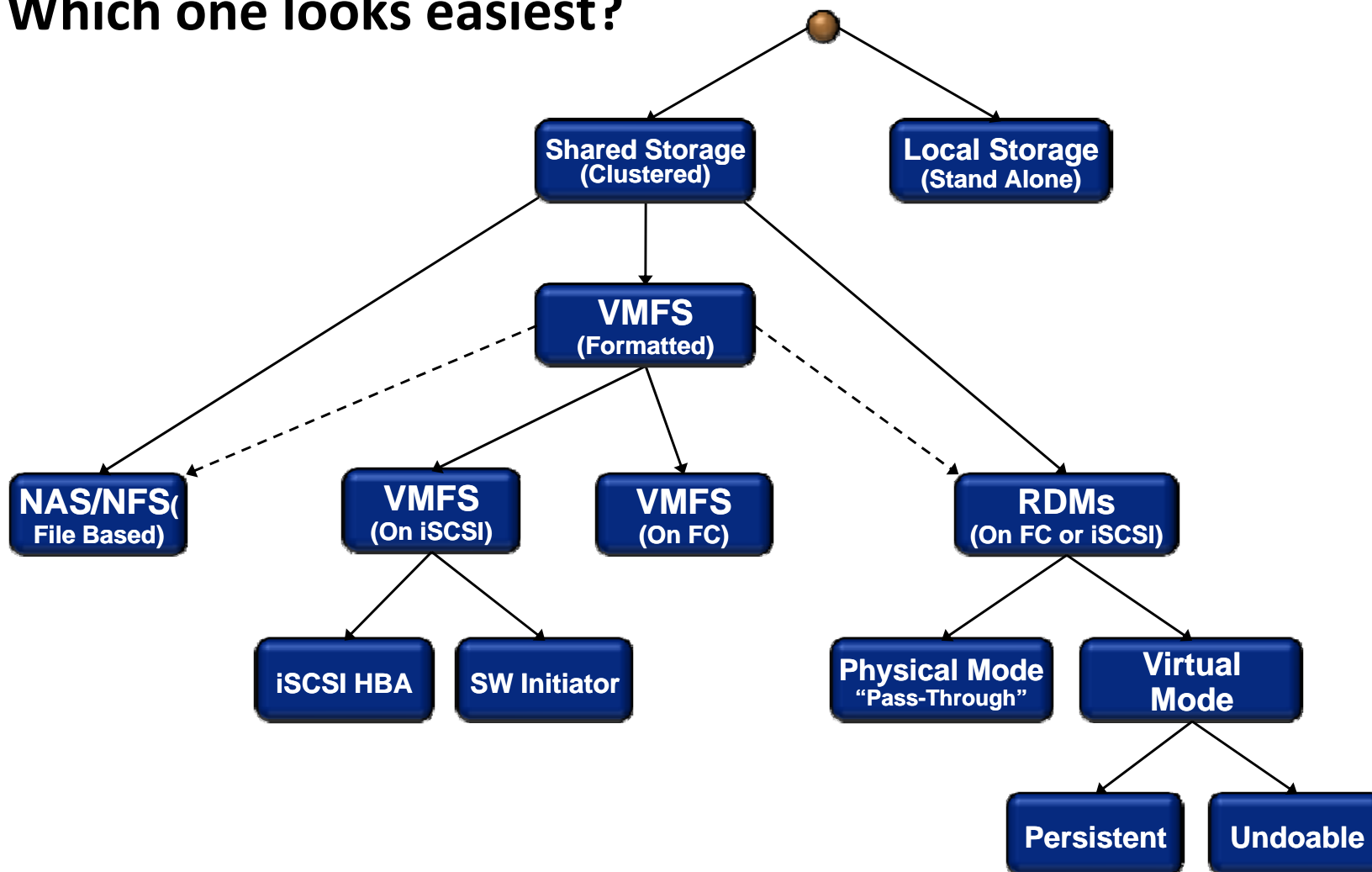   - And the most overlooked & incredibly easy NFS

VMware Shared NFS Storage

- **Incredibly easy provisioning since all VMDKs are just files**
- **VMware VMDK thin provisioning**
  - vSphere 4 default setting
- **Simple online NFS datastore & perf. scaling**
  - Effective w/simple datastore refresh button
  - Perf. Increases w/easy trunking IEEE 802.3ad
- **"0" mgmt or or need for VMFS or RDM**
- **Eliminates constrained or single IO queues**
  - Performance ltd only by server, NAS, or BW IO
- **Greatly simplifies mgmt, ops, & change mgmt**
  - Eliminates FC switch, zone, hard zone, HBA mgmt
  - Eliminates LUN masking, or identical LUN IDs
- **Improves data protection VM BUR**
  - Increased granularity (whole VMs or files within VMs)
- **Optimized random IO response time – ideal for ESX & vSphere**
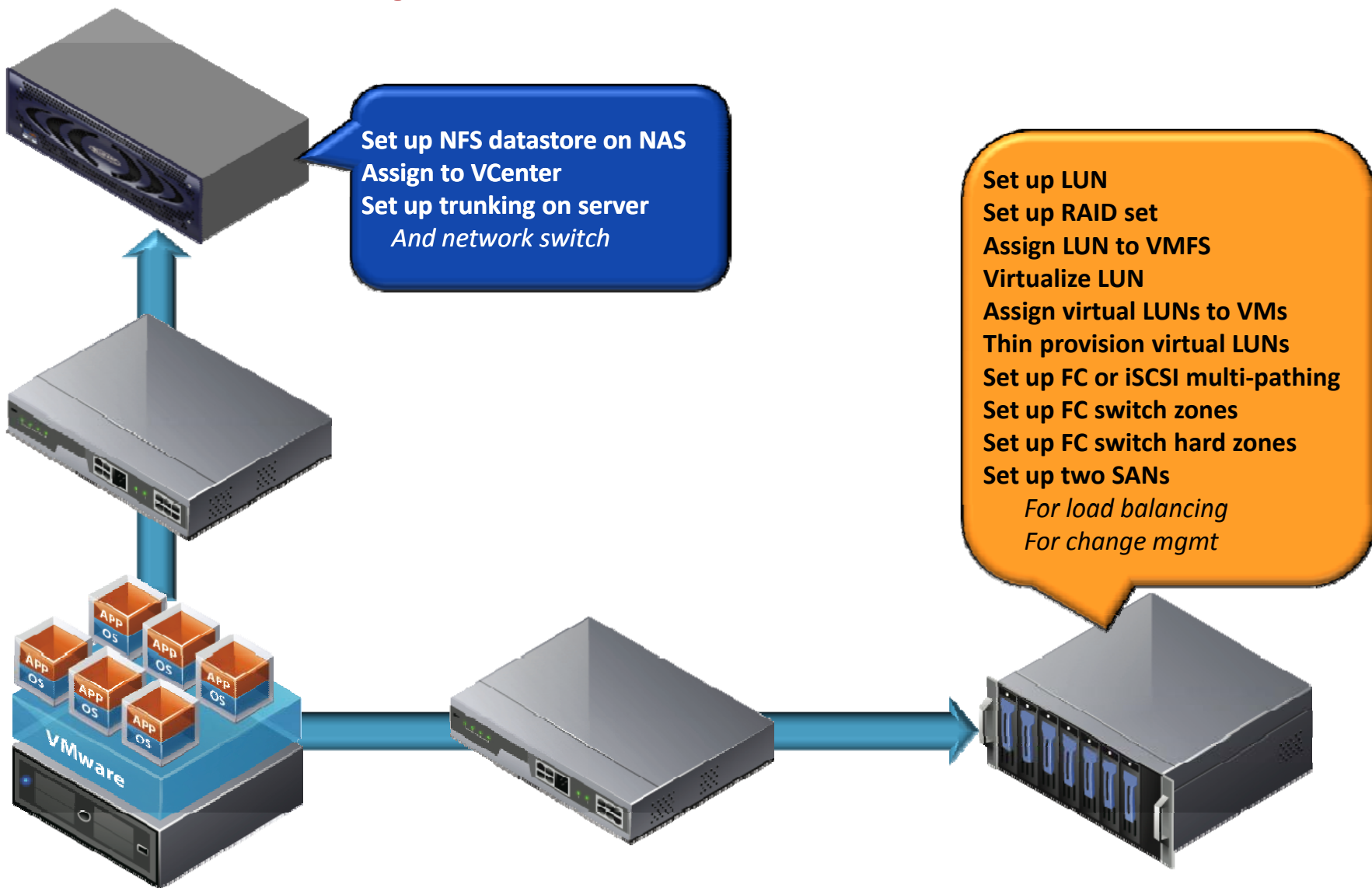- **Single mount across point across multiple IP addresses**
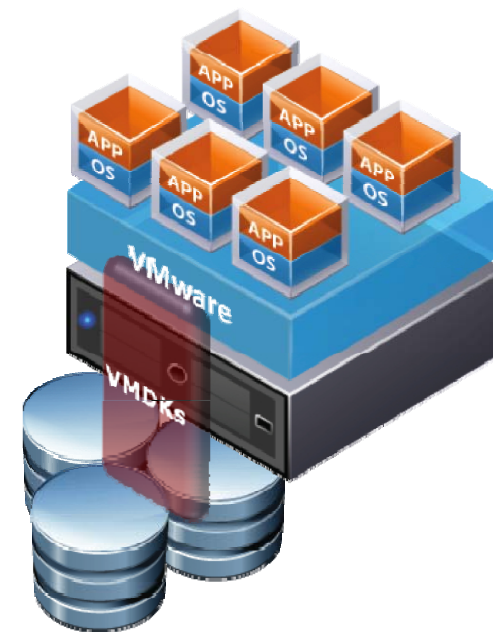
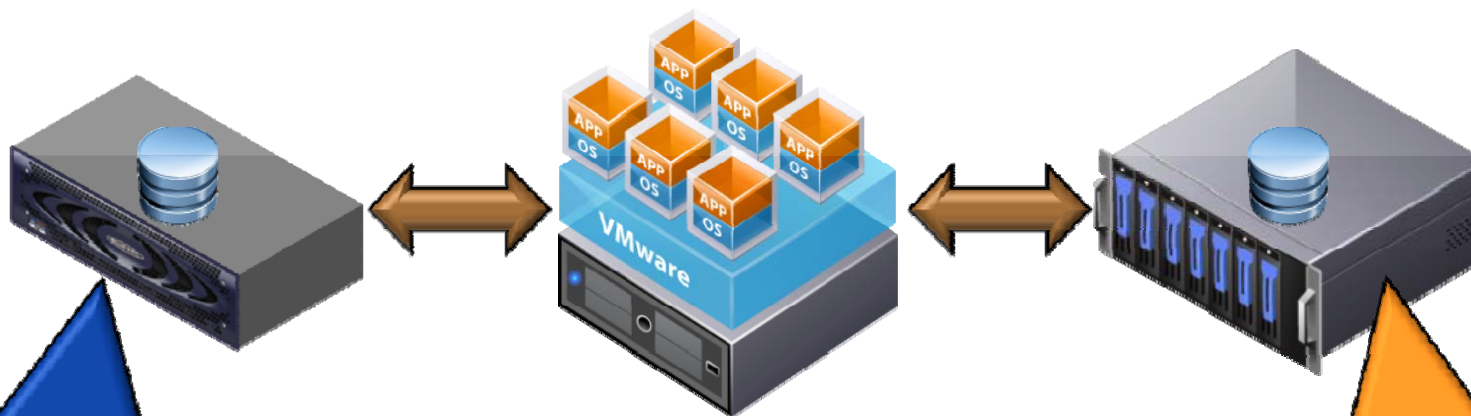# How The Different Storage Options Compare

- **Which one looks easiest?**

```
                                    ( )
                            /                 \
                Shared Storage            Local Storage
                  (Clustered)             (Stand Alone)
                  /      |
                 /       |
              VMFS
            (Formatted)
          /    |    |       \
         /     |    |        \
  NAS/NFS(   VMFS   VMFS      RDMs
 File Based) (On iSCSI)(On FC) (On FC or iSCSI)
            /    \              /         \
      iSCSI HBA  SW Initiator  Physical Mode   Virtual
                              "Pass-Through"    Mode
                                             /       \
                                       Persistent  Undoable
```

How NFS Compares to VMFS

Set up NFS datastore on NAS
Assign to VCenter
Set up trunking on server
  And network switch

Set up LUN
Set up RAID set
Assign LUN to VMFS
Virtualize LUN
Assign virtual LUNs to VMs
Thin provision virtual LUNs
Set up FC or iSCSI multi-pathing
Set up FC switch zones
Set up FC switch hard zones
Set up two SANs
  For load balancing
  For change mgmt

NFS Datastores Simpler Capacity Expansion

VMware

Increase storage on NFS side
Refresh each VM's storage
Use newly provisioned storage

Halt IO to the LUN
Increase storage on SAN side
Increase datastore size
 Including underlying partition
Rescan each VM connected
 To SAN storage
Resume IO to LUN
Use newly provisioned storage

I'm confused. Then Why isn't NFS the Dominant VMware Storage????

VMware Infrastructure NFS Myths

- **VMware doesn't fully support NFS**
  - Or all functions on NFS
- **NFS is too slow for VMware**
- **Few use NFS w/VMware in production**
- **NFS VMware CPU load is heavier**
- **VMware limited to 8 NFS datastores**
- **NFS datastores limited to 16TB**
- **NFS thin provisioned VMDKs auto-rehydrate**
  - When moved or cloned
- **Windows VMs can't boot or utilize NFS datastores**

VMware vSphere 4.1 NFS Support Improved

- **NFS throughput over 10G improved significantly**
  - 12 – 40% for Reads & 32 – 124% for Writes
- **Vs. iSCSI improvements over 10G of only**
  - 6–23% for Reads & 8–19% for Writes
- **Plus better performance monitoring**
  - Providing latency & throughput statistics for
    - Host per NFS datastore
    - VMs per NFS datastore
    - And VMs per VMDK

NAS

Some VMware Storage Conclusions

- **RDMs required for clusters**
- **VMFS SAN storage is the most common**
  - Has some issues to watch out for
- **NFS storage is the easiest**
- **iSCSI storage is VMFS SAN storage**
  - And pretty darn easy as well
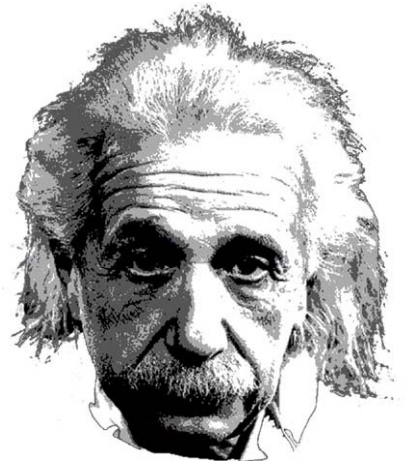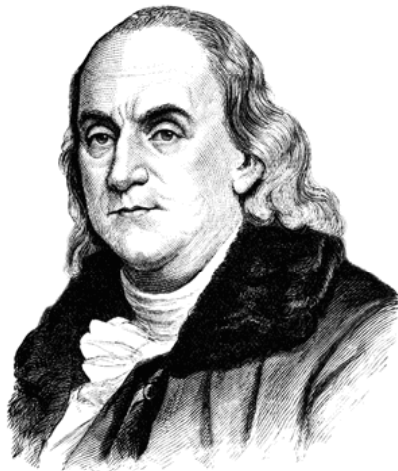   - Depending on vendor
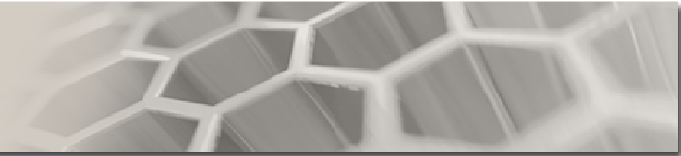
VMware DP Major Issues

- **Consumes major server resources**
  - Significantly more than comparable physical servers
  - Reducing server consolidation benefits
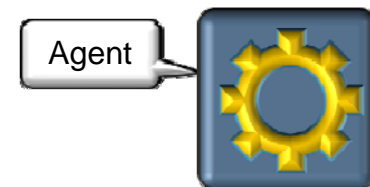- **Difficult to manage**
- **Why????**
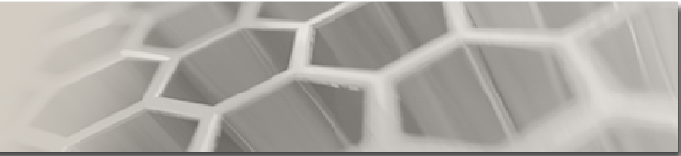
Definition of Insanity is Doing the Same Thing Over & Over & Expecting Different Results

Key DP Vmware Problem – Agents

Agent

- **Agents are software w/admin privileges**
  - A.k.a. plug-ins, lite agents, client software
- **Role is collect data & send it to a backup or media server**
  - Complete files and ongoing incremental changes
- **Separate agents typical / OS, database, ERP, & email app**
  - As well as for BU, CDP, & Archiving / app
  - Can be more than one agent / server (OS, DBMS, Email, etc.)
 - When agents deduplicate and/or encrypt = more resources

Compared to Physical Servers…

- **Instead of 1 or 2 agents per physical server**
  - There are lots of agents per physical server
- **Wasting underutilized server resources is one thing**
  - It's quite another when that server is oversubscribed
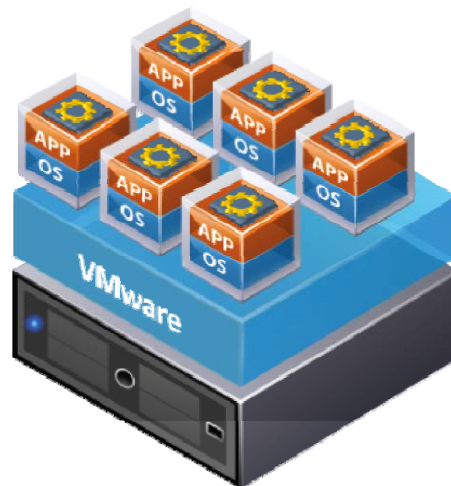
Reduces Server Virtualization Consolidation

- **Agents limit VMs / physical server**
  - Reduces effective consolidation benefits
  - Decreases financial savings, payback, & ROI
- **VM backups bottleneck the IO**
  - Simultaneous backups have bandwidth constraints
  - Backups must be manually scheduled serially
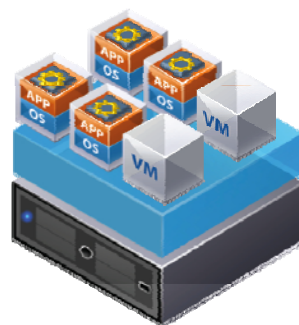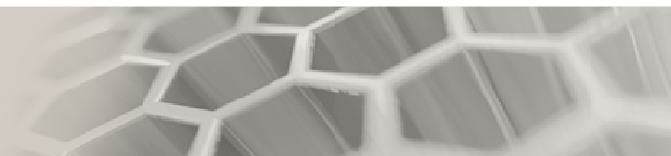
# Agents Have Other Frustrations As Well

- **Agents compromise security**
- **Agents are very difficult to admin & manage**
  - Especially as servers & apps proliferate

Agents Compromise Security

- **A firewall port must be opened per agent**

- **Agents have admin privileges**
  - Creates a backdoor access to everything on the server
  - Hackers target agents – BU data must be important
    - Agents are listening on a port just waiting to be hacked
    - Hacker can try to hack dozens thousands of servers
      - Often w/o being detected - > agents = > attack points
  - Lack of encryption in-flight puts transmitted data at risk
  - Agent encryption wastes even more server resources
    - A no win situation
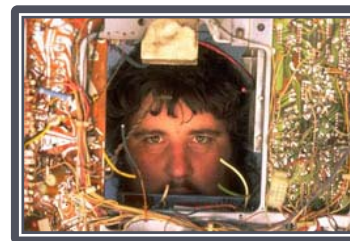
Agents are Difficult to Admin & Manage

- **Installing an agent can be maddeningly frustrating**
  - Requires an app disruptive system reboot to initialize
  - Upgrades patches hot fixes require app disruptive system reboot[1]
  - Agents upgraded when SW is upgraded

[1]Some BU software have an automated upgrade process; however, the reboots are still disruptive

Agent = Too Many Restore Failures

- **Limited rollback capability if at all**
  - No time based versioning
  - Multi-step restores
    - Data has to be restored from backup media
      - To media or backup server before restored to server
      - Usually requires multiple steps & passes
- **Infrastructure complexity = increased failures**
  - More agent software parts increases failure probability
    - Lots of agent flavors Platforms, OS, DBMS (all kinds), & email
- **Complicated Troubleshooting**
  - Particularly aggravating when an agent stops working
  - The larger the environment, the more complicated the trouble shooting

Three Ways to Mitigate VMware DP Issues

- **VADP based Data Protection**
- **Storage based snapshots**
- **Cloud based BUR**

# VADP – vStorage API for Data Protection

- **Significantly reduces agents on VMs**
  - Allows BU software to collect data through VADP
    - Utilizing VMware snapshots & VSS for Windows
- **Best for Windows & Linux guests**
  - All storage architectures – FC, ISCSI & NFS
  - Faster incremental BU
  - Faster differential BU
  - Faster full VM image BU
  - File level backup & restore
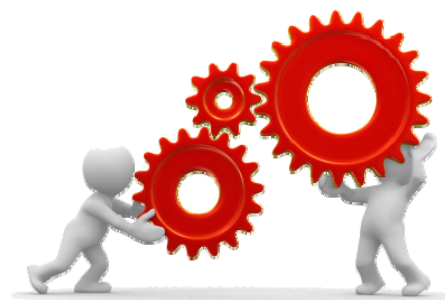- **Available on vSphere 4 & ESX 3.5 (latest)**

VCB (EOL)
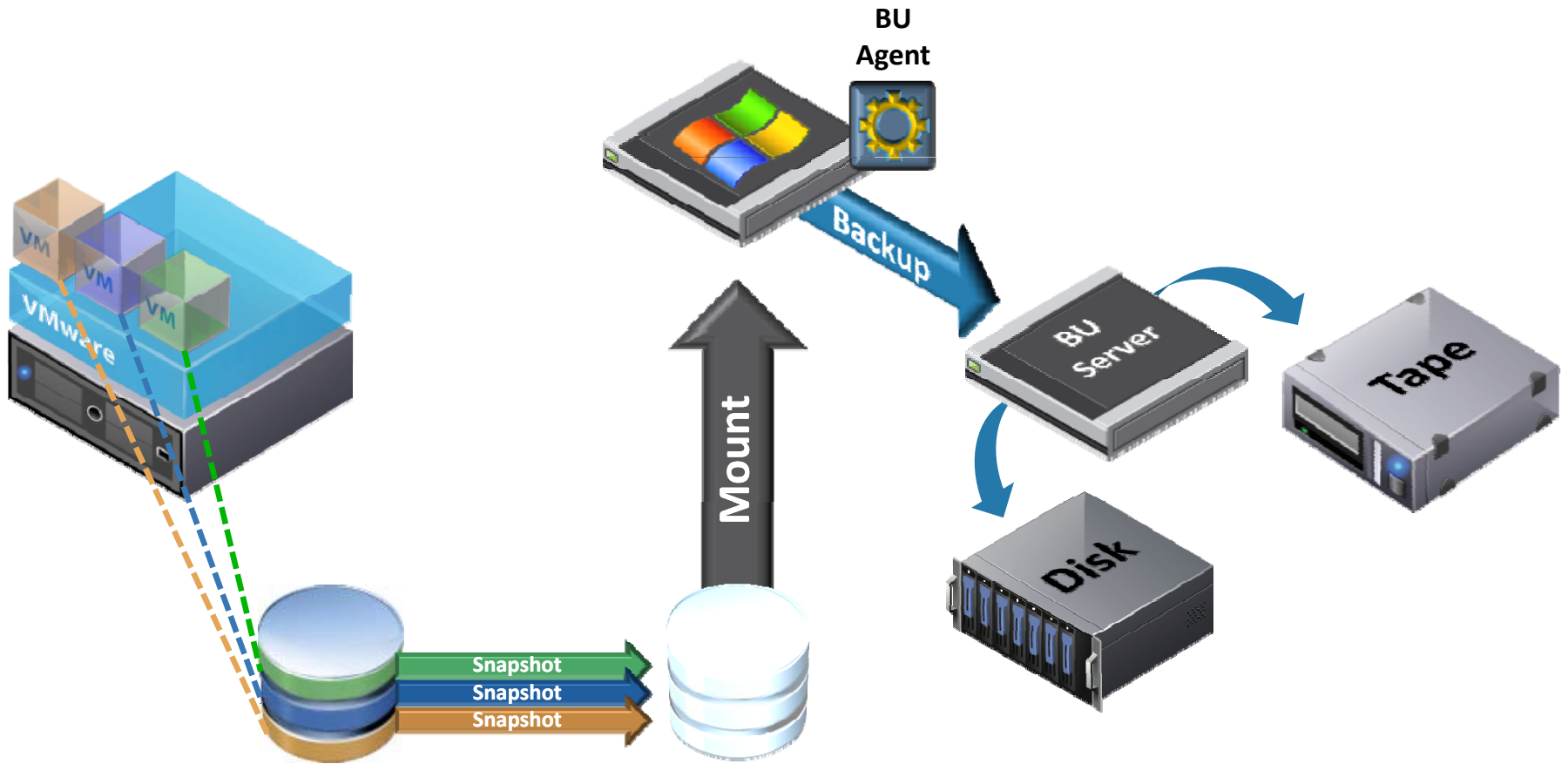
VADP
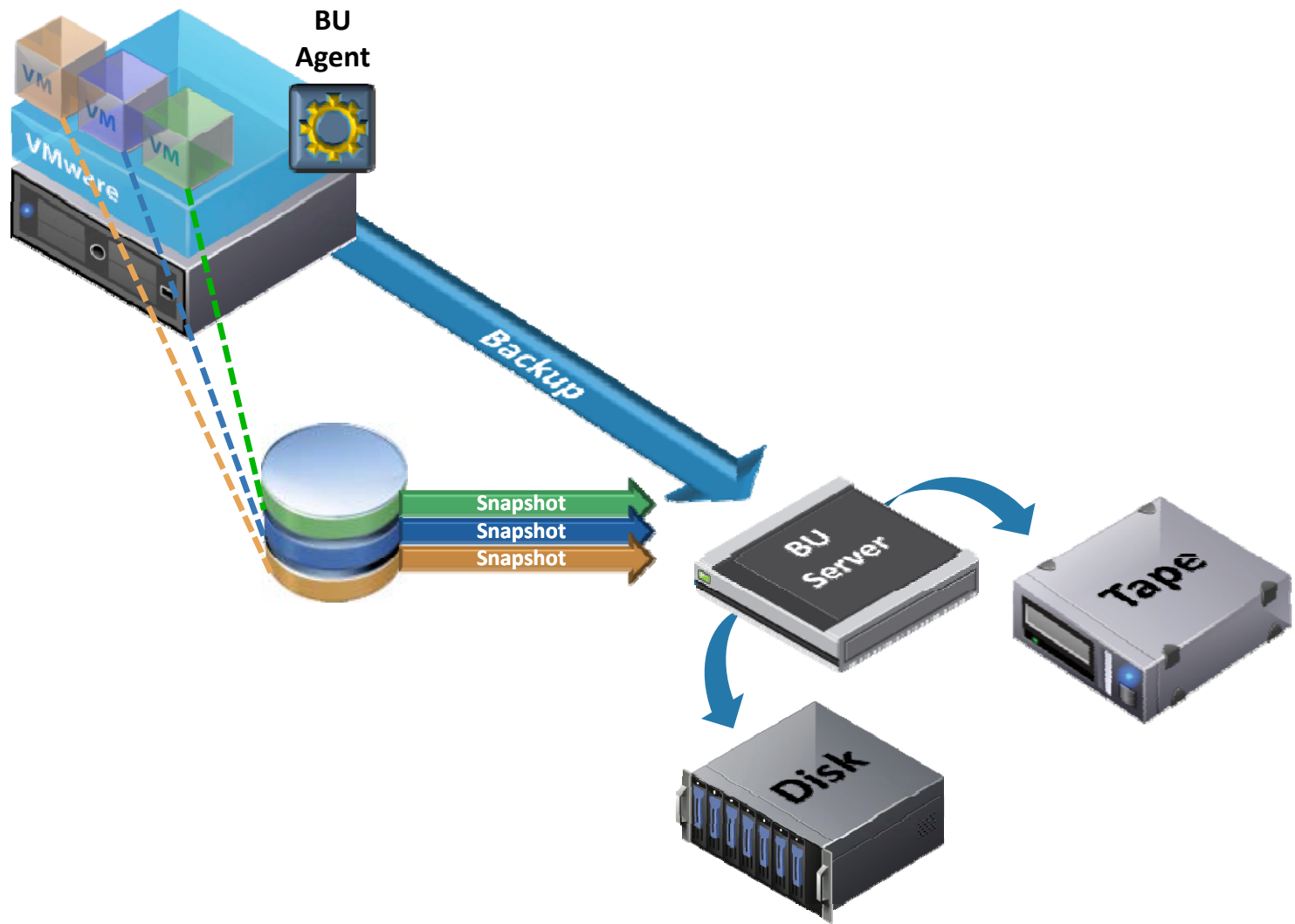
VADP vs VCB

|	**VADP**	**VCB**
Requires add'l download & install	No, built into the data protection software	Yes
Full VM image backup	Yes, single step copy – source to target	No
Incremental VM image BU	Yes	No
File level BU	Yes, Windows & Linux	Yes, Windows only
Full VM image restore	Yes	Yes, by using VMware Converter
Incremental VM image restore	Yes	No
File level restore	Yes, using restore agents	Yes, using restore agents

But VADP Has Limitations

- **Just 1 piece of the puzzle**
  - Requires integration w/BU software
    - VMware's low-end VDR is an example
  - No Unix file level backup or restores
  - Still needs at least one agent on vSphere
  - Not application aware
    - Needs agents for structured apps not on Windows
    - Or on Windows but not VSS supported
  - Requires physical server resources

Storage Based Snapshots
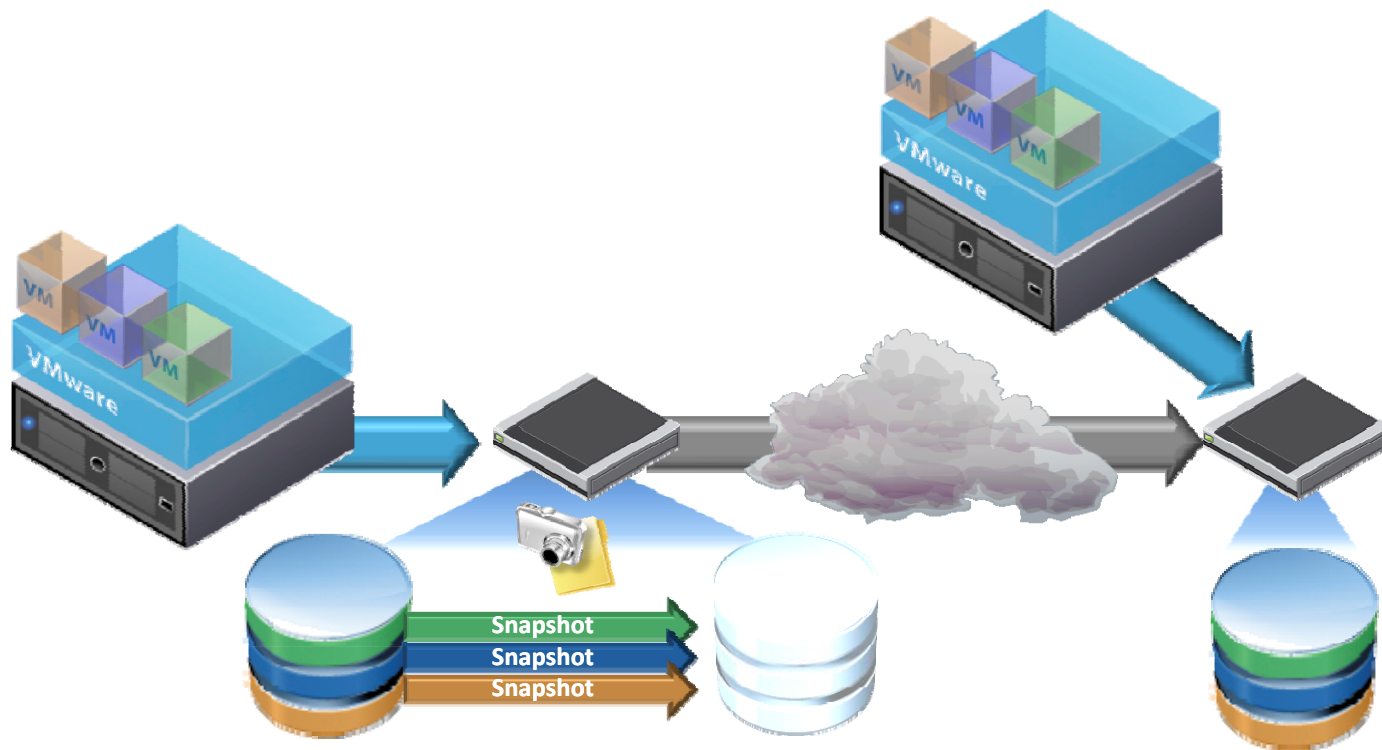
Storage Snapshots Are Generally Pretty Easy

- **No agents on servers or apps**

- **Medium to fine granularity RPO & RTO**

- **Snapshots sent to other site**
  - With possibility of bi-directional

- **Snap restores = mount data, point & done**
  - Remote Snapshot promotable to production volume

- **Fast – virtually instantaneous with no BU Windows**

- **Centrally administer w/storage**

- **In some cases even deduped**
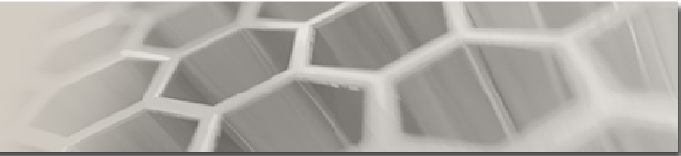
Storage Snapshots Has Its Issues
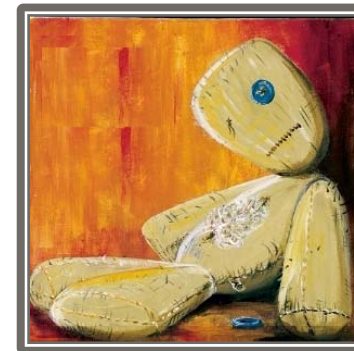
- **Snaps often not structured data crash consistent**
  - Most not integrated w/Windows VSS
  - Structured apps requiring crash consistency = agents
    - Requires integration w/BU SW console
    - Agents quiesce DBMS, providing write consistency
    - BU software tells storage system to snapshot
- **Snapshots / system**
- **High cost w/Capacity based licensing**
  - Storage system tends to be higher cost

Snapshot Workarounds – BU SW + Agents

- **Integrate w/BU Software & BU agents**
  - To quiesce structured apps
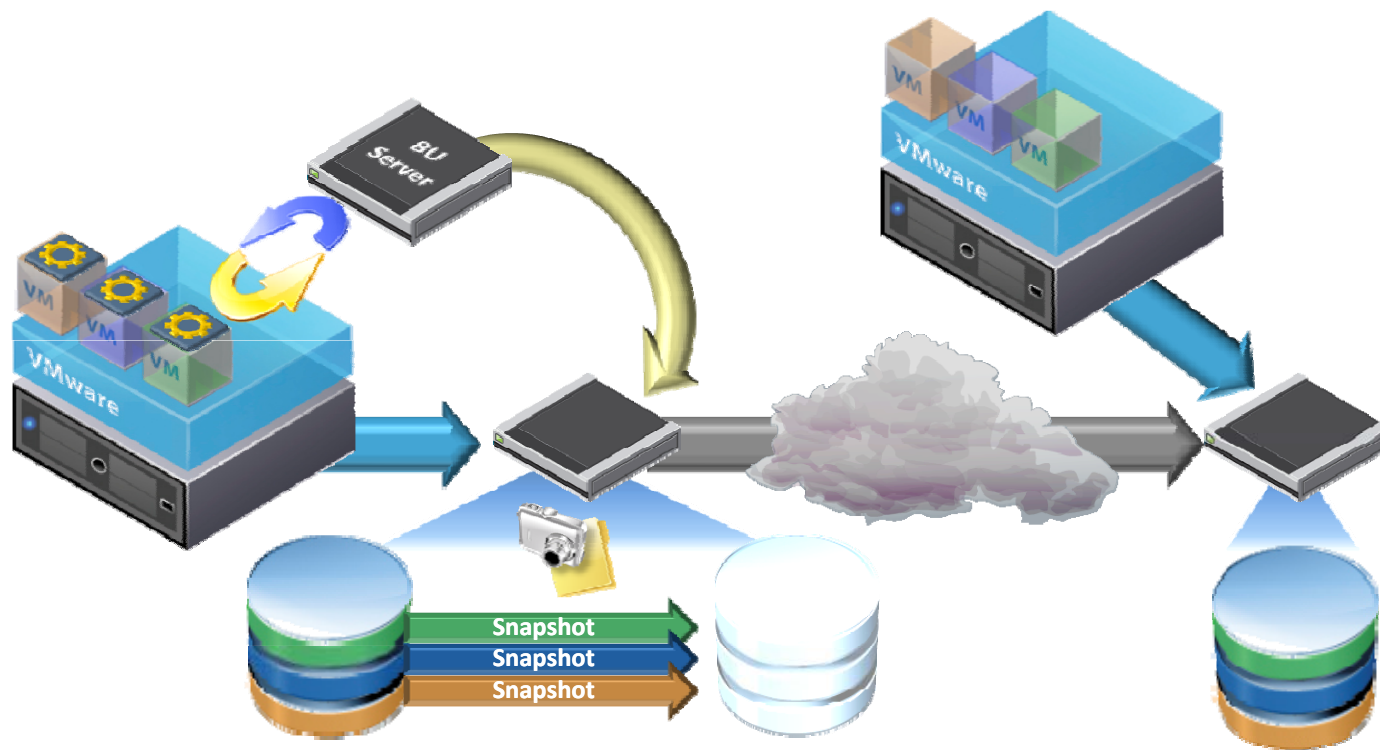
Storage Snapshots w/VMware Should...

- **Be integrated w/VSS**

- **Have nominal operational burden**

- **Eliminate backup window issues**

- **Provide minimal app disruptions**

- **Be automated**

- **Allow for extensive numbers of snapshots**

Cloud Based BUR (Backup & Restore)

- **Agentless or agent limited**
  - App non-disruptive
    - Implementations, ops, mgmt
- **Local recoveries**
  - Onsite physical or virtual appliance
- **VADP integrated**
  - CBT a plus
- **Disaster recover anywhere**
  - Offsite data
- **From fine to coarse grain RPO & RTO**
- **Limited onsite storage & infrastructure requirements**
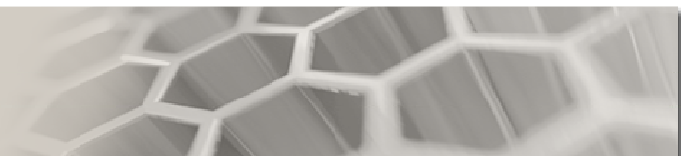
Cloud BUR

- **Very low TCO**
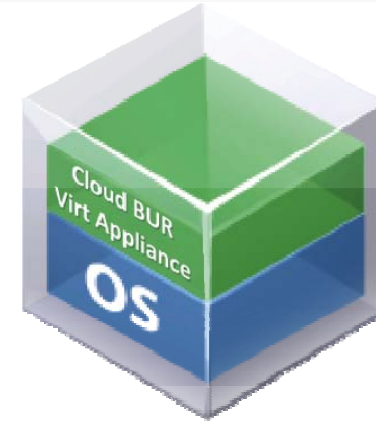
Not All Cloud BUR is Created Equal

- **Make sure Cloud BUR service provider meets your requirements:**
  - Multi-tenant
    - Stores your data encrypted
      - Gives you key mgmt
  - Integrates with VMware VADP
    - CBT is a plus
  - Is as agentless as possible
    - Time based versioning, even CDP
  - Provides both local recoveries & DR
    - File level for all VMs (not just Linux & Windows)
  - Is easy to recover at any granularity
    - One pass recoveries preferred
  - Provides written SLA RPO & RTO guarantees
  - Makes it fairly simple to move to another provider
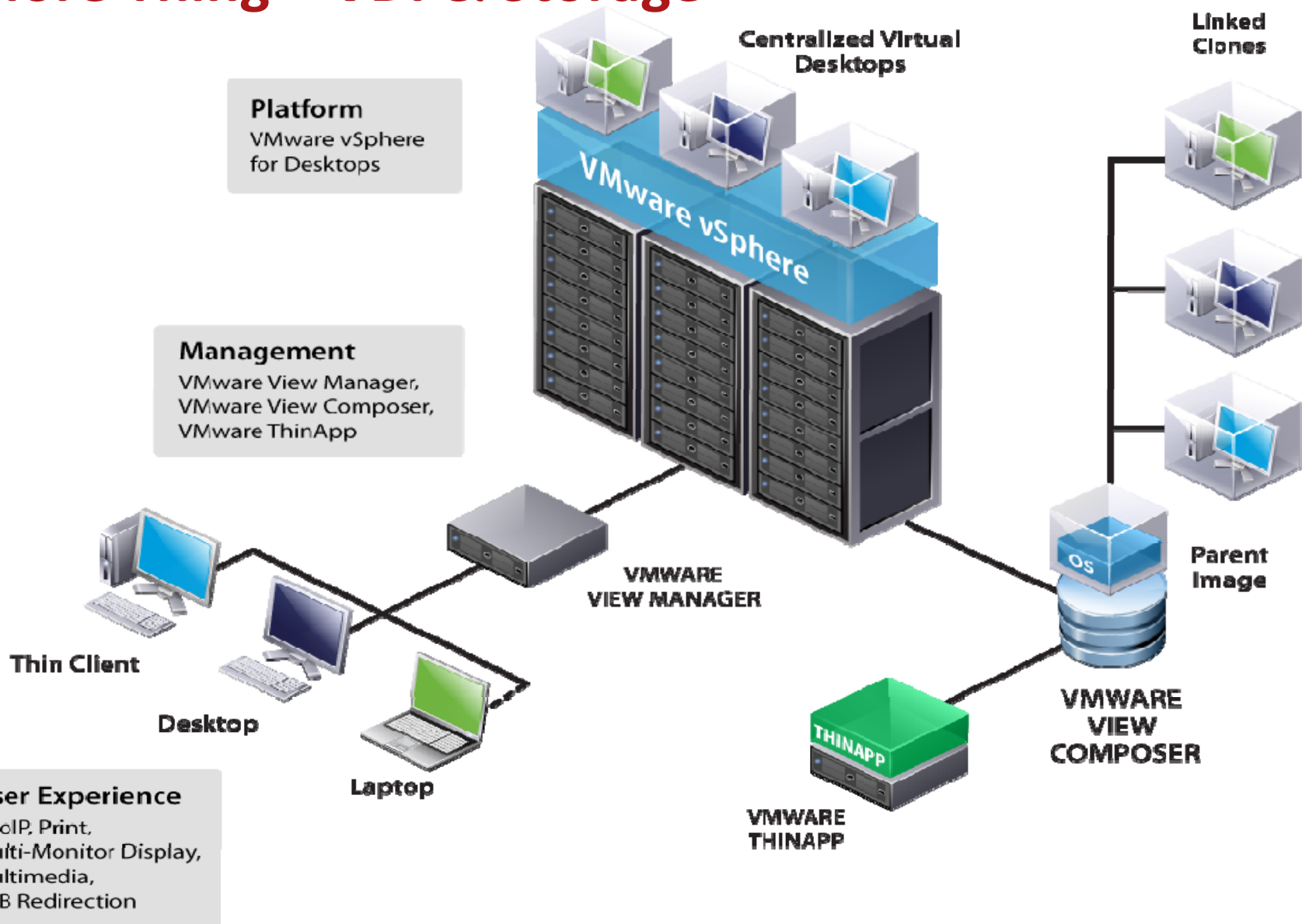   - Or in-house
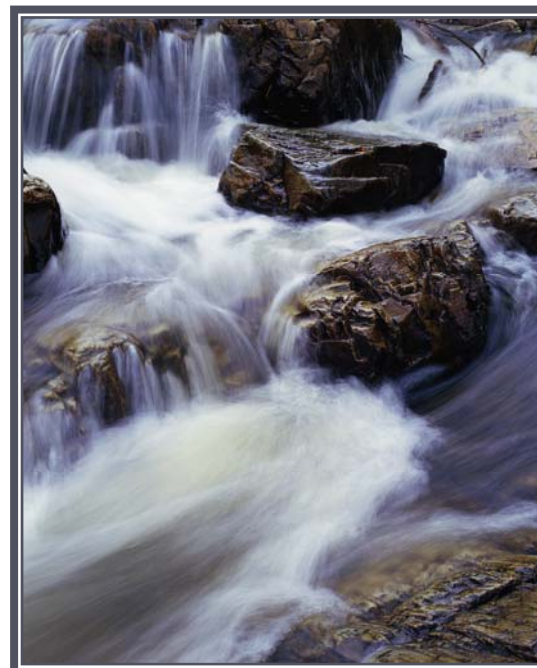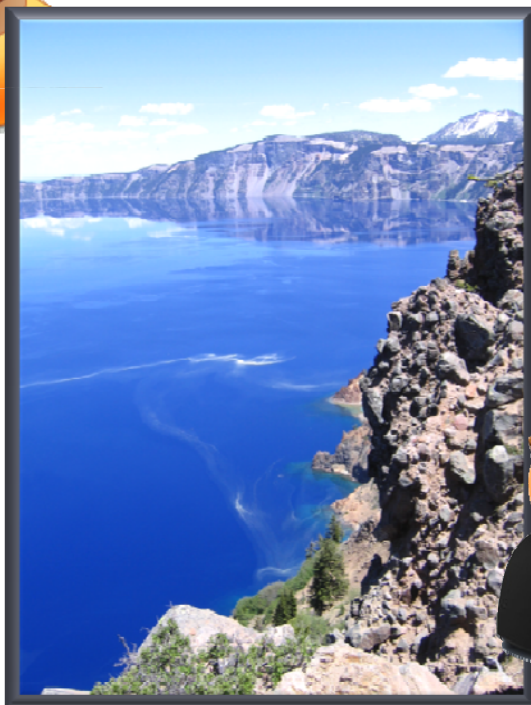
VMware DP Conclusions

- **There are 3 very good alternatives**
  - VADP
    - An enabling tool
  - Storage Snapshot
    - More than just a tool
  - Cloud based BUR
    - Often uses both plus infinitely scalable

1 More Thing – VDI & Storage

Platform
VMware vSphere
for Desktops

Management
VMware View Manager,
VMware View Composer,
VMware ThinApp

User Experience
PCoIP, Print,
Multi-Monitor Display,
Multimedia,
USB Redirection

Centralized Virtual Desktops

Linked Clones

VMware vSphere

Thin Client

Desktop

Laptop

VMWARE VIEW MANAGER

VMWARE THINAPP

VMWARE VIEW COMPOSER

Parent Image

Now You Know Where the Rocks Are

marcstaimer@mac.com

503-579-3763