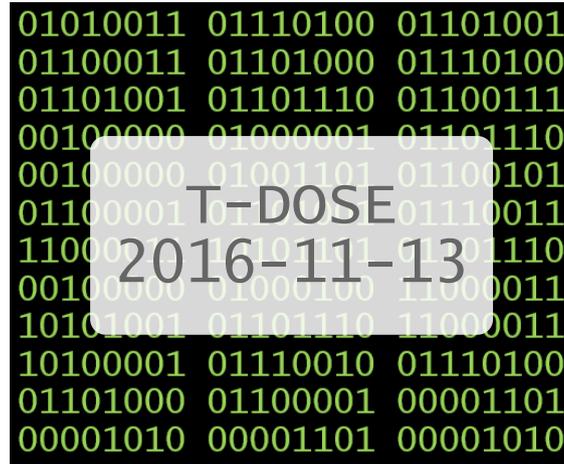


Roghanna, Roghanna

(Choices, Choices)



Harmen de Ruiter

Stichting An Meaisín Dénártha

Let me introduce myself...





Harmen de Ruiter

- Acorn Electron
- 20 years in ICT
- Sr. Sysadmin, ICT “Jack of all trades”
- Co-founder & chairman of Stichting An Meaisín Dénártha
- Strong focus on the human side of ICT
- OSS biased (but not anti-Microsoft)
- Cat personnel with Anne, rock climber



Program

- Stichting An Meaisín Dénártha
- Choices, Choices
 - Virtualization
 - Storage
 - Backups
- Facts & Figures
- Questions
- Demo?
 - At the AnMD booth!



Stichting
An Meaisín Dénártha
– *The Binary Machine* –

Non-Profit
Virtual Private Servers
ICT Services & Education

Supported by YOU!

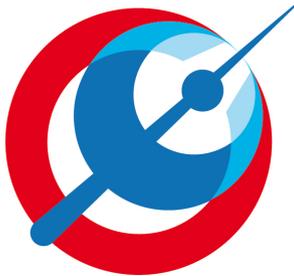
anmd.org

Stichting An Meaisín Dénártha

- Non-profit (foundation)
 - Volunteer-run
 - Founded Jan. 27th 2016
 - Server operational 17th Feb.
- Services
 - Virtual Private Servers (VPS)
 - ICT Services, (web)hosting
- Activities
 - Knowledge-sharing, workshops & events
- Community

Partners & Sponsors

Partners



RTV ARNHEM



Sponsors



DATA CENTER
ARNHEM ^{BV}

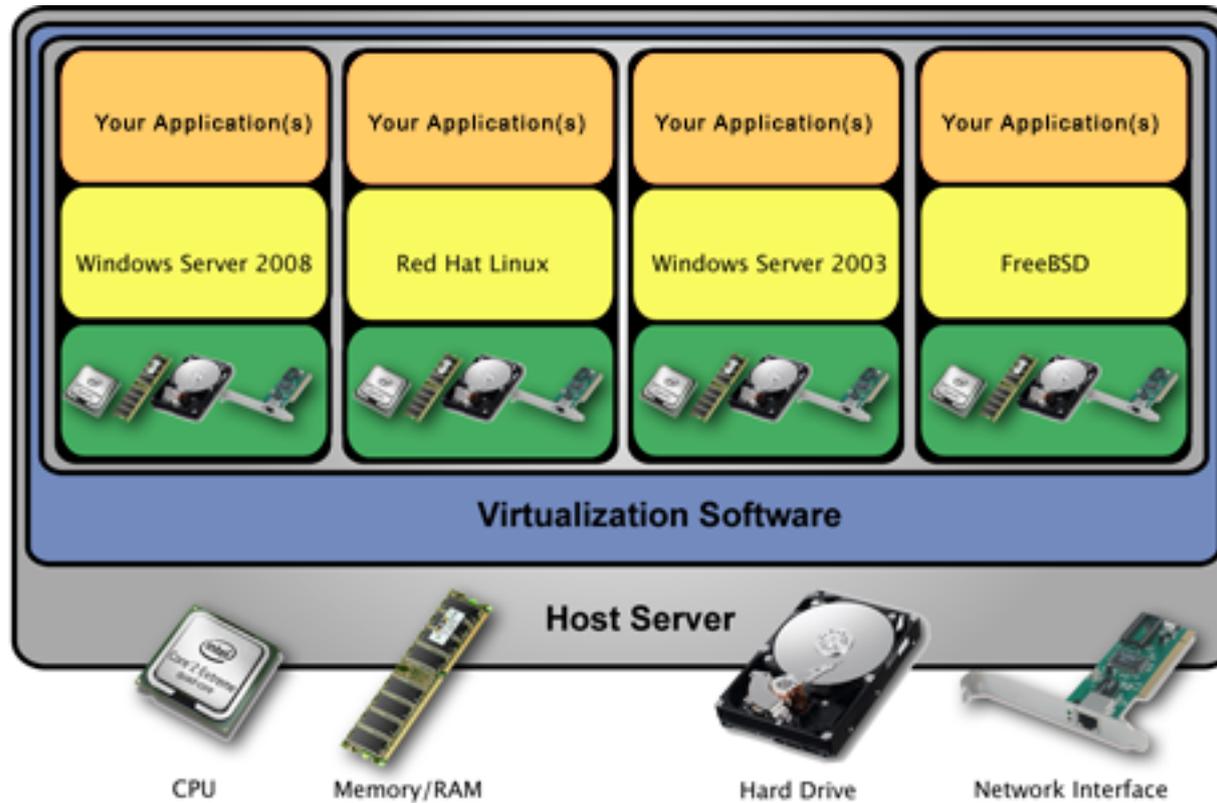


How to get involved?

- Become a volunteer
- Become a partner
- Become a sponsor
- Join our community
- Donate
- Use our paid services
- Spread the word!



virtualisation





Advantages (1/2)

- Easy to run multiple autonomous systems on 1 physical machine
- More efficient use of mostly oversized/overspec'ed hardware
- Flexibility in assigning resources like CPU, memory, disk, network cards, etc.



Advantages (2/2)

- Backups and restores of complete systems are VERY easy
- Cloning of systems is extremely easy
- Higher reliability by means of clustering (HA – High Availability – environment)
- „Appliance principle“ is easy to apply



Considerations

- You can not virtualize each and every server (Tapedrives, dongles, etc.)
- Graphical applications
- For some very heavy tasks, demanding near-100% of your server capacity, it might be better to run bare-metal
- 1 single vm can have massive impact on all vm's running on that node



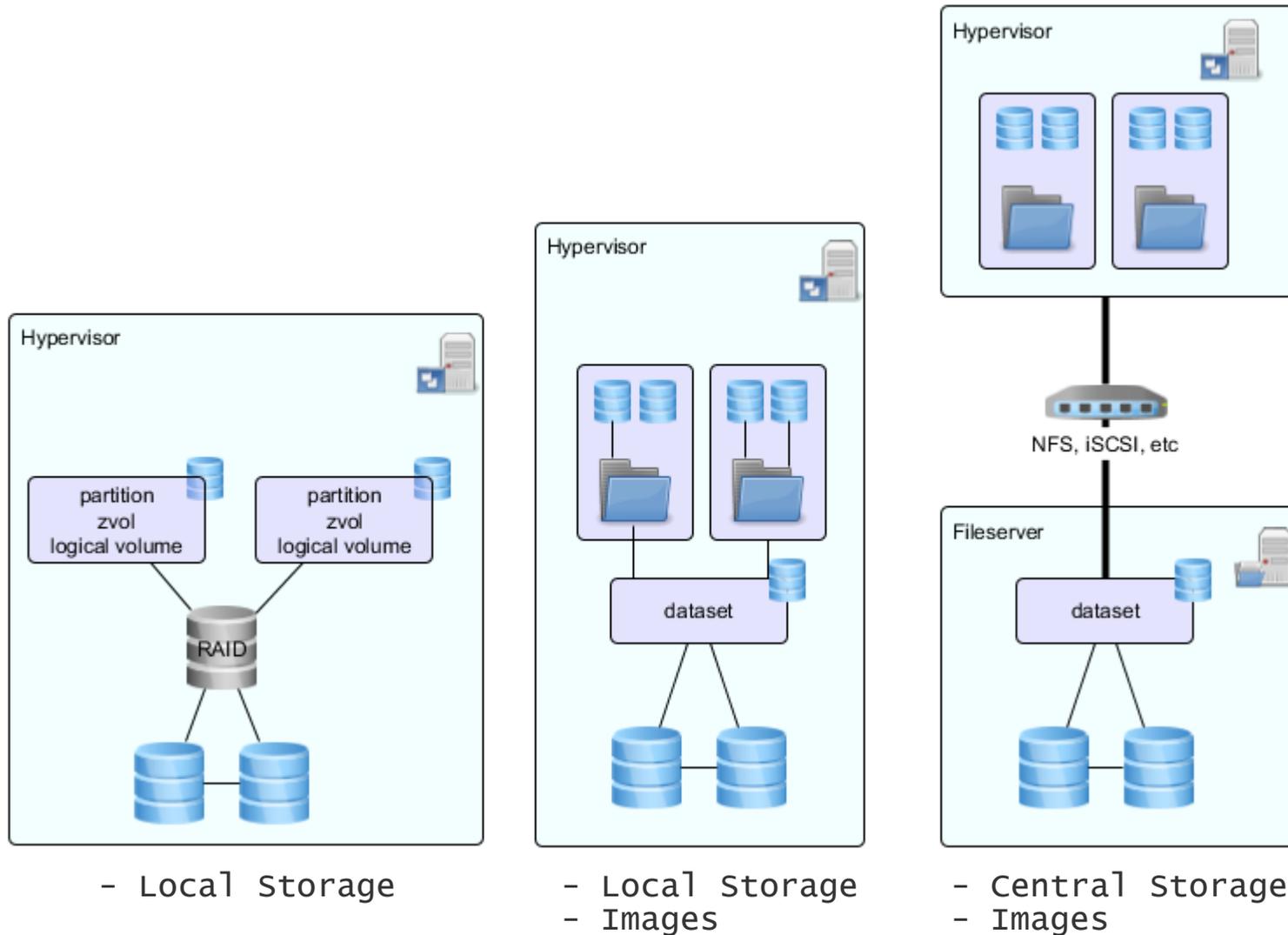
Disadvantages

- You DO lose performance; 2-3 percent
 - Performance benchmarks: KVM vs. Xen - Major Hayden
<https://major.io/2014/06/22/performance-benchmarks-kvm-vs-xen/>
- More complex setup than just a single piece of hardware

Storage



Overview



Local storage



- Advantages
 - Low latency
 - Maximum throughput
- Disadvantages
 - HA setup is possible, but only with replication (DRBD, CEPH)

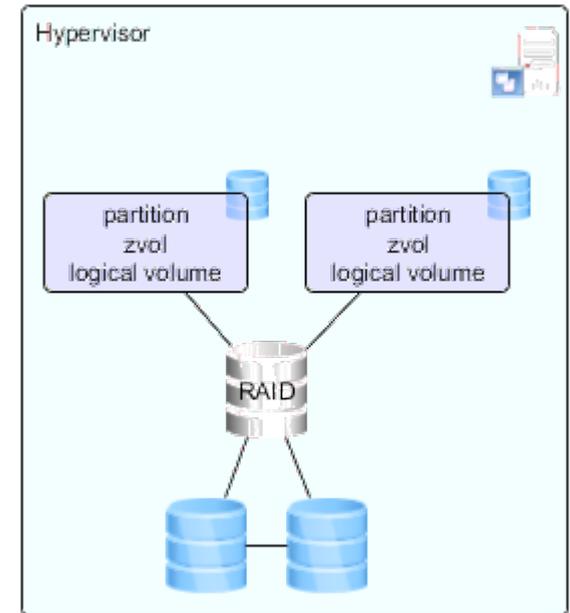


Image + local



- Advantages
 - More flexible when moving vm's
 - Easy migration to central storage
- Disadvantages
 - Performance loss

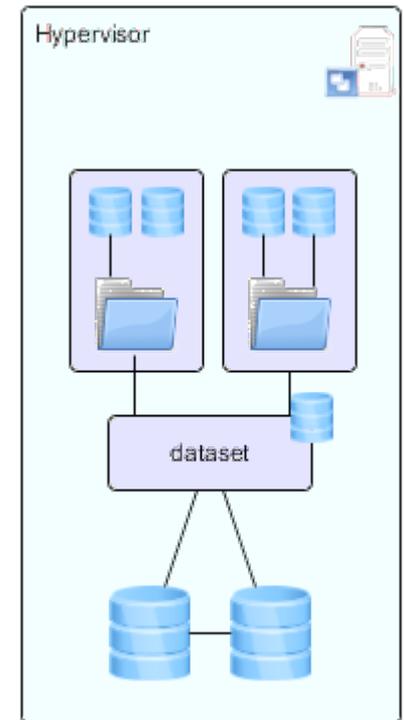
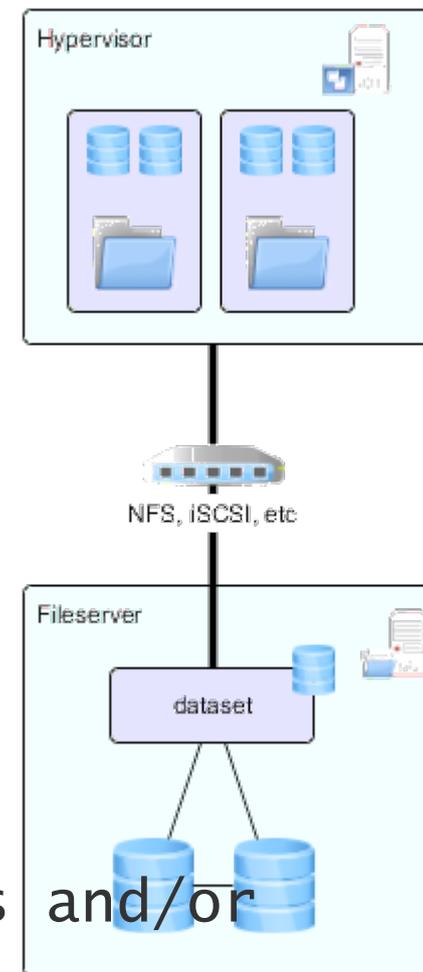


Image + transport + remote

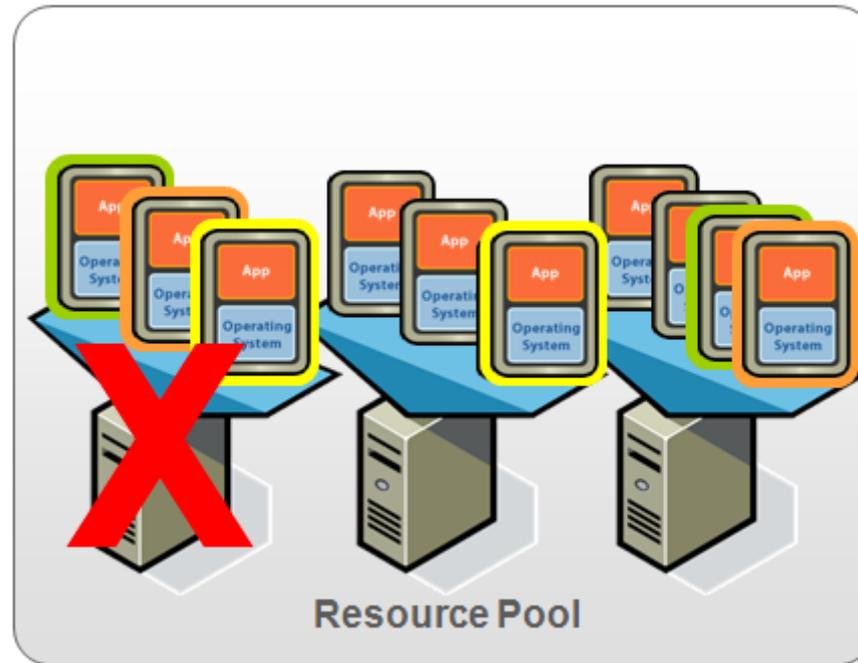


This is how the big boys roll!

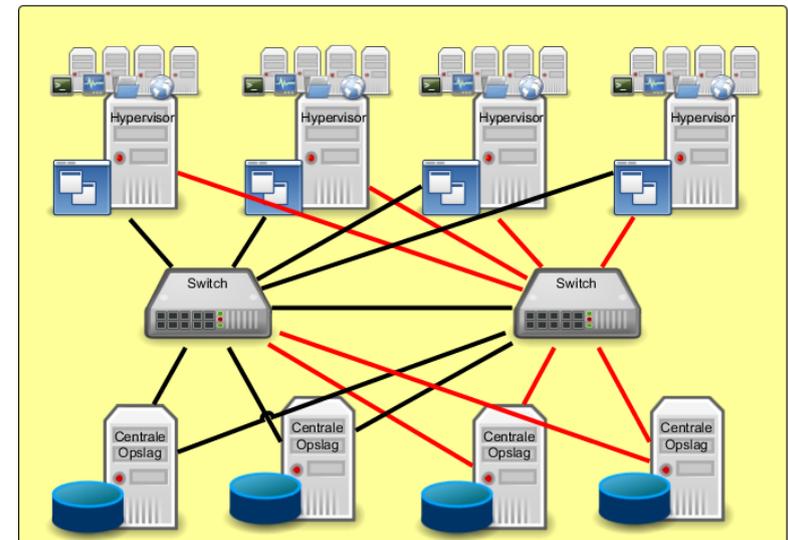
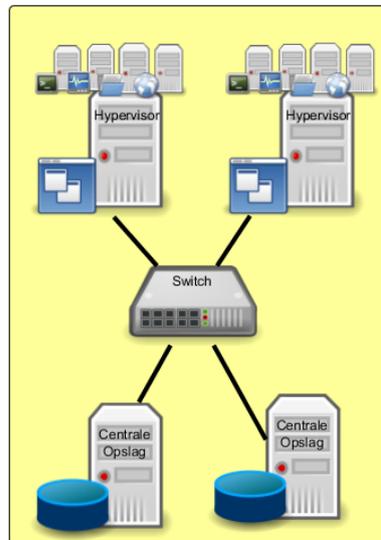
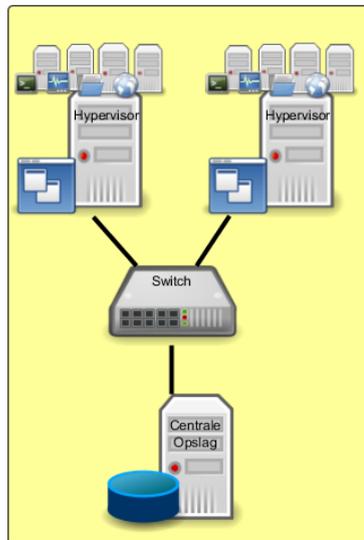
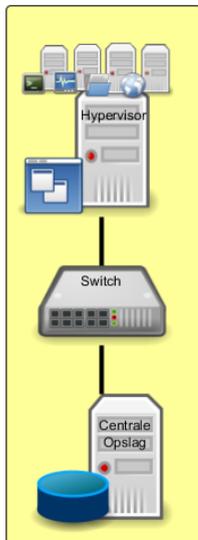
- Advantages
 - HA setup on multiple hypervisors
 - Scalable
- Disadvantages
 - Slower (because network)
 - Network (switches) needed
 - For more performance you need more expensive (L2 managed) switches and/or 10Ge+ NIC's + switches (expensive!)



High Availability & Scalability



Overview



- Hypervisor
- Local storage

- Single hypervisors
- Single central storage

- Multiple hypervisors
- Single switch
- Single central storage

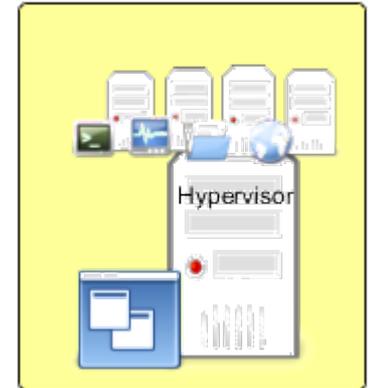
- Multiple hypervisors
- Single switch
- Multiple central storage

- Multiple hypervisors
- Multiple switches
- Multiple central storage

Single System, Local storage



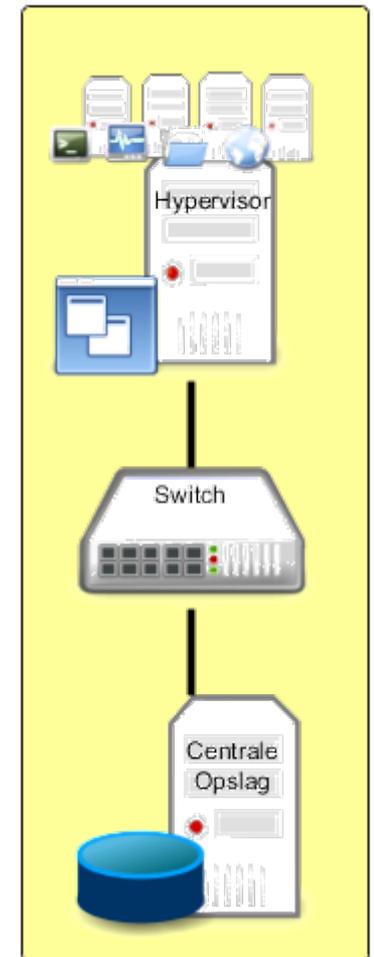
- Advantages
 - Simple installation / setup
 - Cheap (only one server)
- Disadvantages
 - No load-balancing on the hypervisor
 - Complicated expansion
 - If the one system goes down...



Single system, central storage



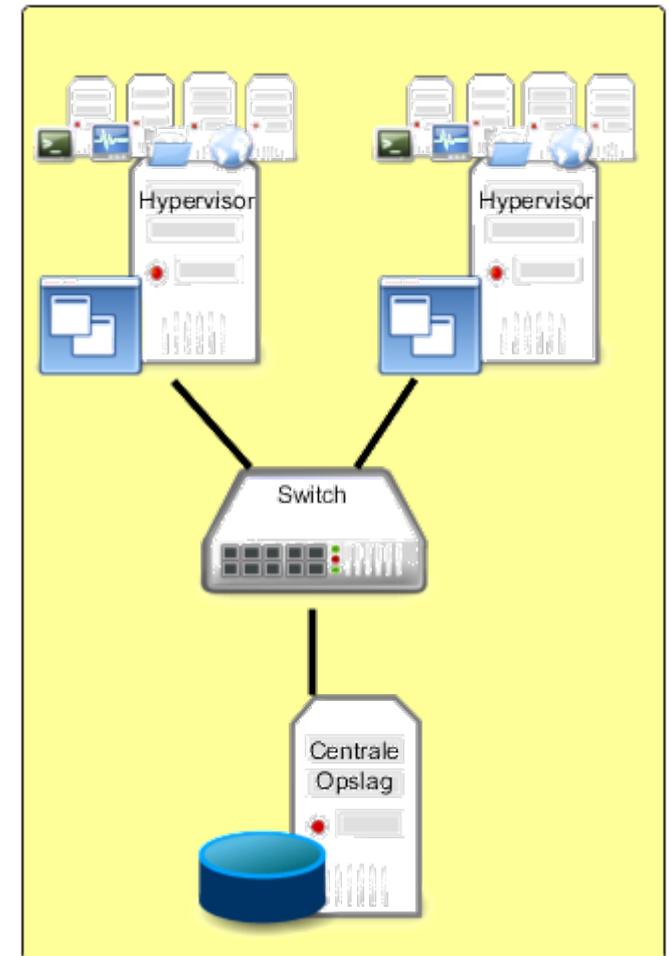
- Advantages
 - Easy to expand to more complex setups
 - Dedicated hypervisor + storage
- Disadvantages
 - 2 Single Point of Failures (SPOF)
 - If either the storage or the hypervisor goes down, you're dead



Multiple hypervisors, central storage



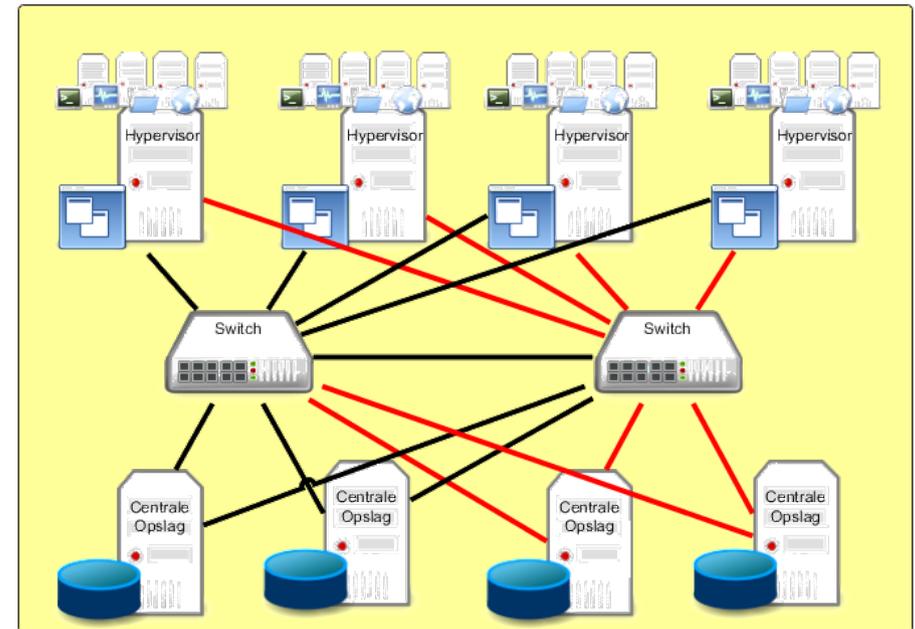
- Advantages
 - Load balancing
 - (Live) Migration of vm's
 - Zero to no downtime when maintenance @ hypervisor
- Disadvantages
 - More complex setup
 - More expensive
 - SPOF on storage



Multiple hypervisors, central HA storage



- Advantages
 - Best of all
 - Enterprise
 - Zero downtime
- Disadvantages
 - Complex setup
 - HUGE costs



Scalability



Small...



- 1x HP Microserver Gen8
 - CPU: Intel Celeron G1610T
 - Memory: 16GByte DDR3
 - Disks: 2-4x HDD
 - Costs: €800



A little bigger...



- 1x HP Microserver Gen8
 - CPU: Intel Celeron G1610T
 - Memory: 16GByte DDR3
 - Disk: 1x SSD
- 1x HP Microserver Gen8
 - CPU: Intel Celeron G1610T
 - Memory: 16GByte DDR3
 - Disk: 2-4x HDD
 - Price (total): €1500





Company-size...

- HP Storageworks P4500 G2 "LeftHand"
- CPU: 2x Intel Xeon E5620 @ 2.40GHz
- Memory: 96GByte DDR3 (max. 144GByte)
- Disk Controller: LSI SAS 9207-8i
- Disks: 2x 128G SSD + 4x 5TByte HDD (max. 12 disks)
- Price (totaal): €1400 - €2250



Humongous !



- Tianhe-2
- CPU: 32.000 Intel Xeon E5-2692 + 48.000 Xeon Phi 31S1P
- Memory: 1,375 TiB
- Storage: 12,4 PB
- Prijs: \$390 million



Solutions





Commercially vs OSS

- Commercially
 - VMware (ESXi, Workstation)
 - Microsoft Hyper-V
 - Citrix XenServer
- OSS
 - VirtualBox
 - XenProject
 - KVM
 - Proxmox (KVM + LXC)

Proxmox VE



Why Proxmox VE

- Easy to install & use
- Enterprise features
 - HA on storage & hypervisors
 - Central management (multi-master)
 - CLI interaction
 - Backup/restore built-in
- Pricing
- 100% OSS
- KVM + LXC

Advantages

- KVM & LXC
- HTML5 GUI (+ console)
- Linux (debian) base-layer
- Out-of-the-box solution (works straight out of the box, without additions)
- Authentication against multiple backends
- Commercially (paid) support possible
- Clustering of hypervisors possible

Disadvantages

- Less „enterprise“ then VMWare or XenServer
 - Less usable in really large environments

Features - Storage

- Local storage
 - LVM Group (block devices, FC devices, DRBD, etc.)
 - Directory (storage on existing filesystem)
 - ZFS
- Network storage
 - LVM Group (network backing with iSCSI targets)
 - iSCSI target
 - NFS Share
 - Ceph RBD
 - Direct to iSCSI LUN
 - GlusterFS

Features – Image formats

- Support for multiple image formats.
 - QCOW2, Native format of KVM
 - VMDK, Native format of VMware
 - RAW

Features - Virtualization

- Full-virtualization (KVM)
 - All operating systems
- Paravirtualisation (KVM)
 - KVM-aware operating systems
- Containervirtualization (LXC)
 - Linux only!
 - No migration to other node possible

Features – Guest OSes

- Runs (thanks to KVM)
 - Linux
 - FreeBSD
 - FreeDOS
 - Windows
 - ... and almost any other x86 operating system

Features - Other

- Live migration (including storage, and changing image-format)
- Snapshots & backups
- Role-based access
- Resource pools (role separation)
- KVM is supported / built by RedHat (VirtIO drivers)
- Native ZFS support (eigen kernel)

Storage



RAID?



- Redundant
- Array
- of
- Independent (Inexpensive, orig.)
- Disks



RAID vs Single disks

- Advantages (RAID vs 1 disk)
 - Less prone to data loss
 - More speed
- Disadvantages (RAID vs 1 disk)
 - Overhead
 - More disks needed
 - More expensive

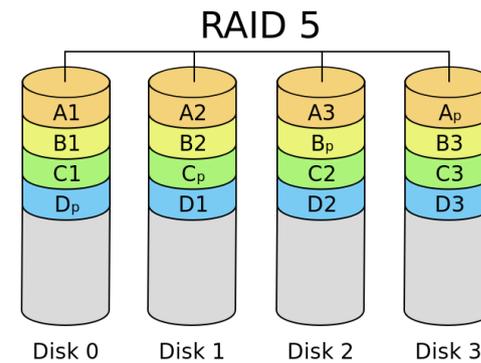
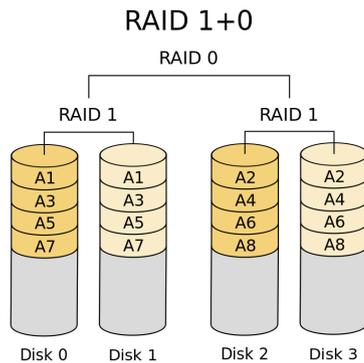
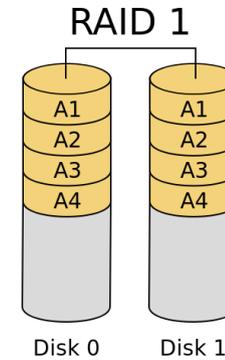
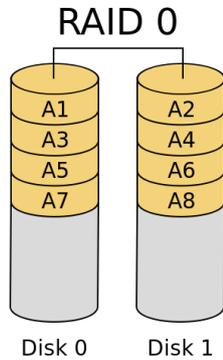


RAID is no backup!

- RAID is no backup!
- It protects against certain ways of hardware failure but it CAN fail!
- Data loss because of viruses or user-error is no different with RAID
- **RAID is *NO* backup!**



RAID Levels



<http://www.acnc.com/raid>

Hardware RAID vs Software RAID



- Hardware RAID
 - RAID calculations done by a dedicated RAID controller
- Software RAID
 - RAID calculations done by the OS
- Next-Gen filesystems (ZFS, BTRFS)
 - Combine software RAID with filesystems

Hardware RAID - Features



- RAID calculations are done by a dedicated RAID controller

Hardware RAID - Advantages



- Less CPU load (albeit negligible with modern hardware)
 - It is NOT faster than software RAID!
- OS independent
 - Data is accessible to all OSes

Hardware RAID - Disadvantage



- Expensive
 - hardware RAID controllers are more expensive than a HBA, especially with more complex RAID levels
- Less flexible
 - Disks / arrays can NOT be easily transported to a different brand or type of RAID controller



Hardware RAID - Remarks

- Even the big boys like NetApp use software RAID (NetApp uses WAFL)
- Hardware RAID is less & less used in small to enterprise environments

Software RAID - Features



- RAID calculations are performed by the OS
- Hardware needed
 - Regular SATA or SAS ports
 - Host Bus Adaptor (HBA) if you need more disks or higher throughput

Software RAID – Advantages



- Flexibility
 - You can easily move an entire diskset to a totally different server without loss of data

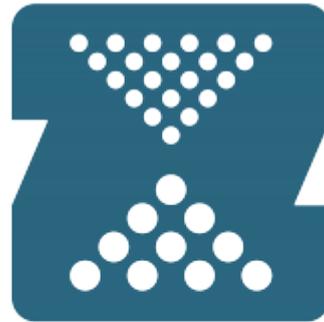
Software RAID – Disadvantage



- It will cost you some CPU power
- OS dependant, you can not move a Linux (md) raidset to a windows machine and vice versa

ZFS

Zetabyte
File
System



Open**ZFS**



1 zettabyte

- 2.000.000.000 (2 billion) years of music
- 250.000.000.000 (250 billion) DVD's
- 36.000.000 (36 million) years of HD video
- 2x the entire internet in 2009
- The great wall of China, filled with coffee, if 1 cup of coffee (32cl) represents 1 gigabyte of data

256 quadrillion zettabyte

- ZFS is capable of addressing 256 quadrillion zettabytes of storage (they say...)
- That's 256.000.000.000.000.000 ZB...

Advantages (1/2)

- md, LVM en ext4 are integrated
- All writes are “atomic”
- Fast, fast, fast, fast, extremely fast
- Bitrot is seen and corrected
- RAID equivalents possible, including RAID levels who don't exist
(RAIDZ-3 == „RAID7“, 3 disk failure possible without data loss)

Advantages (2/2)

- Rebuilds are much faster (only stored data is rebuilt, not the entire disk)
- ZFS send & ZFS receive
- Snapshots



Disadvantages

- You can NOT add disks to a RAIDsets (zpool)
 - You can replace disks with larger ones, without any loss of data and without downtime



Features

- Totally different terminology than you're used to
- ZFS needs to be able to access the hardware **directly**, so you NEED a HBA in IT-mode and ABSOLUTELY NO RAID controller (not even a P410 with a RAID0 array per disk)
- It's quite new on Linux (stable in 2013)



Speed (Caching)

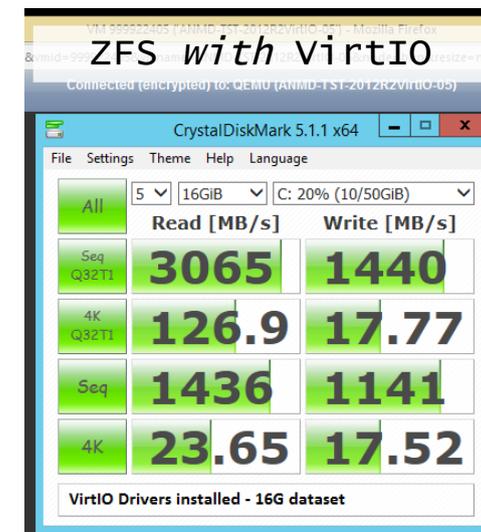
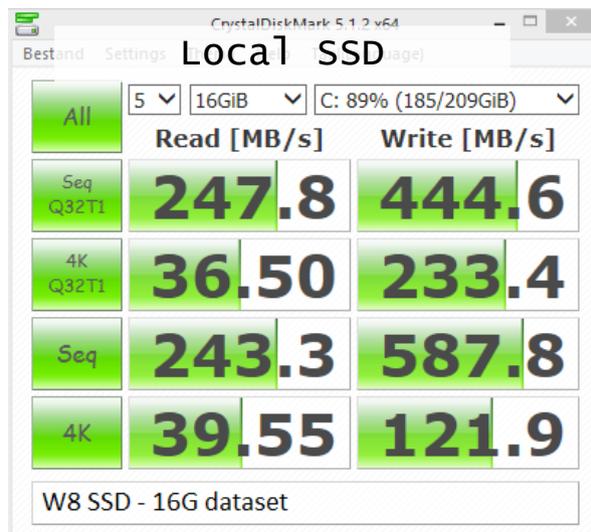
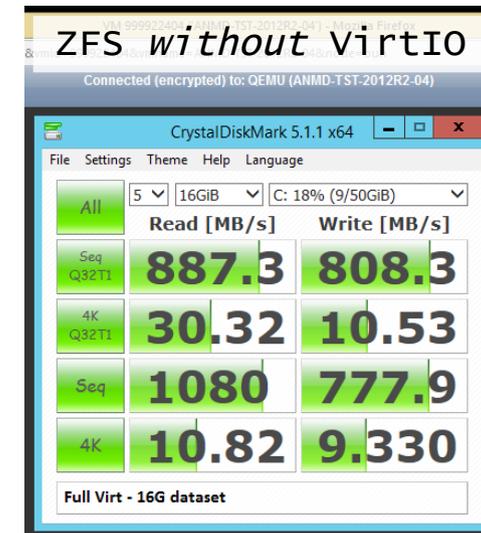
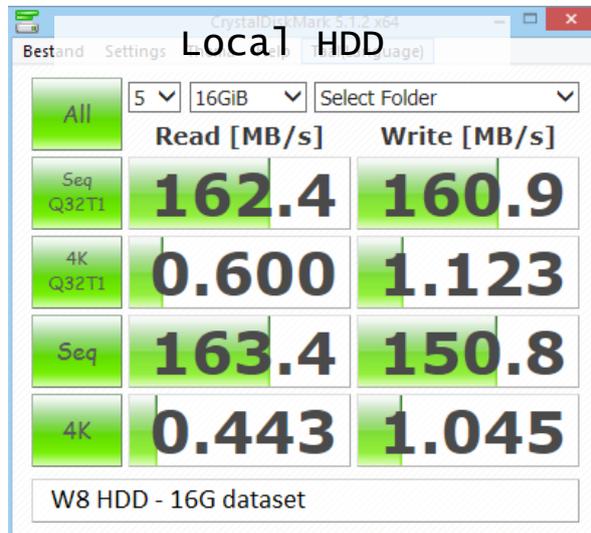
- Caching
 - Read + write caching possible on separate disks
 - Possible to add later without downtime
 - Don't start with caching straight away, only after you've run production for a while and know what kind of caching you need



Speed (Compression)

- Compression
 - By switching on compression, you gain performance
 - Compressing data costs less time than writing the extra bits to disk
 - Saves diskspace

Speed (benchmark)



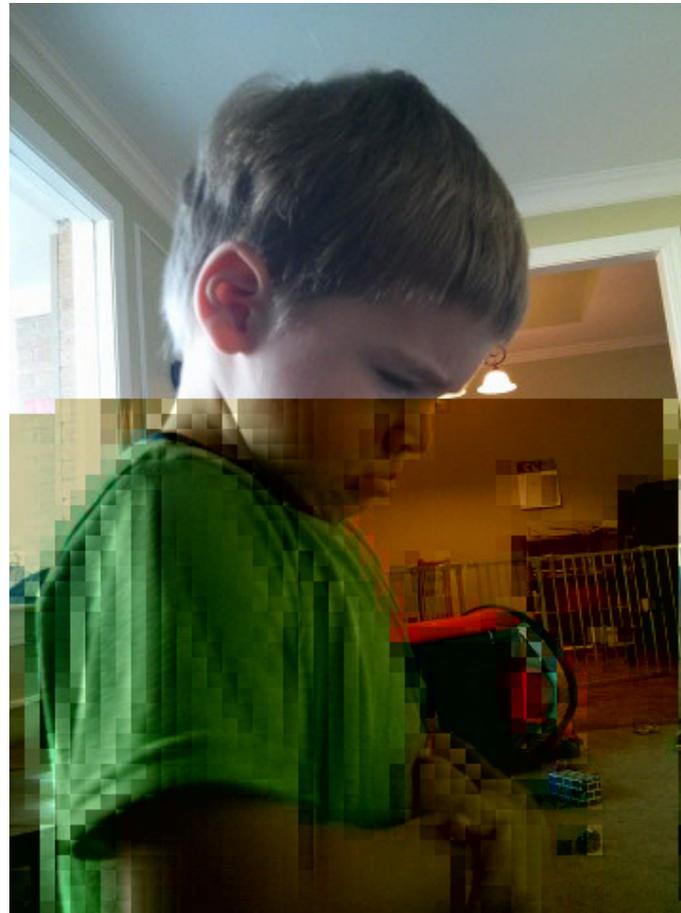
Bitrot

- Examples are from ArsTechnica
„Bitrot and atomic COWs: Inside “next-gen”
filesystems“
<http://arstechnica.com/information-technology/2014/01/bitrot-and-atomic-cows-inside-next-gen-filesystems/3/>

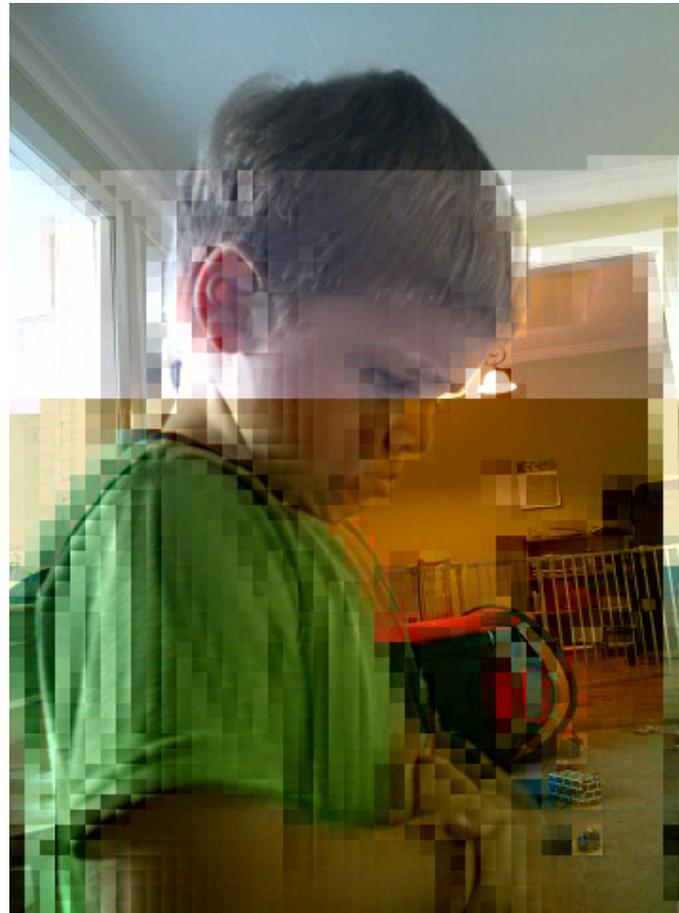
Bitrot - Original



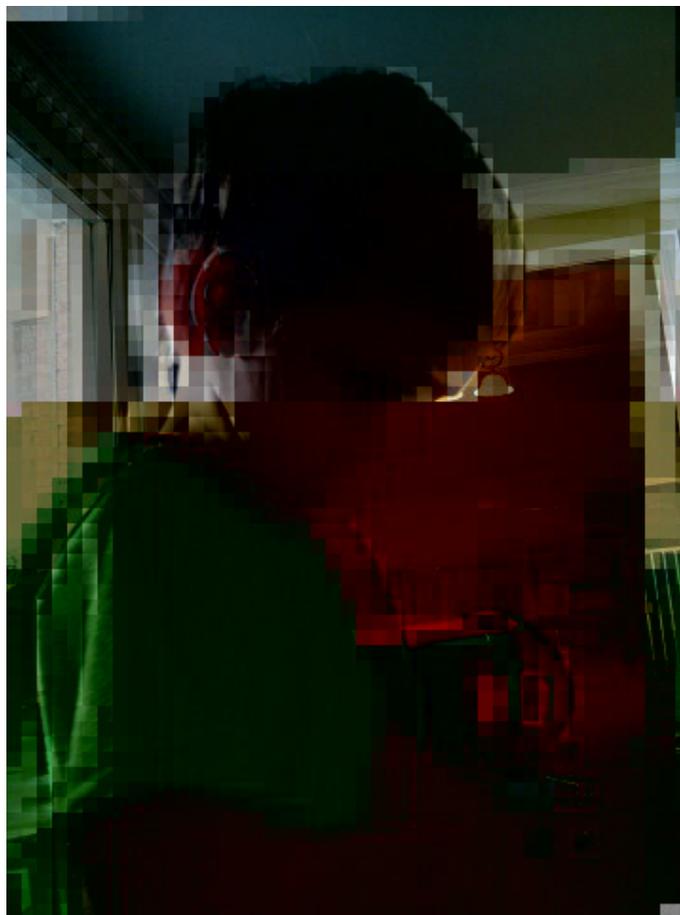
Bitrot - 1 bit flipped



Bitrot - 2 bits flipped



Bitrot – 3 bits flipped



Bitrot – ZFS corrected





Choices

- Striped mirror (RAID10 equivalent)
 - Faster rebuild vs RAIDZ
 - Less stress on the disks then during a RAIDZ rebuild
 - More continuous disk capacity then with mirroring alone
 - Faster than only mirroring or striping (RAID0 of RAID1 equivalent)

Backups (because disks fail...)





Backups

- RAID and snapshots are not backups...
- Backups DO protect against
 - User error (`dd if=/dev/zero of=/dev/sda`)
 - Malware
 - Multiple disk failures / system crashes
 - Fire / robbery
- Backups need to be
 - Offline (not on the same system)
 - Offsite (because of fire / robbery)



(AnMD) Requirements

- Offsite
 - At least on another system
- Delta's
 - Bandwidth-saving
- ZFS-aware / ZFS-friendly
 - Snapshots
- Vm's
 - Bare-metal
- Hypervisor
 - Not bare-metal



Backup (Choices)

- Proxmox
 - Built-in snapshot
 - Built-in backup
- Amanda, backuppc, etc
- Homebrew ZFS scripts
 - snapshots
 - zfs send
- Off-the-shelf ZFS scripts



Sanoïd

- Designed for ZFS
- Easy to understand configuration
- Relatively small script (human readable)



- Sanoïd
 - Creating snapshots (and removing old)
- Syncoid
 - Syncing snapshots to a remote system

<https://github.com/jimsalterjrs/sanoïd>

Choices, so many choices...

ZABBIX



Etc, etc, etc.

Nagios®

openNMS®

icinga

The logo for OSTicket, featuring an orange silhouette of a kangaroo or wallaby to the left of the text "OSTicket" in a bold, sans-serif font, with "Support Ticket System" in a smaller font below it.
OSTicket
Support Ticket System

The inevitable final slide...

- Email : harmen@anmd.org
- WWW : <https://anmd.org>
- FB page : <https://goo.gl/3Qz0AZ>
- FB group : <https://goo.gl/zdG5bG>

