



Solaris und OpenSolaris

Eine sinnvolle Alternative?

Wolfgang Stief
best Systeme GmbH
MUCOSUG, GUUG e. V.
stief@stiefkind.org
2009-11-23

Agenda

- OpenSolaris, Solaris Express, Solaris Community Edition
- Das „Open“ in OpenSolaris Community, Lizenzen, Projects
- Features
Container/Zones, ZFS, DTrace, Crossbow ...
- Und warum dann nicht gleich Linux?

OpenSolaris?

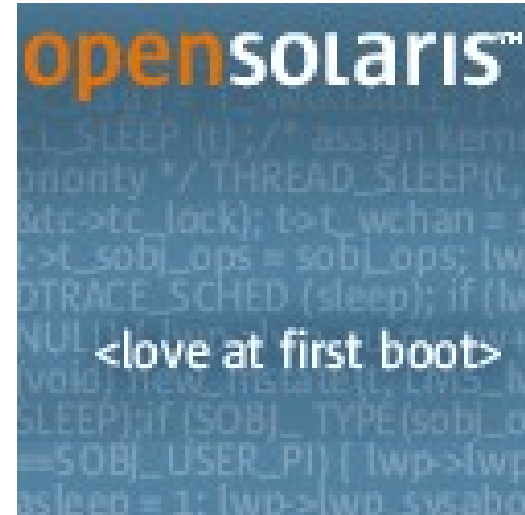


OpenSolaris Binary Distribution

- <http://www.opensolaris.com>
 - stabiler, getester Code
 - Support möglich
 - erscheint ca. 2x jährlich, x86 und SPARC
 - aktuelle Pakete (GNOME etc.), Installer
 - neues Paketformat, Repositories ähnlich Debian
 - aktuell (noch) 2009.06
- 
- opensolaris**

OpenSolaris Source Code

- <http://www.opensolaris.org>
- ab Juni 2005: zunächst DTrace, dann sukzessive weitere Teile
- aktive Community
- Source Browser OpenGrok
<http://src.opensolaris.org/>
- Features werden in „Projects“ entwickelt
- Community Release 2-wöchentlich (b127)
- „BFU“ nach Bedarf (blindingly fast upgrade)



OpenSolaris Community

- Launch am 14. Juni 2005
- mehrere Distributionen aus der Community
Schillix, Belenix, Nexenta, Milax, StormOS, OSUNIX
- Stand Frühjahr 2009 (ca. 3½ Jahre):
 - 116.000 registrierte Mitglieder
 - 350.000 Postings in Mailinglisten + Foren
 - 130+ Usergroups weltweit in 30 Ländern



OpenSolaris Community in Deutschland

- Deutsches Portal
<http://de.opensolaris.org/>
- OpenSolaris Developer Conference
<http://www.osdevcon.org/>
organisiert durch German Unix User Group (GUUG) e. V.
- deutsche User Groups
München, Hamburg, Berlin, Osnabrück, Rhein-Neckar, Rhein-Main, Erlangen/Nürnberg
<http://hub.opensolaris.org/bin/view/Main/usergroups>
- #opensolaris-de @ irc.freenode.net

OpenSolaris Projects

ADSL Modem Enhancement Automatic Data Migration BIND Update Bluetooth Stack & Drivers Brocade FC HBA - Initiator Brocade FC HBA - Target Brussels - unified network link configuration CIFS Server CIFS client for Solaris COMSTAR: Common Multiprotocol SCSI Target CPU Observability Caiman, Solaris Install Revisited Celeste Clearview: Network Interface Coherence Cluster Agent: Informix Dynamic Server Cluster Agent: OpenSolaris Container Cluster Agent: OpenSolaris xVM Cluster Agent: Oracle E-Business Suite Cluster Agent: Samba Cluster Agent: Tomcat Cluster Agent: PostgreSQL Coarse Data Flow Parallelism Colorado: Open HA Cluster on OpenSolaris Common Array Manager Companion - /opt/sfw: Free and Open Source software Country Portals Crossbow: Network Virtualization and Resource Control Crypto KMS Agent Toolkit Cryptographic Framework DHCPv6 Client DTrace Guide Data Migration Manager Data Tethers Device Detection Tool Device Driver Utility Device Manager Device Mapper Direct Rendering Infrastructure & 3D drivers Duckwater: Simplified name services management Easy Tools Emancipation Emulex Fibre Channel Device Driver Enable/Enhance Solaris support for Intel Platform Enhance the support of USB webcams Enhanced SMF Profiles Enhancements for AMD-based Platforms Ethernet bridge module for Solaris Evaluate Conary Events Registry Ext3 file system support F/OSS Package Base Fibre Channel over Ethernet File System Driven Device Naming Fine Grained Access Policy (FGAP) Fingerprint Authentication Flexible Mandatory Access Control Forensic Tools Fully Open X Project Fuse on Solaris Generic Machine Check Architecture Improvements Google Summer of Code HA-JBoss HA-MySQL HPC Stack Hadoop Live CD Hitachi HoneyComb Fixed Content Storage IP Datapath Refactoring IP over Infiniband IPsec Tunnel Reform Image Packaging System Improved Performance MIB Indiana Input Method Intel Graphics JKstat – a java binding for libkstat Java Desktop System: A GNOME-based desktop Journaled File System (JFS) K Desktop Environment Kerberos Kernel Sockets Key Management Framework Korn Shell 93 integration/migration project Labeled IPsec LatencyTOP Layer 2 Filtering Link Layer Discovery Live Media: Technologies for distributions running from CD and other media Locale Data MRSL.NONsharedDevice Media Management System Mega_sas MilaX minimal Live Distribution Mozilla DTrace Multi-lingual Glossary Multi-pathing software (MPxIO) Multiprotocol Label Switching Muskoka: An open repository for OpenSolaris technical content NFS RDMA transport update and performance analysis NFS Server in non-Global Zones NFS version 4.1 pNFS NFSv4 namespace extensions NPort ID Virtualization (NPIV) Navigator Nemo: A Framework for High-Performance Networking Network Auto-Magic Network Data Management Protocol Network MIBs Network Storage Network Time Protocol (NTP) Nevada Globalization New Design of 4over6 Mechanism Based on OpenSolaris Nightingale: Port Songbird to OpenSolaris OHACGE Script Based Plug-in ON/Nevada (ONNV) Project Object Storage Device (OSD) support for Solaris Open Development Infrastructure Open HA Cluster Utilities Open Sound System OpenGroK OpenPegasus CIM Server OpenRTI OpenSolaris Busybox OpenSolaris Community Lending Project OpenSolaris Content Project OpenSolaris Security Audit OpenSolaris support for the QEMU processor emulator -- host and guest PEF: Packet Event Framework Performance Wrappers Pkgfactory Power Management Usability Interfaces Presto: Automatic Printing Configuration Printable Many Page Solaris Manuals Programming Contest Quagga Routing Protocol Suite Integration RBridge (IETF TRILL) support RDMA Offload Framework Reno: Login Process Enhancements for Interop Resource Management SAM/QFS SCM Migration Project SCSI RDMA Protocol SDcard Drivers SFW SM-HBA Based SAS HBA Management SMF Documentation SVR4 Packaging Sensor Abstraction Layer Session Initiation Protocol Shell: bourne shell, korn shell, C shell, etc. Sierra: Intel WiFi Chipsets Support Simple Panels Solaris PowerPC Port Solaris iSCSI Target SourceJuicer Sparks: name service switch/nsd enhancements Squashfs Star integration/migration project Starfish Starter Kit Storage Power Management Sun StorageTek Availability Suite Support for OpenFabrics User Verbs / API on OpenSolaris OS Support gcc4/GCCfss in Solaris Suspend/Resume Systemz TIPC Tadpole Platform Support Tamarack: Removable Media Enhancements in Solaris Tesla: OpenSolaris Enhanced Power Management Test Development Tickless Kernel Architecture Trademark and Branding Trusted Platform Module support Trusted networking interface policy database for Trusted Extensions VRRP: Virtual Router Redundancy Protocol Implementation VSCAN service Validated Execution Project Virtual Console Virtual Network Machines Visual Panels Visualization for HPC Volo Web Stack Website Winchester: Schema mapping and ID mapping for AD Interoperability Wireless USB Support Wireless Wide Area Network Xfce – A lightweight desktop environment ZFS on disk encryption support Zone Manager Zone Statistics gcore iSCSI Extensions for Remote DMA (iSER) iSNS Server libMicro - portable microbenchmarks lofi compression & cryptography support lx64 brand x86 Generic FMA Topology Enumerator

<http://www.opensolaris.org/os/projects/>

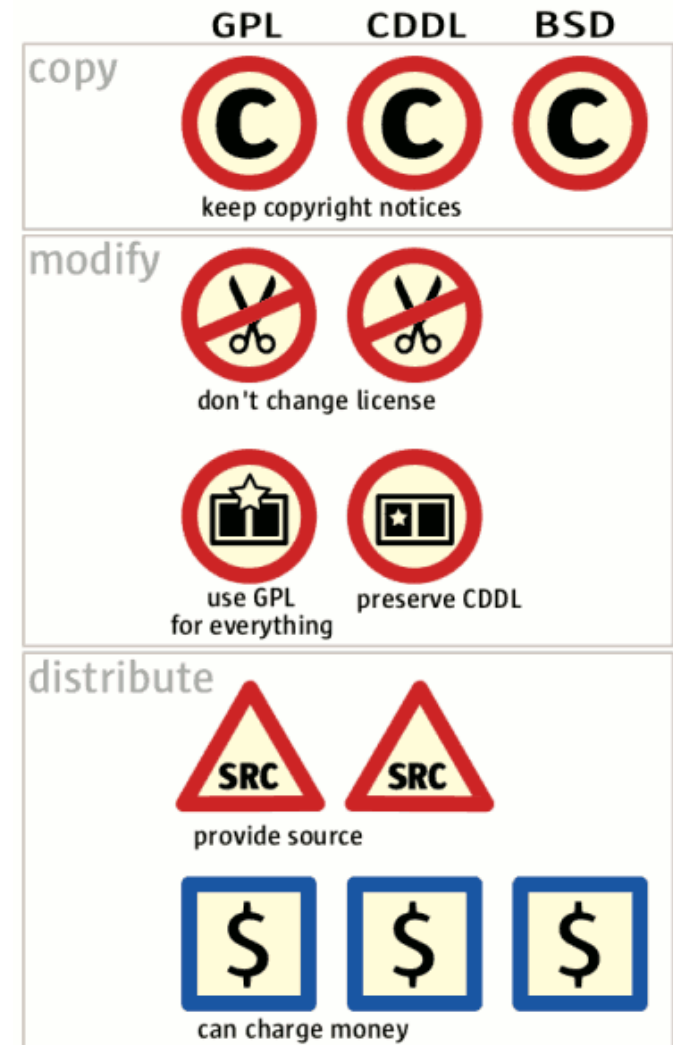
Stand 2009-04, ohne Usergroups und Länderportale

OpenSolaris Projects

- Quellcode, selbst kompilieren
 - aufwändig, Build-Umgebung notwendig
- BFU für Community Edition
 - nicht für alle neuen Features verfügbar
- Community Edition
 - nicht alle neuen Funktionen immer drin
 - abgekündigt
- OpenSolaris Binary Edition
 - Aktualisierung halbjährlich
 - nicht alle neuen Funktionen integriert

OpenSolaris Source Code Lizenz

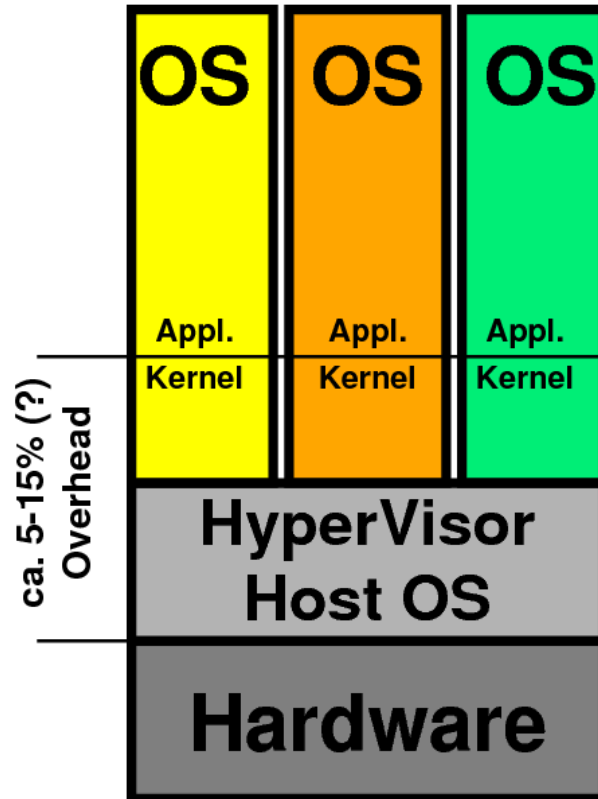
- CDDL
Common Development and Distribution License
<http://opensource.org/licenses/cddl1.php>
- inkompatibel zu GPL
- auch andere Projekte:
NetBeans IDE and RCP,
GlassFish, JWS DP, Project DReaM
- basiert auf Mozilla Public License 1.1



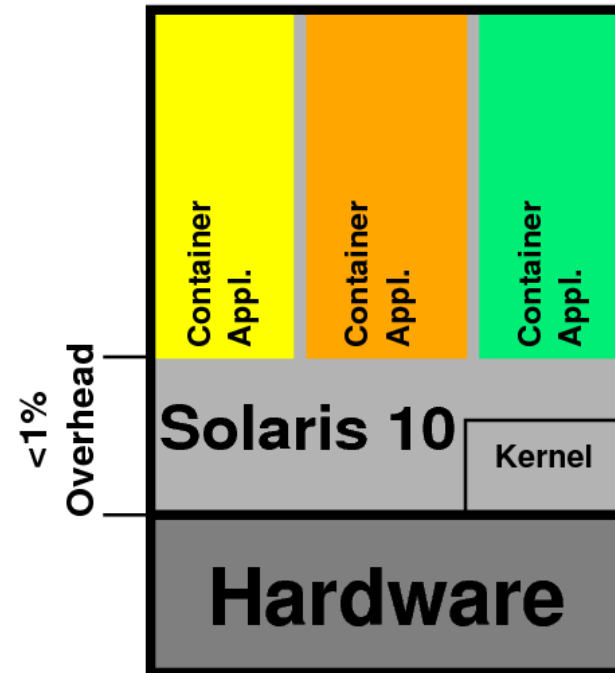
OpenSolaris – Features

Features

Features – Container aka Zones



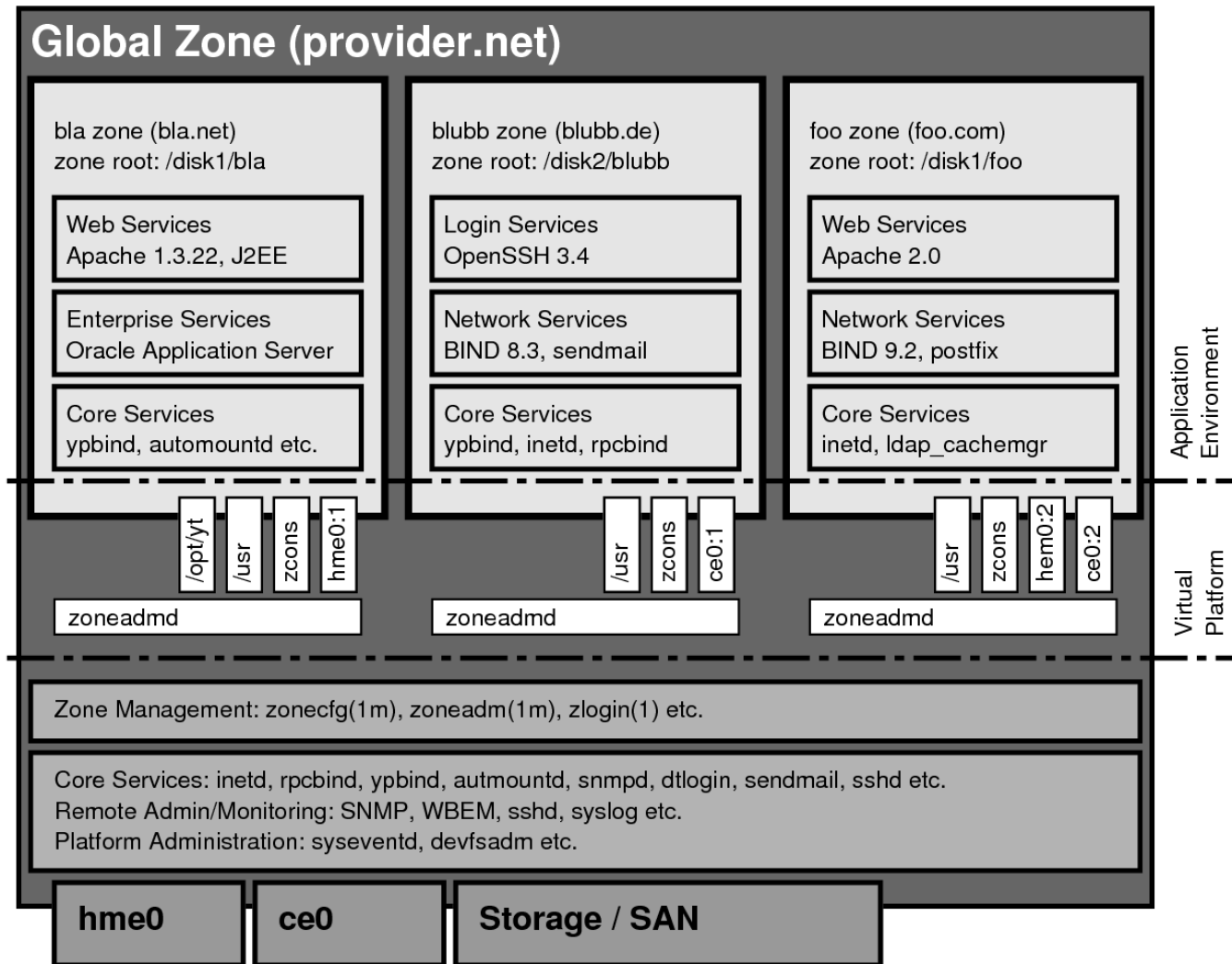
IBM LPAR
HP VPAR
EMC VMware



Features – Container aka Zones

- Virtualisierung mit Ressource Management
- nur eine OS-Instanz
 - wenig Overhead (4000 Zonen @ V880)
 - einfaches, zentrales Patchen
- automatisierbar per Shellscript
- Zonen können geclustert werden
- abgeschottet gegeneinander
 - getrennte /etc/passwd
 - getrennte Konfig-Files
- **zonecfg(1M), zoneadm(1M), zlogin(1)**

Features – Container aka Zones



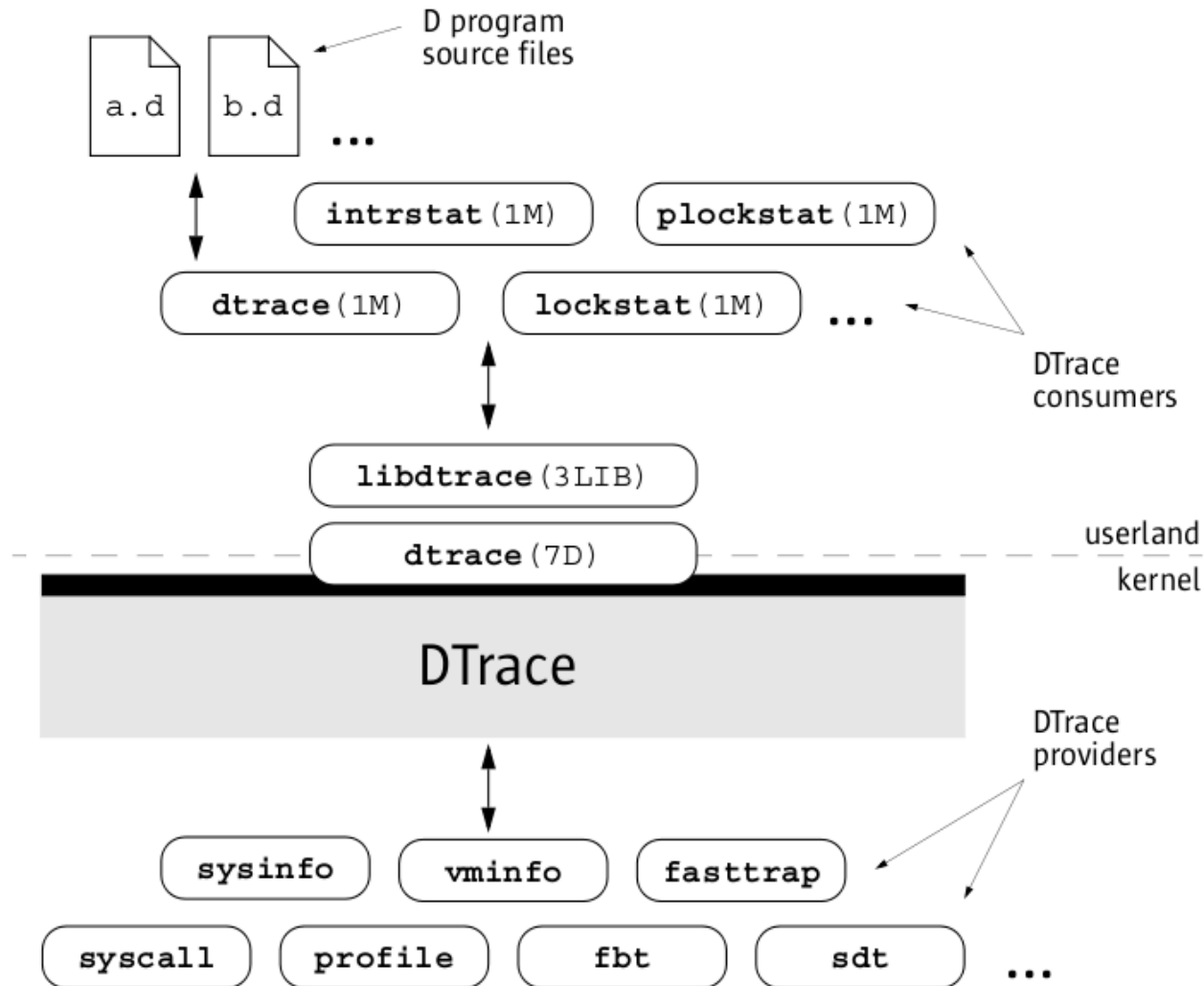
Features – ZFS

- Zettabyte Filesystem → 128bit
http://blogs.sun.com/dcb/entry/zfs_boils_the_ocean_consumes
- „The Last Word in Filesystems“
- Storage Pools, Journaling (transactional), Redirect on Write, Checksummen für Ende-zu-Ende Datenintegrität, RAID-Z und RAID-Z2, dynamic striping, Compression, Flashdisk für ZIL und L2ARC, Verschlüsselung, Snapshots und Clones, Send/Receive, CIFS, ACLs, Quota, Deduplication, Timeslider, Verschlüsselung
- **zpool(1M), zfs(1M)**

Features – DTrace

- Dynamic Tracing Framework
- 50.000+ Probes (Kernel, Applications)
- Scriptsprache D
- instrumentiert Userland und Kernel
- Systemutilities auf Basis von DTrace
lockstat, ps-Utilities, vmstat, iostat ...
- Probes: provider:module:function:name
- DTrace Toolkit, Chime
- <http://www.opensolaris.org/os/community/dtrace/>

Features – DTrace



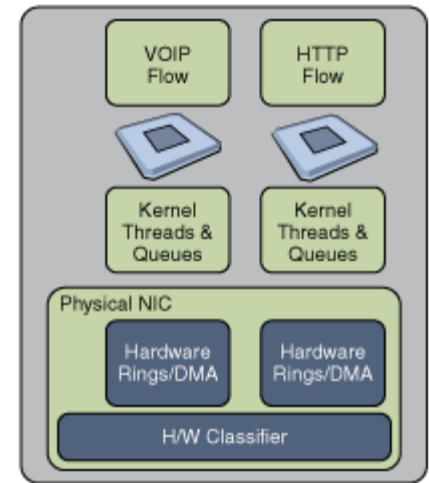
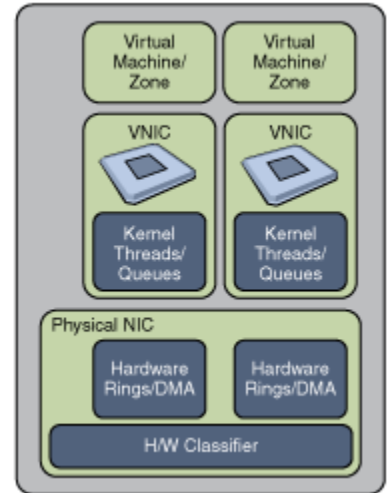
Features – DTrace

○ Verteilung der Blocksize (lesen) je Prozess

```
stief@ra:~$ pfexec dtrace -n \
  'syscall::read:entry{@[execname]=quantize(arg2);}'
dtrace: description 'syscall::read:entry' matched 1 probe
[...]
cat
  value  ----- Distribution ----- count
   4096  |                                                                    0
   8192  | @@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@ 2
  16384  |                                                                    0
[...]
dd
  value  ----- Distribution ----- count
 524288  |                                                                    0
1048576  | @@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@ 120
2097152  |                                                                    0
```

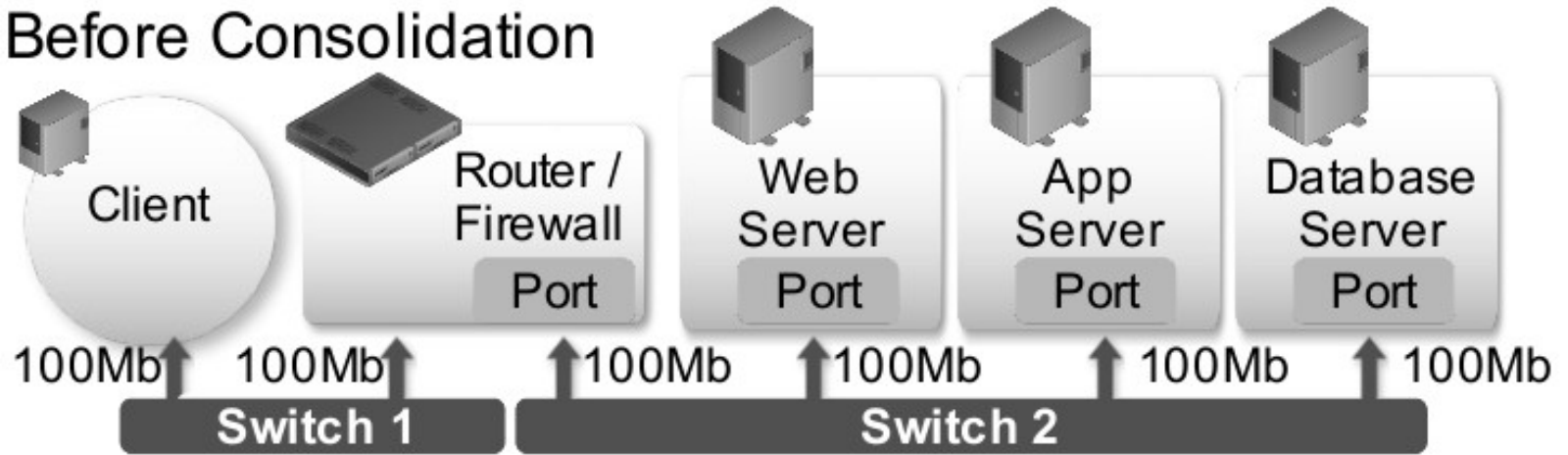
Features – Crossbow

- Virtualisierung von Netzwerkkomponenten: Virtual NICs, Virtual Switching, Verkabelung
- Flows (Quality of Service): Bandbreite, CPU, Protokoll
- **dladm(1M)**

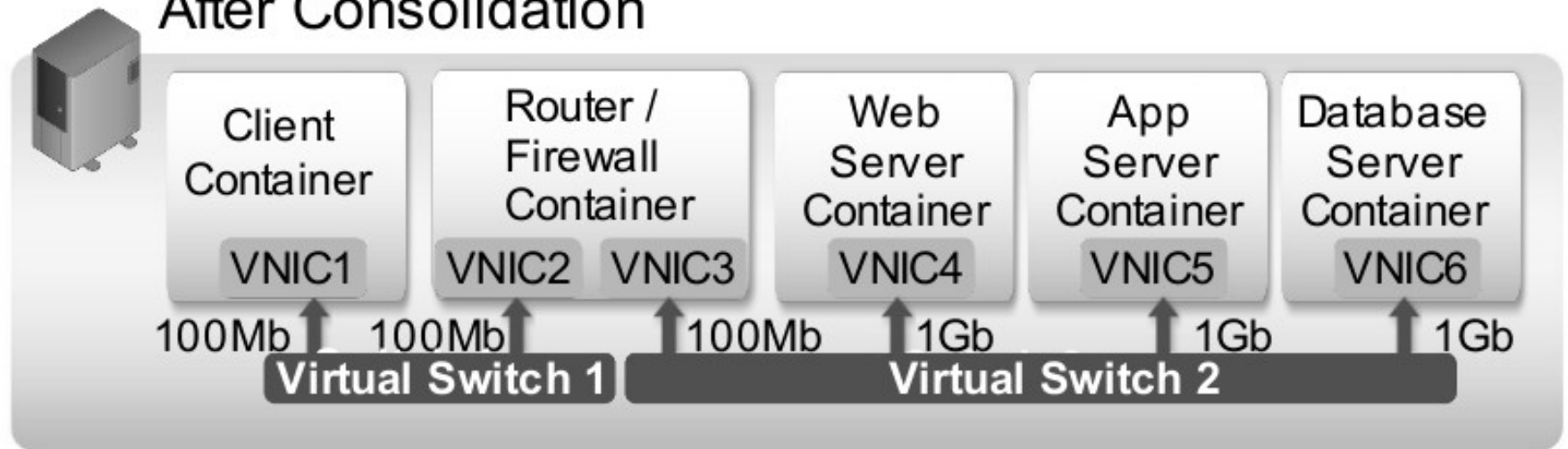


Features – Crossbow

Before Consolidation



After Consolidation



Features – Crossbow

```
stief@etain:~$ dladm show-link
```

LINK	CLASS	MTU	STATE	OVER
e1000g0	phys	1500	up	--

```
stief@etain:~$ pfexec dladm create-vnic -l e1000g0 schorsch1
```

```
stief@etain:~$ dladm show-link
```

LINK	CLASS	MTU	STATE	OVER
e1000g0	phys	1500	up	--
schorsch1	vnuc	1500	up	e1000g0

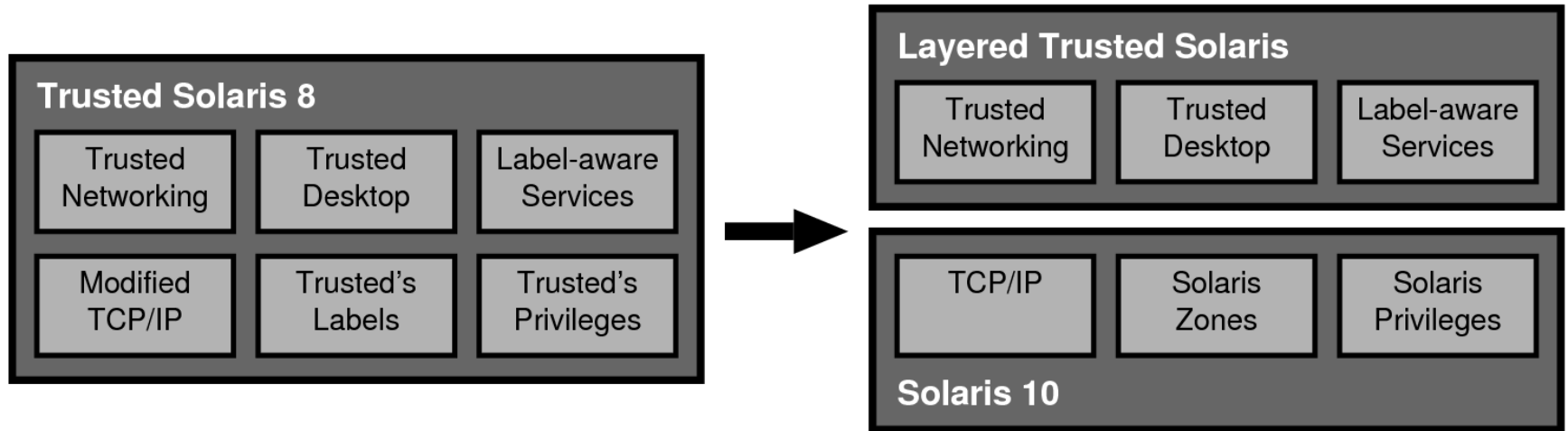
```
stief@etain:~$ pfexec dladm set-linkprop -p maxbw=40m schorsch1
```

```
stief@etain:~$ pfexec dladm show-linkprop schorsch1
```

LINK	PROPERTY	PERM	VALUE	DEFAULT	POSSIBLE
schorsch1	autopush	-w	--	--	--
schorsch1	zone	rw	--	--	--
schorsch1	state	r-	unknown	up	up, down
schorsch1	mtu	r-	1500	1500	--
schorsch1	maxbw	rw	40	--	--
schorsch1	cpus	rw	--	--	--
schorsch1	priority	rw	high	high	low, medium, high
schorsch1	tagmode	rw	vlanonly	vlanonly	normal, vlanonly

 <http://www.opensolaris.org/os/project/crossbow/>

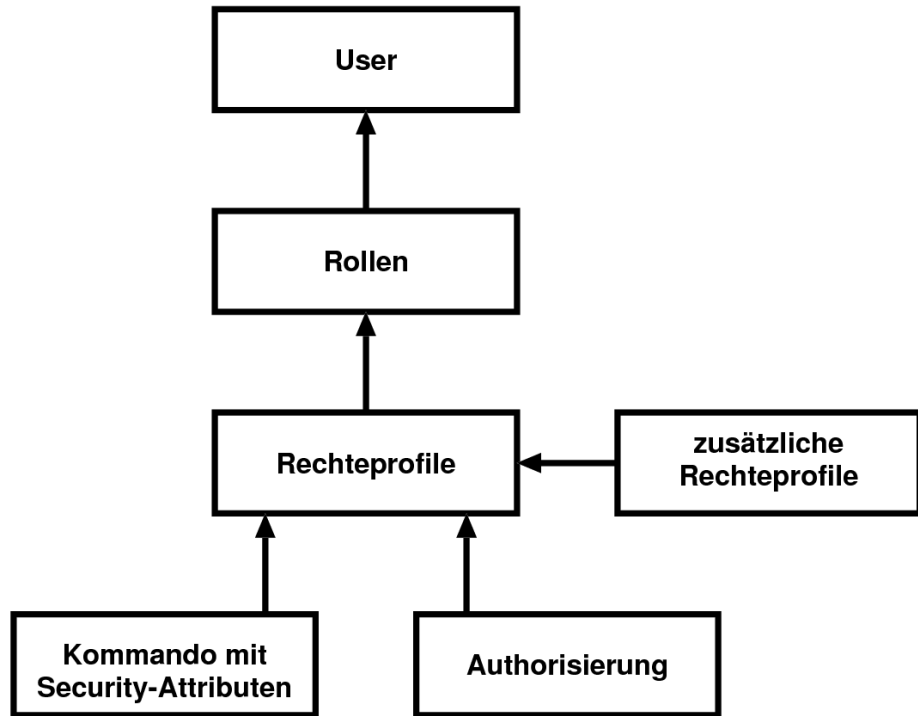
Features – Trusted Extensions



- installierbares Package, OS-Aufsatz
- „labeled security“
- Common Criteria, EAL4+

Features – RBAC

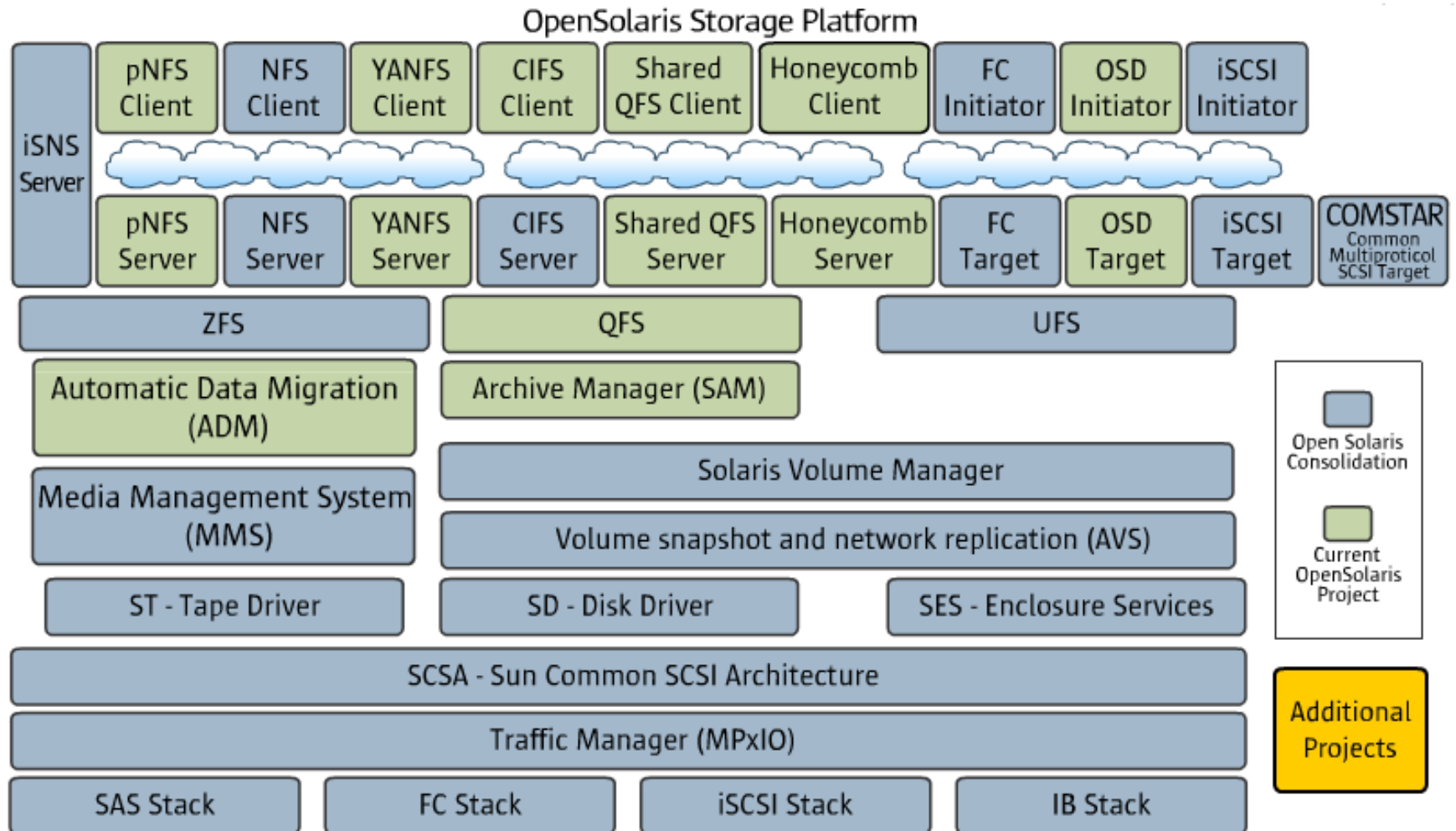
- Role Based Access Control
- Rechteprofile, Rollen
- **root** ist einfach eine spezielle Rolle
- *least privilege model*
- **pfexec** (= **sudo**)
- Administration:
roles, roleadd, smrole, usermod



Features – Process Rights Management

- Privilegien auf einzelne Prozesse
- 40 Privilegien in 5 Gruppen
FILE, IPC, NET, PROC, SYS
- Konfiguration:
ppriv, usermod, privileges
- vererbbar mit **exec(2)**

Features – Storage Stack



Features – Tesla

- OpenSolaris Enhanced Power Management
- **powertop (1M)**
- “tickless” kernel → `clock()` rewrite
- <http://hub.opensolaris.org/bin/view/Project+tesla/>
<http://hub.opensolaris.org/bin/view/Project+tickless/>

Und das war's schon?

- Service Management Framework
- Fault Management Architecture
- OpenHA Cluster, HPC Stack
- Automated Data Migration (HSM)
- CIFS Server (Kernel) u. Client
- NFSv4.1 aka pNFS
- COMSTAR, iSCSI, FC-Target, iSER, MPxIO
- xVM und LDOM Virtualisierung
- ...

Außerhalb des Betriebssystems?

○ SAMP-Stack

- Solaris, Apache, MySQL, PHP
- alternativ Perl, Ruby, Java, jRuby etc.
- vieles davon bereits mit DTrace Probes

○ OpenSource / GNU größtenteils verfügbar

- SourceJuicer, build service
- <http://juicer.opensolaris.org/>

○ SunRay Infrastruktur

- Thin Client, „Green IT“, hot desking, Smartcard

Und warum nicht Linux?

- viel weniger Features
 - ZFS? DTrace? MPxIO? Crossbow? Zones? SMF? FMA?
- *hakelige* Device Driver
- nur rudimentäres Volume Management
- Dokumentation / Manpages nicht konsistent, unvollständig, verstreut
- *Teufelskreis* Kernelversion vs. Modulversion
- Linux ist „work in progress“
 - wenig QA
 - Entwicklung im „produktiven“ Source

Wer mehr wissen will...

- HELDENfunk Podcast
<http://www.heldenfunk.de/>
- Das Solarium – Gruppenblog
<http://blogs.sun.com/solarium/>
- IRC-Channel (international/deutsch)
`#opensolaris @ irc.freenode.net`
`#opensolaris-de @ irc.freenode.net`
- OpenSolaris Forum
<http://www.opensolarisforum.de/>
- OpenSolaris Blogs / RSS-Feed
<http://planet.opensolaris.org/>

The End

Danke.

Fragen?

stief@stiefkind.org

<http://www.synapseninferno.org/>

<http://twitter.com/stiefkind>

stiefkind@irc.freenode.net