Fachbereich Technik
Abteilung Elektrotechnik und Informatik

Bachelor-Thesis

Effiziente Speicherung und Verteilung virtueller Festplatten mit aktueller Open-Source Software

Bastian de Groot

13. Januar 2011

Prüfer	Prof. Dr. Jörg Thomaschewski
Zweitprüfer	Dr. Arvid Requate

Erklärung

Soweit meine Rechte berührt sind, erkläre ich mich einverstanden, dass die Bachelor-Arbeit Angehörigen der Hochschule Emden/Leer für Studium / Lehre / Forschung uneingeschränkt zugänglich gemacht werden kann.

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Bachelor-Arbeit bis auf die offizielle Betreuung selbst und ohne fremde Hilfe angefertigt habe und die benutzten Quellen und Hilfsmittel vollständig angegeben sind.

Bremen, 13. Januar 2011

Bastian de Groot

Danksagung

Zunächst geht mein Dank an das Unternehmen **Univention**, das mich unterstützt hat und mir während der Erstellung der Bachelor-Thesis Virtualisierungshardware zur Verfügung stellte.

Besonders danke ich allen Mitarbeitern von Univention für die sehr gute Zusammenarbeit und das angenehme Arbeitsklima. Hierbei möchte ich mich besonders bei meinem Betreuer **Arvid Requate** bedanken, für die hervorragende Unterstützung während dieser Arbeit. Ebenfalls danke ich **Phillip Hahn** für die erkenntnisreiche Unterstützung in vielen fachlichen Fragen. Auch danke ich **Janek Walkenhorst**, mit dem ich mich nicht nur fachlich über viele Dinge ausgetauscht habe, sondern auch privat einen sehr guten Kontakt pflegte.

Nicht zuletzt gilt mein tiefer Dank **meinen Eltern**, die mir mein Studium durch Ihre finanzielle Unterstützung ermöglicht haben und mir auch jederzeit in allen anderen Belangen mit Rat und Tat zur Seite standen.

Es ist nicht Deine Schuld, dass die Welt ist, wie sie ist.
Es wär nur Deine Schuld, wenn sie so bleibt.
- Die Ärzte - Deine Schuld -

Inhaltsverzeichnis

1. Einleitung

Virtualisierung ist heute ein sehr wichtiger Begriff in der Informatik. Die Virtualisierung verteilt die vorhandenen Ressourcen effizient und hilft somit Prozesse zu optimieren und Kosten zu senken. Es gibt unterschiedliche Formen von Virtualisierung wie zum Beispiel die Anwendungsvirtualisierung und die Betriebsystemvirtualisierung. Diese Arbeit beschäftigt sich mit einem Aspekt der Betriebsystemvirtualisierung.

Von "Betriebsystemvirtualisierung" spricht man, wenn sich mehrere virtuelle Betriebsysteminstanzen Hardwareressourcen wie CPU, RAM oder Festplatten teilen. Der Virtualisierungskern (Hypervisor) stellt den virtuellen Betriebsysteminstanzen eine in Software und Hardware realisierte Umgebung zur Verfügung, die für die darin laufenden Instanzen kaum von einer echten Hardwareumgebung unterscheidbar sind [Prz] [Bau08]. Es gibt unterschiedliche technische Ansätze der Virtualisierung, wie Paravirtualisierung oder Vollvirtualisierung. Diese Kategorisierung bezieht sich darauf, wie der Hypervisor die vorhandene Hardware für die virtuelle Instanz bereitstellt. Auf diesem Gebiet gibt es eine sehr aktive Entwicklung.

Wenig beachtet bei der Entwicklung von Virtualisierungssoftware ist jedoch die Speicherung von virtuellen Festplatten. In dieser Arbeit wird dieser Punkt aufgegriffen und die Möglichkeit der Optimierung mit der Copy-on-Write Strategie beleuchtet.

Copy-on-Write ist eine Optimierungsstrategie, die dazu dient unnötiges Kopieren zu vermeiden. Diese Strategie wird vom Linux-Kernel genutzt um Arbeitsspeicher einzusparen. Aber auch bei der Desktopvirtualisierung wird Copy-on-Write eingesetzt, um die benötigte Zeit für die Bereitstellung einer geklonten virtuellen Maschine zu minimieren. Hierbei wird nicht für jeden Benutzer ein eigenes Image kopiert, sondern alle Benutzer verwenden ein Master-Image. Falls ein Benutzer Änderungen an diesem Master-Image vornimmt, werden die Änderungen separat abgespeichert.

1.1. Zieldefinition

Ziel dieser Arbeit ist es, Möglichkeiten zur effizienten Speicherung von virtuellen Festplatten aufzuzeigen. Hierbei wird ausschließlich auf bestehende Open Source Lösungen zurückgegriffen (siehe Kapitel 1.3). Die freien Open Source Lösungen werden miteinander verglichen und eine effiziente Lösung herausgearbeitet. Außerdem wird betrachtet, wie die für das Copy-on-Write benötigten Master-Images im Netzwerk effizient verteilt werden können.

1.2. Vorgehen und Kurzzusammenfassung

Zunächst werden die vorhandenen Softwarelösungen für Copy-on-Write und für die Verteilung der Master-Images erläutert. Danach werden diese anhand verschiedener anwendungsrelevanter Kriterien miteinander verglichen. Nachdem die besten Lösungen beider Kategorien gefunden wurden, werden Softwaretools erstellt, die die Nutzung der gefundenen Lösung ohne tiefgreifende Vorkenntnisse ermöglicht.

1.3. Anmerkung zur Verwendung von Open-Source

Für die Verwendung von Open-Source gibt es mehrere Gründe. Führende Hersteller von Virtualisierungssoftware wie zum Beispiel Citrix bieten in vielen Bereichen Lösungen an, die auf Open-Source-Technologien basieren. Auch zu beachten ist, dass der finanzielle Rahmen dieser Arbeit den Einsatz von proprietärer Software nicht ermöglicht. Der letzte wichtige Grund sind Lizenzprobleme bei proprietärer Software. Sie unterbinden zum Beispiel das Veröffentlichen von Performance-Tests oder die Distribution mit selbst erstellter Software [Vmw].

1.4. Anmerkung zu den verwendeten Literaturquellen

Diese Arbeit bezieht sich neben den herkömmlichen Literaturquellen auch auf Mailinglisten- und Forenbeiträge, sowie Blogeinträge.

Bei Quellenangaben im Bereich der Open Source Software gibt es einige Punkte die zu beachten sind. Es gibt keine einheitliche Dokumentation der Software. Häufig sind die Informationen nicht an einer zentralen Stelle vereint, sondern liegen verstreut im Internet in Foren, Blogs, Mailinglisten oder auch in Manpages und den Quelltexten selbst. Die Relevanz und die Richtigkeit einer solcher Quellen ist schwer zu bewerten, da Blogs, Mailinglisten und Foren keinen Beschränkungen unterliegen.

Die oben genannte Verstreuung birgt, neben der schwierigen Bewertbarkeit der Richtigkeit und Relevanz, ein weiteres Problem. Da sehr viele Autoren zu einem Thema etwas schreiben, werden unterschiedliche Begriffe synonym verwendet oder sind mehrdeutig.

Alle Quellen sind mit der zu Grunde liegenden Erfahrung des Autors dieser Arbeit ausgewählt und überprüft, können aber aus den oben genannten Gründen keine absolute Richtigkeit für sich beanspruchen.

2. Analyse Copy-on-Write

Für das Erstellen mehrerer gleichartiger virtueller Maschinen benötigt man mehrere virtuelle Festplatten. Das kann man auf herkömmliche Art und Weise lösen, in dem ein vorhandenes Festplattenimage N mal kopiert wird. Durch das häufige Kopieren entstehen allerdings große Mengen an Daten. Außerdem benötigt es viel Zeit Festplattenimages zu kopieren. Um diesen beiden Problemen entgegen zu wirken werden Copy-on-Write-Strategien eingesetzt.

Die Copy-on-Write-Strategie wird von Unix-artigen Betriebssystemen verwendet, um Arbeitsspeicher einzusparen. Sie wird eingesetzt, um nicht den ganzen Speicherbereich eines "geforkten" Prozesses kopieren zu müssen [CB05]. Die Vorteile der Optimierungsstrategie zeigen sich jedoch auch bei der Speicherung virtueller Festplatten.

Wie in Abbildung 2.1 schematisch dargestellt ist, wird bei Copy-on-Write nicht das gesamte Image kopiert. Es werden in dem Copy-on-Write-Image nur die Veränderungen gegenüber dem so genannten Master- oder Quellimage gespeichert. Für die Platzersparnis werden Sparse-Dateien genutzt, welche im Folgenden erklärt werden. Außerdem werden die unterschiedlichen Verfahren zur Verwendung von Copy-on-Write erläutert und analysiert.

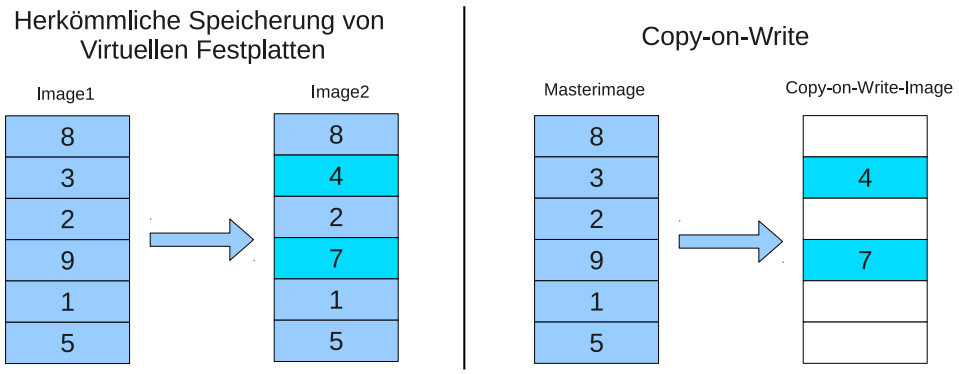

Abbildung 2.1.: Copy-on-Write

2.1. Sparse-Dateien

Eine Sparse-Datei ist eine Datei, die nicht vom Anfang bis zum Ende beschrieben ist. Sie enthält also Lücken. Um Speicherplatz zu sparen, werden diese Lücken bei Sparse-Dateien nicht auf den Datenträger geschrieben. Die Abbildung 2.2 zeigt, dass der tatsächlich benutzte Speicherplatz auf der Festplatte weitaus geringer sein kann als die eigentliche Dateigröße [Lar10].

Eine Sparse-Datei ist kein eigenes Imageformat sondern eine Optimierungsstrategie. Sie verhilft Copy-on-Write-Images zu einer großen Platzersparnis. In Imageformaten wie qcow2 oder VHD ist diese Optimierungsstrategie ein fester Bestandteil.

2.2. Imageformat qcow2

Das Imageformat qcow2 ist im Rahmen des qemu Projekts entwickelt wurde [Qem10]. Es ist der Nachfolger des ebenfalls aus dem qemu Projekt stammenden Formats qcow [McL08].

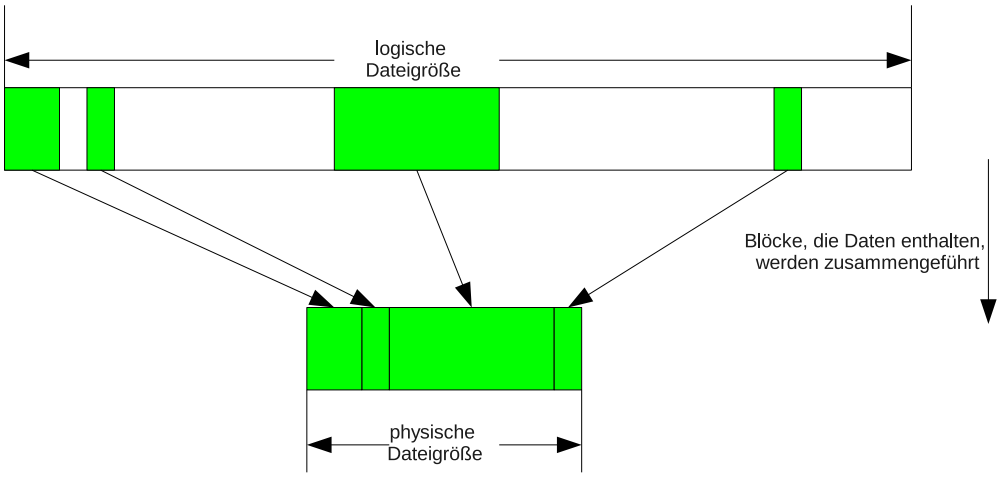

Abbildung 2.2.: Sparse-Datei

Vorteile

- Einfache Einrichtung
- Aktive Entwicklung im Rahmen der Projekte KVM und qemu

Nachteil

- aktuell fehlende Unterstützung durch Xen und andere offene Virtualisierungstechniken (z.B. VirtualBox)

2.3. Imageformat VHD

Das Format VHD ist von Conectix und Microsoft entwickelt worden. Die Spezifikation des Imageformats wurde von Microsoft im Zuge des “Microsoft Open Specification Promise” freigegeben [Mso06] [Vhd09]. Seit der Freigabe der Spezifikation bieten einige Open Source Virtualisierungslö-

sungen wie qemu, Xen oder VirtualBox die Möglichkeit dieses Format zu verwenden.

Vorteile

- Einfache Einrichtung
- Unterstützung durch Softwarehersteller mit hoher Marktakzeptanz

Nachteile

- Weiterentwicklung scheint derzeit fragwürdig
- Verwendung der Copy-on-Write-Funktion von VHD mit KVM derzeit nicht möglich

2.4. dm-Snapshots

Die dm-Snapshots sind eine Funktion des Device Mappers. Device Mapper ist ein Treiber im Linux-Kernel. Er erstellt virtuelle Gerätedateien, die mit bestimmten erweiterten Features wie zum Beispiel Verschlüsselung ausgestattet sind [Bro]. Bei dm-Snapshots wird eine solche virtuelle Gerätedatei erstellt, die aus zwei anderen Gerätedateien zusammengesetzt wird. Die erste Gerätedatei ist der Ausgangspunkt, wenn daran Änderungen vorgenommen werden, werden sie als Differenz in der zweiten Gerätedatei gespeichert [Dmk].

Die von Device Mapper erstellten Gerätedateien benötigen keine Unterstützung der Virtualisierungstechnik, da sie für diese nicht von physikalischen Festplattenpartitionen unterscheidbar sind. Dieses ist nicht nur ein Vorteil, sondern zugleich auch ein Nachteil. Es muss immer vor dem

Starten einer virtuellen Maschine das Copy-on-Write-Image und das Masterimage zu einer Gerätedatei verbunden werden.

Vorteile

- Hohes Entwicklungsstadium
- Gesicherte Weiterentwicklung
- Unabhängig von Virtualisierungstechnik

Nachteile

- Aufwendige Einrichtung
- Erfordert zusätzlichen Programmstart vor dem VM-Start

2.5. LVM-Snapshots

LVM-Snapshots sind ein Teil des Logical Volume Managers. LVM ist eine Software-Schicht die über den eigentlichen Hardware-Festplatten einzuordnen ist [Hof] [Lvm06]. Sie basiert auf Device Mapper [Lvm]. LVM ermöglicht das Anlegen von virtuelle Partitionen (logical volumes). Diese können sich über mehrere Festplatten-Partitionen erstrecken und Funktionen wie Copy-on-Write bereitstellen.

Vorteile

- Hohes Entwicklungsstadium
- Gesicherte Weiterentwicklung
- Unabhängig von Virtualisierungstechnik

Nachteile

- Aufwendige Einrichtung
- Live-Migration nicht möglich
- Nutzung von Sparse-Dateien schwer umsetzbar

2.6. Benchmarks

Ein wichtiger Punkt für die Entscheidung welche Copy-on-Write Implementierung optimal ist, ist die Lese- und Schreibgeschwindigkeit. Hierbei gibt es zwei Zugriffsarten, den sequentiellen und den wahlfreien oder auch zufälligen Zugriff.

2.6.1. Testbedingungen

Das Hostsystem für die Performance-Tests hat einen AMD Athlon II X2 250 Prozessor und 4 GiB RAM. Als Betriebssystem kommt sowohl auf Host- als auch Gastsystem ein 64 bit Debian squeeze zum Einsatz. Bei den KVM-Tests ist 2.6.32-5-amd64 der eingesetzte Kernel, für Xen wird der gleiche Kernel mit Xen-Unterstützung verwendet.

Während der Performance-Tests laufen neben der Virtuellen Maschine auf dem Hostsystem keine anderen aktiven Programme, die das Ergebnis verfälschen könnten. Als Referenz zu den Copy-on-Write-Techniken dient eine echte Festplattenpartition. Zum Testen der Performance werden IOzone und Bonnie++ eingesetzt.

IOzone

IOzone ist ein Tool mit dem in einer Reihe von unterschiedlichen Tests die Lese- und Schreib-Geschwindigkeit überprüft werden kann. Es wird hier zur Überprüfung der sequentiellen Lese- und Schreibgeschwindigkeit verwendet.

Bonnie++

Bonnie++ dient wie IOzone als Tool zum Testen von Festplatten. Es wird hier zur Überprüfung der sequentiellen Lese- und Schreibgeschwindigkeit sowie zum Testen des wahlfreien Zugriffs verwendet.

2.6.2. Testergebnisse

Die Testergebnisse werden in diesem Kapitel zusammenfassend aufgeführt und analysiert. Die kompletten Testergebnisse befinden sich im Anhang.

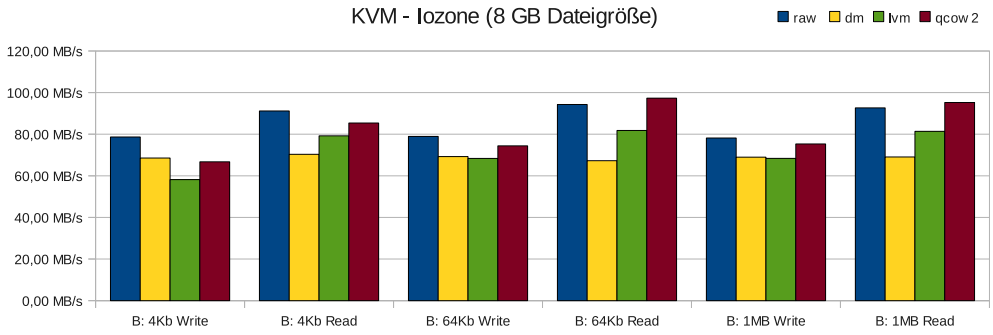

Abbildung 2.3.: Performance-Testergebnisse von Iozone für KVM mit der Dateigröße 8gb

Die Abbildung 2.3 zeigt, dass mit KVM qcow2 gegenüber den anderen Copy-on-Write-Techniken einen Geschwindigkeitsvorteil beim sequentiellen Lesen und Schreiben hat. Insbesondere bei großen Blockgrößen zeigt

sich dieser Vorteil. LVM-Snapshots und dm-Snapshots liegen hingegen ungefähr gleich auf.

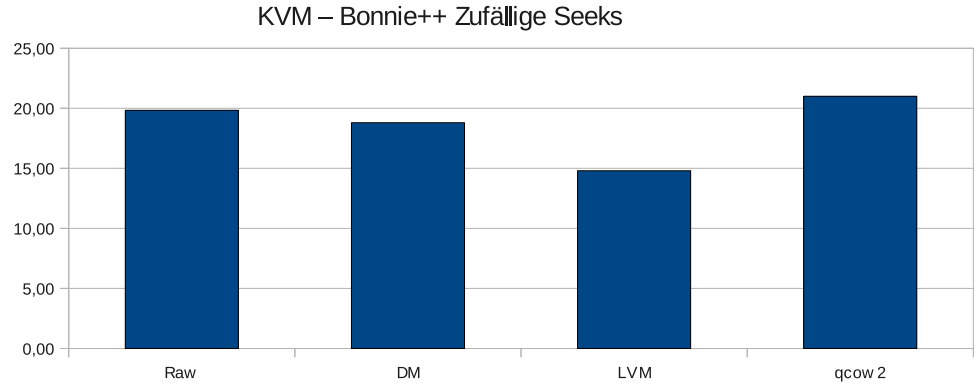

Abbildung 2.4.: Performance-Testergebnisse von bonnie++ für KVM

Abbildung 2.4 ist zu entnehmen, dass qcow2 wie auch bei den sequentiellen Tests vor LVM-Snapshots und dm-Snapshots liegt. Der Unterschied zu der echten Festplattenpartition ist in beiden Tests sehr gering. Die guten Werte von qcow2, sowohl beim sequentiellen als auch beim zufälligem Zugriff auf die Festplatte, hängen mit der guten Integration in KVM zusammen (siehe Kapitel 2.7.1).

Beim sequentiellen Lesen und Schreiben in Xen schneiden die dm-Snapshots besser ab als LVM-Snapshots und VHD , wie in Abbildung 2.5 zu sehen ist.

Bei zufälligem Zugriff auf die Festplatte unter Xen ist VHD langsamer als LVM-Snapshots und dm-Snapshots. Die LVM-Snapshots und dm-Snapshots haben eine ähnliche Geschwindigkeit und keine signifikanten Nachteile gegenüber der Festplattenpartition (siehe Abbildung 2.6). Trotz der von Citrix für Xen eigens entwickelten VHD-Unterstützung, hat VHD nur ein mittelmäßiges Ergebnis erzielt [Cro09].

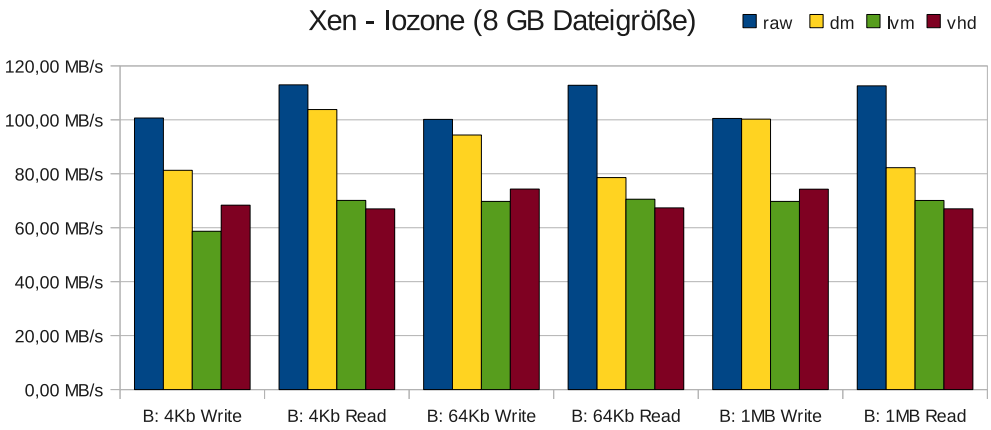

Abbildung 2.5.: Performance-Testergebnisse von Iozone für Xen mit der Dateigröße 8gb

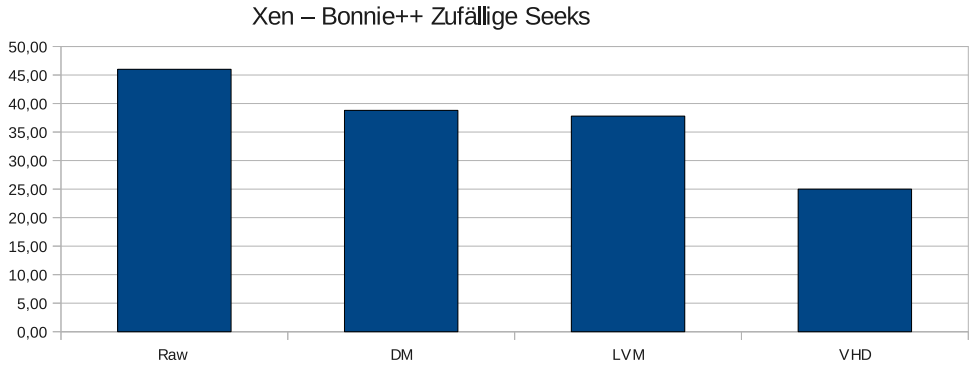

Abbildung 2.6.: Performance-Testergebnisse von bonnie++ Xen

Die Testergebnisse zeigen, dass es Geschwindigkeitsunterschiede zwischen den Copy-on-Write-Techniken gibt. Diese Unterschiede in der Geschwindigkeit sind aber nicht so groß, dass man einzelne Copy-on-Write-Lösungen aufgrund der Performance-Tests kategorisch ausschließen müsste. Dennoch sind besonders die Vorteile von qcow2 in Verbindung mit KVM zu erwähnen. Für Xen gibt es zur Zeit kein Image-Format (siehe Abbildung 2.5), dass ähnliche Testergebnisse wie qcow2 in Verbindung mit KVM (siehe Abbildung 2.4) vorweisen kann.

2.7. Fazit

Es gibt bei den Testergebnissen keine eindeutige Empfehlung für eine Copy-on-Write-Technik aufgrund der Geschwindigkeit. Im Großen und Ganzen fallen bei den Ergebnissen unter den einzelnen Copy-on-Write-Verfahren Unterschiede auf, sie lassen jedoch keine eindeutige Entscheidung zu.

Aufgrund der unterschiedlichen Implementierungen der Copy-on-Write-Techniken in KVM und Xen, wird auch für die beiden Virtualisierungslösungen ein jeweiliges Fazit gezogen.

2.7.1. KVM

Unter KVM gibt es die Alternativen dm-Snapshots, LVM-Snapshots oder qcow2. Das von Microsoft entwickelte VHD kommt nicht in Frage, da KVM zwar das VHD-Format unterstützt, aber nicht die Copy-on-Write-Funktion des Formats.

Die effizienteste Lösung für Copy-on-Write mit KVM ist qcow2. Dafür gibt es mehrere Gründe. Das qcow2-Format ist Teil des qemu-Projekts

und damit sehr gut in dem darauf basierendem KVM integriert. Durch die gute Integration werden sehr gute Performance-Werte erreicht. Außerdem lässt es sich im Gegensatz zu dm-Snapshots und LVM-Snapshots leichter administrieren.

2.7.2. Xen

Die für Xen zur Verfügung stehenden Copy-on-Write-Formate sind dm-Snapshots, LVM-Snapshots und VHD. Xen unterstützte in einigen vergangenen Versionen qcow2, diese Unterstützung ist jedoch nicht in der aktuellen Version 4.0.1 enthalten [Qco10].

Für Xen ist VHD aktuell die attraktivste Copy-on-Write-Lösung. Es ist zwar laut der Performance-Tests nicht die schnellste Lösung, hat aber wesentliche Vorteile gegenüber dm-Snapshots und LVM-Snapshots. Es werden keine Änderungen am Xen-Quelltext benötigt, wie es bei dm-Snapshots der Fall ist [Rac10]. Die Funktion der Live-Migration ist mit VHD leichter zu realisieren als mit LVM-Snapshots und dm-Snapshots. Die im weiteren Verlauf dieser Arbeit verwendete Lösung ist VHD. Falls Xen in den nächsten Versionen wieder qcow2 unterstützt, sollte jedoch dessen Verwendung auch unter Xen geprüft werden.

3. Analyse Verteilung von Images

Der Copy-on-Write-Mechanismus benötigt immer eine Vorlage - das Masterimage. Um es auf mehreren Virtualisierungsservern nutzen zu können, muss es über das Netzwerk verteilt werden oder über ein gemeinsam genutztes Storage-Backend zur Verfügung gestellt werden. Dieses Kapitel soll Wege aufzeigen, diese Verteilung oder Bereitstellung möglichst effizient vorzunehmen.

Die Verteilungslösungen werden darauf überprüft, wie störanfällig sie sind. Ein anderer Punkt für die Entscheidungsfindung ist die benötigte Dauer der Verteilung. Außerdem wird einbezogen, wie skalierbar die Lösungen sind.

3.1. Multicast

Ein Multicast ist eine Mehrpunktverbindung. Der Sender schickt die Daten gleichzeitig an mehrere Empfänger. Durch das einmalige Senden an mehrere Empfänger wird Bandbreite eingespart. Die Daten werden nur an Rechner im Netz versendet, die diese auch angefordert haben, wie in Abbildung 3.1 schematisch dargestellt. Die Ausnahme bilden Switches, die Multicasting nicht unterstützen. Sie versenden die gesendeten Daten an alle damit verbundenen Netzwerkknoten [Mul].

Da es bei den Masterimages darauf ankommt, dass sie komplett und fehlerfrei dupliziert werden, kann der Sender maximal so schnell senden, wie es der langsamste Empfänger entgegen nehmen kann. Dadurch ist die Verwendung von Multicast, in einer heterogenen Umgebung mit einem langsamen oder weit entfernten Empfänger, sehr ineffizient. Anwendung findet Multicast heute vor allem bei der Verteilung von Multimediadaten [Lei00].

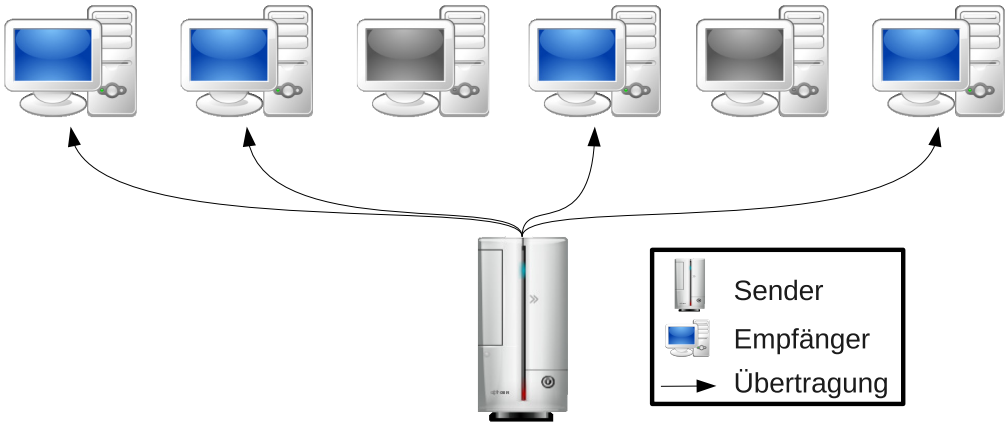

Abbildung 3.1.: Multicast Beispiel

Vorteil

- Sehr hohe Geschwindigkeit durch Parallelität

Nachteile

- Hohe Netzwerklast während der Verteilung
- Geschwindigkeitseinbruch bei heterogener Umgebung oder schlechten Netzanbindungen

3.2. Netzwerkprotokoll BitTorrent

BitTorrent ist ein Netzwerkprotokoll zum effizienten Verteilen großer Dateien oder von Sammlungen großer Dateien. Die Empfänger der Daten sind hierbei gleichzeitig auch Sender, sie werden Peers genannt [Coh08]. Damit wird nicht ein einziger zentraler Sender ausgelastet, sondern die Last wird auch auf alle Empfänger verteilt (zu sehen in Abbildung 3.2). Für die Kontaktaufnahme der Peers untereinander werden sogenannte Tracker eingesetzt. Es wird mindestens ein Tracker für die Kontaktaufnahme benötigt. Aktuellere BitTorrent-Clients können aber auch trackerlos über eine verteilte Hashtabelle (engl. “Distributed Hash Table”; DHT) andere Peers finden [Loe08]. Durch den Einsatz von DHT kann die Einrichtung eines Trackers eingespart werden. Außerdem bringt es zusätzliche Ausfallsichertheit, da die Liste der verfügbaren Peers dezentral gespeichert wird.

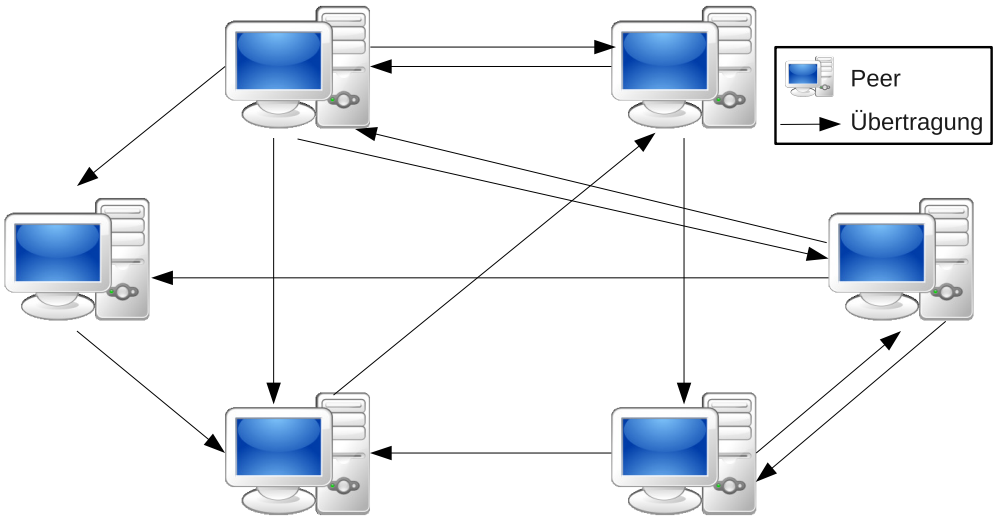

Abbildung 3.2.: Bittorrent Beispiel

Die zu übertragenden Daten werden nicht komplett in einem Stück übermittelt, sondern in Blöcke aufgeteilt. Bei zwischenzeitlichen Netzausfällen müssen somit auch nicht alle Daten noch einmal übertragen werden. Der BitTorrent-Client setzt nach dem Netzwerkausfall die Datenübertragung problemlos fort und muss nur gegebenenfalls die bereits übertragenen Daten einen Blockes verwerfen.

Vorteile

- Hohe Skalierbarkeit
- Netzwerklast auf teilnehmende Netzwerksegmente beschränkt
- Sehr effizient auch in heterogenen Umgebungen

Nachteile

- verringerte Geschwindigkeit bei asymmetrischer Upload- und Downloadgeschwindigkeit

3.3. Netzwerkprotokoll NFS

NFS (Network File System) ist ein Protokoll für das Bereitstellen von Daten über das Netzwerk. Im Gegensatz zu den beiden vorher genannten Technologien stellt das Bereitstellen einen großen Unterschied dar. Die Daten werden nicht von einem Rechner auf den anderen kopiert, sondern über das Netzwerk wie eine lokale Festplatte zur Verfügung gestellt [Nfs03]. Der Server macht hierbei eine Freigabe, die von dem Clientrechner "gemountet" wird [Smi06]. Die vom Clientrechner gemountete Freigabe wird in den Verzeichnisbaum eingebunden und kann wie lokales Verzeichnis angesteuert werden.

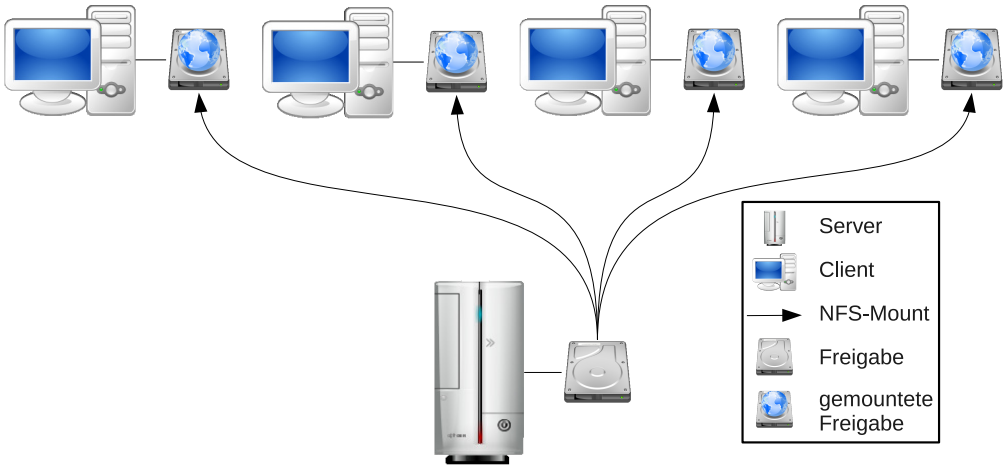

Abbildung 3.3.: NFS Beispiel

Vorteile

- Geringer Administrationsaufwand

Nachteile

- Schlechte Skalierbarkeit, da viele von einer NFS-Freigabe gestartete virtuelle Maschinen, zu einer permanent hohen Netzwerklast führen
- Keine Lastenverteilung

3.4. Vergleich

Im Folgenden werden die Verteilungsalternativen in Hinsicht auf die Kriterien Skalierbarkeit, Netzwerkausfall und Geschwindigkeit untersucht werden.

3.4.1. Skalierbarkeit

Eine gute Skalierbarkeit zeichnet sich dadurch aus, dass der Aufwand nicht signifikant ansteigt oder sich verlangsamt, wenn das Masterimage an einen weiteren Virtualisierungsserver (N+1) verteilt wird. NFS zeigt dabei eine Schwäche, die Last steigt des NFS-Servers stetig mit jedem neuen NFS-Client linear an [Ker00].

Der Aufwand der Verteilung per Multicast steigt bei einem zusätzlichem Empfänger nicht an. Jedoch wird die Übertragung erheblich langsamer, wenn der zusätzliche Empfänger eine langsame Verbindung zu dem Server hat.

Der dezentrale Aufbau des BitTorrent-Netzes macht es sehr skalierbar. Jeder zusätzliche Empfänger des Masterimages, wird auch gleichzeitig zu einem Sender. Wenn die Upload- und die Downloadgeschwindigkeit bei einem zusätzlichen Peer gleich hoch sind, wird das Netz dadurch theoretisch also nicht langsamer. Das BitTorrent-Netz profitiert sogar von zusätzlichen Peers, da sie die Störanfälligkeit des Netzes verringern [EK05].

3.4.2. Störanfälligkeit

Hier wird verglichen, wie sich der Ausfall eines Netzwerkknotens auf die Verteilung auswirkt. BitTorrent ist besonders unanfällig gegen Ausfälle im Netz. Dieses wird durch die dezentrale Struktur ermöglicht. Wenn ein einzelner Netzwerkknoten ausfällt, besteht trotzdem unter den noch verfügbaren Knoten ein Netz.

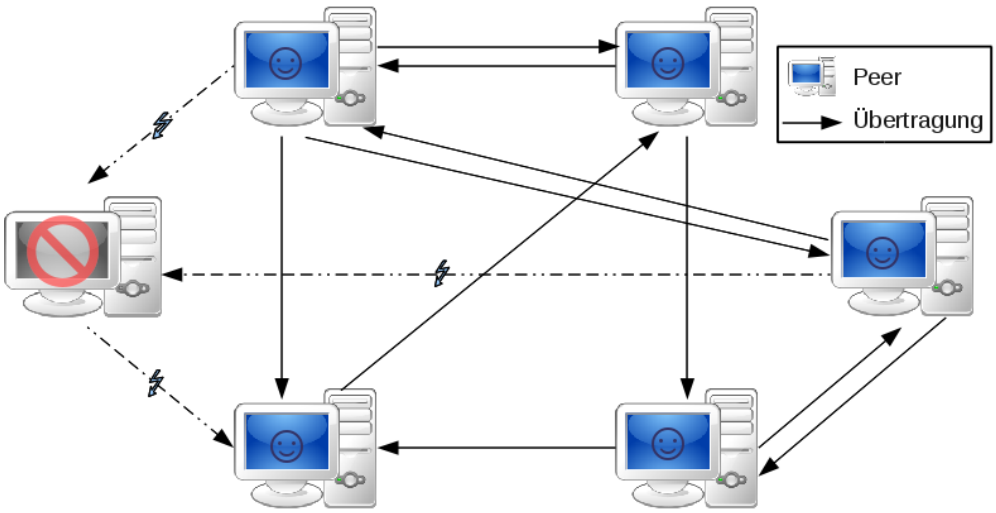

Abbildung 3.4.: Bittorrent Netzwerkausfall

NFS und Multicast haben im Unterschied zu BitTorrent einen großen Nachteil, da die Verteilung über einen einzigen Knoten stattfindet. Der Ausfall eines bestimmten Knotens führt also zum kompletten Abbruch der Verteilung. Man nennt diesen Punkt *Single Point of Failure*.

Bei NFS gibt es beim Bereitstellen der Masterimages zusätzlich die Problematik, dass der Festplattenzugriff der virtuellen Maschinen von der Verfügbarkeit des NFS-Servers abhängt. Ein Ausfall führt damit zum Absturz der virtuellen Maschinen.

3.4.3. Verteilungsdauer

Besonders hervorzuheben ist NFS, da es nicht wie BitTorrent und Multicast die Masterimages verteilt, sondern bereitstellt. Dadurch benötigt es keine Zeit die Masterimages zu verteilen und kann sie direkt zur Verfügung stellen.

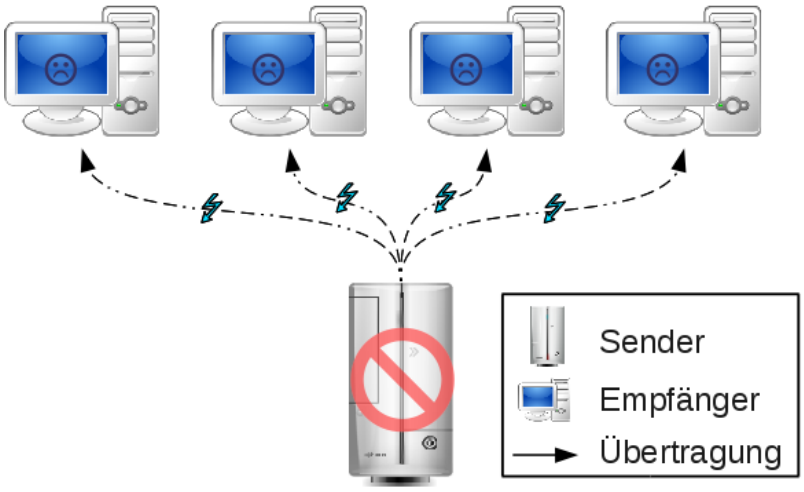

Abbildung 3.5.: Multicast Netzwerkausfall

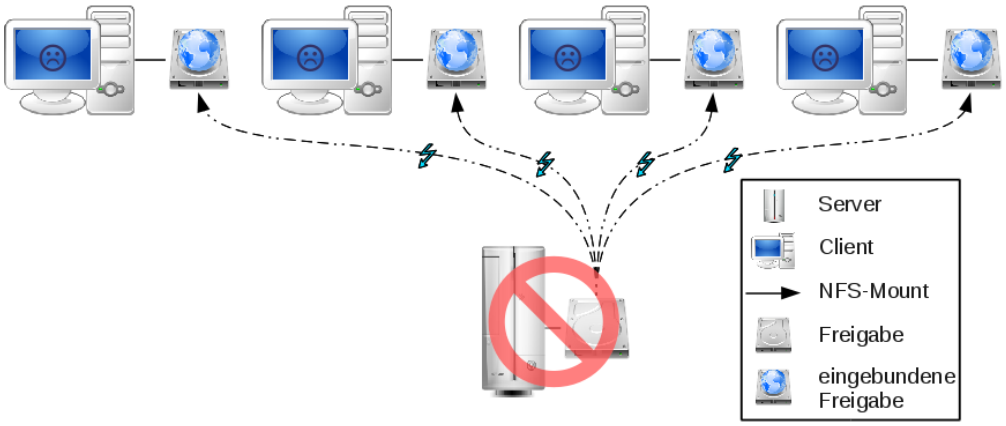

Abbildung 3.6.: NFS Netzwerkausfall

Die Dauer der Übertragung ist bei Multicast vom langsamsten beteiligten Netzwerkknoten abhängig. Ideal ist es, wenn alle Empfänger und der Sender über die gleiche Download- und Upload-Bandbreite verfügen (homogene Umgebung). So kann die gleichzeitige Übertragung an alle Empfänger optimal ausgenutzt werden.

BitTorrent zeichnet sich vor allem dadurch aus, dass es auch gute Ergebnisse erzielt, wenn die Peers über unterschiedliche Download- und Upload-Geschwindigkeiten verfügen. In einer homogenen Umgebung benötigt es mehr Zeit für die Verteilung als Multicast.

3.5. Fazit

Alle aufgezeigten Lösungen für das Verteilen von Masterimages haben ihre Vor- und Nachteile. Jedoch zeigt sich, dass BitTorrent wesentliche Vorteile gegenüber den anderen beiden Lösungen hat. Eine geringe Störanfälligkeit ist im produktiven Einsatz sehr wichtig. Auf diesem Gebiet liegt BitTorrent weit vor NFS und Multicast. Multicast ist ein sehr effizientes Protokoll und es wird produktiv eingesetzt, zum Beispiel bei der Übertragung von Multimediadaten. Allerdings ist es ungeeignet, wenn der Empfänger alle Daten erhalten muss. Auch die Erweiterbarkeit um zusätzliche Virtualisierungsserver unterstützt die Schlussfolgerung, dass BitTorrent für den hier diskutierten Einsatz die effizienteste Lösung ist.

4. Konzeptentwicklung und Realisierung

Die Erkenntnisse des vorhergehenden Kapitels werden in diesem Kapitel aufgegriffen und zu einem Konzept für die Implementierung zusammengefügt. Dieses Konzept wird im Anschluss auf geeignetem Wege umgesetzt.

4.1. Konzept

Die zu entwickelnde Verwaltungslösung dient dem Zweck virtuelle Maschinen sehr schnell zu replizieren und die dafür benötigten Vorlagen schnell auf die Virtualisierungsserver zu verteilen. Die nötigen Voraussetzungen dafür sind:

- Steuerung und Kommunikation
- Verteilung
- Klonen.

Im Folgenden werden diese Voraussetzungen im Einzelnen erläutert.

4.1.1. Steuerung und Kommunikation

Um Masterimages in einem Netz mit mehreren Virtualisierungsservern zu verteilen und zu klonen, bedarf es einer Kommunikation zwischen den Rechnern. Diese Kommunikation sollte von einem zentralen Server gesteuert werden können. Dieser Verwaltungsserver kann selbst ein Virtualisierungsserver sein oder ausschließlich mit der Verwaltung beschäftigt sein.

Die Virtualisierungstechniken sollen über eine einheitliche Schnittstelle verwaltet werden können. Durch die einheitliche Schnittstelle wird die Verwaltung vereinfacht und zusätzlicher Aufwand vermieden. Das Starten Stoppen und das Definieren virtueller Maschinen erfolgt über eine zentrale Schnittstelle. Auch die Möglichkeit, neue Virtualisierungstechniken zu integrieren, soll bei einer Weiterentwicklung des Programms gegeben sein.

4.1.2. Verteilung

Die Verteilung der Masterimages findet über das BitTorrent-Protokoll statt (siehe Kapitel 4.2.4). Der BitTorrent-Client muss für eine einfache und automatisierte Verteilung über die Kommandozeile bedienbar sein. Eine weitere Voraussetzung ist die Unterstützung des Protokolls DHT. DHT ermöglicht das Finden anderer Peers ohne zentralen Tracker.

Zum Starten der Verteilung der Masterimages wird zunächst eine Torrent-Datei erstellt und an alle Virtualisierungsserver gesendet, die es erhalten sollen. Danach wird der BitTorrent-Client gestartet und der Download initiiert.

Nicht jeder Virtualisierungsserver kann das Verteilen initiieren, sondern nur das Verwaltungsprogramm des Verwaltungsservers. Dies gewährleistet, dass nicht jeder Virtualisierungsserver auf jeden anderen zugreifen können muss.

4.1.3. Klonen

Das Klonen wird, wie auch die Verteilung, von dem zentralen Verwaltungsserver initiiert. Für das eigentliche Klonen der virtuellen Festplatten werden die in den Virtualisierungstechniken integrierten Programme eingesetzt.

4.2. Implementierung

In diesem Unterkapitel werden die oben genannten Punkte aufgegriffen und deren Umsetzung beschrieben. Im Einzelnen wird hier auf die Punkte wie die Verwaltung der virtuellen Maschinen, das Klonen und die Verteilung eingegangen. Eine Auflistung der entwickelten Skripte findet sich im Anhang (siehe Kapitel A.1).

4.2.1. Rahmenbedingungen

Die Verwaltungslösung wird auf einem Debian squeeze System implementiert. In der Implementierung werden ein paar wenige debianspezifische Befehle wie zum Beispiel *apt-get* verwendet. Diese können aber leicht für andere Linux-Distributionen portiert werden. Neben der oben genannten Software kommen ssh und rsync zum Einsatz.

Für die Programmierung wird die Skriptsprache Python eingesetzt. Da das hier entwickelte Verwaltungsprogramm nicht zeitkritisch ist, hat die Performanz keine hohe Priorität. Viel wichtiger ist es, den Wartungsaufwand niedrig zu halten. Mit diesen Bedingungen ist die Skriptsprache Python eine sehr gute Wahl.

4.2.2. Steuerung und Kommunikation

Um die Steuerung der Virtualisierungsserver zu vereinfachen und zu vereinheitlichen wird in dieser Arbeit die Virtualisierungs-API libvirt verwendet. Die Virtualisierungstechniken Xen und KVM können beide mit libvirt verwaltet werden. Die Fähigkeiten von libvirt umfassen zum Beispiel das Erstellen, Starten, Stoppen, Pausieren sowie die Migration von virtuellen Maschinen.

Alle virtuellen Maschinen werden von libvirt als XML-Beschreibung verwaltet. Sie enthalten Informationen zu der virtuellen Hardware und eine eindeutige Identifikationsnummer. Eine solche XML-Beschreibung ist beispielhaft im folgenden Listing 4.1 dargestellt.

```
1  <domain type='kvm'>
2    <name>debian</name>
3    <memory>512000</memory>
4    <currentMemory>512000</currentMemory>
5    <vcpu>1</vcpu>
6    <os>
7      <type>hvm</type>
8      <boot dev='hd'/>
9    </os>
10   <features>
11     <acpi/>
12   </features>
13   <clock offset='utc'/>
14   <on_poweroff>destroy</on_poweroff>
15   <on_crash>destroy</on_crash>
```

```
16   <devices>
17     <emulator>/usr/bin/kvm</emulator>
18     <disk type='file' device='disk'>
19      <driver name='qemu' type='qcow2'/>
20       <source file='/var/lib/libvirt/images/debian.qcow2'/>
21       <target dev='hda'/>
22     </disk>
23     <interface type='network'>
24       <source network='default'/>
25     </interface>
26     <input type='mouse' bus='ps2'/>
27     <graphics type='vnc' port='-1' listen='0.0.0.0'/>
28   </devices>
29 </domain>
```

Listing 4.1: libvirt-XML Beispiel

Libvirt bietet die Möglichkeit über das Netzwerk angesprochen zu werden. Außerdem unterstützt libvirt neben Xen und KVM noch andere Virtualisierungstechniken, die bei einer weiteren Entwicklung in die Softwarelösung integriert werden können.

Die einzelnen Aufgaben wie das Verteilen und das Klonen werden auf den Virtualisierungsservern von lokal installierten Skripten erledigt. So kann vermieden werden, dass unnötig viele Befehle über das Netzwerk gesendet werden müssen. Die Skripte werden über das Netzwerkprotokoll ssh gestartet.

4.2.3. Einrichtung eines Virtualisierungshosts

Die Einrichtung wird durchgeführt, um auf allen verwalteten Virtualisierungshosts die Grundvoraussetzungen für das Klonen und Verteilen zu schaffen. Während der Einrichtung wird die nötige Software installiert und es werden Einstellungen vorgenommen. Sie ermöglichen das einfache

Kopieren und Klonen von virtuellen Maschinen. Der Ablauf der Einrichtung wird im Folgenden dargelegt.

Ablauf

Der Benutzer gibt zunächst die IP-Adresse des Virtualisierungshosts an. Ebenfalls wird die Virtualisierungstechnik des neuen Hosts abgefragt. Nach der Eingabe wird über die IP-Adresse der Rechnername erfragt. Für die einfache Kommunikation wird der ssh-key des Verwaltungsservers auf dem zu verwaltenden Virtualisierungshost hinzugefügt (siehe Listing 4.2).

```
1 #!/bin/bash
2 ip="$1"
3 host=$(ssh root@$ip 'echo "$HOSTNAME"')
4 ssh-copy-id root@$host > /dev/null
```

Listing 4.2: Abruf des Rechnernamens und Kopieren des ssh-keys (hostname.sh)

Um nicht alle Aktionen remote über das Netzwerk ausführen zu müssen, werden die Funktion des Klonens und der Verteilung in Skripte ausgelagert. Diese Skripte werden von dem Verwaltungsserver auf den neuen Host übertragen (siehe Listing 4.3).

```
1 command = ['rsync', '-r', 'client-scripts/' ,'root@' + hostName +
      ':' + binDir]
2 execute(command)
```

Listing 4.3: Übertragung der Client-Skripte (cow.py)

Im Anschluss folgt die Installation der benötigten Software-Pakete. Es werden die Pakete für libvirt, deluge (siehe Kapitel 4.2.4), sowie für administrative Tools installiert, wie das Listing 4.3 zeigt.

```
command = ['ssh', 'root@' + hostName, '/opt/cow/packageinstall.py
    "deluged deluge-console mktorrent libvirt-bin python-libvirt
    "' ]
execute(command)
```

Listing 4.4: Paketinstallation auf dem Virtualisierungsserver (cow.py)

Für die Abfrage über das Netzwerk verwendet libvirt X.509-Zertifikate. Es gibt drei unterschiedliche Zertifikate. Das Server-Zertifikat dient dazu die Echtheit des Virtualisierungsservers zu validieren. Das Client-Zertifikat wird von dem Server dazu verwendet, den Client zu authentifizieren und ihm dann Zugriff zu gewähren. Das CA-Zertifikat wird benötigt um das Server-Zertifikat und das Client-Zertifikat zu generieren und zu signieren.

In der Implementierung wird das Server-Zertifikat für den Virtualisierungsserver auf dem Verwaltungsserver generiert. Aus Gründen der Einfachheit wird es vom Verwaltungsserver generiert. In der Regel generiert sonst jeder Server sein eigenes Zertifikat. Nach der Erstellung wird das Server-Zertifikat, wie in Listing 4.5 zu sehen, auf dem Virtualisierungsserver abgelegt. (**Hinweis:** Dies ist nur eine vereinfachte Darstellung des Zertifikate-Infrastruktur. Eine ausführliche Beschreibung ist unter [Lib] zu finden.)

```
certtool --generate-privkey > "$tempdir/serverkey.pem"
certtool --generate-certificate --load-privkey "$tempdir/
    serverkey.pem" \
  --load-ca-certificate /etc/pki/CA/cacert.pem --load-ca-privkey
     /etc/pki/CA/private/cakey.pem \
  --template "$tempdir/server.info" --outfile "$tempdir/
     servercert.pem" 2> "/var/log/cow.log"

ssh "root@$host" "mkdir -p /etc/pki/libvirt/private/"
rsync "$tempdir/serverkey.pem" "root@$host:/etc/pki/libvirt/
```

```
    private/serverkey.pem"
rsync "$tempdir/servercert.pem" "root@$host:/etc/pki/libvirt/
    servercert.pem"
```

Listing 4.5: Erstellung des Server-Zertifikats für den jeweiligen Virtualisierungsserver (servercert.sh)

Die verwalteten Virtualisierungsserver werden als Liste in einer Klartextdatei abgespeichert. Sie enthält zu jedem Virtualisierungsserver den Rechnernamen sowie die Virtualisierungstechnik.

4.2.4. Verteilung

Für den Zweck der Verteilung, kommt in dieser Arbeit *deluge* als BitTorrent-Client zum Einsatz. Er kann komplett über die Kommandozeile gesteuert werden und hat die Möglichkeit per DHT andere Peers zu finden.

Ablauf

Zunächst wählt der Benutzer einen Virtualisierungshost aus, der die zu verteilende virtuelle Maschine beherbergt. Die Virtualisierungshosts werden aus der zuvor abgelegten Klartextdatei ausgelesen (siehe Listing 4.6).

```
def hostList():
  hostList = []
  hosts = open(os.path.expanduser('~/.cow/vhosts'), 'r').
      readlines()
  for i in range(0,len(hosts)):
    host = hosts[i].split('\t')
    if len(host) == 2: #ignore malformed rows
      hostList.append([len(hostList), host[0], host[1].strip()])
  return hostList
```

Listing 4.6: Auslesen der registrierten Virtualisierungshosts (cow.py)

Die Funktion für die Auswahl lässt den Benutzer zwischen allen ausgeschalteten virtuellen Maschinen auswählen. Diese werden wie im folgenden Listing 4.7 über die Virtualisierungs-API abgerufen.

```
1  def vmOffList(hostName, vType):
2    vOffList = []
3    if vType == 'xen':
4      conn = libvirt.open('xen://' + hostName + '/')
5    else:
6      conn = libvirt.open('qemu://' + hostName + '/system')
7
8    for name in conn.listDefinedDomains():
9      vOffList.append(conn.lookupByName(name))
10
11   return vOffList
```

Listing 4.7: Abruf der ausgeschalteten virtuellen Maschine mit libvirt (cow.py)

Außerdem gibt der Benutzer an, auf welche Virtualisierungsserver die virtuelle Maschine verteilt werden soll. Nach der Auswahl der Server und der VM erstellt das Skript `maketorrent.py` (siehe Listing 4.8) eine Torrent-Datei aus der XML-Beschreibung von libvirt und den virtuellen Festplatten. Sie wird an alle ausgewählten Virtualisierungsserver mit rsync weitergegeben.

```
1  command = ['mktorrent', '-a', config.ip, '-o' , torrentFileName,
       torrentDir]
2  execute(command)
```

Listing 4.8: Erstellen der Torrent-Datei (maketorrent.py)

Zuletzt werden alle BitTorrent-Clients gestartet und die erstellte torrent-Datei hinzugefügt. Durch das Hinzufügen wird automatisch der Download bzw. die Verteilung gestartet.

4.2.5. Klonen

Für das Klonen der virtuellen Maschinen werden die von den Virtualisierungstechniken mitgebrachten Tools verwendet. Auf einem Xen-Server ist es das Tool *vhd-util*, bei KVM *kvm-img*. Um die virtuelle Maschine zu klonen, müssen Änderungen an der XML-Beschreibung vorgenommen werden und die Festplatten mit den Tools der Virtualisierungstechniken von der Vorlage abgeleitet werden. Der Ablauf des Klonens wird im Folgenden beschrieben.

Ablauf

Beim Klonen einer virtuellen Maschine wählt der Benutzer, wie bei der Verteilung, einen Virtualisierungs-Host und eine virtuelle Maschine aus. Zusätzlich dazu wird die Anzahl der Klone und die Option alle Klone sofort zu starten abgefragt. Nach den erfolgten Benutzereingaben ruft das Verwaltungsprogramm das auf dem Virtualisierungsserver befindliche Skript `clone.py` zum Klonen auf.

```
1 command = ['ssh', 'root@' + hostName, binDir + '/clone.py ' + vm.
    name() + ' ' + cloneCount + ' ' + autostart + ' ' + str(debug
    )]
2 stdout, stderr = execute(command)
```

Listing 4.9: Starten des Klonvorgangs (cow.py)

Im ersten Schritt des Klonvorgangs generiert das Skript einen neuen Namen (siehe Listing 4.10). Der Name setzt sich aus dem alten Namen und sechs zufälligen und Buchstaben zusammen.

```
1 def randomName(vmName):
2   length = len(vmName) + 6
3   chars = string.letters+string.digits
4   name = vmName
5   while(len(name) < int(length)):
```

```
6       name += random.choice(chars)
7     return name
```

Listing 4.10: Erstellen des Namens der VM (clone.py)

Der nächste Schritt ist es die Beschreibung der Vorlage aus libvirt zu laden. Aus ihr werden die Festplatten der Vorlage ausgelesen und geklont. Wie in dem folgenden Listing 4.11 gezeigt wird, gibt es unterschiedliche Klon-Funktionen für Xen und KVM. Sie rufen die Klonwerkzeuge der jeweiligen Virtualisierungstechnik auf.

```
1   def cloneHddKvm(hdd, newHddPath):
2     command = ['kvm-img', 'info' , hdd]
3     baseFormat = re.search('file␣format:␣(?P<format>[\S]*)',
          execute(command)).groupdict()['format']
4     command = ['kvm-img', 'create', '-f', 'qcow2', '-b', hdd, '-o',
          'backing_fmt=' + baseFormat, newHddPath]
5     execute(command)
6
7   def cloneHddXen(hdd, newHddPath):
8     command = ['vhd-util', 'snapshot', '-n', newHddPath, '-p', hdd]
9     execute(command)
```

Listing 4.11: Klonfunktionen für Xen und KVM (clone.py)

Die Identifikationsnummer und die MAC-Adresse aus der XML-Beschreibung werden gelöscht und der neue Name eingetragen. Die MAC-Adresse und die Identifikationsnummer generiert libvirt neu beim Anlegen der geklonten virtuellen Maschine. Die Änderungen an der XML-Beschreibung sind in dem Listing 4.12 zu sehen. Alle entfernten Zeilen sind rot markiert, alle hinzugefügten grün.

```
1   <domain type='kvm'>
2 -   <name>debian</name>
3 -   <uuid>a6a02d47-6255-7ca7-79e7-22b2cde046a7</uuid>
4 +   <name>debianVxyIZ5</name>
5 +
6     <memory>512000</memory>
```

```
7     <currentMemory>512000</currentMemory>
8     <vcpu>1</vcpu>
9     [...]
10    <devices>
11      <emulator>/usr/bin/kvm</emulator>
12      <disk type='file' device='disk'>
13       <driver name='qemu' type='qcow2'/>
14  -      <source file='/var/lib/libvirt/images/debian.qcow2'/>
15  +      <source file='/var/lib/libvirt/images/debianVxyIZ5-debian.
        qcow2'/>
16         <target dev='hda'/>
17      </disk>
18      <interface type='network'>
19  -      <mac address='52:54:00:3d:eb:4b'/>
20  +
21         <source network='default'/>
22      </interface>
23      <input type='mouse' bus='ps2'/>
24      <graphics type='vnc' port='-1' listen='0.0.0.0'/>
25    </devices>
```

Listing 4.12: modifizierte XML-Beschreibung

4.2.6. Gesamtübersicht der Implementierung

Die Softwarelösung unterscheidet grundsätzlich zwischen zwei unterschiedlichen Knotenpunkten. Zum einen gibt es den Virtualisierungsserver oder auch Virtualisierungshost. Er beherbergt die virtuellen Maschinen. Zum anderen gibt es den Verwaltungsserver. Auf ihm findet die Verwaltung statt und es wird von dem Verwaltungsserver ausgehend das Klonen und das Verteilen der virtuellen auf den Virtualisierungshosts initiiert (siehe Abbildung 4.1).

Die aufwendigeren Aktionen wie das Klonen und erstellen der Torrent-Datei werden nicht über das Netzwerk vom Virtualisierungsserver erledigt. Stattdessen sind sie in Skripte auf den Virtualisierungsservern aus-

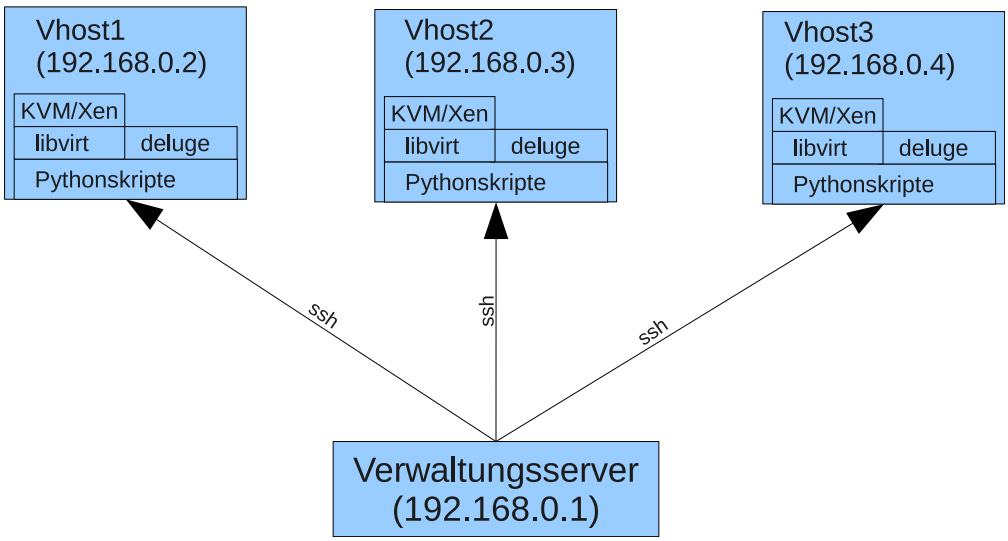

Abbildung 4.1.: Kommunikation

gelagert. Der Verwaltungsserver übernimmt das sammeln der Parameter und übergibt sie den lokalen Skripten auf den Virtualisierungsservern beim Starten per ssh.

5. Zusammenfassung und Ausblick

In dem folgenden Kapitel werden Schlussfolgerungen über die erzielten Ergebnisse der Analyse gezogen und ein Fazit der Implementierung dargelegt. Abschließend werden in einem Ausblick mögliche Ansätze für die weitere Entwicklung der Implementierung gegeben.

5.1. Zusammenfassung

Es konnte mit objektiven Analysemethoden gezeigt werden, dass qcow2 und BitTorrent die geeignetsten Technologien für das schnelle Replizieren und Verteilen von virtuellen Maschinen sind. Mit der an die Analyse anschließenden Implementierung wurden diese so zur Verfügung gestellt, dass sie auch ohne Vorkenntnisse bedienbar sind. Außerdem wurde aufgezeigt, dass die analysierten Technologien praxistauglich sind.

Das Imageformat VHD wird, im Gegensatz zu den beiden oben genannten Technologien, nur als Übergangslösung zum Einsatz kommen. Es ist zwar praxistauglich, jedoch hat es Performancenachteile im Vergleich zu qcow2. Der Schritt wurde notwendig, da es während der Xen-Entwicklung zu einer fragwürdigen Umstrukturierung kam. Die Umstrukturierung führt dazu, dass Xen bestimmte Imageformate wie qcow2 zur Zeit nicht unter-

stützt. Dieser Punkt und die Tatsache das Xen nicht ohne selbst eingespielte Patches fehlerfrei einsetzbar ist, führen zu dem Urteil, dass der Einsatz von Xen für den Zweck dieser Arbeit aktuell nicht zu empfehlen ist. Die entwickelte Softwarelösung ermöglicht das Replizieren und Verteilen. Somit erfüllt die Softwarelösung die gestellten Erwartungen, virtuelle Maschinen schnell zu Klonen und über das Netzwerk zu verteilen. Hierbei war der Einsatz von bereits vorhandener Open Source Software sehr hilfreich. Ohne die eingebundene freie Software wäre dieses Projekt zeitlich nicht in dem Rahmen dieser Arbeit realisierbar gewesen.

In dieser Arbeit wird der wichtige Punkt deutlich, dass bei einer Softwareentwicklung nicht immer das Rad neu erfunden werden muss. Dieses gilt vor allem im Bereich von Open Source. Hier gibt es bereits sehr viele freie Bibliotheken und Programme für alle Themengebiete. Sie können problemlos in einer Eigenentwicklung verwendet und eingebunden werden. Trotz der nicht vorhandenen Lizenzkosten, ist Open Source Software in der Regel nicht funktionell eingeschränkter oder langsamer als proprietäre Software.

5.2. Ausblick

Für die weitere Entwicklung wäre es möglich, die Bedienung für den Benutzer durch eine grafische Oberfläche zu vereinfachen. Hierfür könnte das grafische Framework QT eingesetzt werden, dass auch Bibliotheken für die in der Implementierung verwendete Sprache Python bereitstellt.

Eine andere denkbare Weiterentwicklung wäre die Integration in libvirt. Hierbei könnten dann direkt Klone mit der Virtualisierungs-API erstellt werden.

A. Anhang

A.1. Skripte/Programme

Die einzelnen Aufgaben der Verwaltungslösung werden nicht in einem einzigen Skript abgearbeitet, sondern in einzelne übersichtliche Skripte aufgeteilt. In diesem Unterkapitel werden sie aufgelistet und ihre Funktion beschrieben.

A.1.1. Verwaltungsserver

Die Skripte des Verwaltungsservers nehmen die Verwaltungs- und Einrichtungsaufgaben wahr. Ihre Aufgaben werden hier kurz erklärt.

cow.py - Hauptprogramm der Verwaltungslösung: Es nimmt die Benutzereingaben entgegen und ruft die anderen Skripte auf.

hostname.sh - fragt den Rechnernamen des Virtualisierungsservers ab und trägt dort den ssh-key des Verwaltungsservers ein.

servercert.sh - erstellt ein Server-Zertifikat für den Virtualisierungsserver.

cacert.sh - erstellt das CA-Zertifikat wenn es nicht vorhanden ist und überträgt den öffentlichen Schlüssel des CA-Zertifikats an den jeweiligen Virtualisierungsserver.

clientcert.sh - erstellt ein Client-Zertifikat, dass dem Verwaltungsserver Zugriff auf die libvirt-Installationen der Virtualisierungsserver ermöglicht.

A.1.2. Virtualisierungsserver

Diese Skripte liegen auf den Virtualisierungsserven. Sie werden während der Einrichtung vom Verwaltungsserver installiert (siehe 4.2.3) , um lokal abrufbar zu sein. Ihre Funktionen werden im Folgenden kurz erläutert.

clone.py - Klonen der virtuellen Maschinen; Das Programm nimmt die nötigen Modifikationen an der XML-Beschreibung der Vorlage vor und klont die virtuellen Festplatten.

maketorrent.py - erzeugt eine torrent-Datei und startet das Verteilen mit dem BitTorrent-Client.

packageinstall.py - installiert die benötigten Software-Pakete auf dem Virtualisierungsserver.

whoami.py - legt eine Konfigurationsdatei mit Informationen zu dem Virtualisierungsserver an.

xenprep.py - nimmt Einstellungen an dem Xen-Daemon vor, damit libvirt den Xen-Daemon abfragen kann.

A.2. Testergebnisse

Im Folgenden sind die Testergebnisse von bonnie++ und Iozone dargestellt. Die Werte sind gemittelt. Alle Messwerte finden sich auf der CD.

Abbildung A.1.: Iozone Testergebnisse für Xen

Abbildung A.2.: Bonnie++ Testergebnisse für Xen

Abbildung A.3.: Bonnie++ Testergebnisse für KVM

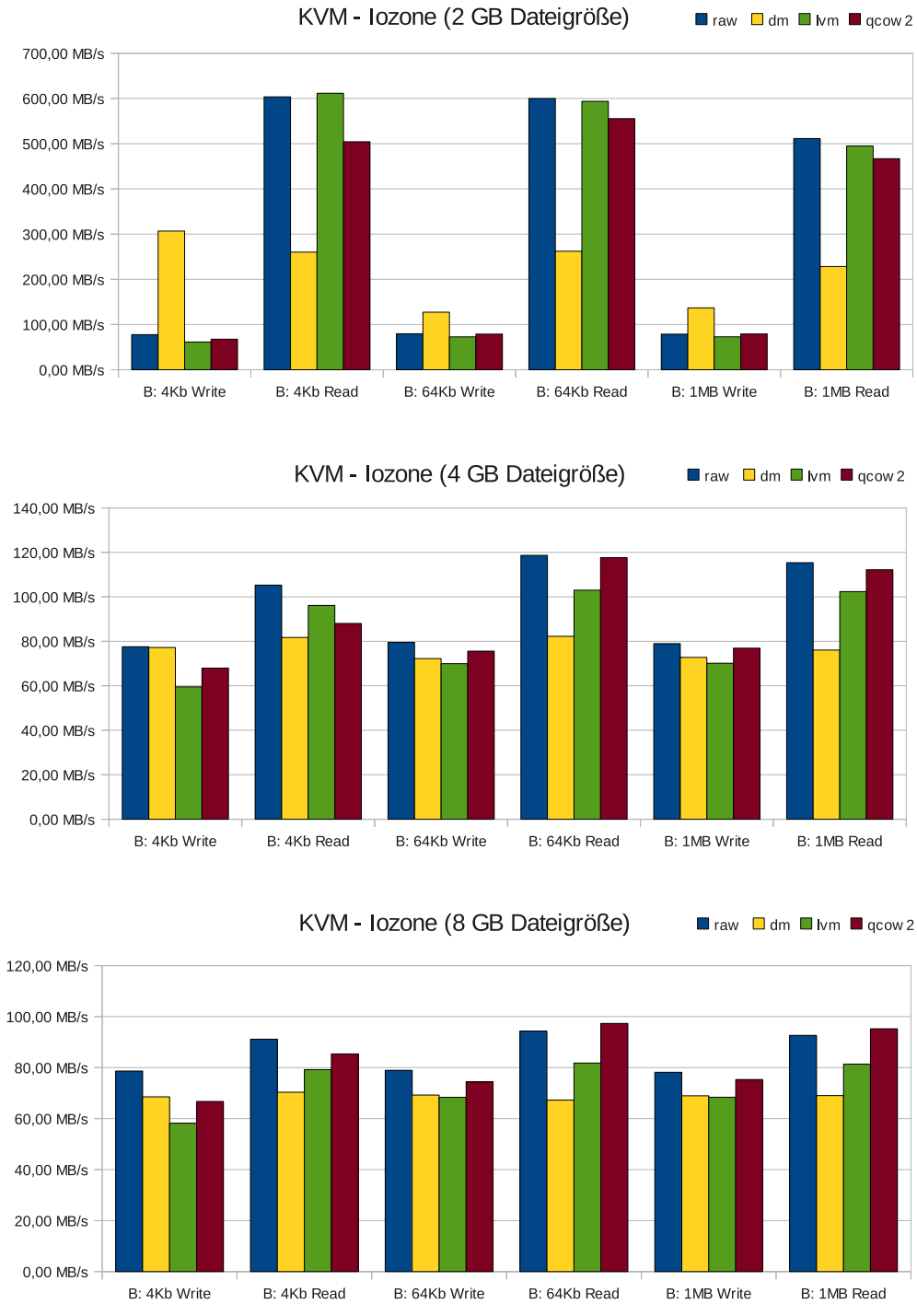

Abbildung A.4.: Iozone Testergebnisse für KVM

Literaturverzeichnis

[Bau08] Baun, Christian: *Vorlesung Systemsoftware*. http://jonathan.sv.hs-mannheim.de/~c.baun/SYS0708/Skript/folien_sys_vorlesung_13_WS0708.pdf. Version: 2008, Abruf: 31.10.2010

[Bro] Brož, Milan: *Device mapper*. http://mbroz.fedorapeople.org/talks/DeviceMapperBasics/dm.pdf, Abruf: 17.10.2010

[CB05] Cesati, Marco ; Bovet, Daniel P.: *Understanding the Linux Kernel*. dritte Ausgabe. Linux-Server-Praxis, 2005

[Coh08] Cohen, Bram: *The BitTorrent Protocol Specification*. http://www.bittorrent.org/beps/bep_0003.html. Version: 2008, Abruf: 01.01.2011

[Cro09] Crosby, Simon: *We've Open Sourced Our Optimized VHD Support*. http://community.citrix.com/x/OYKiAw. Version: 2009, Abruf: 11.11.2010

[Dmk] *Device-mapper snapshot support*. http://www.kernel.org/doc/Documentation/device-mapper/snapshot.txt, Abruf: 17.10.2010

[EK05] Eger, Kolja ; Killat, Ulrich: *Scalability of the BitTorrent P2P Application*. http://www3.informatik.uni-wuerzburg.de/

ITG/2005/presentations/kolja.eger.pdf. Version: 2005, Abruf: 01.01.2011

[Hof] Hoffmann, Markus: *Linux LVM-HOWTO*. http://www.selflinux.org/selflinux/html/lvm01.html, Abruf: 18.10.2010

[Ker00] Kerr, Shane: *Use of NFS Considered Harmful*. http://www.time-travellers.org/shane/papers/NFS_considered_harmful.html. Version: 2000, Abruf: 01.01.2011

[Lar10] Laroche, Eric: *Sparse files*. http://www.lrdev.com/lr/unix/sparsefile.html. Version: 2010, Abruf: 18.10.2010

[Lei00] Leitner, Felix von: *Wir erfinden IP Multicasting*. http://www.fefe.de/multicast/multicast.pdf. Version: 2000, Abruf: 01.01.2011

[Lib] *Setting up libvirt for TLS*. http://wiki.libvirt.org/page/TLSSetup, Abruf: 10.01.2011

[Loe08] Loewenstern, Andrew: *DHT Protocol*. http://www.bittorrent.org/beps/bep_0005.html. Version: 2008, Abruf: 01.01.2011

[Lvm] *LVM2 Resource Page*. http://sourceware.org/lvm2/, Abruf: 18.10.2010

[Lvm06] *What is Logical Volume Management?* http://tldp.org/HOWTO/LVM-HOWTO/whatisvolman.html. Version: 2006, Abruf: 18.10.2010

[McL08] McLoughlin, Mark: *The QCOW2 Image Format*. http:

//people.gnome.org/~markmc/qcow-image-format.html. Version: 2008, Abruf: 18.10.2010

[Mso06] Microsoft: *Microsoft Open Specification Promise*. https://www.microsoft.com/interop/osp/default.mspx. Version: 2006, Abruf: 18.10.2010

[Mul] Multicast Technologies, Inc.: *Multicast FAQ File*. http://www.multicasttech.com/faq/, Abruf: 01.01.2011

[Nfs03] The Internet Society: *Network File System (NFS) version 4 Protocol*. http://tools.ietf.org/html/rfc3530. Version: 2003, Abruf: 01.01.2011

[Prz] Przywara, André: *Virtualization Primer*. http://www.andrep.de/virtual/, Abruf: 01.11.2010

[Qco10] *Qcow2 Support*. http://lists.xensource.com/archives/html/xen-devel/2010-11/msg00256.html. Version: 2010, Abruf: 14.11.2010

[Qem10] *QEMU Emulator User Documentation*. http://wiki.qemu.org/download/qemu-doc.html#disk_005fimages. Version: 2010, Abruf: 18.10.2010

[Rac10] *Race condition in /etc/xen/scripts/block*. http://lists.xensource.com/archives/html/xen-devel/2010-07/msg00827.html. Version: 2010, Abruf: 14.11.2010

[Smi06] Smith, Christopher: *4. Setting up an NFS Client*. http://nfs.sourceforge.net/nfs-howto/ar01s04.html#mounting_remote_dirs. Version: 2006, Abruf: 01.01.2011

[Vhd09] Microsoft: *Virtual Hard Disk Image Format Specification*.

http://technet.microsoft.com/en-us/virtualserver/bb676673.aspx. Version: 2009, Abruf: 18.10.2010

[Vmw] VMware: *VMware Benchmarking Approval Process*. http://www.vmware.com/pdf/benchmarking_approval_process.pdf, Abruf: 05.11.2010

Abbildungsverzeichnis

Listings