

Vortrag: Cluster Dateisysteme

im Rahmen des Proseminars :
„Speicher- und Dateisysteme“

22.6.2012

Marcus Struwe

Themenübersicht

- 1) Wozu braucht man Cluster Dateisysteme (Cluster FS)?
- 2) Exkurs: Einige Begriffe zum Thema Cluster
- 3) Welche Funktionen bietet ein Cluster Dateisystem?
- 4) Beispiele zu Cluster Dateisystemen
- 5) Backup
- 6) Quellen

Wozu braucht man Cluster Dateisysteme?

Motivation:

- Hochleistungsrechner brauchen schnelle Speichersysteme
 - Netzwerkdateisysteme sind bei parallelem Zugriff, der Belastung der Bandbreite des Netzwerk und der Aufteilung der Daten auf die Server nicht optimal
- Vorteilhaft für große Datenmengen
 - z.B. Datenbanken, Videos, Messwerte
- Ressourcen sparen

Wozu braucht man Cluster Dateisysteme?

Motivation:

- gleichzeitiger Zugriff auf eine Datei muss koordiniert werden – Konsistenz

Beispiel:

Client A öffnet Datei x und schreibt den Wert 5 hinein.

Client B öffnet gleichzeitig Datei x und schreibt den Wert 7 hinein.

Problem: Welche Version soll beibehalten werden?

Wozu braucht man Cluster Dateisysteme?

Lernziele:

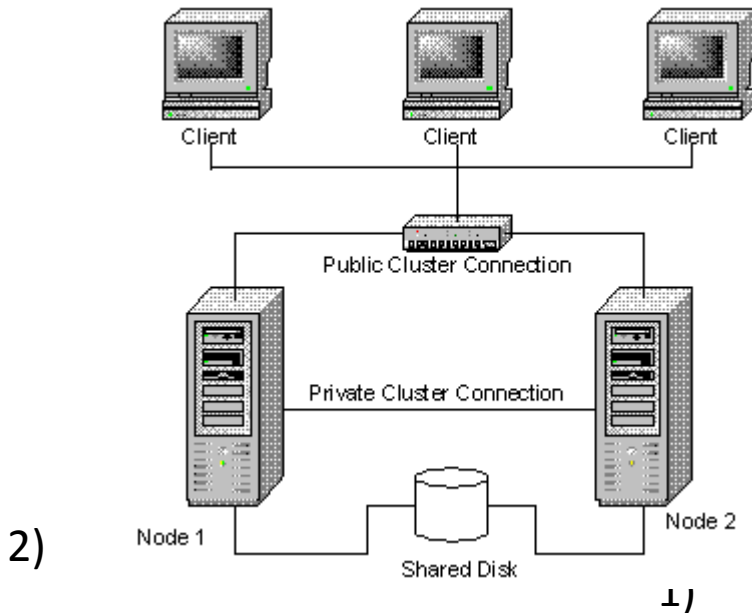
- Was sind die Kennzeichen von Cluster Dateisystemen?
 - Welche Techniken werden eingesetzt?
 - Wozu dienen diese Techniken?
- Was ist der Nutzen beim Einsatz von Cluster Dateisystemen?

Exkurs: Einige Begriffe zum Thema Cluster

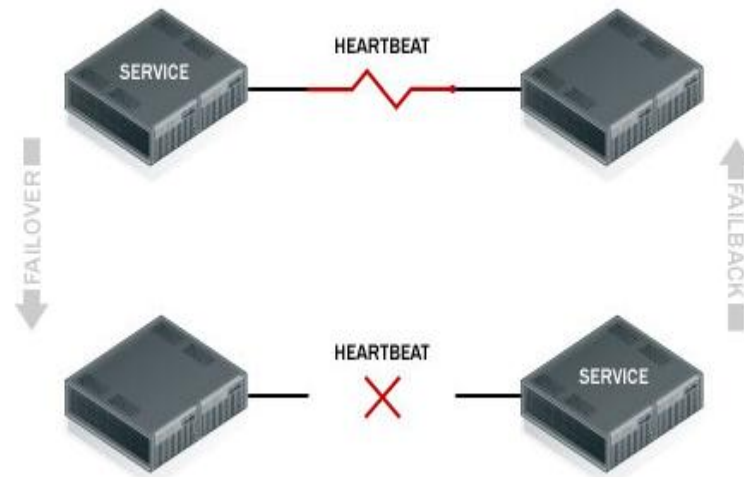
- Begriff: Cluster meint eine Anzahl von Rechnern, die über ein Netzwerk miteinander verbunden sind
- Arten von Clustern
 - Hochverfügbarkeitscluster
 - Load - balancing Cluster
 - High Performance Computing Cluster

Exkurs: Einige Begriffe zum Thema Cluster

- Node, Interconnect, split-brain, single point of failure, etc

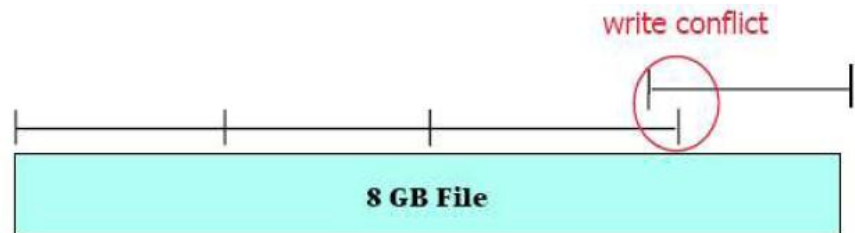


Active / Passiv



Welche Funktionen bietet ein Cluster Dateisystem?

- Locking → Konsistenz
 - Locks werden wie die Metadaten von Metadatenservern (MDS) verwaltet
 - Byterange-level, block-level und file-level locks
 - Überlappende Zugriffe → sequenziell



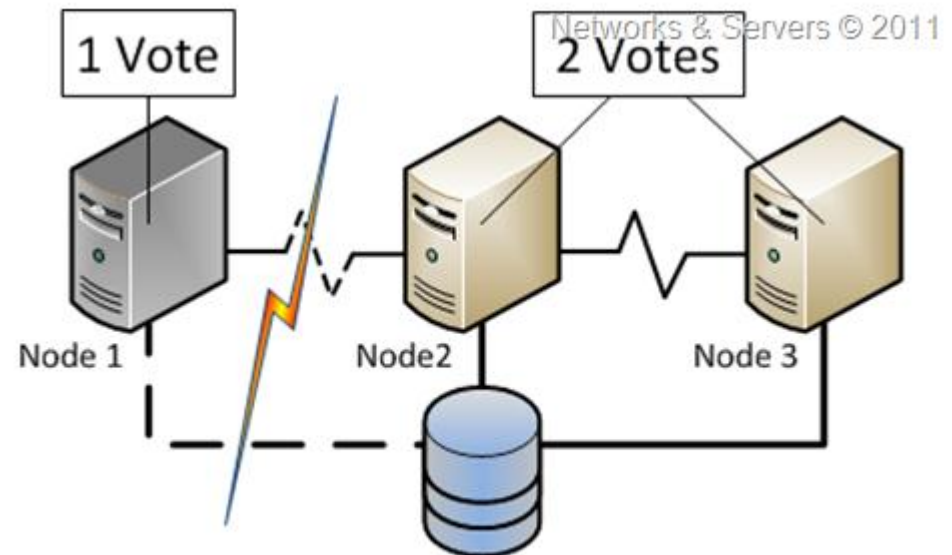
6)

node 1	node 2	node 3	node 4
locks	locks	locks	locks
offsets	offsets	offsets	offsets
0-2 GB	2-4 GB	4-6.5 GB	6-8 GB

- Performanz

Welche Funktionen bietet ein Cluster Dateisystem?

- Fencing bei shared disk Dateisystemen
 - MDS übernimmt bei einem Ausfall auch das Fencing
 - Quorum
 - Tie-breaker



3)

Majority of votes in communication:
The cluster runs

Beispiele zu Cluster Dateisystemen

Typische Cluster Dateisysteme

- RMS (OpenVMS)
- AdvFS
- Veritas Cluster File System
- OCFS2
- GPFS
- GlusterFS
- CXFS
- StorNextFS
- Polyserve Matrix Server
- Lustre
- Global File System
- MelioFS
- Ceph

Beispiele zu Cluster Dateisystemen - GPFS

GPFS – General Parallel File System

- Entwickelt von IBM
- Plattformen: AIX, Linux, Windows Server
- Nachfolger von Tiger Shark file system 1993
- Als GPFS seit 1998 auf dem Markt
- Aktuell in der Version 3.5 (Release: April 2012)

Beispiele zu Cluster Dateisystemen - GPFS

- Eingesetzt auf: ASC Purple
 - 2009 : Platz 66 in den Top500
 - 12k Prozessoren, 2 petabytes auf 11k Festplatten

4)



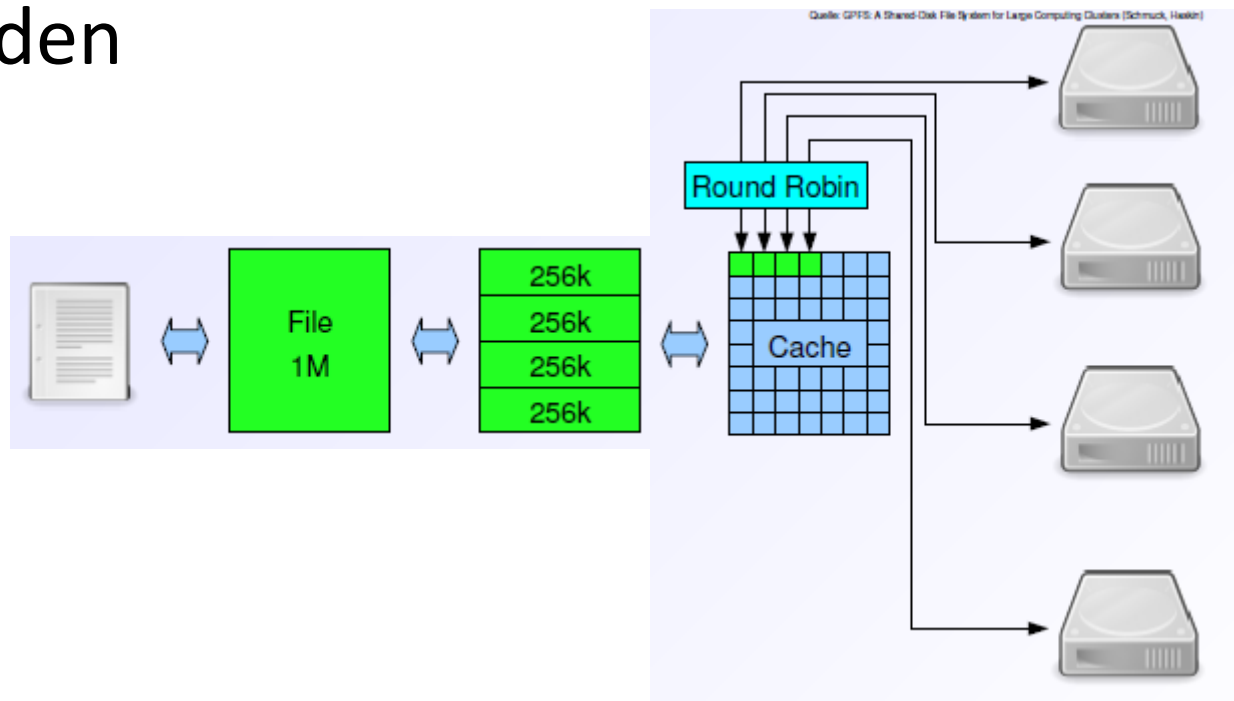
Beispiele zu Cluster Dateisystemen - GPFS

- Maximale Dateigröße 2^{99} bytes
- Maximale Anzahl von Dateien: 2 Milliarden
- Maximale Länge der Dateinamen 256 UTF-8
- In Verbindung mit Linux bis zu 2441 Nodes
- Ursprünglich für Multimedia Daten gedacht
– die meisten Kommandozeilen Befehle beginnen noch mit mm....

Beispiele zu Cluster Dateisystemen - GPFS

- Durch striping kann ein deutlicher Performancegewinn erreicht werden
- Die Art des striping kann konfiguriert werden

5)



Beispiele zu Cluster Dateisystemen - GPFS

- Striping wird für Daten und Metadaten angewandt → Ausfall einer Festplatte bringt großen Datenverlust
 - Lösung: In der Praxis werden Raidsysteme eingesetzt
 - Alternativ bietet GPFS auch von sich aus Reproduktion an (Für Daten und Metadaten werden zwei Kopien angelegt und bei Bedarf zur Wiederherstellung genutzt)

Beispiele zu Cluster Dateisystemen - GPFS

- **Vorteile:**
 - Namhafter Hersteller : IBM
 - Flexibel skalierbar
 - Durch den Einsatz in Forschungsbereichen gut an diesen Bereich angepasst
 - Sehr gute Dokumentation
- **Nachteile:**
 - Proprietäre Lizenz
 - z.T. Hardwareabhängig
 - nur auf wenigen Plattformen einsetzbar

Backup

- Was sind die Kennzeichen von Cluster Dateisystemen?
 - Locking
 - Fencing
- Was ist der Nutzen beim Einsatz von Cluster Dateisystemen?
 - Große Datenmengen gleichzeitig bearbeitbar bei Sicherstellung von konsistenten Zuständen
 - Performanz

Bildnachweis

1. http://www.happyware.de/_serverstorage/IT-Loesungen/Clusterloesungen/Clustertypen/HA-Cluster/
2. <http://technet.microsoft.com/en-us/library/bb727114.aspx>
3. <http://networksandservers.blogspot.de/2011/09/failover-clustering-iii.html>
4. https://asc.llnl.gov/computing_resources/purple/photogallery/
5. [http://www.ipvs.uni-stuttgart.de/abteilungen/sgs/lehre/lehveranstaltungen/hauptseminare/SS07/Cluster%20Computing_terminen/dateien/Verteilte Dateisysteme GPFS CXFS.pdf](http://www.ipvs.uni-stuttgart.de/abteilungen/sgs/lehre/lehveranstaltungen/hauptseminare/SS07/Cluster%20Computing_terminen/dateien/Verteilte_Dateisysteme_GPFS_CXFS.pdf)
6. http://berrendorf.inf.fh-bonn-rheinsieg.de/lehre/ss05/parsys/s_GPFS.pdf

Quellen

<http://wiki.techstories.de/display/IT/Clusterfilesystem+GPFS>

http://en.wikipedia.org/wiki/Computer_cluster

http://en.wikipedia.org/wiki/IBM_General_Parallel_File_System

<http://www.youtube.com/watch?v=2g5rx4gP6yU>

<http://de.wikipedia.org/wiki/Cluster-Dateisystem>

http://en.wikipedia.org/wiki/Clustered_file_system

<http://www.ipvs.uni->

[stuttgart.de/abteilungen/sgs/lehre/lehrveranstaltungen/hauptseminare/SS07/Cluster%20Computing_termine/dateien/Verteilte_Dateisysteme_GPFS_CXFS.pdf](http://www.ipvs.uni-stuttgart.de/abteilungen/sgs/lehre/lehrveranstaltungen/hauptseminare/SS07/Cluster%20Computing_termine/dateien/Verteilte_Dateisysteme_GPFS_CXFS.pdf)

<http://www.sourcetalk.de/2010/downloads/presentations/landhaeusser.pdf>

<http://pc2.uni->

[paderborn.de/fileadmin/pc2/media/staffweb/Andre_Brinkmann/Courses/Speichersysteme/Kapitel_13_-_Parallele_Dateisysteme.pdf](http://pc2.uni-paderborn.de/fileadmin/pc2/media/staffweb/Andre_Brinkmann/Courses/Speichersysteme/Kapitel_13_-_Parallele_Dateisysteme.pdf)

http://de.wikipedia.org/wiki/Cluster_Interconnect

<http://networksandservers.blogspot.de/2011/09/failover-clustering-iii.html>

http://berrendorf.inf.fh-bonn-rhein-sieg.de/lehre/ss05/parsys/s_GPFS.pdf

Bookman Charles, Linux Clustering, New Riders

Tzong-Jye Liu et al., Organize Metadata Servers by Using Quorum System, System Integration (SII), 2011 IEEE/SICE International Symposium on 20-22 Dec. 2011