



Erfahrungsbericht: Portierung von Open HA Cluster auf OpenSolaris

SourceTalk 2008 in Göttingen

Thorsten Früauf
Availability Engineering
Sun Microsystems

Agenda

- Rückblick – **Open HA Cluster**
- Was ist OpenSolaris?
 - Veränderungen mit OpenSolaris
 - Solaris Express vs OpenSolaris
- Projekt Colorado
 - Ziele
 - Herausforderungen und mögliche Lösungen
- Referenzen
- Fragen / Antworten

Rückblick – Open HA Cluster

- Open Source von Solaris Cluster
 - Data Services (Juni 2007)
 - Geographic Edition (Dezember 2007)
 - Core Framework (Mai 2008)
- Einführung in Open HA Cluster incl. Live Demo letztes Jahr bei SourceTalk 2007
 - http://blogs.sun.com/TF/entry/sourcetalk_2007_reloaded
- Solaris Cluster Express läuft auf Solaris Express Community Edition (Nevada)
- ... und was ist mit OpenSolaris?

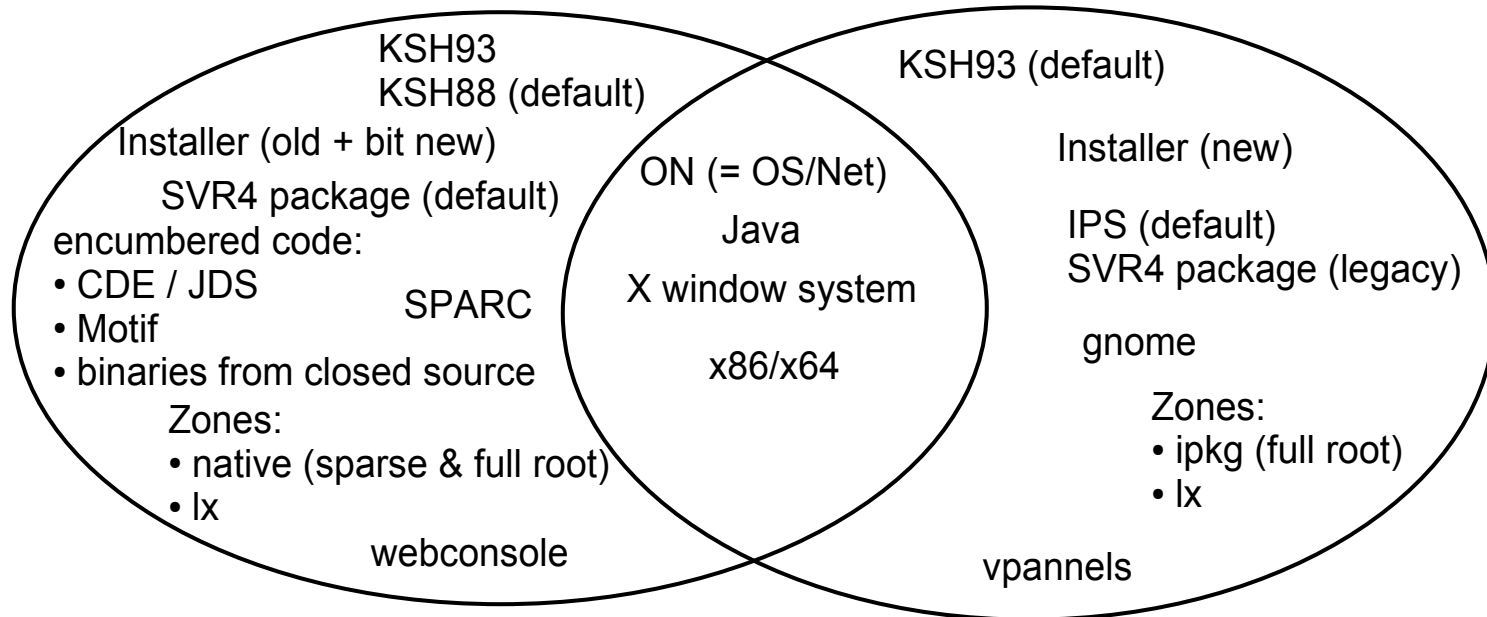
Was ist OpenSolaris?

- OpenSolaris – der Quellcode
 - Open Source Quellcode von Solaris und anderen Betriebssystem nahen Schichten wie Open HA Cluster
- OpenSolaris – die Distribution(en)
 - Solaris Express (Developer und Community Edition)
 - **OpenSolaris 2008.05 (Indiana)**
 - NexentaOS, BeleniX, MartUX, SchilliX, etc
- OpenSolaris – die Gemeinschaft (Community)
 - Portal für Community Groups und Projects
 - Infrastruktur (Diskussionsgruppen, Code Verwaltung, Webseiten, Wahlsystem, Bugtracking, etc)
 - Portal für lokale User Gruppen

Veränderungen mit OpenSolaris

- ZFS als root Filesystem
- Image Packaging System (IPS)
- Neuer Installer
- ipkg brand type für Zonen
- KSH93 ersetzt KSH88
- Frei neu verteilbar
 - keine Binaries basierend auf „closed source“
- Live CD
- SunStudioExpress als Compiler

Solaris Express vs OpenSolaris



Solaris Express (Nevada)

Binärdistribution aus
usr/src und usr/closed
nicht frei neu verteilbar

OpenSolaris (Indiana)

Binärdistribution aus
usr/src
frei neu verteilbar

Projekt Colorado - Ziele

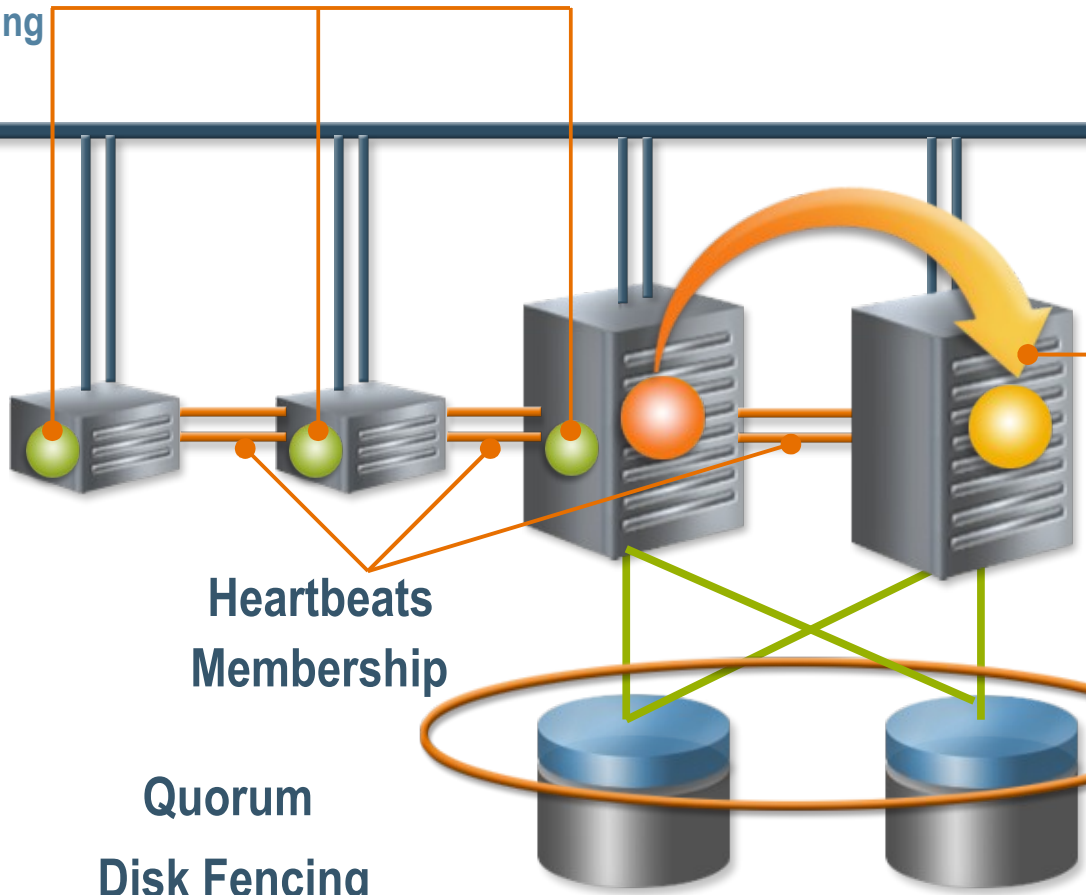
- Teile von Solaris Cluster laufen auf der OpenSolaris Binärdistribution
- Verteilung als IPS Pakete von IPS package repositories (z.B. pkg.opensolaris.org)
- Hardware Minimierung: Off-the-shelf low-end Workstations und PCs nutzbar
- Software Minimierung: Installation und Aktivierung eines Features nur wenn es benutzt wird. Reduzierung des statischen und dynamischen Speicherverbrauchs.

Solaris Cluster Architektur

Scalable Service
Software Load
Balancing

Global Network Service
Provides Global IP address
with failure protection

Monitoring



Heartbeats
Membership

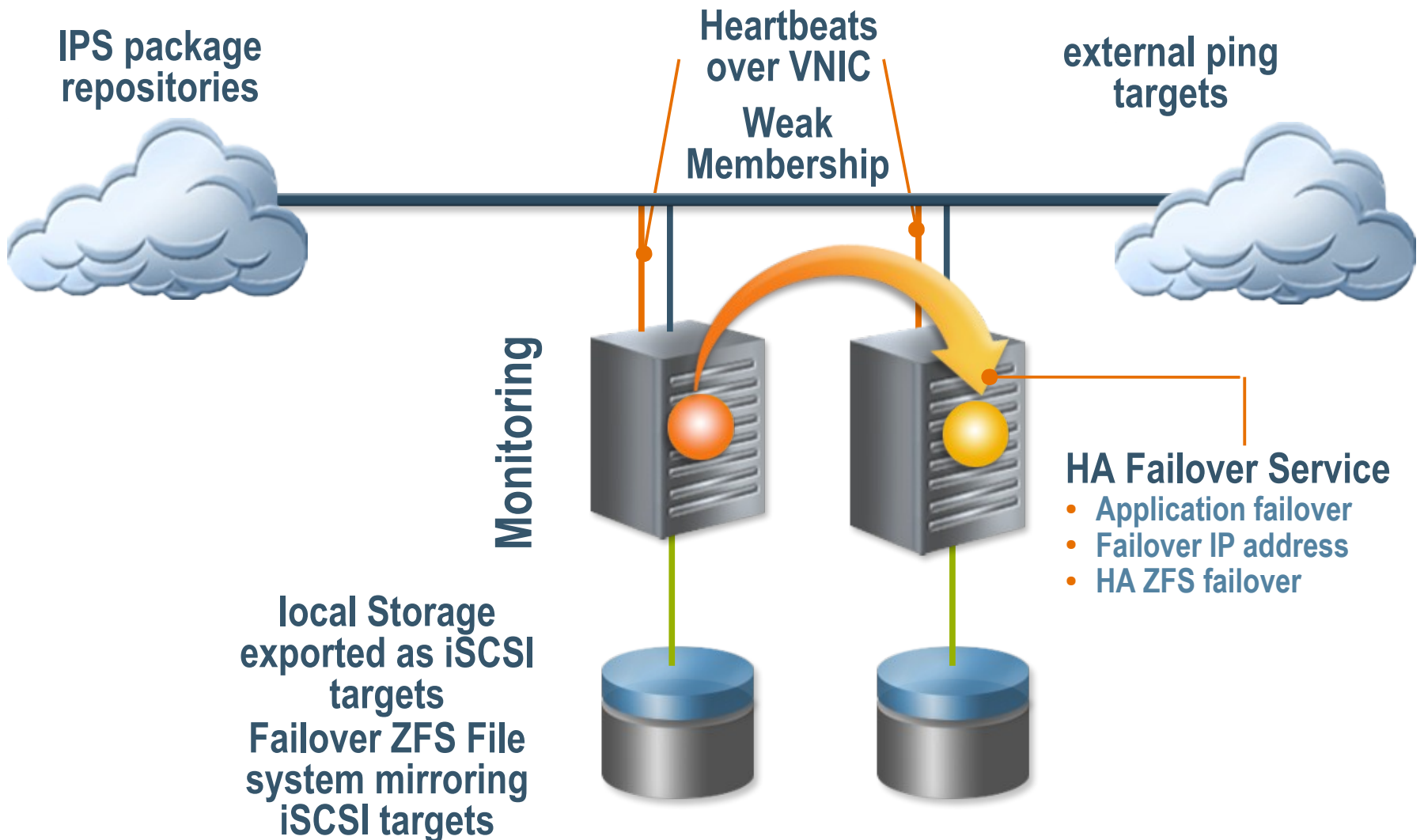
Quorum
Disk Fencing

HA Failover Service

- Application failover
 - Within nodes or container
- Containers failover
- Failover IP address

Global File Service
Failover File service

Minimale HA Colorado Konfiguration



Herausforderungen und mögliche Lösungen

ZFS als default root Filesystem

○ Herausforderung

- Installer lässt keine Wahl – ZFS root ist default
- keine zusätzlichen Filesysteme/Slices konfigurierbar

○ Lösung

- optionaler /globaldevices Slice
 - lofi mount eines Files
 - scinstall schlägt diese Option vor

✓ bereits erledigt :-)

Image Packaging System (IPS) (1)

○ Herausforderung

- Entwicklung noch nicht abgeschlossen
- keine Skripten für pre/post-install oder pre/post-remove (Designentscheidung)
- Handhabung von optionalen Funktionalitäten

○ Lösung (unvollständig)

- Erweiterungen anregen für allgemeingültiges
 - SMF action
 - spezifische action für z.B. crontab, /etc/system, /etc/rpc, RBAC, etc
- Aktionen in eigenes Tool integrieren

Image Packaging System (IPS) (2)

○ Herausforderungen

- IPS Pakete haben keine Filesystemdarstellung
 - Build Prozess stark auf SVR4 Pakete abgestimmt
- Unterscheidung i386 / SPARC Pakete
- Unterscheidung debug / non-debug Pakete
- jeder Entwickler braucht eigenes Repository

○ Lösung

- Build Prozess anpassen (Makefile, Skripte)
- IPS Repository Infrastruktur bereitstellen

Image Packaging System (IPS) (3)

○ Herausforderungen

- bisherige Paketabhängigkeiten nicht ausreichend oder nur sehr grob definiert
- implizite Abhängigkeiten durch Meta Paketcluster abgebildet

○ Lösung

- möglichst automatisiert fein granulare Abhängigkeiten feststellen und definieren

ipkg Zonen

○ Herausforderung

- Entwicklung noch nicht abgeschlossen
- Pakete werden nicht mehr vererbt oder automatisch installiert bei Zoneninstallation
- neuer brand-type ipkg
- keine sparse root Zonen

○ Lösung (unvollständig)

- Anpassung HA Container Agent (Phase 1)
- Mechanismus zur Installation/Konfiguration der notwendigen Cluster Pakete in der Zone entwickeln

KSH93 vs KSH88

○ Herausforderung

- KSH93 handhabt lokale/globale Variablen anders als KSH88

- dazu noch unterschiedliche Optionen für `typeset`

- komplexe Skripte laufen nicht 1:1

○ Lösung (sehr grob)

- Verwendung von `/usr/xpg4/bin/sh`

- Anpassung auf KSH93 sehr aufwendig und zeitintensiv

StudioExpress Compiler

○ Herausforderung

- Debugging Format Umstellung von `stabs` auf `DWARF`
 - `ctfconvert / ctfmerge` kann noch kein `DWARF`
- C++ kernel Module funktionieren nicht
 - vermutlich Fehler (CR 6747192), existiert auch mit Studio 12

○ Lösung (unvollständig)

- Anpassung von `ctfconvert / ctfmerge`
- Verwendung von Studio 11 bis Fehler behoben

Sun Java Webconsole

○ Herausforderung

- Webconsole mit Lockhart Framework (und andere wie JATO) nicht mehr Teil von OpenSolaris
- Umstellung auf anderes Framework nicht trivial
 - damit keine GUI!

○ Lösung

- noch in der Evaluierung
 - ggf. Neuimplementierung
 - AJAX
 - Visual Pannel

Closed Source

○ Herausforderung

- CDE (und damit Motif) nicht Teil von OpenSolaris
 - adminconsole hat Abhängigkeit auf Motif / ToolTalk
- einige „shared Components“ nicht verfügbar für OpenSolaris
- einige Treiber für Hardware sind nicht verfügbar für OpenSolaris

○ Lösung (unvollständig)

- statt adminconsole Open Source Tool pconsole verwenden

Verschiedenes

- Herausforderung
 - derzeit keine Netzwerkinstallation
 - derzeit kein SPARC Support
- Lösung
 - man arbeitet daran
 - erste Besserung mit OpenSolaris 2008.11 erwartet

Referenzen (1)

- HA Clusters Community Group
 - <http://opensolaris.org/os/community/ha-clusters/>
- Projekt Colorado
 - <http://opensolaris.org/os/project/colorado/>
- Image Packaging System (IPS)
 - <http://opensolaris.org/os/project/pkg/>
- Projekt Crossbow (VNICs)
 - <http://opensolaris.org/os/project/crossbow/>
- KSH93 Information
 - <http://opensolaris.org/os/project/ksh93-integration/>

Referenzen (2)

- Blog zum Zonen ipkg brand type
 - http://blogs.sun.com/dp/entry/a_field_guide_to_zones
- Roadmap OpenSolaris Town Hall Meetings
 - http://www.genunix.org/wiki/index.php/Indiana#OpenSolaris_Community_Strategy_Planning_Meeting
- Projekt Visual Pannels
 - <http://www.opensolaris.org/os/project/vpanels/>
- Solaris Cluster Blog
 - <http://blogs.sun.com/SC/>



Vielen Dank!
Fragen?

thorsten.frueauf@sun.com

Thorsten Früauf
Availability Engineering
Sun Microsystems
