

Private Cloud Devotional



Sacristan **Gordon** is the consultant's answer to The Swiss Army knife - versatile and always at hand. Gordon's core competency is to undertake projects that no one else dare, will, or can. Titled CTO on the business card he has the responsibility to see through the fog of technological B & B (Bull Shit & Buzzwords).



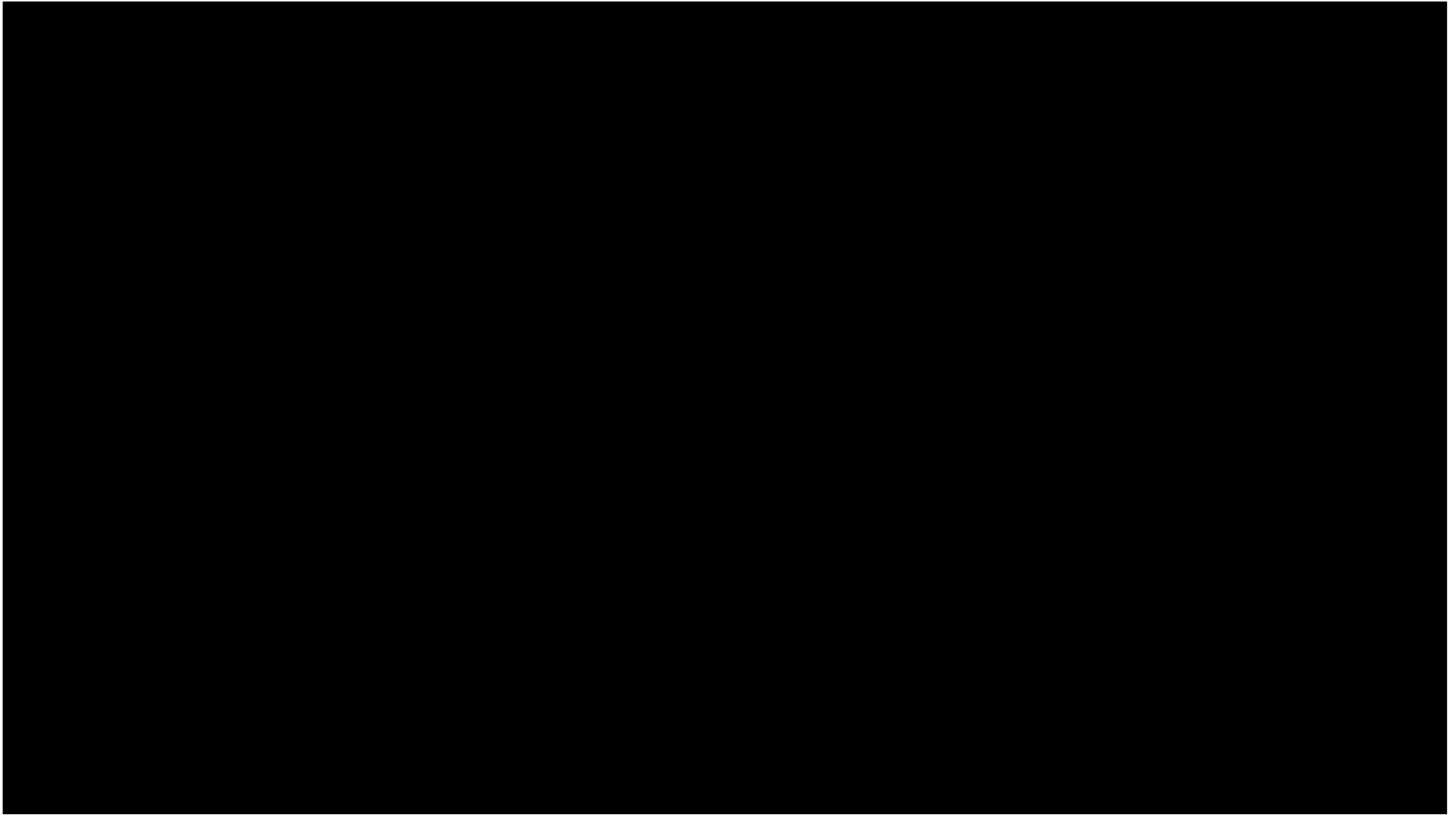
High Priest **Jørgen** is one of the greatest authorities on Solaris, and he has been in the business since DASK (If anyone can remember this!). The seminar is a dear-come opportunity to ask questions about this legend in Danish IT. Jørgen was originally trained as a carpenter, like his biblical colleague Jesus.



IT KONGERNE

IT Kongerne is considered one of the world's leading authorities in terms of expertise in database-centric platforms. We are best known for our ability to rescue systems and databases where others have had to give up..

The cloud has made it to Hollywood and “nobody understands it”

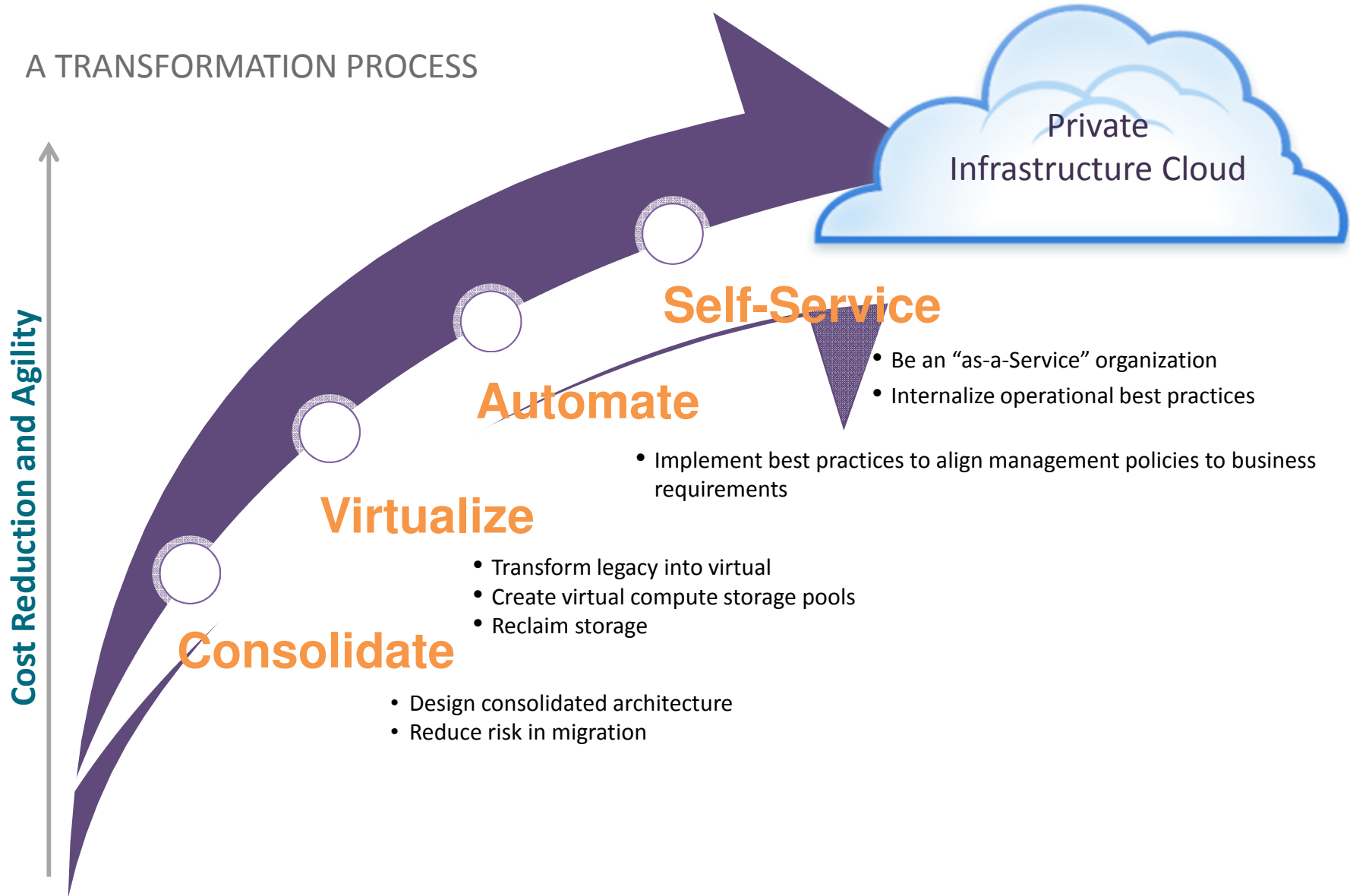


What's cloud – What's not



Your way to ~~Private~~ Cloud Salvation

A TRANSFORMATION PROCESS



JESUS COMMANDS...



SmartOS

ZFS



dtrace

nodeJS



Marka Hoglund 09

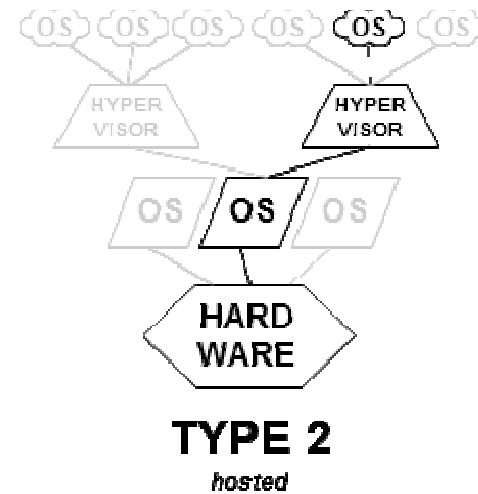
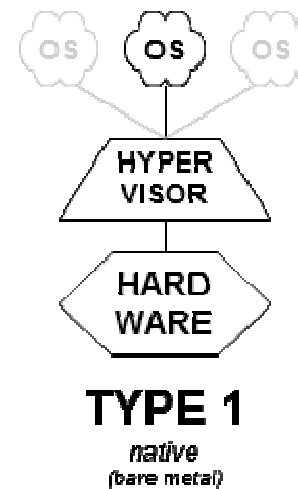


SmartOS

SmartOS is an open source **Type 1** hypervisor platform based on Illumos, a descendant of OpenSolaris, and developed by Joyent

SmartOS offers

- Performance
- Observability
- Reliability



SmartOS offers two types of Virtualization


Container-based virtualization (Zones)

- A container is the combination of resource controls and Solaris zones.
- Provides a lightweight solution and a complete and secure user space environment on a single global kernel.
- Can easily scale vertically, something that is more difficult with KVM virtualization.

KVM virtual machines

- KVM virtual machines provide a solution for running a variety of guest OS's, including Linux and Windows, in a full, hardware-assisted virtualization.
- KVM images on SmartOS run as a process inside of a zone.

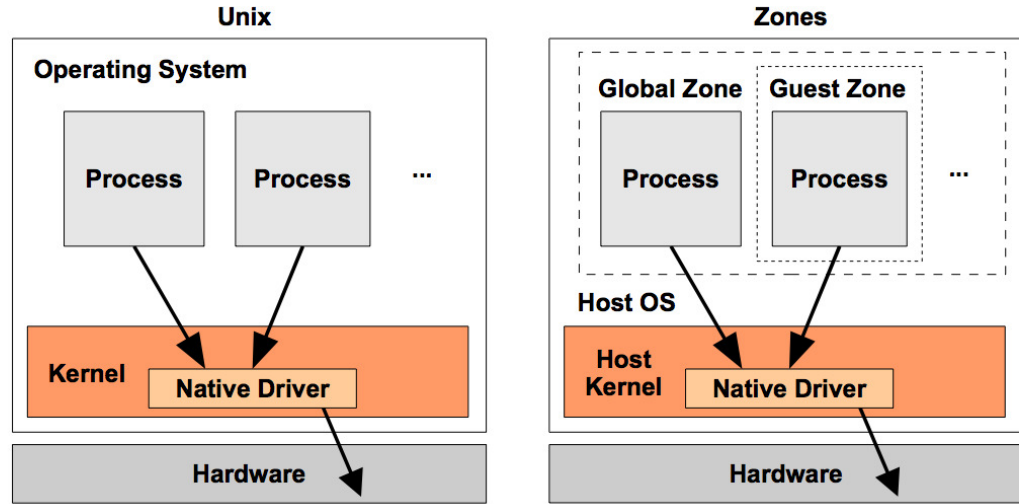
Performance

- Networking, SmartOS makes use of a kernel  mechanism called crossbow. When running on a Zone the code path for doing network I/O is directly to bare metal.
- ZFS, every Zone runs in its own ZFS dataset, and each VM gets its own ZFS volume. As with the network, a Zone runs directly on SmartOS to do disk I/O

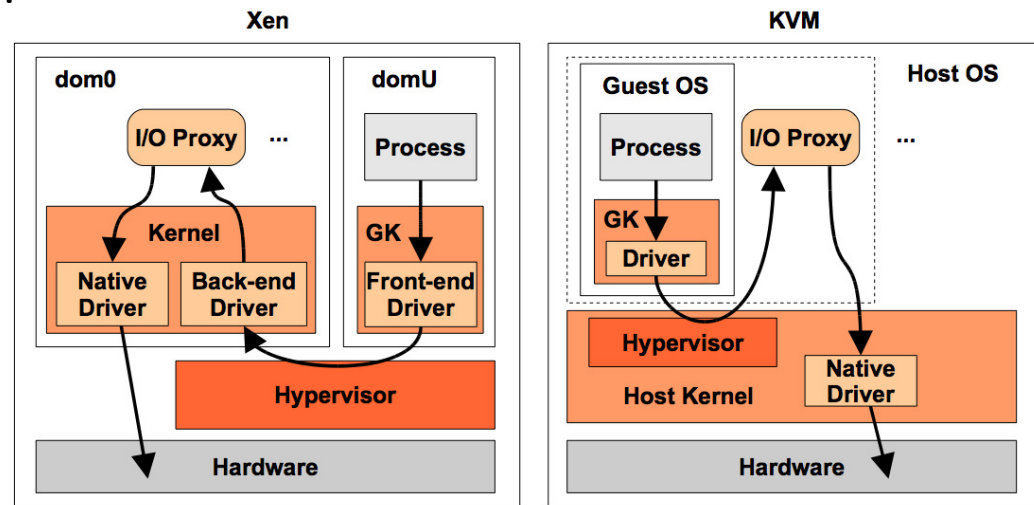


Virtualization Performance: Zones, KVM, Xen

How does I/O differ between traditional Unix and Zones?



Now for Xen and KVM (simplified!):



Virtualization Performance: Throughput

Throughput

Searching for the highest Gbits/sec:

source	dest	threads	result	suspected limiter
SmartOS 1 GB	SmartOS 1 GB	1	2.75 Gbits/sec	client iperf @80% CPU, and network latency
SmartOS 1 GB	SmartOS 1 GB	2	3.32 Gbits/sec	dest iperf up to 19% LAT, and network latency
SmartOS 1 GB	SmartOS 1 GB	4	4.54 Gbits/sec	client iperf over 10% LAT, hitting CPU caps
SmartOS 1 GB	SmartOS 1 GB	8	1.96 Gbits/sec	client iperf LAT, hitting CPU caps
KVM CentOS 1 GB	KVM CentOS 1 GB	1	400 Mbits/sec	network/KVM latency (dest 60% of the 1 VCPU)
KVM CentOS 1 GB	KVM CentOS 1 GB	2	394 Mbits/sec	network/KVM latency (dest 60% of the 1 VCPU)
KVM CentOS 1 GB	KVM CentOS 1 GB	4	388 Mbits/sec	network/KVM latency (dest 60% of the 1 VCPU)
KVM CentOS 1 GB	KVM CentOS 1 GB	8	389 Mbits/sec	network/KVM latency (dest 70% of the 1 VCPU)

Virtualization Performance: IOPS

IOPS

Searching for the highest packets/sec:

source	dest	threads	result	suspected limiter
SmartOS 1 GB	SmartOS 1 GB	1	14000 packets/sec	client/dest thread count (each thread about 18% CPU total)
SmartOS 1 GB	SmartOS 1 GB	2	23000 packets/sec	client/dest thread count
SmartOS 1 GB	SmartOS 1 GB	4	36000 packets/sec	client/dest thread count
SmartOS 1 GB	SmartOS 1 GB	8	60000 packets/sec	client/dest thread count
SmartOS 1 GB	SmartOS 1 GB	16	78000 packets/sec	both client & dest CPU cap
KVM Centos 1 GB	KVM Centos 1 GB	1	1180 packets/sec	network/KVM latency, thread count (client thread about 10% CPU)
KVM Centos 1 GB	KVM Centos 1 GB	2	2300 packets/sec	network/KVM latency, thread count
KVM Centos 1 GB	KVM Centos 1 GB	4	4400 packets/sec	network/KVM latency, thread count
KVM Centos 1 GB	KVM Centos 1 GB	8	7900 packets/sec	network/KVM latency, thread count (threads now using about 30% CPU each; plenty idle)
KVM Centos 1 GB	KVM Centos 1 GB	16	13500 packets/sec	network/KVM latency, thread count (~50% idle on both)
KVM Centos 1 GB	KVM Centos 1 GB	32	18000 packets/sec	CPU (dest >90% of the 1 VCPU)

Observability

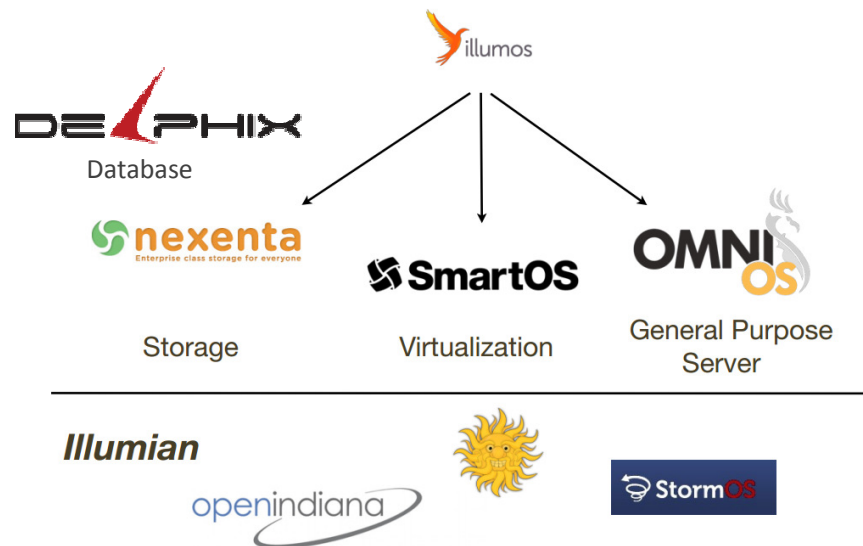


The ability to see what is being done by the entire software stack from application to hardware. SmartOS provides a tool called Dtrace that can be used for:

- **Debugging** - trace entry and return from functions
- **Performance Analysis** - get nanosecond timing information.
- **Code Coverage** - determine whether or not code is getting executed.

Reliability

Pedigree



Features

- **Fault Management Architecture (FMA)**
Helps detect, report and diagnose any fault or defect that can occur on a SmartOS system.
- **Service Management Facility (SMF)**
Dependencies between services
Parallel starting
Automatic restart upon
Delegation of services to non-root users.
- **Virtual Router Redundancy Protocol (VRRP)**
Highly available load balancing.
Provides a means to implement hot-failover via virtual IP sharing.

Designed for Virtualization

Lightweight Deployment via USB or PXE

- No “on disk” installation
- No disk wasted for root disks
- No patching
- Fast and Reliable Upgrades (just reboot)
- Enhanced Security.
- No ZFS Boot Environments
- No Image Package System

Designed for Virtualization

Abstraction toolset for unified management of Zones...

vmadm: VM Administration tool

- Create, Start, Stop, Modify, etc.
- Both Zones & KVM instances

imgadm: Image Administration tool

- Find, Download & Install Images
- Integrates with “dataset server” for easy distribution

Installation and configuration

```
Live 64-bit (text)
Live 64-bit (noinstall)
Live 64-bit (text) +kmdb
Live 64-bit (ttya) +kmdb
Live 64-bit (ttya)

-----
SmartOS Setup
Verify Configuration http://wiki.smartos.org/install
-----

Verify that the following values are correct:

MAC address: 0:c:29:73:59:c9
IP address: 192.168.146.132
Netmask: 255.255.255.0
Gateway router IP address: 192.168.146.2
DNS servers: 8.8.8.8,8.8.4.4
Default DNS search domain: com
NTP server: pool.ntp.org
Domain name: odd.job

Is this correct? [y]: y

Your configuration is about to be
Would you like to edit the final

-----
      / _ |  / _ |  / _ |  / _ |  / _ |  / _ |  / _ |  / _ |  / _ |  / _ |
     / ___|\_/ / ___|\_/ / ___|\_/ / ___|\_/ / ___|\_/ / ___|\_/ / ___|
    /___ \/_ \ /___ \/_ \ /___ \/_ \ /___ \/_ \ /___ \/_ \ /___ \/_ \
   / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___|
  / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___|
 / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___|
/ ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___|
/ ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___|
/ ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___|
/ ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___|
/ ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___|
/ ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___|
/ ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___|
/ ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___|
/ ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___| / ___|
/___ \ /___ \ /___ \ /___ \ /___ \ /___ \ /___ \ /___ \ /___ \ /___ \
 \___) \___) \___) \___) \___) \___) \___) \___) \___) \___) \___)
  \___) \___) \___) \___) \___) \___) \___) \___) \___) \___) \___)

    / ; Joyent Live Image v0.147+
    \ ; build: 20130419T073558Z

00-0c-29-73-59-c9 wscons login: _

00-0c-29-73-59-c9
login as: root
Using keyboard-interactive authentication.
Password:
Last login: Sun Jun 2 16:49:08 2013
- SmartOS Live Image v0.147+ build: 20130419T073558Z
[root@00-0c-29-73-59-c9 ~]#
```


Zones workflow

1. Import Image

If you want minimal, use “base”

It is not possible to create a zone without a template (image)

2. Write the JSON description

3. Create the zone

KVM workflow

Same process as Zone, but with 2 options:

1. Create an empty instance and boot from ISO
2. Import remote dataset from server

QEMU Runs inside minimal zone

- Logs are in `/zones/UUID/root/tmp/`

KVM Example

```
[root@00-0c-29-73-59-c9
Importing image 8700b668
Imported image 8700b668-
[root@00-0c-29-73-59-c9
{
  "brand": "kvm",
  "resolvers": [
    "208.67.222.222",
    "8.8.4.4"
  ],
  "default-gateway": "192.
  "ram": "512",
  "vcpus": "1",
  "nics": [
    {
      "nic_tag": "admin",
      "ip": "192.168.146.150",
      "netmask": "255.255.255.
      "gateway": "192.168.146.
      "model": "virtio",
      "primary": true
    }
  ],
  "disks": [
    {
      "image_uuid": "8700b668-
      "boot": true,
      "model": "virtio",
      "size": 10240,
      "image_size": 10240
    }
  ]
}

[root@00-0c-29-73-59-c9
Successfully created VM
]0;00-0c-29-73-59-c9 [ro
Usage: grep [-c|-l|-q] [
]0;00-0c-29-73-59-c9 [ro
root 4189 4138 24

[root@00-0c-29-73-59-c9
{
  "vnc": {
    "host": "192.168.146
    "port": 57806,
    "display": 51906
  }
}
]0;00-0c-29-73-59-c9 [ro
logout

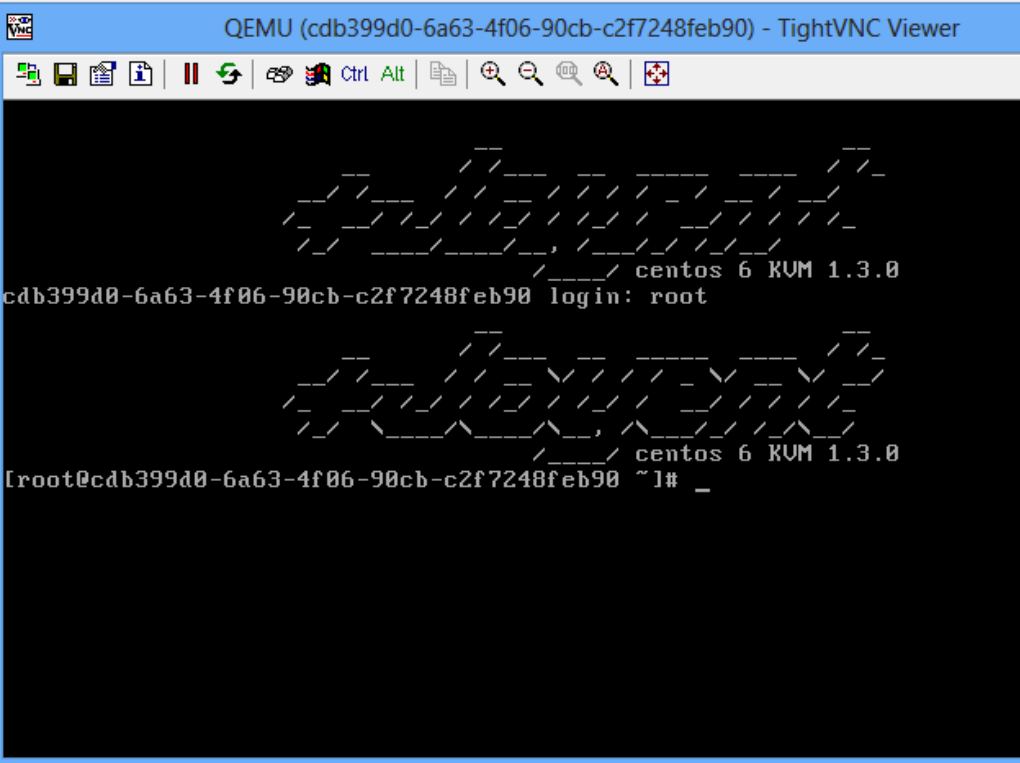
{
  "brand": "kvm",
  "resolvers": [
    "208.67.222.222",
    "8.8.4.4"
  ],
  "default-gateway": "192.168.146.2",
  "ram": "512",
  "vcpus": "1",
  "nics": [
    {
      "nic_tag": "admin",
      "ip": "192.168.146.150",
      "netmask": "255.255.255.0",
      "gateway": "192.168.146.2",
      "model": "virtio",
      "primary": true
    }
  ],
  "disks": [
    {
      "image_uuid": "8700b668-0da4-11e2-bde4-17221283a2f4",
      "boot": true,
      "model": "virtio",
      "size": 10240,
      "image_size": 10240
    }
  ]
}
```

KVM Example

```
00-0c-29-73-59-c9
[root@00-0c-29-73-59-c9 ~]# ps -ef | grep qem | grep -v
Usage: grep [-c|-l|-q] [-r|-R] -hbsvwi pattern file . . .
[root@00-0c-29-73-59-c9 ~]# ps -ef | grep qem | grep -v grep
root 4189 4138 24 17:54:58 ?        1:12 /smartdc/bin/qemu-system-x86_64 -m 512 -name cdb399d0-6a63-4f06-90cb-c2f7248feb90 -s -hda /smartdc/images/centos6-kvm-disk.img -m 512 -name cdb399d0-6a63-4f06-90cb-c2f7248feb90 -s -hda /smartdc/images/centos6-kvm-disk.img -m 512 -name cdb399d0-6a63-4f06-90cb-c2f7248feb90 -s -hda /smartdc/images/centos6-kvm-disk.img
[root@00-0c-29-73-59-c9 ~]# vmadm
Invalid or missing UUID for info
Usage: /usr/sbin/vmadm <command>

create [-f <filename>]
create-snapshot <uuid> <snapname>
console <uuid>
delete <uuid>
delete-snapshot <uuid> <snapname>
get <uuid>
info <uuid> [type,...]
install <uuid>
list [-p] [-H] [-o field,...] [-s]
lookup [-j|-l] [-o field,...] [file]
reboot <uuid> [-F]
receive [-f <filename>]
rollback-snapshot <uuid> <snapname>
send <uuid> [target]
start <uuid> [option=value ...]
stop <uuid> [-F]
sysrq <uuid> <nmi|screenshot>
update <uuid> [-f <filename>]
  -or- update <uuid> property=value
validate create [-f <filename>]
validate update <brand> [-f <filename>]

For more detailed information on the usage of vmadm, see the man page.
[root@00-0c-29-73-59-c9 ~]# vmadm info cdb399d0-6a63-4f06-90cb-c2f7248feb90
{
  "vnc": {
    "host": "192.168.146.132",
    "port": 57806,
    "display": 51906
  }
}
[root@00-0c-29-73-59-c9 ~]#
```



Big Data and SmartOS

Convert data into actionable insight

The SmartOS is the ideal platform for applications requiring Big Data storage and processing.

Use Cases

- High-performance-compute projects
- Distributed analytics projects
- Map reduce processes

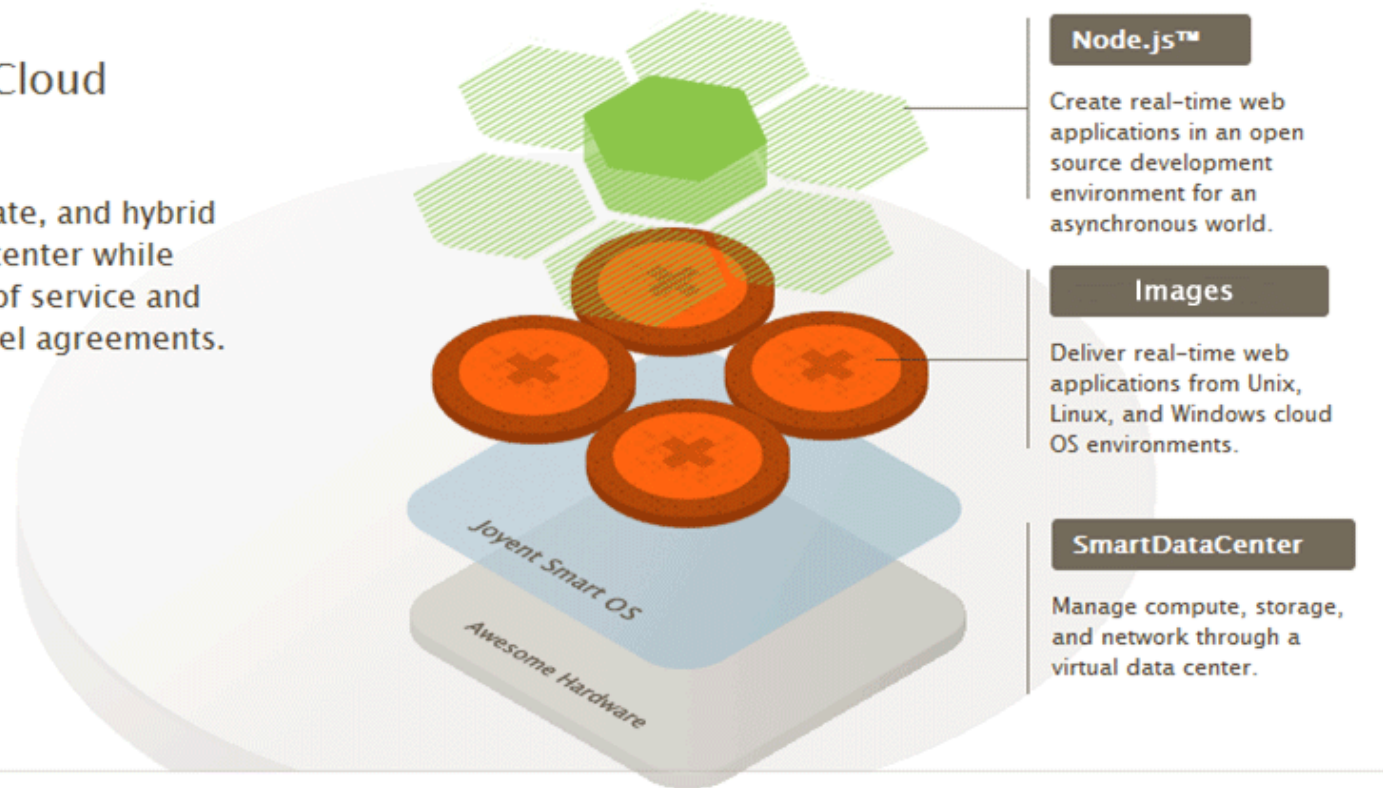
Getting Started in two easy steps

1. Deploy a standard image of your choice, and directly install and configure your database OR use a pre-configured database appliances, such as Riak, MongoDB, and Percona, that are made specifically for the SmartOS.
2. Then deploy the preconfigured Hadoop appliance and Bob's your uncle

Building high performance cloud applications

The Complete Cloud Software Stack

Operate public, private, and hybrid clouds in your data center while maintaining quality of service and honoring service level agreements.



Cloud Management and Orchestration

OpenNebula.org
The Open Source Toolkit for Cloud Computing

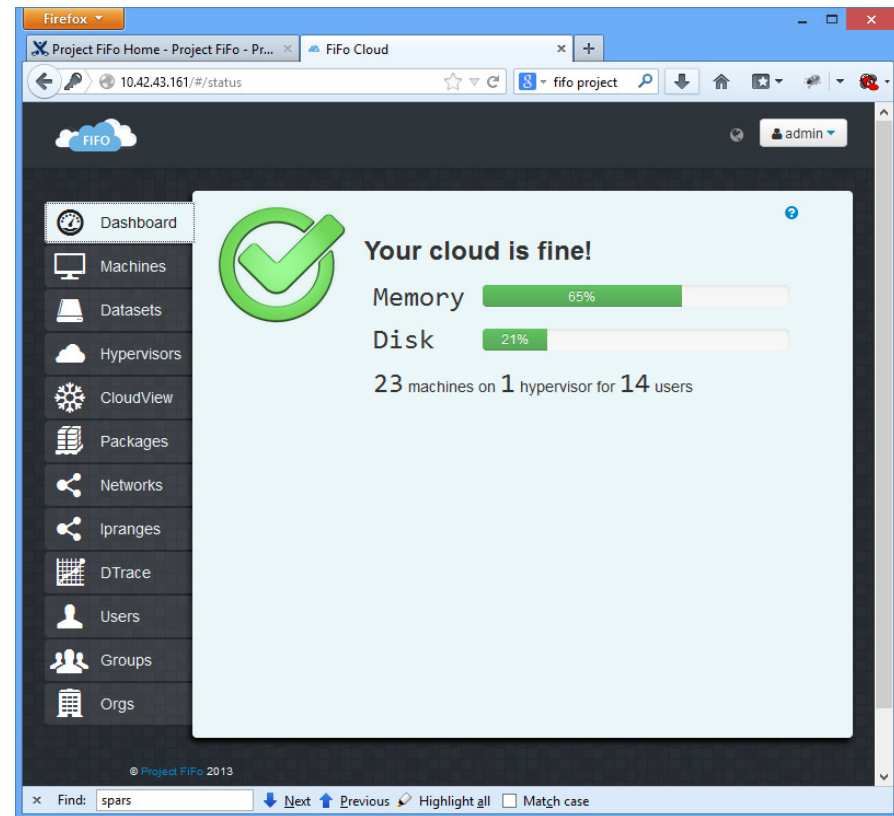
cloudstack
open source cloud computing



PROJECTFIFO
SmartCloud Orchestration

Cloud Management and Orchestration

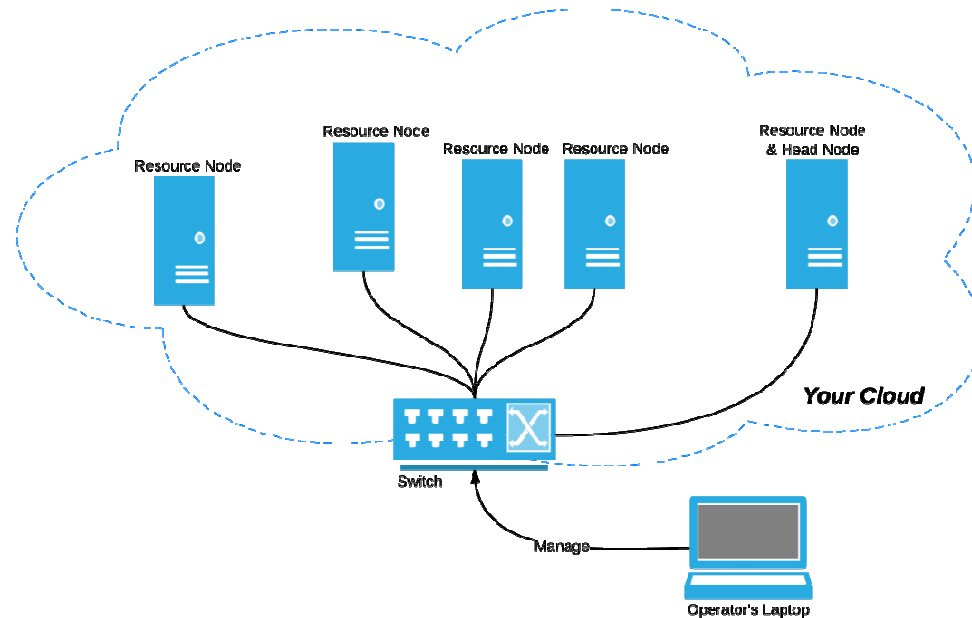
- Project FiFo is an open-source Cloud Management and Orchestration system for SmartOS virtualisation environments.
- The components of FiFo are written entirely in Erlang which gives the suite excellent stability and fault recovery as it continues maturing to a production quality release.
- 0.4.5 - "Whirling Whippet"
Current Release



DogeOS, the 2-in-1 bundle..

DogeOS is a distribution based on SmartOS and FIFO project. It is made to be the ultimate cloud OS for data center.

- All industry proven features of SmartOS: ZFS, Dtrace, KVM, Zones and Crossbow.
- Ready-to-use management console from FIFO.
- Nearly 100% resource utilization of hardware.
- No installation time for Resource Node (a.k.a chunter node).
- Guided, fast (< 10min) provision of management VM for Head Node (a.k.a fifo zone node), and works even without Internet access.



- Simply connect all your servers to your switch. They are all Resource Node, which will be VM farm.
- Select anyone of your servers to be Head Node, which will host a special VM(FiFo zone) for management system.



- For anyone interested in following the latest developments in the SmartOS community
- Join the group on LinkedIn
<http://www.linkedin.com/groups?gid=6520548>

Read all about it!

- All Things SmartOS
<http://wiki.smartos.org/display/DOC/Home>
- SmartOS home
<http://smartos.org/>
- Joyent
<http://www.joyent.com/>
- Node.js
<http://nodejs.org/>
- Misc.
<http://www.joyent.com/blog/magical-block-store-when-abstractions-fail-us>
<https://www.joyent.com/blog/network-storage-in-the-cloud-delicious-but-deadly/>
<https://www.joyent.com/blog/on-cascading-failures-and-amazons-elastic-block-store/>
<http://blog.hendrikvolkmer.de/2013/04/03/there-will-be-no-reliable-cloud-part-1/>
<http://blog.hendrikvolkmer.de/2013/04/09/there-will-be-no-reliable-cloud-part-2/>
<http://blog.hendrikvolkmer.de/2013/04/12/there-will-be-no-reliable-cloud-part-3/>