

Ace Crngarov

Evaluierung von Cluster-Dateisystemen für den Einsatz auf Parallelrechnern

Bachelorarbeit

BEI GRIN MACHT SICH IHR WISSEN BEZAHLT



- Wir veröffentlichen Ihre Hausarbeit, Bachelor- und Masterarbeit
- Ihr eigenes eBook und Buch - weltweit in allen wichtigen Shops
- Verdienen Sie an jedem Verkauf

Jetzt bei www.GRIN.com hochladen
und kostenlos publizieren





**Fachhochschule
Bonn-Rhein-Sieg**
University of Applied Sciences

Fachbereich Informatik
Department of Computer Science

Abschlussarbeit

im Bachelor Studiengang

Evaluierung von Cluster-Dateisystemen für den Einsatz auf Parallelrechnern

von
Ace Crngarov

Eingereicht am: 07.08.2006

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Bachelorarbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Die Arbeit wurde bisher keiner Prüfungsbehörde vorgelegt und noch nicht veröffentlicht.

Sankt Augustin, 07.08.2006

(Ace Crngarov)

Zusammenfassung

Zur Ein- und Ausgabe großer Datenmengen auf Parallelrechnern und Cluster-Systemen werden spezielle Dateisysteme eingesetzt, die den parallelen Zugriff von mehreren Rechnern gleichzeitig effizient unterstützen. Beispiele für solche Cluster-Dateisysteme sind Parallel Virtual File System (PVFS / PVFS2), Oracle Cluster File System (OCFS2), Red Hat Global File System (GFS), IBM General Parallel File System (GPFS) und Lustre.

Diese Arbeit evaluiert die oben genannten Produkte hinsichtlich ihrer Effizienz und praktischen Einsetzbarkeit in einer Parallelrechnerumgebung. Zu Beginn werden die verschiedenen Cluster-Dateisysteme vorgestellt, der Schwerpunkt liegt hier auf frei verfügbaren Dateisystemen. Weiter legt diese Arbeit geeignete Effizienzkriterien und Strategien zur Beurteilung solcher Dateisysteme fest. Zu den Kriterien zählen neben einer hohen Transferrate auch eine geringe Latenzzeit bei Zugriffen auf kleine Dateien. Ein Benchmark-Programm setzt die festgelegten Kriterien und die Strategie um. Dazu wird ein neues Benchmark-Programm entwickelt, da kein verfügbares Programm alle festgelegten Kriterien untersuchen kann.

Die vorgestellte Auswahl an Cluster-Dateisystemen wird mit dem entwickelten Benchmark-Programm untersucht. Diese Arbeit präsentiert und vergleicht die gewonnenen Ergebnisse. Abschließend werden die resultierenden Leistungsdaten analysiert und darauf aufbauend Empfehlungen gegeben.

Inhaltsverzeichnis

1	Einleitung.....	1
1.1	Ziel und Motivation dieser Arbeit	1
1.2	Vorgehensweise.....	2
1.3	Abgrenzung	2
2	Dateisysteme für Parallelrechner-Cluster.....	3
2.1	Begriffsdefinitionen und Bedeutung	3
2.1.1	Cluster	3
2.1.2	Cluster-Dateisystem	4
2.1.3	Datei-intensive Anwendungen	4
2.2	Architekturen von Cluster-Dateisystemen.....	5
2.2.1	Shared Storage	5
2.2.2	Intelligente Server	5
2.3	Technische Übersicht	6
2.3.1	Parallel Virtual File System 2	6
2.3.2	Lustre.....	8
2.3.3	Oracle Cluster File System 2.....	9
2.3.4	General Parallel File System	9
2.3.5	Global File System.....	10
2.3.6	Weitere Cluster-Dateisysteme.....	11
2.3.7	Techniken zur Optimierung von Zugriffen	12
3	Kriterien zur Beurteilung von Cluster-Dateisystemen	14
3.1	Zugriffsmuster wissenschaftlicher Anwendungen	14
3.2	Festlegung der zu untersuchenden Leistungsdaten.....	16
3.3	Erhebung der festgelegten Leistungsdaten in Form von Testszenarien	17
3.3.1	Sequentieller Zugriff, gemeinsame Datei.....	18
3.3.2	Sequentieller Zugriff, unterschiedliche Dateien.....	18
3.3.3	Schrittweiser Zugriff	19
3.3.4	Metadaten.....	19
3.3.5	Pufferfähigkeit.....	20
3.3.6	Systemauslastung	20
4	Umsetzung der konzipierten Szenarien und Testvorbereitung.....	21
4.1	Benchmark-Programm zur Umsetzung der Szenarien	21
4.1.1	Anforderungen an ein Benchmark-Programm	21
4.1.2	Auswahl eines Benchmark-Programms	22
4.1.3	Entwicklung und Umsetzung eines Benchmark-Programms	23
4.2	Testvorbereitung.....	27
4.2.1	Testumgebung	27
4.2.2	Gruppierung der Rechner.....	28
4.2.3	Installation und Konfiguration	30
5	Ergebnisse der Testdurchführung.....	32
5.1	Testergebnisse der Szenarien.....	32
5.1.1	Sequentieller Zugriff, gemeinsame Datei.....	32
5.1.2	Sequentieller Zugriff, unterschiedliche Dateien.....	36
5.1.3	Schrittweiser Zugriff	42
5.1.4	Metadaten.....	49

5.1.5	Pufferfähigkeit.....	56
5.1.6	Systemauslastung.....	60
5.2	Diskussion der Ergebnisse.....	62
5.2.1	Sequentieller Zugriff, gemeinsame Datei.....	62
5.2.2	Sequentieller Zugriff, unterschiedliche Dateien.....	63
5.2.3	Schrittweiser Zugriff.....	65
5.2.4	Metadaten.....	67
5.2.5	Pufferfähigkeit.....	69
6	Fazit der Untersuchung und Handlungsempfehlungen.....	71
	Literaturverzeichnis	74
	Anhang.....	77

1 Einleitung

Die Geschwindigkeit, die Arbeitsspeichergröße und die Speicherkapazität von Parallelrechnern nehmen rasant zu, wohingegen sich die Geschwindigkeit von einzelnen Festplatten jedoch nicht in dem selben Maße erhöht [May01]. Das hat zur Konsequenz, dass Parallelrechner Daten nicht so schnell lesen und schreiben können, wie sie Daten erzeugen. Die Folge davon ist, dass die Leistungsfähigkeit von Anwendungen auf Parallelrechnern dadurch beschränkt sein kann. In Zukunft wird das Problem durch die sich immer weiter öffnende Schere zwischen Rechen- und Festplattengeschwindigkeit noch größer.

Cluster-Dateisysteme sollen diese Schere wieder zusammenführen, indem sie Daten über mehrere Knoten eines Parallelrechners verteilen. Die Knoten schreiben ihren zugewiesenen Teil der Daten auf die lokale Festplatte; die akkumulierte Leistungsfähigkeit der einzelnen Knoten lässt sich dadurch ausnutzen. Der Benutzer eines Cluster-Dateisystems erhält trotz der physikalischen Verteilung der Daten eine transparente Datensicht auf das Dateisystem, ohne zu wissen, dass die Daten physikalisch verteilt liegen [Bau06].

1.1 Ziel und Motivation dieser Arbeit

Das Ziel dieser Arbeit ist die systematische Untersuchung von verschiedenen Cluster-Dateisystemen hinsichtlich ihrer Effizienz und der praktischen Einsetzbarkeit in einer Parallelrechnerumgebung. Diese Arbeit evaluiert unterschiedliche Produkte aus dem Bereich der Cluster-Dateisysteme und vergleicht sie anhand von festgelegten Kriterien. Der Fokus liegt dabei auf Produkten, die in einer Umgebung ohne gemeinsamen Festplattenspeicher (engl. shared storage [Gro03]) funktionieren. Systematische Untersuchungen in der Form existieren dazu bislang nicht. In [Cop05] und [Mau05] finden sich zwei Studien zu Cluster-Dateisystemen, die einen Leistungsvergleich zwischen unterschiedlichen Systemen anstellen. Dies geschieht jedoch in beiden Fällen auf Grundlage einer Umgebung mit Zugriff auf einen gemeinsamen Festplattenspeicher.

Einige Hersteller von Cluster-Dateisystemen haben selber Benchmarks mit ihren eigenen Produkten durchgeführt und teilweise mit anderen Cluster-Dateisystemen verglichen ([CLUb], [IBMb]). So stellt es sich in den Ergebnissen der Hersteller dar, dass das Produkt aus dem eigenen Haus am besten abschneidet. Um einen unabhängigen Vergleich durchzuführen, sollten alle Produkte jedoch auf der selben Hardware und unter dem selben Betriebssystem getestet werden.

Diese Arbeit untersucht das Parallel Virtual File System 2 (PVFS2) [PVFS2a], das Oracle Cluster File System (OCFS) [ORAA], das Red Hat Global File System (GFS) [REDA], das IBM General Parallel File System (GPFS) [IBMa] und Lustre von Cluster File Systems [CLUa].

1.2 Vorgehensweise

Die Arbeit gliedert sich im Weiteren wie folgt. Kapitel zwei gibt einen Überblick über die wichtigsten Cluster-Dateisysteme; der Schwerpunkt liegt hier auf frei verfügbaren Dateisystemen. Des Weiteren werden geeignete Effizienzkriterien zur Beurteilung solcher Dateisysteme in einer Cluster-Umgebung festgelegt und Strategien zur Messung dieser Kriterien entwickelt (Kapitel drei). Ein entsprechendes Benchmark-Programm wird hierfür in Kapitel vier konzipiert und umgesetzt. Außerdem wird in diesem Kapitel die Testumgebung vorgestellt. Kapitel fünf präsentiert und diskutiert die Ergebnisse des Benchmark-Programms. Darauf aufbauend gibt das Fazit der Arbeit Empfehlungen für den Einsatz der verschiedenen Produkte.

1.3 Abgrenzung

Diese Arbeit untersucht nicht die speziellen Funktionen, die jedes Produkt mit sich bringt, wie z. B. die Replikation bei dem Cluster-Dateisystem GPFS. Da diese Funktionen nicht in allen Produkten enthalten sind, können sie deswegen nicht miteinander verglichen werden.

Ferner nennt diese Arbeit zwar die speziellen Schnittstellen der Cluster-Dateisysteme (siehe Kapitel 2.3.7), die eine Optimierung von Zugriffen ermöglichen, untersucht sie aber nicht weiter, da nicht jedes Cluster-Dateisystem diese Schnittstellen unterstützt.

Auch stellt diese Arbeit keine Installationsanleitung für die Cluster-Dateisysteme dar. Kapitel 4.2.3 gibt eine kurze Darstellung über die verwendeten Parameter bei der Testdurchführung.