



Virtual Storage Domains im Schnelldurchlauf

OSL Technologietag

Schöneiche / Berlin • 29. November 2017

info@osl.eu

Heutige I/O-Treiber

Heutige I/O-Treiber

Auswahl nachteiliger Aspekte typischer Implementierungen



- Design meist älter als 15 Jahre
- Typisch komplett im Kernel → extreme Spezialisierung → eingeschränkte Portabilität
- Typischerweise keine Abbildung von Partitionen / Zonen in modernen Systemen (anonymer, globaler Ausführungskontext)

Doch die Welt hat sich geändert

Dinge neu denken – worum geht es heute?



- Performancesteigerung heißt heute maßgeblich Parallelisierung
- WIRKLICH!
- Segmentierung von Administration und Lasten wichtiger als die letzten 5% Durchsatz
- Virtualisierung überall
- Non-Stop-Betrieb
- Isolation und korrekte Behandlung von Fehlern
- Flexibilität und Agilität

OSL RSIO

OSL RSIO - Remote Storage I/O

Prototyp einer neuen Architektur



- Kombination aus Kernel-Minimodul und Userspace-Driver
- Kernel-Modul:
 - Implementierung plattformspezifischer Interfaces
 - RSIO-spezifischer Code als Cross-Platform Common Code
- Userspace-Driver → RSIO-spezifischer Code als Cross-Platform Common Code
- Vorteile:
 - Portabilität
 - Systemstabilität
 - bessere Fehlerbehandlung und leichteres Debugging
 - Wahlfreiheit bei der Entwicklungsplattform

OSL Virtual Storage Domains



- Weiterentwicklung der erstmals mit RSIO implementierten Treiberarchitektur
- Kernelmodul jetzt als generisches Micro-Module
- Spezifische Funktionen komplett im Driver
- Raus aus der "Kernel-Falle":
 - Ablauf im Wesentlichen in einem Applikationskontext
 - Bestmögliche Loslösung von proprietären Interfaces
- Hoch portables, umfassendes Framework für universelle Verwendbarkeit
- Segmentierbar (Applikationen / Partitionen / Zonen)
- Exzellente Performance (niedrige Latenzen / maximale Parallelität)

OSL Virtual Storage Domains

Eckdaten

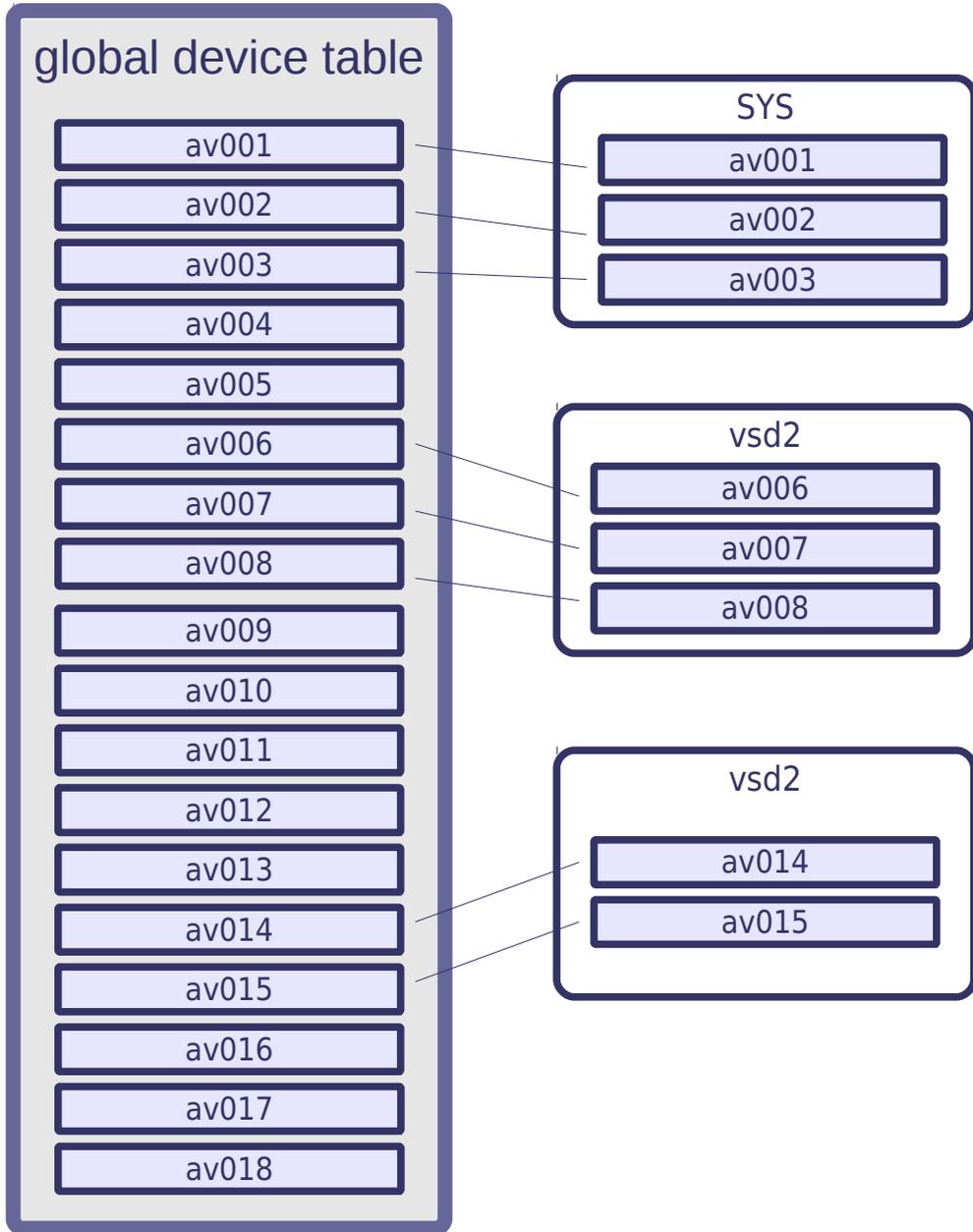


- Bis zu 256 Domains je Server
- Insgesamt bis zu 1 Mio. (Linux) bzw. 4 Mrd. Devices (Solaris) zuordenbar (effektive Begrenzung derzeit auf 65.536 je Server)

OSL VSD und OSL Storage Cluster / OSL UVE



Abbildung von Geräten



- Typisch exklusives Mapping auf eine VSD
- Mit OSL SC 4.8/4.9 zunächst nur "SYS"-Domain, weitere Domains nur experimentell
- Treiber beherrscht Zuordnung laufender Applikationen / VMs zu separaten VSDs, damit zukünftig:
 - Entlastung der globalen Zone
 - Mandanten tragen eigene I/O-Kosten!
- Physical Volumes entfallen, bzw. werden nur noch in einer Diagnose-VSD angeboten
- Globale Gerätetabelle stellt identische Gerätenummern über Rechengrenzen hinweg sicher
- Sonderoperationen (etwa atomarer Split usw. möglich)
- Teilweise Neuimplementierung XDM und Multipathing bereits für ovvemu

Die nächsten Schritte

Differenziertes Vorgehen



- Pilotierung mit 4.8 / allg. Freigabe mit 4.9
- Weiterentwicklung zunächst im Rahmen von OSL SC 4.9
- Anpassungen für OSL UVE 4.9 (All-Linux-UVE)
- Technologie-Demonstrator für Linux (single-threaded File- und Memory-Backend)
- Vorbereitungen für OSL SC und OSL UVE 5.0
- Aufnahme von RSIO 2



virtualization and clustering - made simple