



OSL Virtual Storage Domains

OSL Technologietag

Schöneiche bei Berlin - 2. November 2016

bert.miemietz@osl.eu

Heutige I/O-Treiber

Heutige I/O-Treiber

Typische Merkmale



- Design meist älter als 15 Jahre
- Laufen komplett im Kernel
- Sind kaum portabel
- Sind extrem spezialisiert
- Können Partitionen / Zonen moderner Systeme nicht abbilden (global)
- Sind fest im I/O-Stack verankert (blockiert?)
- Arbeiten mit Interrupts und Interrupt-Prioritäten

OSL RSIO

OSL RSIO - Remote Storage I/O

Erster Prototyp einer neuen Architektur



Diese Seiten wurden nur auf der Veranstaltung gezeigt und unterliegen im Übrigen einer Vertraulichkeitsvereinbarung.

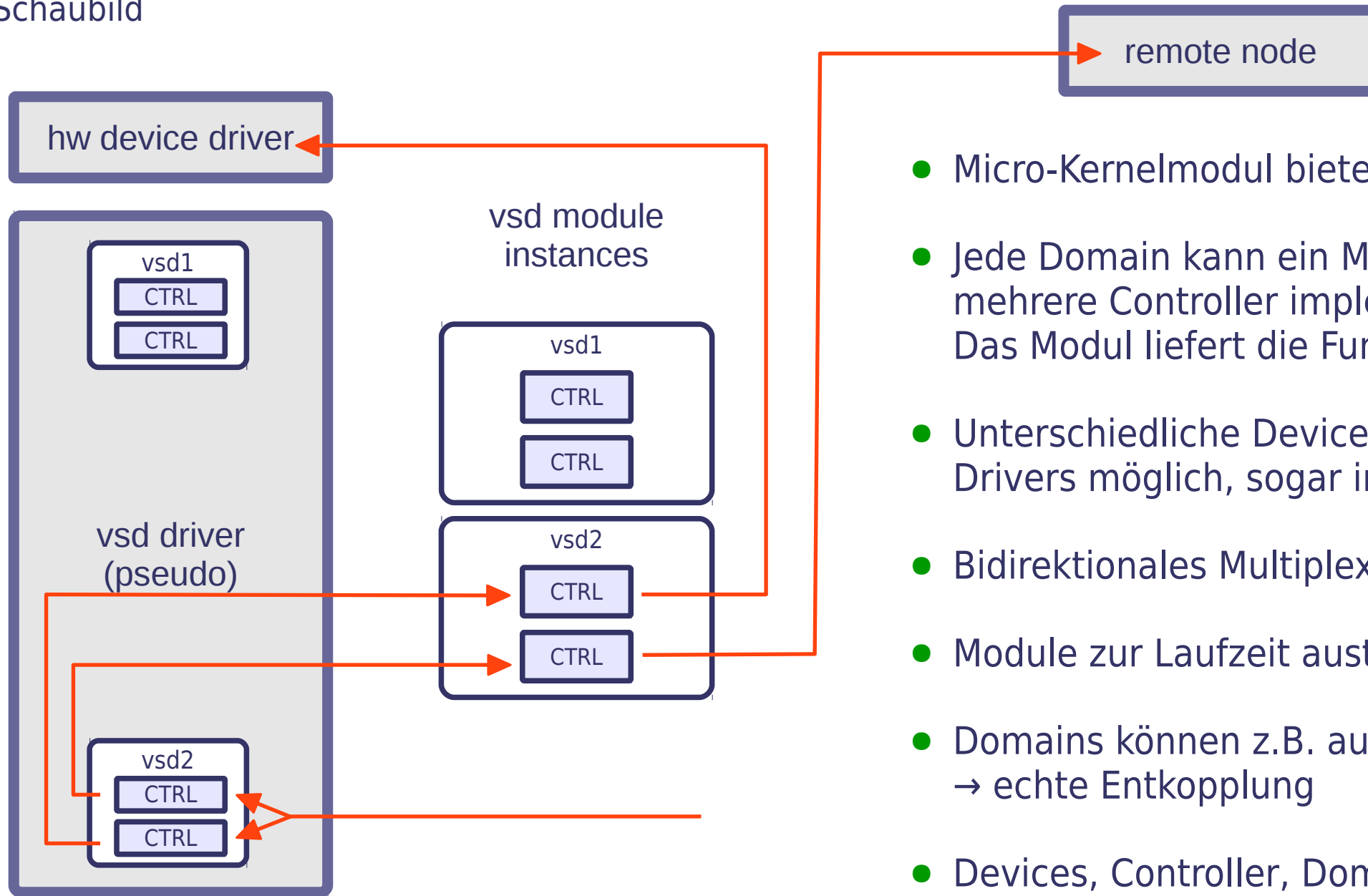
OSL Virtual Storage Domains



- Brandneue Implementierung unter Berücksichtigung der Erfahrungen mit RSIO
- Micro-Kernelmodul weiter geschrumpft
- Extrem portabel
- Universelle Verwendbarkeit
- Können Partitionen / Zonen moderner Systeme abbilden
- Exzellente Performance (ultrakurze Latenzen / maximale Parallelität)
- Pseudo-Device-Treiber ohne Interrupt-Kontext

OSL Virtual Storage Domains

Schaubild



- Micro-Kernelmodul bietet extrem einfaches Interface
- Jede Domain kann ein Modul ansprechen, das mehrere Controller implementieren kann. Das Modul liefert die Funktionalität
- Unterschiedliche Device-Typen innerhalb des VSD-Drivers möglich, sogar innerhalb einer Domain
- Bidirektionales Multiplexing des I/O
- Module zur Laufzeit austauschbar
- Domains können z.B. auf Zonen abgebildet werden → echte Entkopplung
- Devices, Controller, Domains zur Laufzeit konfigurierbar und austauschbar

OSL Virtual Storage Domains

Eckdaten

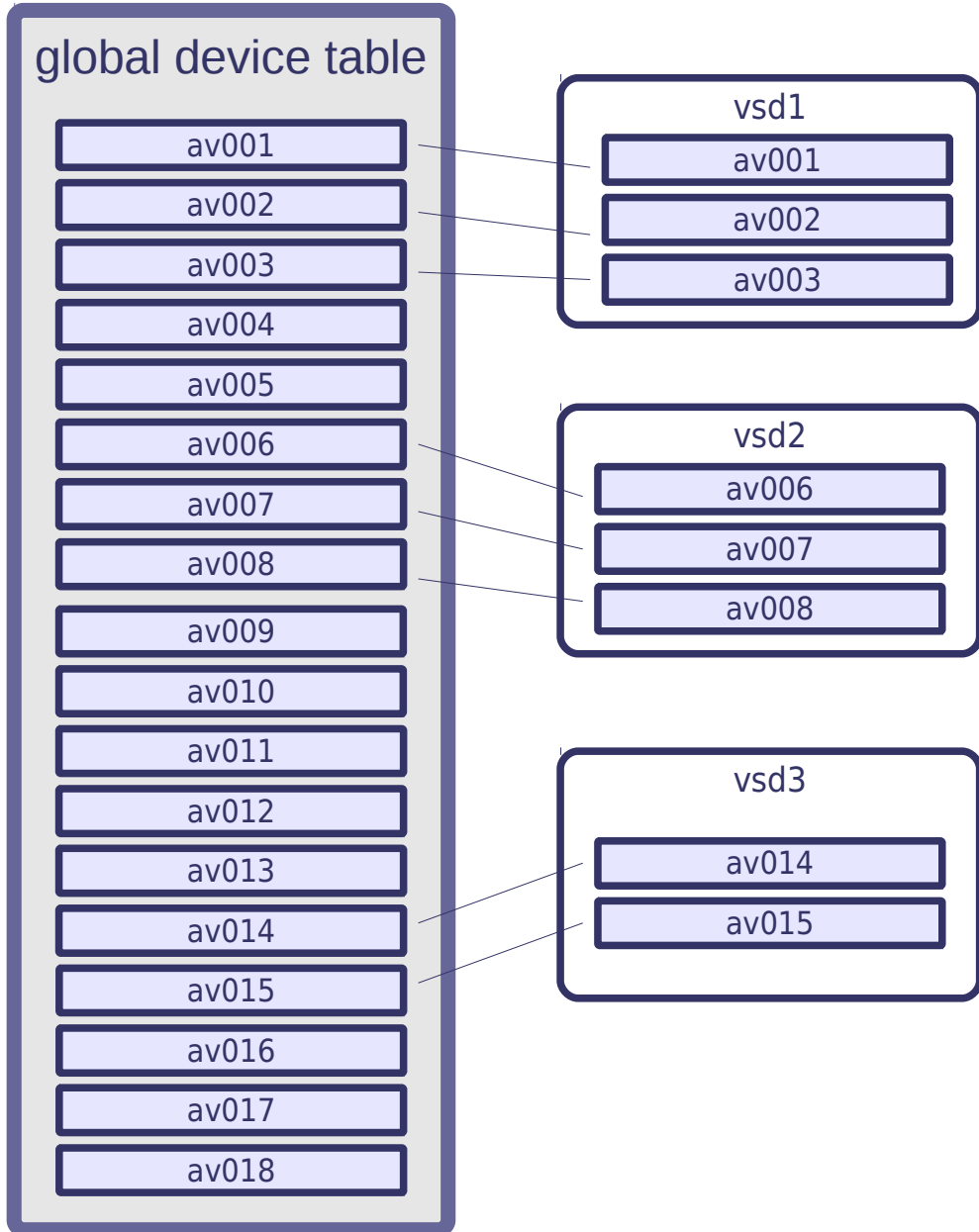


- Bis zu 256 Domains je Server möglich
- Bis zu 8 virtuelle Controller je Domain
- Insgesamt bis zu 1 Mio. (Linux) bzw. 4 Mrd. Devices (Solaris) zuordenbar
- Multiple und Exklusive Mappings möglich
- Module zur Laufzeit tauschbar
- Irreguläre Zustände / Panic auf Ebene Device, Domain und schließlich global behandelbar
- Geräteindividuelles Queue-Management, Timer usw. realisierbar

OSL VSD und OSL Storage Cluster / OSL UVE



Abbildung von Geräten



- Typisch exklusives Mapping auf eine VSD
- Jede laufende Applikation / Zone kann eine eigene VSD erhalten (vnodes!).
Damit bei Zonen 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
- Geräte werden erst beim Start der Applikation / VSD angelegt → signifikante Verkürzung Boot-Zeit
- Domainweise sind Sonderoperationen möglich (etwa atomarer Split usw.)
- Neuimplementierung XDM und Multipathing

Weitere Planungen

Differenziertes Vorgehen



- QI/2017 Preview-Release (orientiert an Anforderungen aus OSL SC OSL UVE)
- Weiterentwicklung zunächst für OSL SC 4.9
- Anpassungen für OSL UVE 4.9
- Module Closed Source
- VSD-Treiber für Linux als Open Source (nach closed beta)
- Technologie-Demonstrator für Linux
- Vorbereitungen für OSL SC und OSL UVE 5.0
- Aufnahme von RSIO 2



virtualization and clustering - made simple