

# SmartOS


Zones + ZFS + KVM + DTrace = Awesome

# # whoami

- Frubar Network seit 2005
- Server Ninja / Gesellschafter bei SkyLime seit 2007
- Administration von ~200 Linux / Unix Server (Gentoo, Ubuntu, Debian, SmartOS)
- Fokus auf Monitoring, Webserver Cluster, Script Entwicklung

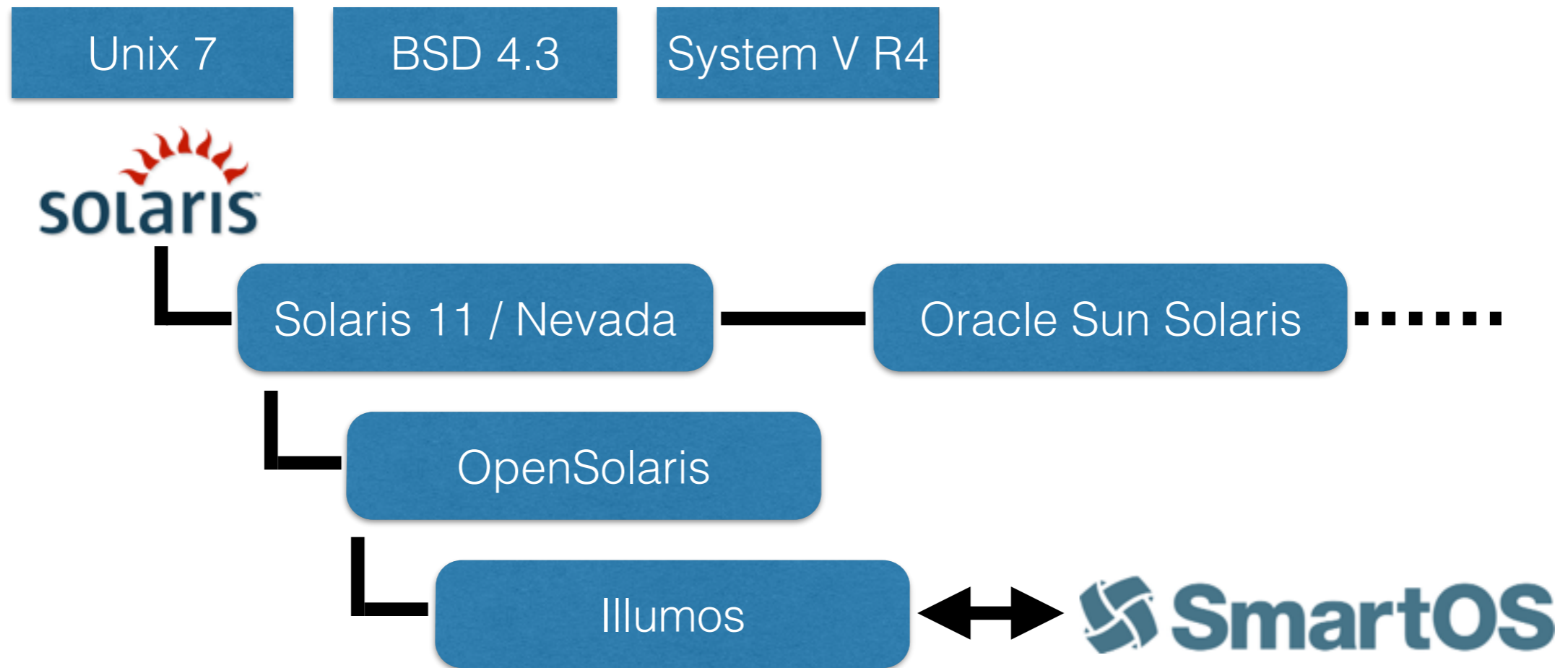
# Agenda

- Was ist SmartOS?
  - Zones
  - Crossbow
  - ZFS
  - KVM
- Thinking SmartOS
  - Installation
  - Tipps
  - Konfiguration der Globalen Zone
  - Datasets
  - Verwaltung von Zones und KVM Images
  - Paketmanagement in Zones
  - Update SmartOS
- Project FiFo
- Links

A large, complex industrial steam engine, likely a multi-cylinder marine engine, is shown in a museum or factory setting. The engine features a series of cylinders connected by a common crankshaft, with various valves, pistons, and connecting rods visible. The metal is dark and polished, with some brass components. The background shows a dark industrial environment with some lighting fixtures.

# Was ist SmartOS?

# Was ist SmartOS?



# Was ist SmartOS?

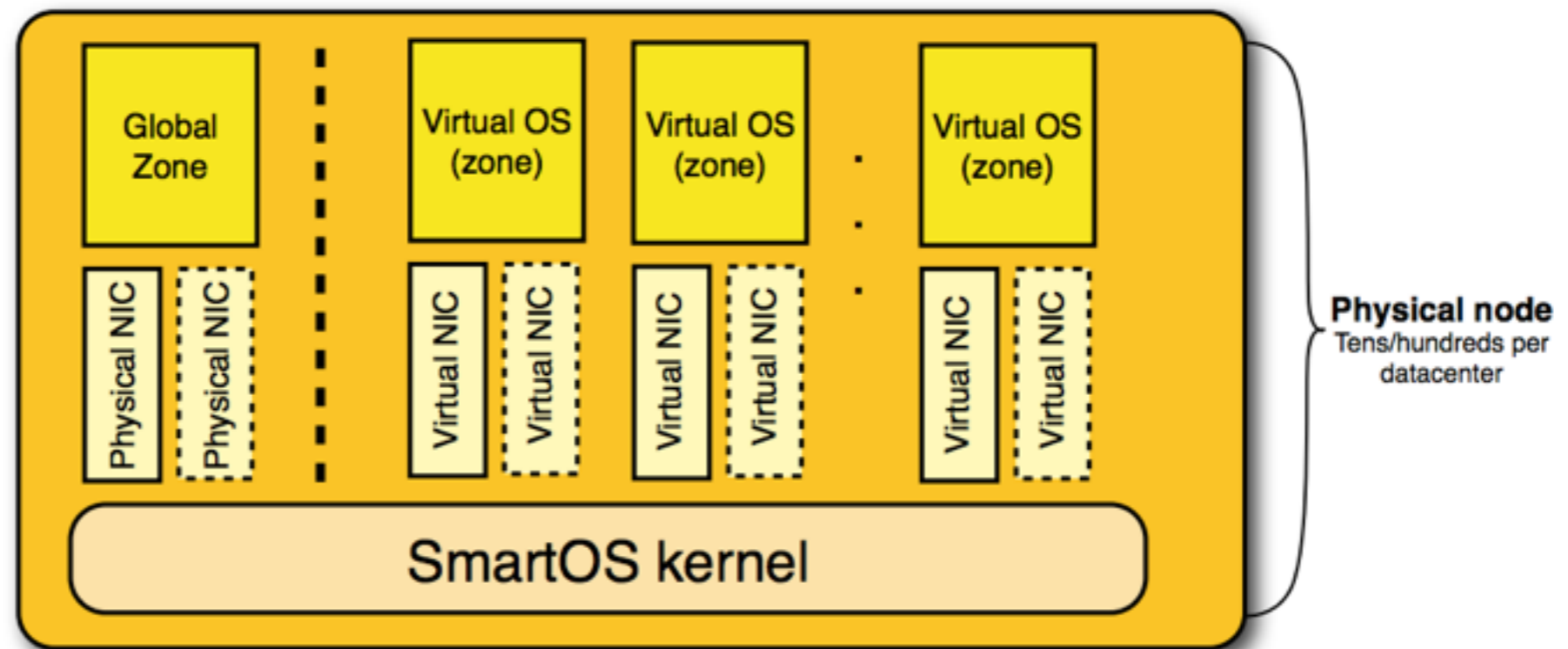
- Entwickelt von Joyent (Open Source)
- Solaris Features
  - Zones
  - Crossbow
  - ZFS
  - DTrace
- Fokus auf Virtualisierung
  - Boot via externer Medien (USB, Netboot)
  - Minimale Konfiguration
  - Verwaltungstools
- KVM Support

# Zones - OS Virtualisierung

- Selbstverwalteter Container
  - Konfiguration eigener Benutzer, Festplatten, Netzwerk und Dienste
  - „Gefühlt“ wie ein OS
- Isoliert
  - Zone sieht nur sich selbst
  - Globale Zone überwacht lokale Zonen
  - Eigenes Netzwerk und isoliertes Dateisystem
- Ressourcen Verwaltung
  - Arbeitsspeicher, Festplatte und Netzwerk I/O
  - CPU Verteilung

# Zones - OS Virtualisierung

- Minimaler Overhead
- Hardware muss nicht emuliert werden
- Dienste können aus der globalen Zone überwacht werden (z.B. via DTrace)





# Crossbow - Virtual NICs

- Erstellung von virtuellen Netzwerkinterfaces oder Switches
- Verbindung von VNICs zu
  - Physikalischen NICs
  - Virtuellen Switches
- Antispoofing
  - MAC Adressen
  - IP Adressen
  - DHCP
- Kontrolle der Bandbreite

# ZFS

- Copy on write Dateisystem
- Pool Storage
  - Partitionierung spielt keine Rolle
  - Verwaltung via Datasets
  - Live Änderungen an Quotas
- RAID
  - RAID 1 (Mirror)
  - RAIDZ, Z2, Z3, ... (Single, Double oder Triple parity)
  - RAID 0 (Striping)

# ZFS

- 128bit Checksum für alles
- Komprimierung
- Deduplizierung
- ZIL (Intent Log)
- L2ARC (Cache)
- Snapshot und Clone Support

# KVM

- Virtualisierung anderer Betriebssysteme
- Joyent unterstützt KVM nur bei Intel CPUs mit EPT
  - AMD Unterstützung durch Community
- Administrationsbefehle für KVM Image oder Zone sind gleich

A detailed LEGO Technic model of a multi-story industrial building. The top level features a row of white rectangular units with grey grilles. Below this, a complex network of white pipes and black valves is visible. The middle level contains two large white rectangular units with red lights on top, supported by grey pillars. The bottom level shows two yellow engines or pumps connected to a grey piping system. The entire model is built on a grey Technic base.

# Thinking SmartOS

# Installation

- Download und `dd` des Images auf USB Stick
  - <http://wiki.smartos.org/display/DOC/Download+SmartOS>
- Server via USB Stick Booten und Installer folgen
  - Netzwerk Informationen, Zpool anlegen, Root Kennwort
- Zusätzliche Konfiguration falls nötig unter `/opt/custom` und `/usbkey/config`
  - Puppet, Bash Scripts, eigene SVCs
- Und los gehts :-)

# Tipps

- Globale Zone
  - RAM Disk nur zur Zonenadministration
  - Daten in **/etc** und **/root**: nicht dauerhaft
  - Daten in **/opt** und **/var**: dauerhaft
  - Basis Konfiguration in **/usbkey**
  - SMF Manifest Dateien in **/opt/custom/smf/** werden beim Booten geladen

# Konfiguration Global Zone

- /usbkey/config

```
# admin_nic is the nic admin_ip will be connected to for headnode.
admin_nic=00:26:b9:87:47:6c
admin_ip=80.190.131.134
admin_netmask=255.255.255.128
admin_network=...
admin_gateway=80.190.131.134

headnode_default_gateway=80.190.131.129

# create a virtual switch for an internal NAT solution.
etherstub="switch0"

dns_resolvers=8.8.8.8,8.8.4.4
dns_domain=srv.skylime.net

ntp_hosts=pool.ntp.org
```



# Konfiguration Global Zone

- /opt/custom/

```
├─ cfg
│   ├── datacenter
│   ├── global
│   │   ├── crontab
│   │   │   ├── fmadm
│   │   │   └─ zpool
│   │   ├── root
│   │   │   └─ root
│   │   └─ script
│   │       ├── 10-hostname.sh
│   │       ├── 15-ipv6.sh
│   │       └─ 20-sendmail.sh
├─ script
│   └─ postboot.sh
└─ smf
    └─ postboot.xml
```

# Konfiguration Global Zone

- /opt/custom/smf/postboot.xml

```
<?xml version="1.0"?>
<!DOCTYPE service_bundle SYSTEM [...]>
<service_bundle type='manifest' name='site:postboot'>
<service name='site/postboot' type='service' version='1'>
  <create_default_instance enabled='true' />
  <single_instance />
  <dependency name="network" grouping="require_all"
    restart_on="error" type="service">
    <service_fmri value="svc:/milestone/network:default"/>
  </dependency>
  <exec_method type='method' name='start' timeout_seconds='0'
    exec=' /opt/custom/script/postboot.sh '>
  </exec_method>
  [...]
</service>
</service_bundle>
```

# Datasets

- Bestehende Images für Zones oder KVM
- Community Images bei <http://datasets.at>

```
# imgadm avail
```

UUID	NAME	VERSION	OS	PUBLISHED
d2ba0f30-bbe8-11e2-a9a2-6bc116856d85	ubuntu-12.04	2.4.2	linux	2013-05-13
30e9e4c8-bbf2-11e2-ac3b-3b598ee13393	centos-6	2.4.2	linux	2013-05-13
e42f8c84-bbea-11e2-b920-078fab2aab1f	fedora	2.4.2	linux	2013-05-17
df8d2ee6-d87f-11e2-b257-2f02c6f6ce80	freebsd	1.0.0	bsd	2013-06-19
bae3f528-e01f-11e2-b2cb-1360087a7d5f	cassandra	13.1.0	smartos	2013-07-02
274bc2bc-d919-11e2-b797-83245409fbeb	hadoop	13.1.0	smartos	2013-07-02
17c98640-1fdb-11e3-bf51-3708ce78e75a	base64	13.2.1	smartos	2013-09-20
b779b49a-29e4-11e3-9b1d-0b0b41ccdcad	standard64	13.2.1	smartos	2013-09-30

```
[...]
```

# Zonenerstellung (1)

- Gewünschtes Dataset importieren

```
# imgadm import 56a0655c-3cc6-11e3-9c79-5701599fdf05  
Importing image 56a0655c-3cc... (standard 13.2.3) from "https://images.joyent.com"  
56a0655c-3cc6-11e3-9c79-5701599fdf05 [ ] 0% 2.83MB 592.83KB/s 19m28s
```

# Zonenerstellung (2)

- JSON Beschreibungsdatei erstellen
- Generator auf <http://datasets.at>

```
{
  "brand": "joyent",
  "image_uuid": "56a0655c-3cc6-11e3-9c79-5701599fdf05",
  "autoboot": true,
  "alias": "nsa-box",
  "hostname": "mirror-customer-information",
  "dns_domain": "nsa.gov",
  "resolvers": [ "8.8.8.8" ],
  "max_physical_memory": 512,
  "max_swap": 512,
  "nics": [ { "nic_tag": "admin", "ip": "dhcp", "primary": true } ]
}
```

# Zonenerstellung (3)

- Zone erstellen via **vmadm**

```
# vmadm create -f nsa-box.json  
Successfully created 59aaecb6-467d-4116-92bd-f31842d44c90
```

- Zone anzeigen

```
# vmadm list
```

UUID	TYPE	RAM	STATE	ALIAS
59aaecb6-467d-4116-92bd-f31842d44c90	OS	512	running	nsa-box

# KVM

- Das selbe wie bei Zones aber mit zusätzlichen Möglichkeiten
  - Erstellen einer leeren Instanz und booten von ISO
  - Verwendung bestehender Datasets
- QEMU läuft in einer minimalen Zone
  - Logdateien unter `/zones/UUID/root/tmp`

# Paketmanagement

- Kein Paketmanagement in der Globalen Zone
- Paketverwaltung erfolgt durch **pkgsrc** von NetBSD
- Binary Pakete (32bit und 64bit)
- Releases jedes Quartal (2013Q4)
- Layout
  - PREFIX=/opt/local
  - PKG\_SYSCONFDIR=/opt/local/etc
  - PKG\_DBDIR=/opt/local/pkg



# Paketmanagement

- Bootstrap **pkgsrc** falls nicht vorhanden

```
# cd /
# curl -k \
    http://pkgsrc.smartos.skylime.net/[...]/bootstrap-2013Q2-x86_64.tar.gz \
    | gzcattar -xf -
# echo http://pkgsrc.smartos.skylime.net/packages/SmartOS/2013Q2/x86_64/All/ \
    > /opt/local/etc/pkgin/repositories.conf
# pkg_admin rebuild
# pkgin -y update
```

# Paketmanagement

- Datenbank update

```
# pkgin update
pkg_summary.bz2
processing remote summary (http://pkgsrc.smartos.skylime.net/[...]/x86_64/All)...
updating database: 100%
```

- Suche

```
# pkgin search bash
bash-doc-2.05.2      Documentation for the GNU Bourne Again Shell
bash-completion-2.1 Programmable completion specifications for bash
bash-4.2nb3 =       The GNU Bourne Again Shell

=: package is installed and up-to-date
<: package is installed but newer version is available
>: installed package has a greater version than available package
```

# Paketmanagement

- Installation

```
# pkgin install nginx-1.5.0
calculating dependencies... done.

nothing to upgrade.
1 packages to be installed: nginx-1.5.0 (342K to download, 777K to install)

proceed ? [Y/n]
```

- Einige Pakete installieren SMF (Service Management Facility)
- Alternativ NetBSD rc.d-Scripte

# Update SmartOS

- Ganz einfach :-)
- USB Stick
  - **smartos-platform-upgrade** Script  
<https://github.com/calmh/smartos-platform-upgrade>
  - Ersetzt den alten **platform** Ordner mit neuem **platform** Ordner

# Update SmartOS

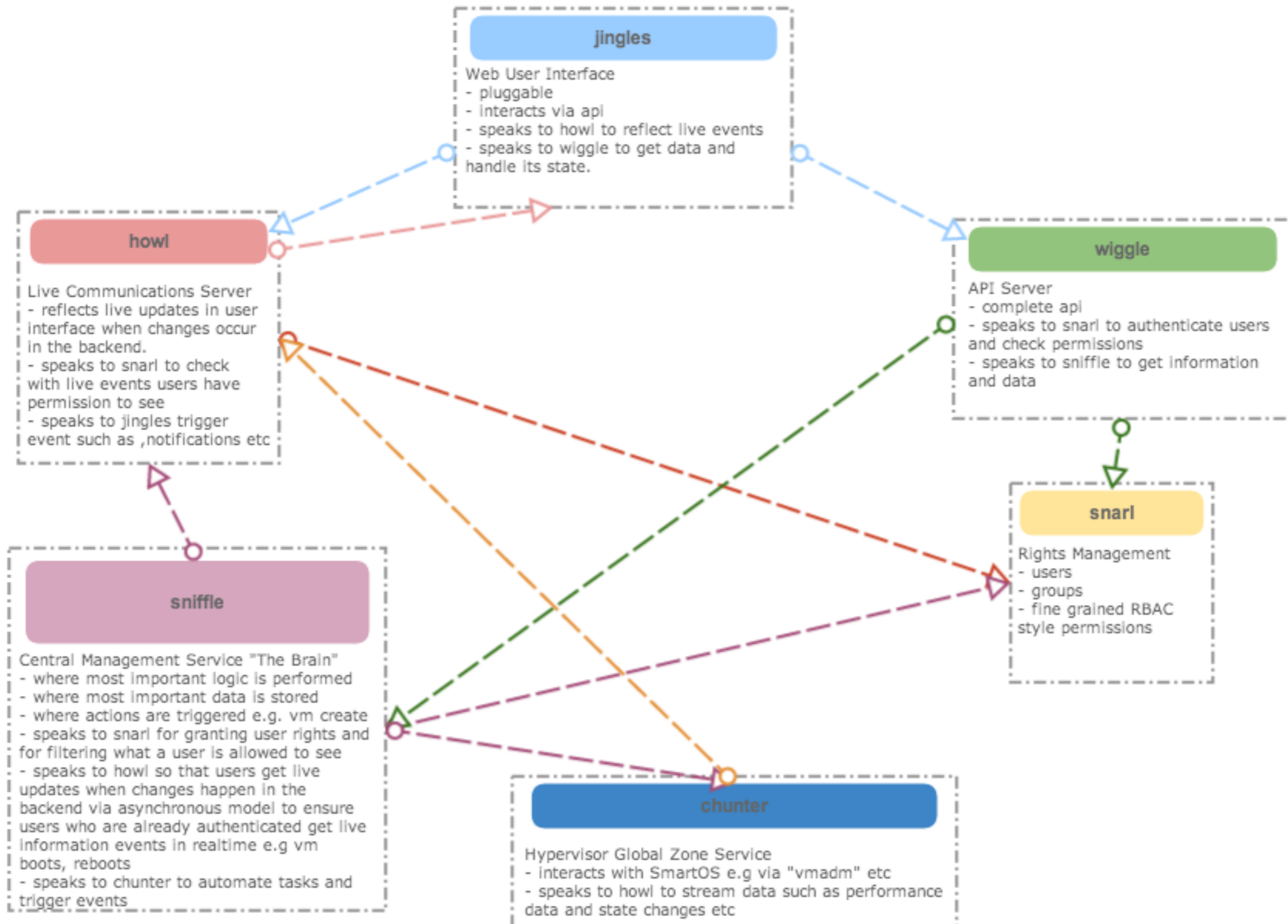
- TFTP Server
  - Platform Archiv herunterladen  
[https://us-east.manta.joyent.com/Joyent\\_Dev/public/SmartOS/platform-latest.tgz](https://us-east.manta.joyent.com/Joyent_Dev/public/SmartOS/platform-latest.tgz)
  - **IPXE Konfiguration** anpassen
- Server rebooten

# Projekt FiFo

# Projekt FiFo

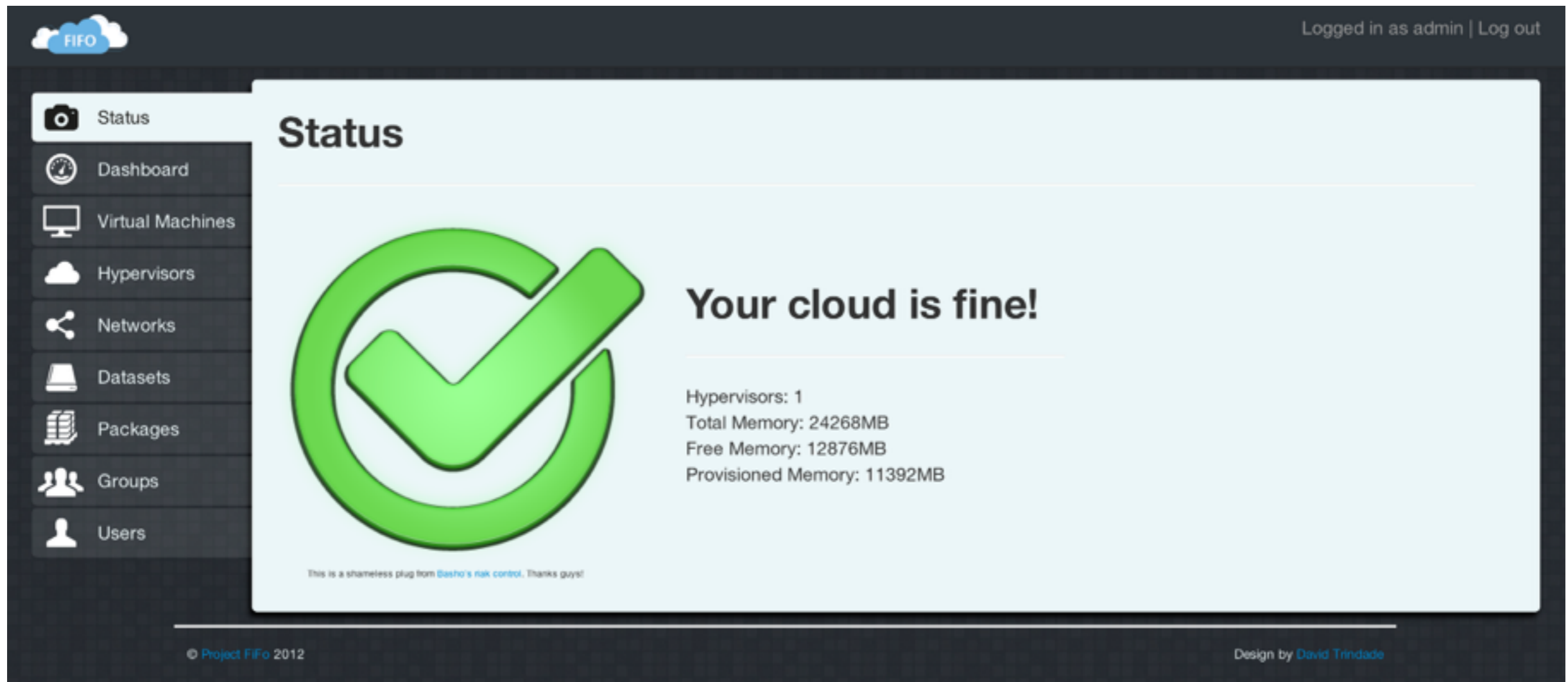
- <http://project-fifo.net>
- Open Source Cloud Management für SmartOS
- Entwicklung in Erlang
- Development Version
- mDNS Service Discovery
- VNC Support via noVNC

# Projekt FiFo - Architektur





# Projekt FiFo - Status



The screenshot displays the Project FiFo Status dashboard. The top right corner shows the user is logged in as 'admin' with a 'Log out' link. The left sidebar contains navigation links for Status, Dashboard, Virtual Machines, Hypervisors, Networks, Datasets, Packages, Groups, and Users. The main content area features a large green checkmark icon and the text 'Your cloud is fine!'. Below this, system metrics are listed: Hypervisors: 1, Total Memory: 24268MB, Free Memory: 12876MB, and Provisioned Memory: 11392MB. A small note at the bottom of the main area reads: 'This is a shameless plug from Basho's rak control. Thanks guys!'. The footer contains the copyright notice '© Project FiFo 2012' and the design credit 'Design by David Trindade'.

Logged in as admin | Log out

**Status**

**Your cloud is fine!**

Hypervisors: 1  
Total Memory: 24268MB  
Free Memory: 12876MB  
Provisioned Memory: 11392MB

This is a shameless plug from Basho's rak control. Thanks guys!

© Project FiFo 2012 Design by David Trindade

# Projekt FiFo - VMs

The screenshot displays the Project FiFo web interface. At the top right, it shows 'Logged in as admin | Log out'. The left sidebar contains navigation links: Status, Dashboard, Virtual Machines (selected), Hypervisors, Networks, Datasets, Packages, Groups, and Users. The main content area is titled 'Virtual Machines' and features a table with the following columns: Alias, UUID, IPs, Memory, CPU, Quota, Hypervisor, State, and Actions. The table lists several VMs, all in a 'running' state. A green '+' button is located in the top right corner of the table area.

Alias	UUID	IPs	Memory	CPU	Quota	Hypervisor	State	Actions
europa	0c91bc4e-0676-4c06-86ea-570e6ef61a47	172.16.0.16	2048MB	sh :100	20GB	00-15-17-19-34-c8	running	⊙ ⏻ ⏪ ⏩
proxy2	2cd647e6-8311-4868-ae61-2167bbd37f33	172.16.0.24	384MB	sh :100	10GB	00-15-17-19-34-c8	running	⊙ ⏻ ⏪ ⏩
io	35c812e7-31f3-4f0c-92bf-242c85a165d3	172.16.0.9	1024MB	sh :100	51GB	00-15-17-19-34-c8	running	⊙ ⏻ ⏪ ⏩
mercury	5ef38abd-a2d4-4bba-a677-9d8af388e482	dhcp,	256MB	sh :100	8GB	00-15-17-19-34-c8	running	⊙ ⏻ ⏪ ⏩
ceres	64ca42ab-5387-45e3-9bc4-8982ae766281	172.16.0.3	512MB	sh :100	10GB	00-15-17-19-34-c8	running	⊙ ⏻ ⏪ ⏩
callisto	91a10b3f-476f-46d6-afb2-660dbcbb35d4	172.16.0.6	2048MB	sh :100	34GB	00-15-17-19-34-c8	running	⊙ ⏻ ⏪ ⏩
-	ae1e2705-c96b-4acc-873c-dd46c8116692		0MB	sh :undefined	-	00-15-17-19-34-c8	running	⊙ ⏻ ⏪ ⏩
mars	f88b108e-4009-4915-bacd-3dc224a39ed9	172.16.0.23	512MB	sh :100	10GB	00-15-17-19-34-c8	running	⊙ ⏻ ⏪ ⏩

© Project FiFo 2012 Design by David Trindade

# Links

- SkyLime /opt/custom Scripts
  - <https://github.com/skylime/smartos-config>
- Joyent Wiki
  - <http://wiki.smartos.org/>
  - <https://github.com/joyent/smartos-live>
- Blog von Jonathan Perkin (Joyent Entwickler / pkgsrc)
  - <http://www.perkin.org.uk/>
- SmartOS mit IPv6 und AMD Support
  - <http://imgapi.uqcloud.net/builds>
- Community Datasets
  - <http://datasets.at>

# Community

- IRC ([irc.freenode.net](http://irc.freenode.net))
  - #smartos
  - #illumos
  - #joyent
  - #project-fifo
- Mailinglist
  - <http://smartos.org/smartos-mailing-list/>
  - <https://groups.google.com/forum/?fromgroups#!forum/project-fifo>



**Fragen?**