Filesystem Sharing

Velocity Software Inc.
196-D Castro Street
Mountain View CA 94041
650-964-8867

Velocity Software GmbH
Max-Joseph-Str. 5
D-68167 Mannheim
Germany
+49 (0)621 373844

Rick Troth
Velocity Software
<rickt@velocitysoftware.com>
http://www.velocitysoftware.com/

VM and Linux Workshop 2012
University of Kentucky

Disclaimer

The content of this presentation is informational only and is not intended to be an endorsement by Velocity Software. (ie: I am speaking only for myself.) The reader or attendee is responsible for his/her own use of the concepts and examples presented herein.

In other words: Your mileage may vary. “It Depends.” Results not typical. Actual mileage will probably be less. Use only as directed. Do not fold, spindle, or mutilate. Not to be taken on an empty stomach. Refrigerate after opening.

In all cases, *“If you can't measure it, I'm just not interested.”*

VELOCITY SOFTWARE

Filesystem Sharing

Some history of shared content

Some ways of sharing content

Some reasons for sharing content

Some solutions to sharing content

Focus: Files and Filesystems

Option: Op-Sys Update and Maint

Perspective: systems, what helps sys admin

VELOCITY SOFTWARE

History of Shared Digital Data

Tapes

Disks

Network

social/consumer

excessive duplication

Only wimps use tape backup: real men just upload their important stuff on ftp, and let the rest of the world mirror it

-- Linus

VELOCITY SOFTWARE

Data Sharing Methods

Tape, Cards

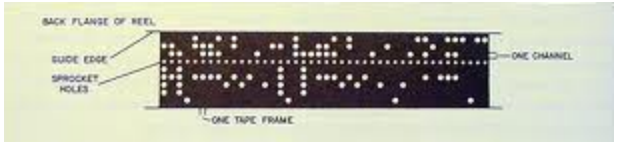
Packs, Floppies

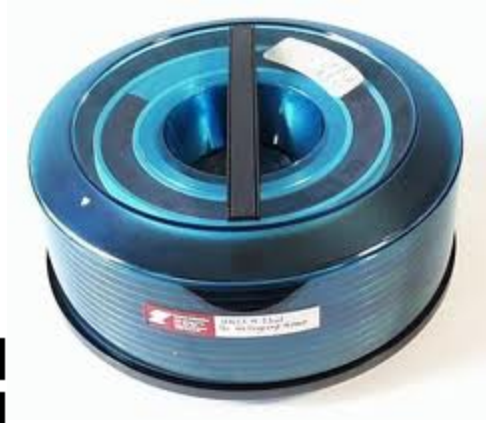
Network Filesystems

CD ROM, Flash

Scan Codes

Network Synch





VELOCITY SOFTWARE

What does “sharing data” mean?

Input/Output

Immediacy

Reliability

Viability

Security

Online -vs- Offline / Dynamic -vs- Resting

VELOCITY SOFTWARE

Filesystem Sharing Rationale

Distribution

Collaboration

Recovery

Control

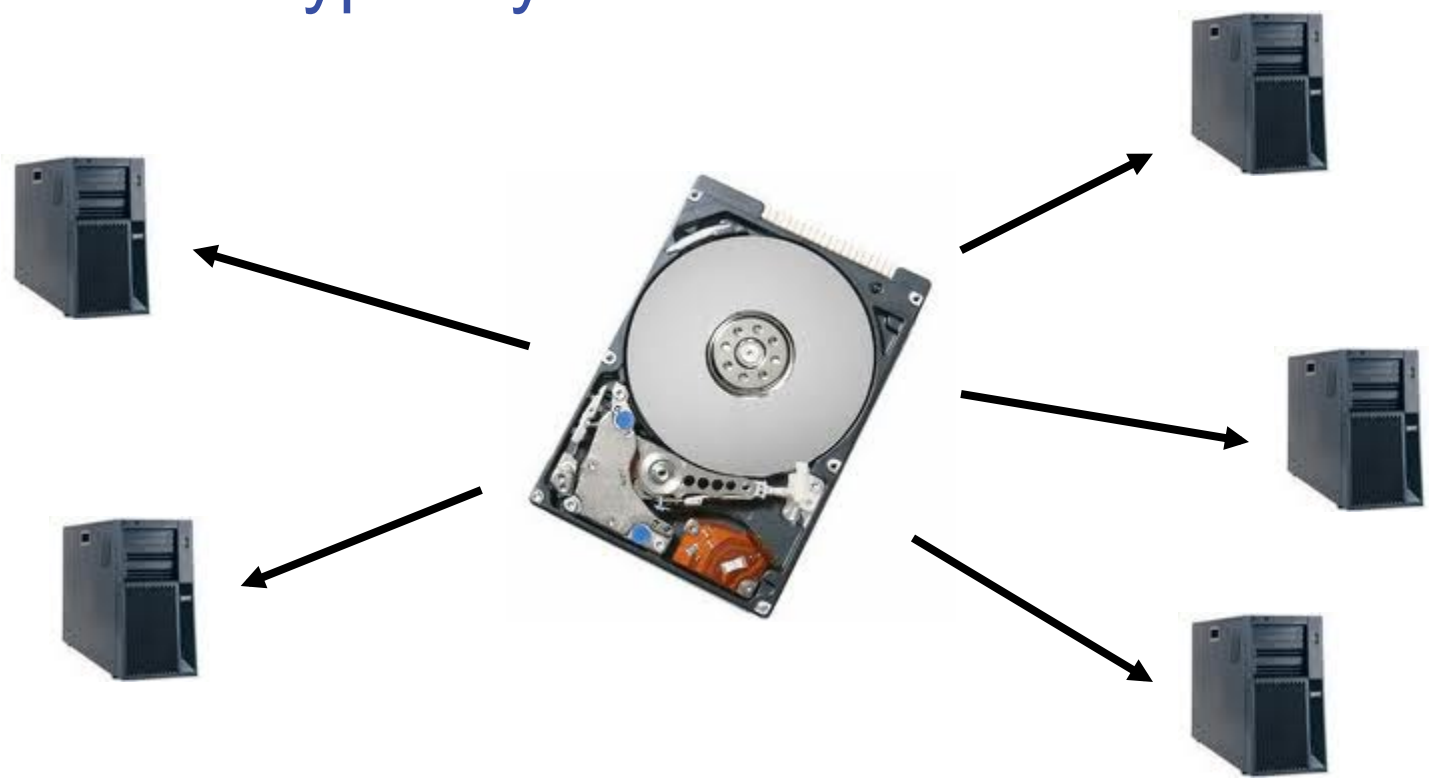
Deduplication

Scalability

VELOCITY SOFTWARE

Shared FS on Disk

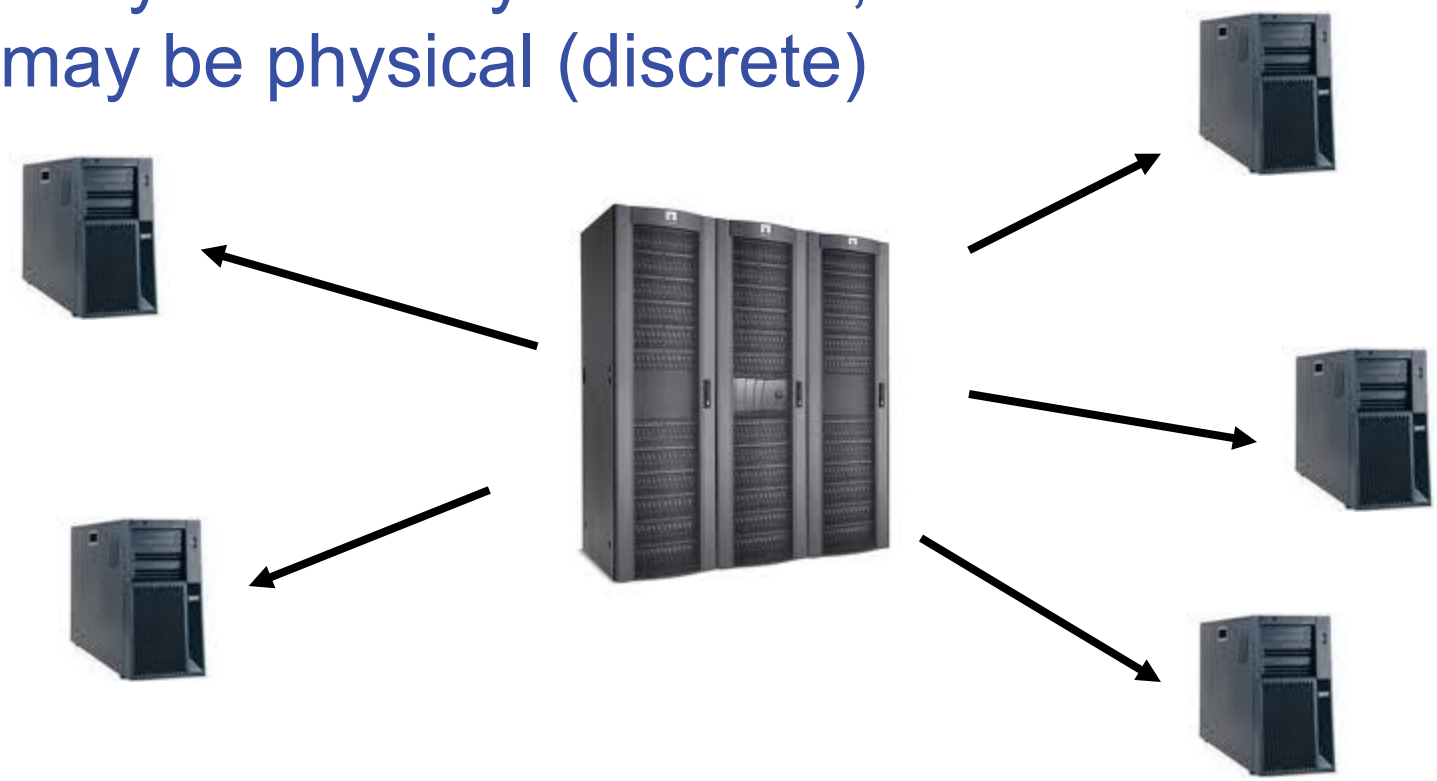
"clients" are typically virtual



VELOCITY SOFTWARE

Shared FS in SAN

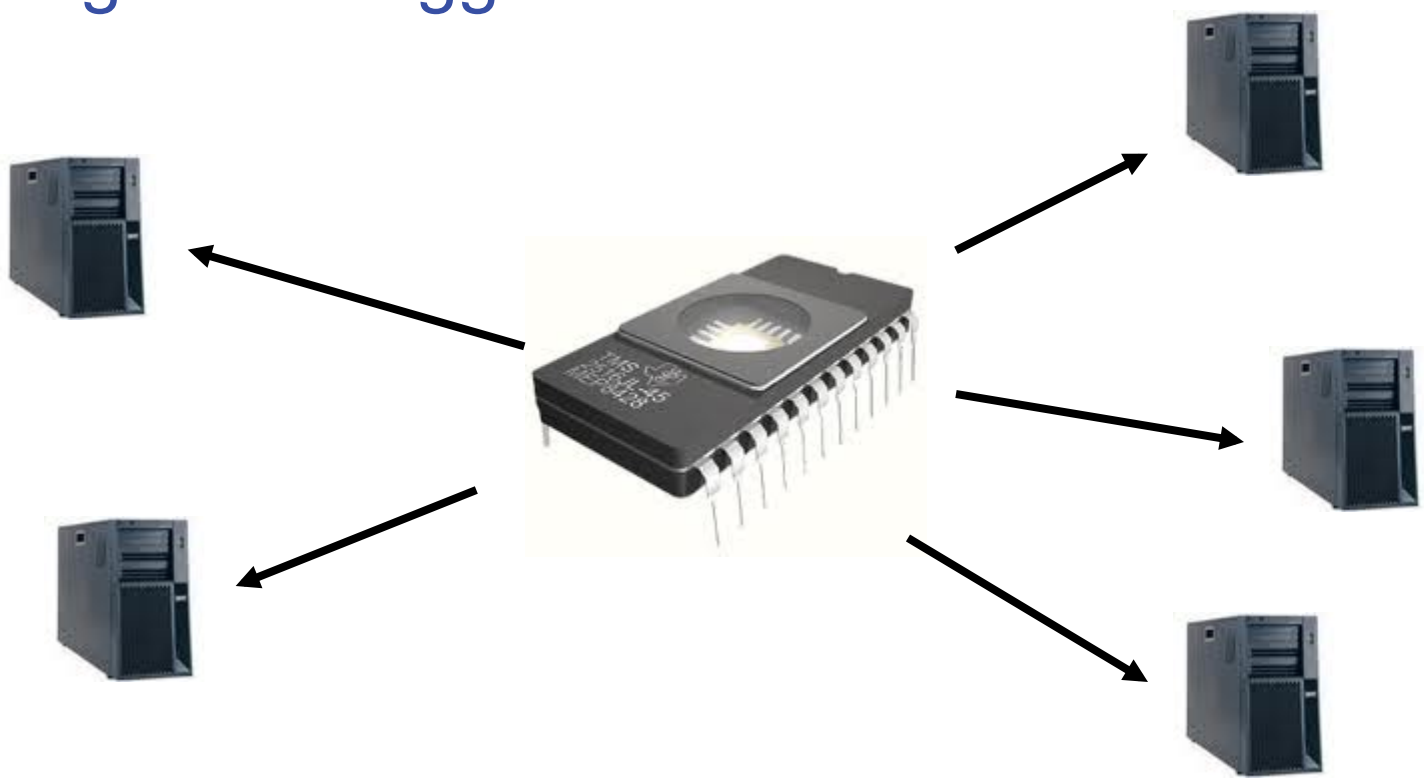
Client systems may be virtual, or may be physical (discrete)



VELOCITY SOFTWARE

Shared FS in ROM

Sharing ROM suggests virtual



VELOCITY SOFTWARE

Shared Disk

Standard for z/VM (minidisks)

Must be R/O (block cache)

Candidate FS:

- EXT2 (no journal)
- ISO-9660 (CD-ROM)

VFAT tends to want partitioning

GFS, OCFS2

Shared SAN too (works for physical)

VELOCITY SOFTWARE

Shared Disk

```
# df
Filesystem            1K-blocks      Used Available Use% Mounted on
/dev/dasda               476104    394940     56588  88% /Linux-s390
/dev/dasdb               126960     26544     93864  23% /
/dev/dasda               476104    394940     56588  88% /lib
/dev/dasda               476104    394940     56588  88% /bin
/dev/dasda               476104    394940     56588  88% /sbin
/dev/dasda               476104    394940     56588  88% /usr
/dev/dasda               476104    394940     56588  88% /boot
udev                      30580         0     30580   0% /dev
/dev/dasdk               253920    112932    127884  47% /opt/CD2
/dev/dasdm               476104    302828    148700  68% /usr/src
tmpfs                     30580         0     30580   0% /tmp
```

VELOCITY SOFTWARE

Mount by Label

Standard for z/VM (host disks or “full pack”)

Increasingly popular with Linux

Also mount-by-uuid (works for swap)

Does not require partitioning

Consistent across architectures

More Secure, not less

R/O media is immutable

Shared media may be R/O

... ergo ... shared *may* be immutable

VELOCITY SOFTWARE

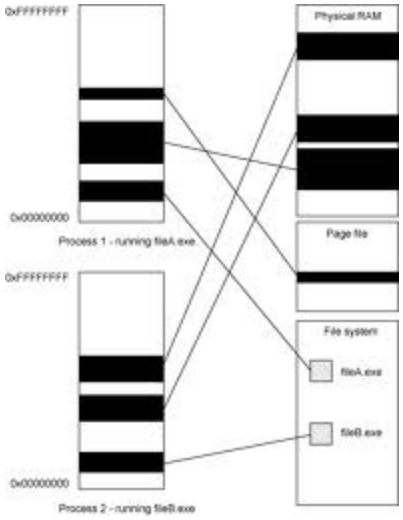
Shared Memory

Shared memory is common

DCSS – variable modes

- Restricted – maybe
- `TYPE SR`

Big boost for CMS
"back in the day"



VELOCITY SOFTWARE

Shared Memory

```
# df
Filesystem            1K-blocks      Used Available Use% Mounted on
/dev/dcssblk0            380888    321900     39328  90% /Linux-s390
/dev/dasdb               126960     24652     95756  21% /
udev                      22448         0     22448   0% /dev
/dev/dasdk               253920    112932    127884  47% /opt/CD2
tmpfs                     22448         0     22448   0% /tmp
/Linux-s390/usr          380888    321900     39328  90% /usr
/Linux-s390/bin          380888    321900     39328  90% /bin
/Linux-s390/boot         380888    321900     39328  90% /boot
/Linux-s390/sbin         380888    321900     39328  90% /sbin
/Linux-s390/lib          380888    321900     39328  90% /lib
```

VELOCITY SOFTWARE

XIP

The “extreme sport” ... execute-in-place

- No copying of content (disk to memory)
- No I/O
- Just point to it and go!

But ... “binaries are small,
thus the savings are mediocre at best.”

VELOCITY SOFTWARE

Automounter

NFS ... and/or SMB

CD-ROM

USB, flash

'vmlink'

DCSS

VELOCITY SOFTWARE

About Partitioning

- Partitioning is another layer, added complexity
- Partitioning may not be needed, find out if it is ... or not
- Certain (non-Linux and non-VM) systems or environments expect it

VELOCITY SOFTWARE

About Partitioning

CDL if you need to share with z/OS

"CMS RESERVE" for direct sharing with CMS

Traditional (PC) partition table makes Windows happier

VELOCITY SOFTWARE

About Partitioning

```
# ls -lad *.fba
-rw-rw---- 1 rmt  root 402653184 2011-09-18 19:41 01b0.fba
-rw-rw---- 1 rmt  root  67108864 2012-05-30 14:48 01b1.fba
-rw-rw---- 1 rmt  root  33554432 2011-09-18 19:42 01b8.fba
-rw-rw---- 1 rmt  root 402653184 2011-09-18 19:42 01b9.fba
-rw-rw---- 1 rmt  root 268435456 2011-09-18 19:42 01ba.fba
-rw-rw---- 1 rmt  root  16777216 2011-09-18 19:43 01bb.fba
-rw-rw---- 1 rmt  root  33554432 2011-09-18 19:43 01bc.fba
-rw-rw---- 1 rmt  root         0 2011-07-11 17:24 01bd.fba
-rw-rw---- 1 rmt  root   1474560 2011-09-18 19:43 01be.fba
lrwxrwxrwx 1 root root         8 2012-02-26 21:00 01bf.fba -> /
 dev/sda

# mount -o loop 01b1.fba /mnt
```

VELOCITY SOFTWARE

About Partitioning

```
# ls -la /mnt
dr-xr-xr-x  2 root root  4096 2010-09-19 20:58 bin
drwxr-xr-x 11 root root  4096 2012-05-30 13:50 etc
dr-xr-xr-x  2 root root  4096 2010-09-19 20:58 lib
dr-xr-xr-x  2 root root  4096 2010-09-19 21:07 Linux-s390
dr-xr-xr-x  2 root root  4096 2010-09-19 20:58 proc
drwx------  6 root root  4096 2010-11-30 18:40 root
dr-xr-xr-x  2 root root  4096 2010-09-19 20:58 sbin
dr-xr-xr-x  2 root root  4096 2010-09-19 20:58 sys
drwxrwxrwt  2 root root  4096 2010-09-19 20:58 tmp
dr-xr-xr-x  2 root root  4096 2010-09-19 20:58 usr
drwxr-xr-x 11 root root  4096 2010-09-19 22:47 var

# umount /mnt
```

VELOCITY SOFTWARE

About Partitioning

```
... start Hercules ...

# ls -la /
drwxr-xr-x  2 root root  1024 Jan  1  2011 bin
drwxr-xr-x 11 root root  4096 Jun 29 09:12 etc
drwxr-xr-x  3 root root  2048 Jan  1  2011 lib
drwxr-xr-x 23 root root  1024 Dec 27  2010 Linux-s390
dr-xr-xr-x 40 root root     0 Jun 29 09:11 proc
drwx------  6 root root  4096 Nov 30  2010 root
drwxr-xr-x  3 root root  3072 Jan  1  2011 sbin
drwxr-xr-x 12 root root     0 Jun 29 09:12 sys
drwxrwxrwt  2 root root    40 Jun 29 09:12 tmp
drwxr-xr-x 12 root root  1024 Jan  1  2011 usr
drwxr-xr-x 11 root root  4096 Sep 19  2010 var
```

VELOCITY SOFTWARE

About Backups

Use 'rsync'

Could replace all other Unix backup tools

Filesystem Sharing

CMS sharing 190, 19E, others

Solaris sharing of /usr

academic work (AIX/370 and UTS)

Linux/390 and shared /usr

Linux/390 at NW and shared root

RW root with shared op sys
(bind mount selected directories)

VELOCITY SOFTWARE

Filesystem Sharing

Shared /usr and others

R/O root with R/W /etc

R/O op sys with R/W root

System maint and package management

Relocatable Packages

DASD on Demand – Disk Automounter

VELOCITY SOFTWARE

Shared op sys or root

Install Once, Run Many
(isn't that why they pitched Java?)

Sharing /usr, /opt, and others,
so why not also share the root?

Sharing /bin, /lib, and standard op sys
works and may be more appealing

VELOCITY SOFTWARE

Untouchable Root

Solaris/SunOS supports NFS root including read-only /usr content

"Live CD" Linux uses bulk R/O content

- Knoppix, Ubuntu, Kubuntu, recovery tools

USS supports ROR already (Unix on z/OS)

Not weird, Not even new

Many uses, but not widely understood

VELOCITY SOFTWARE

Stability and Manageability

R/O media is incorruptible

R/O content is centrally maintained

R/O packages are available on-demand

Better D/R – less per-server replication

R/O zLinux no different from R/O PC Linux

VELOCITY SOFTWARE

How to Build Read-Only OS

Start with standard installation

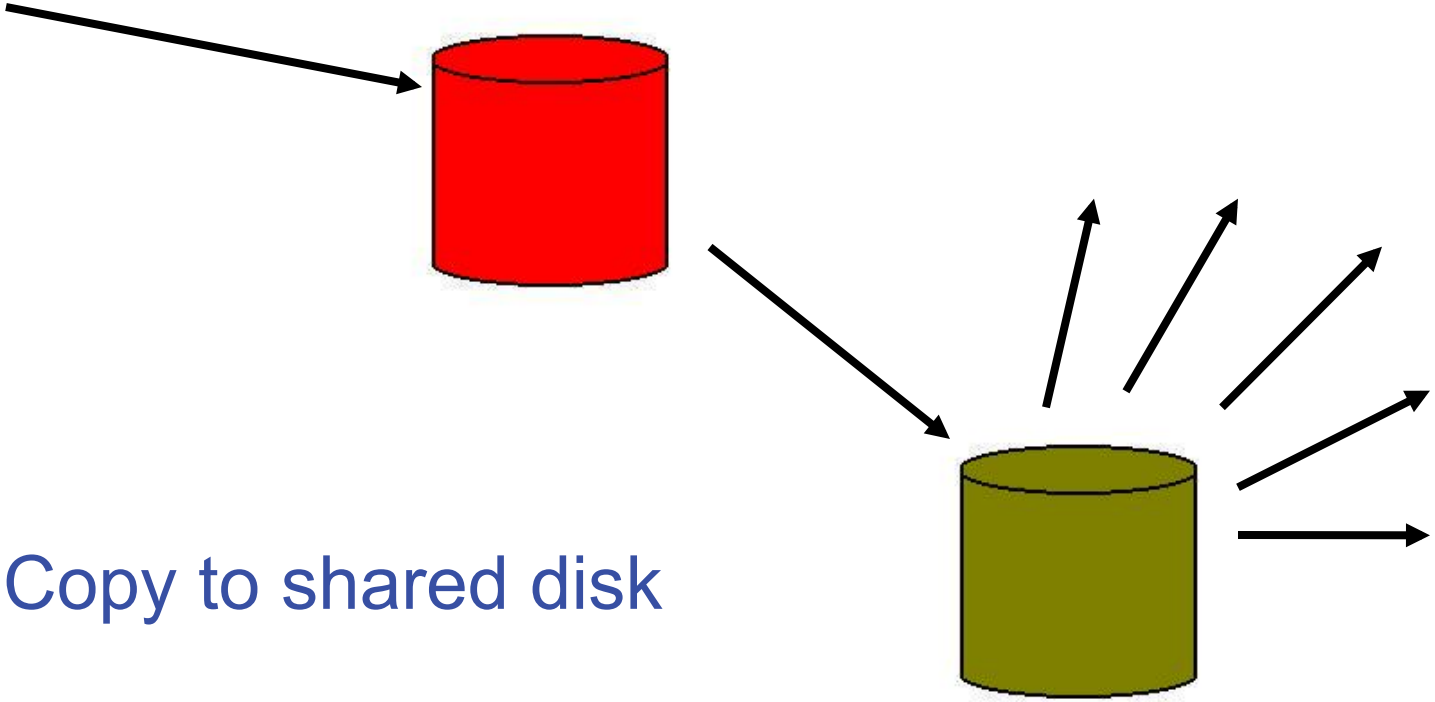
Copy /etc and /var to “run root”

Create other root mount points

Insert /sbin/init+vol script to boot parm

How to Build Read-Only OS

Start with standard installation

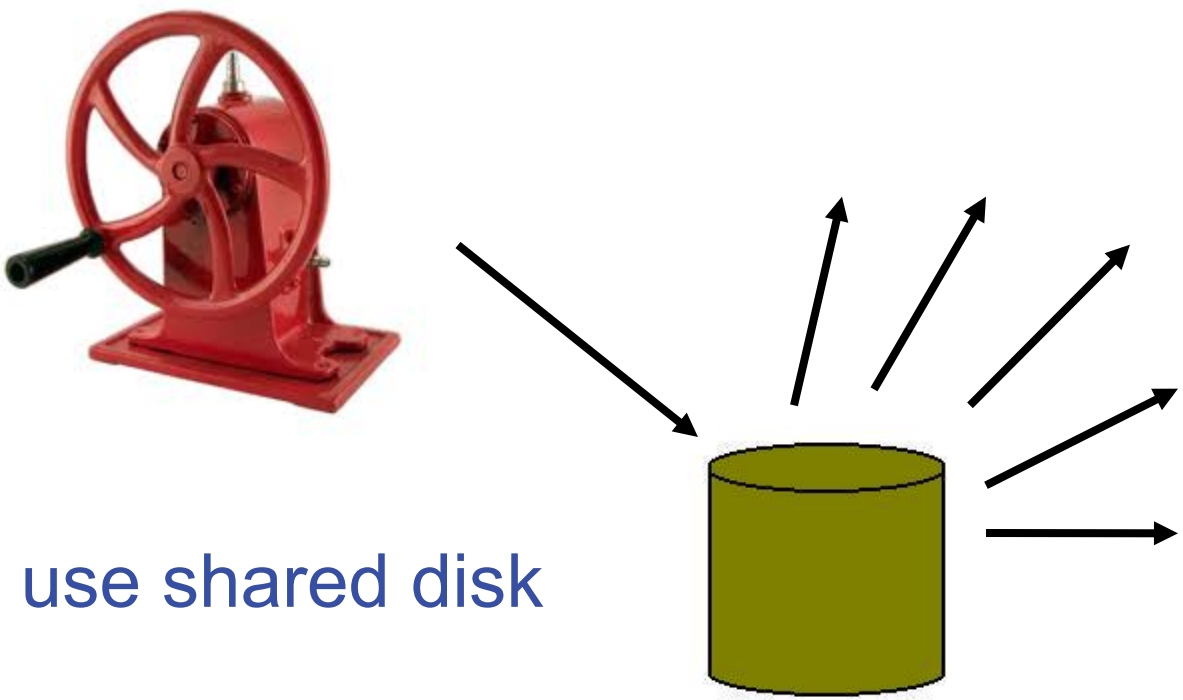


Copy to shared disk

VELOCITY SOFTWARE

How to Build Read-Only OS

Do a bunch of prep work ...



... then use shared disk

VELOCITY SOFTWARE

/sbin/init+vol Startup Script

```
#!/bin/sh
mount -r $_RUNFS /mnt
for D in lib bin sbin usr ; do
  mount -o bind /$D /mnt/$D
done
pivot_root /mnt /mnt/$SYSTEM
cd /
exec /sbin/init $*
```

VELOCITY SOFTWARE

Reconciling RPM Database

Initial RPM DB matches master

“Client” systems may vary

Master may get updates

… now what?

VELOCITY SOFTWARE

Reconciling RPM Database

Extract master package list

```
# rpm -q -a > master.rpml
```

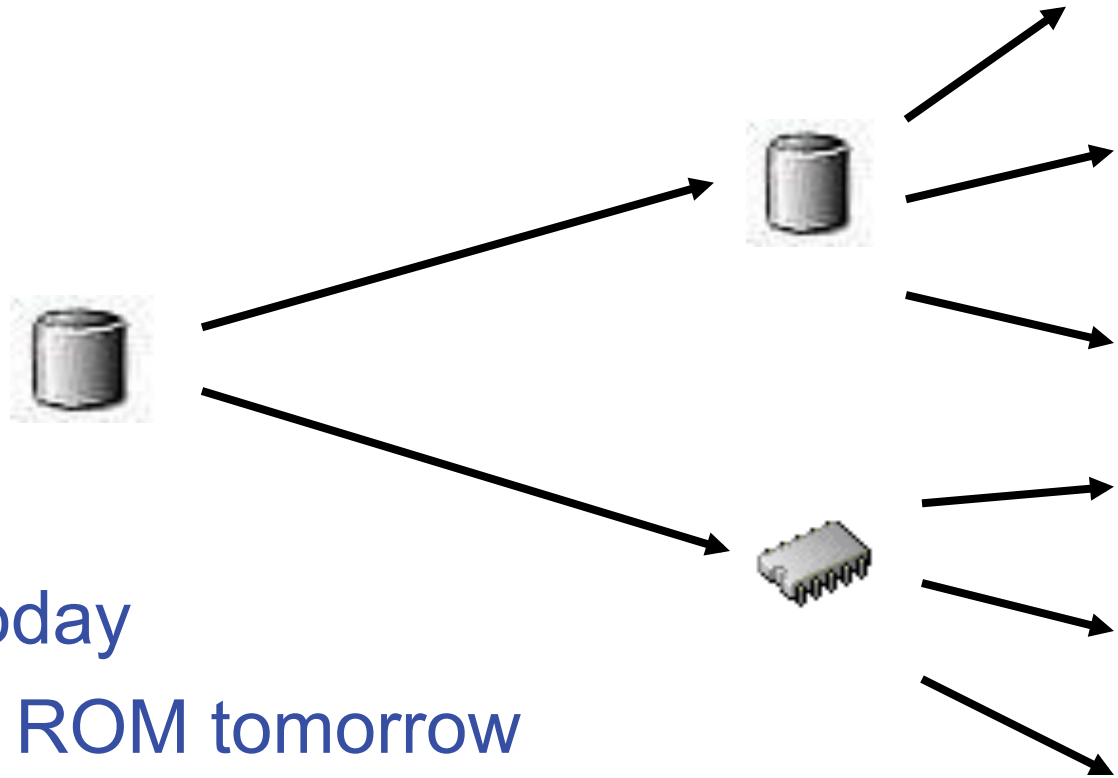
Update client RPM database

```
# for P in ` cat master.rpml ` ; do
  rpm -U --justdb $P.rpm ; done
```

VELOCITY SOFTWARE

How to Share R/O Content

Standard installation



Disk today

Virtual ROM tomorrow

VELOCITY SOFTWARE

How to … reference

`1b0` == boot and op sys root

`1b1` == “run root” with `/bin`, `/lib`, ... bound

`1b5` == `/local`

`1be` == `/usr`

`1bf` == `/opt`

`2b0-2bf` == LVM phys vols and/or maint

`320-33f` == “User Space” LVM phys vols

`100,200` == FCP “HBAs” for SAN

VELOCITY SOFTWARE

How to … reference

`1b0` == boot and op sys root

`1b1` == “run root” with `/bin`, `/lib`, ... bound

~~`1b5` == `/local`~~

~~`1be` == `/usr`~~

~~`1bf` == `/opt`~~

~~`2b0-2bf` == LVM phys vols and/or maint~~

~~`320-33f` == “User Space” LVM phys vols~~

~~`100,200` == FCP “HBAs” for SAN~~

VELOCITY SOFTWARE

R/O OS with Xen

```
nehemiah:~ # df
Filesystem      1K-blocks      Used Available Use% Mounted on
/dev/xvdb         5160576   1427492   3523372  29% /
udev               131168       112    131056   1% /dev
tmpfs              131168         8    131160   1% /tmp
/dev/xvdj        20642428  10102248   9491604  52% /export/home
/dev/xvdk        20642428    176320  19417532   1% /export/opt
/dev/xvdl        30963708  20238400   9152444  69% /export/media
```

VELOCITY SOFTWARE

R/O OS with Xen

```
nehemiah:~ # df
Filesystem      1K-blocks      Used Available Use% Mounted on
/dev/xvda         4127076   1951568   1965864  50% /Linux-i386
/Linux-i386/lib   4127076   1951568   1965864  50% /lib
/Linux-i386/bin   4127076   1951568   1965864  50% /bin
/Linux-i386/sbin  4127076   1951568   1965864  50% /sbin
/Linux-i386/usr   4127076   1951568   1965864  50% /usr
/dev/xvdb         5160576   1427500   3523364  29% /
udev               131168       112    131056   1% /dev
tmpfs              131168         8    131160   1% /tmp
/dev/xvdj        20642428  10102248   9491604  52% /export/home
/dev/xvdk        20642428    176320  19417532   1% /export/opt
/dev/xvdl        30963708  20238400   9152444  69% /export/media
```

VELOCITY SOFTWARE

R/O OS with Xen

```
nehemiah:~ # df
Filesystem      1K-blocks      Used Available Use% Mounted on
/dev/xvda         4127076   1951568   1965864  50% /Linux-i386
/dev/xvdb         5160576   1427500   3523364  29% /
udev               131168       112    131056   1% /dev
tmpfs              131168         8    131160   1% /tmp
/dev/xvdj        20642428  10102248   9491604  52% /export/home
/dev/xvdk        20642428    176320  19417532   1% /export/opt
/dev/xvdl        30963708  20238400   9152444  69% /export/media
```

VELOCITY SOFTWARE

R/O OS with Xen

```
obadiah:~ # df
Filesystem      1K-blocks      Used Available Use% Mounted on
/dev/xvda         4127076   1951568   1965864  50% /Linux-i386
/dev/xvdb         4128448   1927680   1991056  50% /
udev                32864       104     32760   1% /dev
tmpfs               32864        16     32848   1% /tmp
```

VELOCITY SOFTWARE

R/O OS with Xen

```
disk=[ 'file:/var/vmachine/nehemiah/disk0.xvd,xvda,r',
        'phy:/dev/sysvg1/nehemiah,xvdb,w',
        ... ]
```

```
-rw------- 5 root root 4294967296 2011-03-25 09:07
                        /var/vmachine/nehemiah/disk0.xvd
```

VELOCITY SOFTWARE

Relocatable Packages

Deploy instantly

Good candidates for shared FS

- Less content to be backed up

Good candidates for R/O media

- Protected copies (R/O to each client)

Non-intrusive (to the guest op sys)

Non-disruptive (to the users and work)

Mixed releases as needed

VELOCITY SOFTWARE

Automating Disk Attachment

```
#
# /etc/auto.master
#
/home    /etc/auto.home
/misc    /etc/auto.misc
/dasd    /etc/auto.dasd
```

VELOCITY SOFTWARE

Automating DCSS Attachment

```
#
# /etc/auto.master
#
/home    /etc/auto.home
/misc    /etc/auto.misc
/dasd    /etc/auto.dasd
/dcss    /etc/auto.dcss
```

VELOCITY SOFTWARE

Summary

Wide spectrum of data sharing options

File and Filesystem Sharing is rock solid

Condsider your needs, familiarize the team, make a plan and execute

The real advantage is not storage savings but management of myriad systems

VELOCITY SOFTWARE