

ORACLE®



ORACLE[®]

Einführung in das OCFS2 Cluster-Dateisystem für Linux

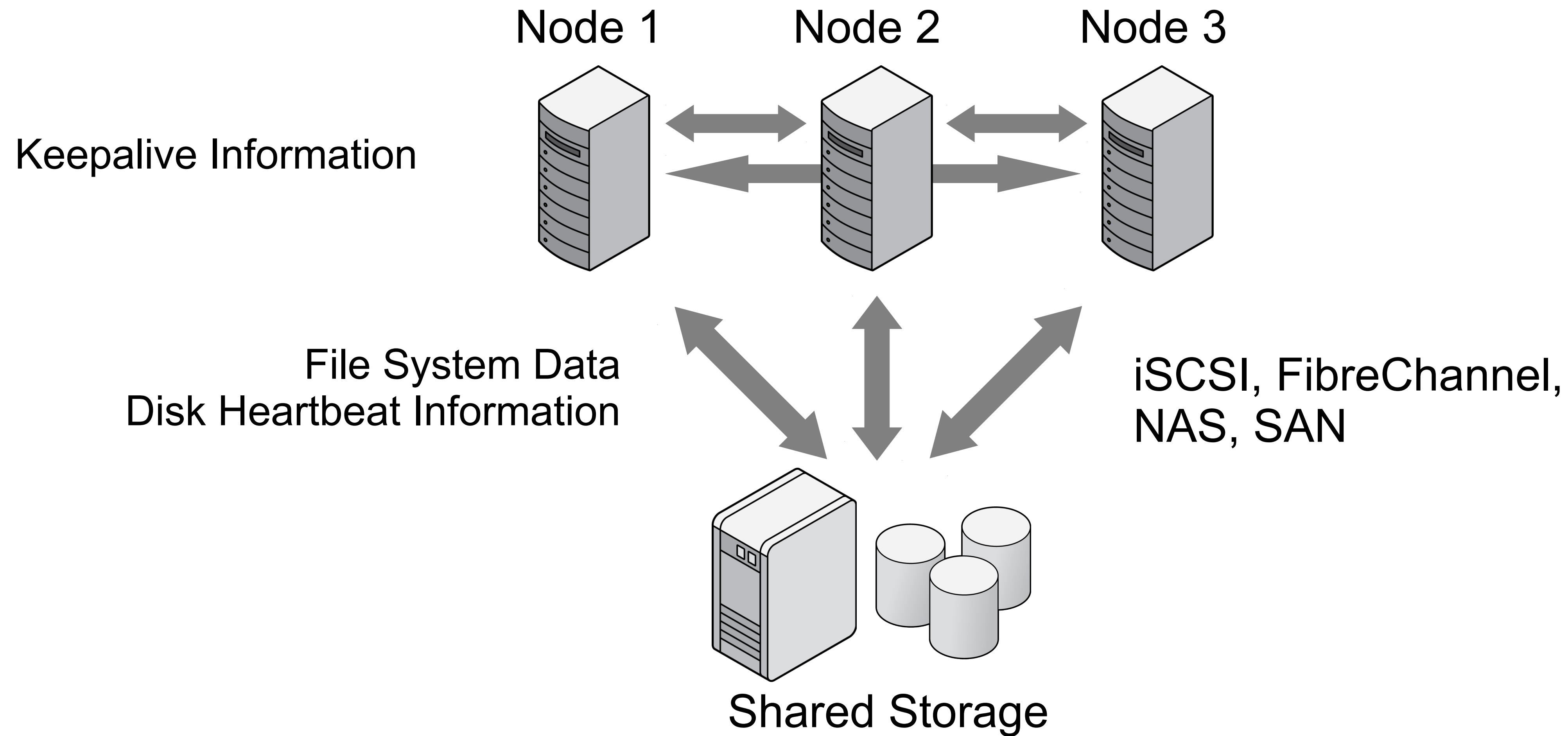
Lenz Grimmer
Senior Product Manager, Oracle Linux



OCFS2 – Überblick

- Shared-Disk Cluster-Dateisystem für Linux
- POSIX+ konform
- Natives Journaling-Dateisystem für Linux
 - 2003: Als Nachfolger von OCFS begonnen
 - Januar 2006: Aufnahme in den Linux-Kernel (2.6.16)
- Architektur- und Endian-Neutral
 - Parallele Mounts an x86, x86-64, IA64 und PPC
 - Big- und Little-Endian, 32-bit und 64-bit
- Open Source (GPL)

Shared Disk Cluster-Dateisystem



OCFS2 – Eigenschaften

- Effizient
 - Benötigt wenig Ressourcen (CPU und RAM)
 - Kompaktes Datenformat (Effiziente Speicherverwaltung)
- Hochverfügbarkeit **und** Skalierbarkeit
- Leichte Konfiguration
 - `/etc/ocfs/cluster.conf` – definiert Cluster-Layout
 - `/etc/sysconfig/o2cb` – definiert Cluster-Timeouts
- Unabhängig von Oracle Software
 - Kann für beliebige Software/Applikationen eingesetzt werden

OCFS2 – Zwei Komponenten

- Im Linux-Kernel enthalten
 - Dateisystemtreiber (`ocfs2.ko`)
 - o2cb Cluster-Stack
- Werkzeuge zur Administration
 - Auf allen gängigen Linux-Distributionen enthalten
 - Paket `ocfs2-tools`
 - Formatieren
 - Konfigurieren/Tunen
 - Einhängen (Mounten)
 - Check und Fehlersuche

OCFS2 – Features

- Indizierte Verzeichnisse
- POSIX ACLs
- Security Attributes (für SELinux, SMACK, AppArmor)
- Standard-Dateiattribute
- Erweiterte Attribute
- User und Group-Quotas
- Checksummen auf Metadaten

OCFS2 – Flexibel konfigurierbar

- Veränderbare Cluster-/Blockgröße
- Verschiedene I/O Modi
- Data Journaling Modi: Ordered (default), Writeback
- Online Resize (Dynamische Vergrößerung)
- Hinzufügen und Entfernen von Nodes zur Laufzeit
- Auswahl des Cluster-Stacks (o2cb, PaceMaker, CMAN, keiner)

OCFS2 – Speichermanagement

- Extents
- Cluster Allocations
- Sparse Files
- Unwritten Extents
 - fallocate(2)
- Hole Punching
- Inline-Daten
 - Daten im Inode
- REFLINKs
 - Snapshot/Clones von Dateien
- Allocation Reservations

OCFS2 – Verschiedene I/O-Modi

- Buffered I/O
- Asynchronous I/O
- Memory Mapped I/O
- Direct I/O*
- Splice I/O*

* Nicht Bestandteil der POSIX-Spezifikation

OCFS2 – Dateisystem-Werkzeuge

- `mkfs.ocfs2(8)` – Formatierung eines Volumes
- `tunefs.ocfs2(8)` – Tuning, Änderung von Parametern
- `fsck.ocfs2(8)` – Dateisystem-Check, Reparatur
- `mount.ocfs2(8)` – Einhängen des Dateisystems
- `o2info(8)` – Information über Dateisysteme
- `debugfs.ocfs2(8)` – Analyse der Dateisystemstruktur
- `o2image(8)` – Kopie der Metadaten (Analyse)
- `mounted.ocfs2(8)` – Erkennung aller OCFS2-Volumes

OCFS2 – Cluster Stack Werkzeuge

- `o2cb(8)` – Cluster-Konfigurationswerkzeug
 - Ersetzt `o2cb_ctl(8)`
- `ocfs2_hb_ctl(8)` – Heartbeat-Kontrolle
 - Start/Stop des lokalen Heartbeats
- `o2hbmonitor(8)` – Disk Heartbeat-Monitor
 - Überwachung und Protokollierung von Zeitüberschreitungen

o2cb – OCFS2 Cluster Stack Komponenten

- In-Kernel Cluster Stack
- o2nm – Node Manager
- o2hb – Disk Heartbeat Agent
- o2net – Network Agent
- o2dlm – Distributed Lock Manager

OCFS2 – Weitere Informationen

- OCFS2 Technical Information:
<http://www.oracle.com/us/technologies/linux/025995.htm>
- OCFS2 Project Page: <http://oss.oracle.com/projects/ocfs2/>
- Mailing Lists: <http://oss.oracle.com/projects/ocfs2/mailman/>
- User's Guide:
http://oss.oracle.com/projects/ocfs2/dist/documentation/v1.6/ocfs2-1_6-usersguide.pdf

Hardware and Software

ORACLE®

Engineered to Work Together

ORACLE®