

RAM basierte Virtualisierung mit ArchivistaVM 2012/XI

Vortrag linuxday.at 2012, 24.11.2012, Dornbirn

Contents

1	Einleitung	2			
1.1	Ausgangslage	2	3.4	Hochfahren eines Clusters 16	
1.2	Neuerungen 2012	3	3.5	Installation von Gästen 16	
2	Wahl der Hardware	5	3.6	Migration von Gästen 17	
2.1	ArchivistaBox Tiny	5	3.7	Datensicherung 17	
2.2	ArchivistaVM Budget	5	3.8	Neustart und Updates 19	
2.3	Antec IKS 310-150	6	3.9	Restore-on-the-fly 19	
2.4	Chenbro PC78131	7	4	RAM-Modus und Ausblick	20
2.5	500-Euro-Cluster	9	4.1	Historisch gesehen	20
2.6	Baukasten != Business-Lösung	10	4.2	Rückmeldungen vom Support	20
3	ArchivistaVM live!	12	4.3	RAM für alle Fälle	21
3.1	Demo mit 'custom ramonly' .	12	4.4	Mehr, nicht weniger Flexibilität	21
3.2	Vorbereitungen für Cluster . .	12	4.5	Mehr Tempo mit weniger Hardware	22
3.3	Lizenz und Copyright	13	4.6	Blick in die Zukunft	22

© 23.11.2012 by Urs Pfister, www.archivista.ch

1 Einleitung

ArchivistaVM ist eine Debian basierende schlanke Linux-Distribution für die Virtualisierung mit KVM. Die ISO-Datei liegt unter 100 MByte (ohne Stand-Alone-Modus 72 MByte), das System wird direkt on-the-fly ins RAM hochgefahren, nach ca. 20 Sekunden kann produktiv gearbeitet werden. Client-seitig kann mit jedem einigermaßen modernen Browser über eine Web-Applikation auf die Gäste zugegriffen werden.

Dank der Integration von noVNC sind seit diesem Jahr keine Plugins mehr im Browser notwendig. Eine ausgefeilte Benutzer-Verwaltung ermöglicht es, exakt festzulegen, wer auf welche Ressourcen (Gäste) Zugriff hat. Der Vortrag bietet eine Übersicht über ArchivistaVM, zeigt aber auch auf, inwiefern sich ArchivistaVM in den letzten 12 Monaten weiterentwickelt hat und wagt einen Blick in die "Kristallkugel", d.h. es wird erörtert, wohin die Reise mit ArchivistaVM in Zukunft gehen wird.

Soweit die Ankündigung zu diesem Vortrag. Es ist immer etwas schwierig, bereits im September (beim Einreichen eines Vortrages) das Thema für den November passgenau zu treffen. So auch dieses Jahr. Zwar arbeitete der Autor bereits seit vielen Monaten mit der RAM basierten Version der ArchivistaBox, und doch benötigten wir deutlich mehr Zeit als geplant, ehe die ArchivistaBox 2012/XI fertiggestellt werden konnte.

Im Prinzip gibt es immer viele neue Ideen, einige konnten realisiert werden, aber letztlich zeigt jeweils erst die Praxis, ob der Glaube, ein Feature sei fertiggestellt, zutreffend ist. Und sollte im Rahmen dieses Vortrages etwas nicht funktionieren, dann war der Glaube stärker als das Feature. In diesem Sinne ist dieser Vortrag explizit nicht redundant ausgelegt. Dafür dürfen wir zusammen mit ArchivistaVM spielen, und dazu sind Sie herzlich eingeladen.

Nebenbemerkung: Dieser Vortrag stellt primär ArchivistaVM vor, all jene, die mit ArchivistaDMS arbeiten, erhalten hier aber dennoch viele wertvollen Hinweise, weil letztlich ArchivistaDMS immer auch ArchivistaVM enthält, und weil beide mit dem gleichen Konzept arbeiten (z.B. ist der Installer zu 100 Prozent in allen ArchivistaBox-Versionen identisch).

1.1 Ausgangslage

ArchivistaVM ist zunächst einmal eine Software, aber ArchivistaVM ist vorallem ein Konzept, um Server-Virtualisierung zu realisieren. Dazu gehört neben der Software (dazu später mehr) die Wahl der Hardware. Und wenn der Autor nachfolgend preisgünstige Hardware für einen Mini-Cluster vorschlägt, so vorallem deshalb, weil er im Rahmen dieses Vortrages aufzeigen möchte, mit wie wenig Ressourcen ganz ansprechende Resultate erzielt werden können. Eben-

falls reizt den Autor die Frage, ob es möglich ist, einen Cluster (selbst mit einer vergleichsweise gemächlichen Infrastruktur) in zwei bis drei Minuten aufzubauen.

1.2 Neuerungen 2012

Vielleicht erinnern sich einige noch an die Vorträge anlässlich des linuxday.at in den Jahren 2009, 2010 und 2011. Es mag ja sein, dass ArchivistaVM im Jahre 2009 noch recht bescheiden und nicht so schlank daherkam (alleine die Hardware-RAID-Geschichten waren nicht wirklich einfach). Doch bereits die ArchivistaBox im Jahre 2010 zeigte mit dem neuen Installer, dass es einfacher und schneller geht. Die DRDB-Cluster im Jahre 2011 waren an sich einfach gestrickt, auch wenn das Vorbereiten eines Clusters noch viel Handarbeit erforderte.

Dieses Jahr haben wir daher versucht, die Dinge vorallem nochmals einfacher zu machen. Das beginnt damit, dass mit einer ISO neu mehrere Knoten verwaltet werden können. Weiter benötigen wir dank dem RAM-Modus (dazu später mehr) auch keinen Zugriff mehr auf die internen Platten, ehe das System hochgefahren ist. Ebenfalls kann dank dem RAM-Modus ArchivistaVM direkt aus dem Stand heraus produktiv eingesetzt werden (der obligate Neustart nach der Installation bzw. diese selber entfällt).

Die ISO-Datei von ArchivistaVM ist noch ganze 100 MByte gross (ohne X-Server und Firefox sogar nur 72 MByte). Dies wiederum ermöglicht es uns, das System selbst in Plattenverbänden von bis zu 24 Festplatten in ca. 1 bis 2 Minuten auf sämtliche Platten aufzuspielen. Damit kann ArchivistaVM jederzeit von jeder Platte aus gebootet werden, auch bzw. gerade wenn der Stick nicht eingelegt ist. Ebenfalls können ArchivistaVM-Server über die Netzwerkkarte mit PXE gebootet werden.



Und weil ArchivistaVM noch ganze 100 MByte benötigt, ist irgendwann die Idee entstanden, ArchivistaVM mit in jede Festplatte der Datensicherung zu spielen. Damit kann eine jede Sicherungsplatte selber produktiv (z.B. autonom) auf beinahe jedem beliebigen 64-Bit-Rechner hochgefahren werden. Aber auch Kleinigkeiten wurden geändert, so wird z.B. der Befehl zum

Anlegen einer Instanz nicht mehr dargestellt (das Interesse daran war bisher gleich Null), vielmehr wird direkt das Formular zum Starten der Instanz angezeigt.

Für den Autor das wichtigste Highlight stellt aber zweifelsohne die Integration von noVNC dar. Damit entfällt die Abhängigkeit zu Java. Dies ist nicht nur deshalb toll, weil wir damit viel Platz auf den ISO-Dateien sparen, sondern auch, weil die HTML5-basierte Lösung deutlich schneller startet und innerhalb des Browsers auch beliebig skaliert werden kann. Ebenfalls entfernt wurde das Java-Applet auch im Monitor. Dank einer terminalbasierten Lösung können die Befehle neu mit der Maus heraus- und hineinkopiert werden.

Im Jahre 2012 (auch wenn es noch nicht beendet ist) konnten die bisher umfangreichsten Fortschritte in ArchivistaVM erzielt werden. Offen gestanden ist dies auch deshalb der Fall, weil wir dank ArchivistaVM deutlich effizienter arbeiten können. Entwickeln bedeutet ja nicht nur Code schreiben, sondern vorallem auch testen und nochmals testen. Bei Installationszeiten von einigen Sekunden, virtualisiert ist derzeit ein Cluster in ca. 60 Sekunden aufgebaut, bleibt ganz einfach mehr Zeit für die Entwicklung übrig.

Natürlich mag der Unterschied zwischen 30 Sekunden zu 10 Minuten für die meisten Anwender/innen nicht ultimativ entscheidend sein. Wird der Vorgang aber (wie in unserem Falle bzw. generell in der Entwicklung) pro Tag im zwei- bis dreistelligen Bereich benötigt, so können dabei pro Tag bereits ca. zwischen 100 und 1000 Minuten an Rechenzeit gespart werden. Sämtliche ISO-Dateien werden virtualisiert zusammengestellt, getestet und ausgeliefert, auch bei diesem Durchlauf sparen wir im Vergleich zum letzten Jahr jeweils ca. 20 Minuten an Arbeitszeit.

2 Wahl der Hardware

Ziel für den diesjährigen Vortrag war es, den Aufbau eines Clusters live zu demonstrieren. Das wäre kein Problem, wenn da nicht die Hardware mit an den Vortrag geschleppt werden müsste. Da der Vortragende per Bahn anreist (und überhaupt schleppt ja niemand gerne Rack-Rechner herum), ging es zunächst einmal darum, eine geeignete leichtgewichtige Hardware auszuwählen.

Nebenbemerkung: Wer die Skripte zu den Vorträgen zur ArchivistaBox durchgeht (die Vorträge sind unter www.archivista.ch abrufbar), wird feststellen, dass der Autor gut und gerne jeweils ein paar Geräte ordert, um was passendes zusammenzustellen. Die dazu notwendigen Moneten gehen zulasten der Archivista GmbH, für den Vortrag selber ist alleine der Autor verantwortlich, der dafür etwas dezidiert(er) ist, als wenn eine Firma sprechen würde. Nun aber zurück zur Hardware.

2.1 ArchivistaBox Tiny

Die äusserst leichten (ca. 500 Gramm wiegenden) Kleinst-Rechner, die für das DMS-System ArchivistaDMS verwendet werden, können auch zur Virtualisierung herangezogen werden, doch passt in das Gehäuse nur eine Festplatte. Und weil in unserem Cluster die Rechner je zwei Festplatten aufnehmen müssen, liesse sich dies einzig mit einer USB-Platte bewerkstelligen, aber dies wäre unschön. Weiter liegt der Verkaufspreis dieser Kleinst-Rechner bei ca. 700 Euro, was angesichts einer DualCore CPU (Intel lässt grüssen) ein stolzer Preis ist.



2.2 ArchivistaVM Budget

Zunächst wollte der Autor zwei ArchivistaVM Budget verwenden. Für 1000 Euro gibt es zwei Hot-Swap-Platten, eine QuadCore (optional auch SixCore) CPU, 16 GByte RAM (bis 32 GByte erweiterbar). Kurzum ein tolles Gerät, mit welchem der Autor privat seit Jahren arbeitet. In

normalen Verkaufspreis von ArchivistaVM ist der Wartungsvertrag für das erste Jahr (200 Euro) sowie ArchivistaVM Light (300 Euro) enthalten, sodass für die Hardware 500 Euro zu kalkulieren wären, bzw. 1000 Euro, weil ja zwei Geräte für den Cluster benötigt werden.



Das Problem bei diesen beiden Modellen war aber, dass die Maschinen (MATX-Hauptplatine) nicht in einen Rucksack passen. Zwei Boxen liessen sich zwar in Tragtaschen transportieren, doch bei insgesamt 8-maligem Umsteigen (Bahnreise hin und zurück) fand der Autor diese Variante nun nicht unsäglich verlockend. Kurz und gut, es musste was Kleineres her.

2.3 Antec IKS 310-150

Das Antec-Gehäuse IKS 310-150 bietet auf kleinstem Raum (ITX-Hauptplatine) die Möglichkeit, Cluster mit QuadCore-CPUs, redundanten Netzwerkkarten und gespiegelten Platten zu realisieren. Ein Gerät dürfte für ca. 500 Euro machbar sein. Im Vergleich zur Hardware, die bei ArchivistaVM Budget verwendet wird, fallen bei der Antec IKS 310-150 kaum Mehrkosten an. Dies beruht daher, dass ITX-Systeme nur noch unwesentlich teurer als (M)ATX-Geräte sind, und dem Komfortverzicht, dass die Festplatten nicht hot-swap fähig nach aussen geführt sind.



Mitte Oktober erfolgte eine Testreise zu einem Kunden mit zwei Geräten. Dabei stellte sich heraus, dass die Geräte zwar im Rucksack Platz finden, aber ganz schön den Rücken belasten bzw. an der oberen 'Schmerzgrenze' angesiedelt sind. Die Suche nach einem 'Leichtgewicht' mit Cluster-Ambitionen ging somit weiter.



2.4 Chenbro PC78131

Es gibt viele kleine ITX-Gehäuse. Oft sind sie schlecht verarbeitet, d.h. die Platine muss irgendwie ins Gehäuse reingepresst werden. Für den Autor gilt dabei die einfache (wenn auch etwas schmerzhaft) Testmethode. Immer wenn die Finger leiden (Schürfwunden mit/ohne Blut, Kratzer, zerstörte Fingernägel), wandern die Gehäuse in den Müll. Das 3-Liter-Gehäuse von Chenbro bildet hier eine löbliche Ausnahme, weil es extrem sauber verarbeitet ist. Dank VESA-Halterung kann das Gehäuse zudem an viele Bildschirme geschraubt werden. Ebenfalls mitgeliefert wird ein Standfuss.



Etwas schade beim Chenbro PC78131 ist, dass das Netzteil nur 60 Watt liefert. Es gibt zwar Netzteile mit 96 Watt, womit sich auch QuadCore CPUs verbauen liessen, aber diese Adapter sind schwer erhältlich. Somit verbleiben für unseren Mini-Cluster einzig AMD CPUs der E-Serie (bzw. C-Serie, diese ist aber ca. 20 Prozent langsamer). Diese Dual-Core-CPU's verbrauchen ca. 18 Watt, womit 'Saft' für schnelle Festplatten übrigbleibt. Offen gestanden wollte der Autor zunächst externe USB-to-Ethernet-Adapter verwenden, um die für den Cluster-Aufbau erforderlichen Netzwerkkarten abzubilden, denn ein ArchivistaVM-Cluster bedingt an sich zwingend redundante Punkt-zu-Punkt-Netzwerkkarten.



Aber, einerseits waren dem Autor die externen USB-Adapter ein Dorn im Auge, geschweige denn, dass eine solche Lösung empfohlen werden könnte. Auf der anderen Seite sind dem Autor dieses Jahr einige Support-Fälle aus der Community in Erinnerung geblieben, bei der das Verkabelungskonzept doch erhebliche Schwierigkeiten verursachte – denn, falsch verkabelt, kann der Cluster gar nicht in Betrieb genommen werden (untenstehend links wäre falsch, rechts richtig).



Aus diesem Grunde können Cluster neu auch mit nur einer Netzwerkkarte eingerichtet werden (Version 2012/XI bzw. dort ab 22.11.2012). Damit sinkt zwar die Schreibgeschwindigkeit aller Maschinen auf den maximalen Durchsatz eines 1-GBit-Switches (etwas über 100 MByte), dafür fällt die Verkabelung weg, und der Aufbau eines Clusters gestaltet sich urplötzlich viel einfacher.

2.5 500-Euro-Cluster

Im letzten Jahr hat der Autor den Aufbau eines günstigen 3er-Clusters aufgezeigt. Daraus ist in diesem Jahr die Anfrage für einen 2er-Cluster entstanden, der sich mit einem Stick (in zweifacher Ausführung) komplett selber konfigurieren sollte. Aus diesem Auftrag ist das Thema für diesen Vortrag entstanden, allerdings mit der zusätzlichen Idee, die Hardware in minimalistischer Weise (es sollte ja portabel werden) aufzubauen. Lange schien es, dass irgendwo um die 1000 Euro für die reinen Hardware-Komponenten notwendig sein würden.

Dank dem Verzicht auf die redundanten Netzwerkkarten (bei einem 2er-Cluster sind es vier zusätzliche Karten) kam plötzlich eine ganz andere Marke in greifbare Nähe, jene der 500-Euro-Grenze. Anbei die aktuellen Preise für die verwendeten Komponenten nach:

www.heise.de/preisvergleich

1. Gehäuse Chenbro PC78131: 58,04
2. AMD E350-Platine: 61,10
3. Kingston RAM 8GB: 30,48
4. Kingston RAM 8GB: 30,48
5. 1. Festplatte: 34,71
6. 2. Festplatte: 34,71

Dies ergibt 249,54 bzw. 499,08 Euro für beide Rechner. Bei den Festplatten musste am untere Ende (160 GByte) geordert werden, um die 500 Euro nicht zu überschreiten. Für ca. 8 Euro mehr pro Platte wären bereits Disks im Bereich um die 500 GByte zu haben. Theoretisch hätten auch nur 8 GByte RAM pro Maschine sowie grössere Platten verbaut werden können, aber 16 GByte schienen dem Autor 'schmackhafter'. Nicht weil ArchivistaVM viel Platz benötigt (es sind ca. 400 MByte), sondern weil ein paar GByte pro Gast heute nicht mehr die Ausnahme darstellen. Wer den Preis für die CPU vermisst, dem sei gesagt, dass die CPU fix auf der AMD E350-Platine sitzt.



Ich würde jetzt nicht sagen wollen, dass die verwendete Konfiguration der Virtualisierungstraum schlechthin darstellt. Aber, ITX-Platinen für Quad-Core CPUs gibt es mittlerweile viele. Für etwas weniger knappe Budgets bzw. etwas schwergewichtigere Lösungen wären der Antec IKS 310-150 oder ArchivistaVM Budget aber sicher dazu prädestiniert, mit bescheidenem finanziellem Aufwand eine deutlich schnellere redundante Virtualisierungs-Infrastruktur aufzubauen. Es muss bekanntlich nicht immer Kaviar sein, ein frisches Seelachs-Filet bietet gemessen am Preis deutlich mehr und längeren Genuss.

2.6 Baukasten != Business-Lösung

Weiter sei hier angefügt, dass wenn der Autor in diesem Vortrag von einem 500-Euro-Cluster spricht, dass dabei der reine Baukasten gemeint ist. Dies nach dem Prinzip: DIY bzw. do it yourself! Diese Komponenten müssen bestellt, zusammengebaut, getestet und mit dem ArchivistaVM Community Release bespielt werden (letzteres sollte kein allzu grosser Aufwand darstellen). Ebenfalls enthält die Community-Version der ArchivistaBox explizit keinen Support. Der Autor würde hier komplett falsch verstanden, wenn im beruflichen Umfeld nicht eine etwas höherwertigere Lösung mit Support zum Einsatz kommen sollte. Diese Empfehlung gilt explizit für ArchivistaVM, sie gilt aber für alle Virtualisierungslösungen.

Es ist naiv und geradezu fahrlässig zu glauben, dass Lösungen ohne Support im Business-Umfeld Sinn ergeben. Natürlich bestehen bei einer Open Source Lösung wie ArchivistaVM bessere Chancen bei einem Ausfall über die Runden zu kommen als dies bei den Gratis-Versionen namhafter Hersteller der Fall ist. Immerhin kann im Code nach den Ursachen allfälliger Probleme gesucht werden. Aber, macht es wirklich Sinn, das Rad ein zweites Mal (selber) zu erfinden? Welcher Zeitaufwand kann/soll in einem Problemfall betrieben werden?

Selbst wenn Linux und Virtualisierung bereits eine spannende Freizeitbeschäftigung darstellen, es wird keinen Sinn ergeben, derart tief in die 'Abgründe' hineinzuschauen. Offen gestanden, der Autor ist nicht traurig, wenn im Support derartige Fälle auf ein absolutes Minimum beschränkt sind, auch wenn es letztlich bezahlte Arbeit ist. Unter Zeitdruck nach einer Lösung zu suchen. von der noch nicht einmal die genaue Problemstellung bekannt sein muss, dies macht nicht immer nur Freude. Bereits das Herausfiltern, ob ein Problem eher bei der Hardware als bei der Software liegt, kann mitunter äusserst nervenaufreibend sein.

Und zum Schluss noch dies: ArchivistaVM kann nur weiterentwickelt werden, weil jemand (sprich meist Firmenkunden) bereit sind, die Weiterentwicklung zu finanzieren. Die Kosten bei der Entwicklung sind erklecklich (in etwa 1000 Arbeitsstunden dürften es dieses Jahr schon sein). Die Kosten für den einzelnen sind moderat, ArchivistaVM Light kostet im mehrjährigen Einsatz pro Jahr um die 100 Euro. Dafür gibt es jederzeit Update und Support, nicht frei Haus, aber in garantierten 8 Stunden Businesszeit.

Dies war jetzt ein kurzer Werbespot – es wird der einzige im ganzen Skript sein. Der Autor sieht sich dazu aber veranlasst, weil er fast täglich Mails oder Telefonate erhält, die selbst diesen bescheidenen Obolus sparen möchten, indem sie es als Selbstverständlichkeit ansehen, dass es zu einem Open Source Produkt absolut kostenfreien Support gibt, und zwar unbegrenzt und in jedwelcher Form (z.B. inklusive lange Telefonate ins Ausland, um nur ein Beispiel zu nennen).

3 ArchivistaVM live!

Im Rahmen einer Live-Sitzung kann nicht alles gezeigt werden, was machbar ist. Die Live-Demo soll eine Übersicht über ArchivistaVM bieten, nicht mehr und nicht weniger.

3.1 Demo mit 'custom ramonly'

Die einfachste Art, einen ArchivistaVM-Rechner hochzufahren, stellt das Starten mit 'custom ramonly' dar. Dabei müssen einzig die IP-Adressen eingegeben werden, und schon steht der ArchivistaVM-Server zum Arbeiten bereit.

3.2 Vorbereitungen für Cluster

Um einen Cluster zu erstellen, kann in unserem Shop kostenfrei eine CD erstellt werden:
shop.archivista.ch/oscommunity/catalog

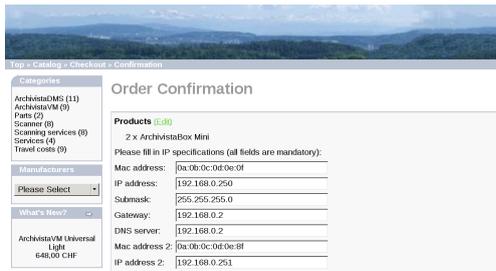
Wählen Sie zunächst bei ArchivistaVM die ArchivistaBox Mini aus:



Wichtig ist weiter, dass Sie zwei Exemplare bestellen:



Bevor Sie die kostenfreie Bestellung absenden, geben Sie die gewünschten IP-Kenndaten für die beiden gewünschten Rechner aus:



Ihre persönliche ArchivistaBox Mini CD wird nun erstellt. Per Mail erhalten Sie den Download-Link.

3.3 Lizenz und Copyright

Die Sourcen von ArchivistaDMS, ArchivistaERP und ArchivistaVM unterstehen der GPLv2-Lizenz. Diese können bezogen werden unter:

www.archivista.ch/de/media/archivista-gpl.tgz

Weiter sei an dieser Stelle kurz erläutert, unter welcher Lizenz bzw. Copyright die ArchivistaBox-ISO-Dateien verfügbar sind. Dazu finden sich auf der Download-Seite die folgenden Ausführungen:

Im Unterschied zu den Sourcen von ArchivistaDMS, ArchivistaERP und ArchivistaVM, die der GPL-Lizenz unterstehen, ist dies beim Handbuch, dem Installer sowie den Logos nicht der Fall. Sie dürfen das Handbuch, den Installer, die Strings und/oder unsere Logos weder kopieren, verändern noch weiterverteilen. Archivista ist eine registrierte Wort-/Bild-Marke. Es ist daher nicht gestattet, diese in anderer Form als auf den unmodifizierten ISO-Datei(en) zu verwenden.

Die Community-Version ist ausschliesslich für den privaten Gebrauch gedacht, eine gewerbliche Nutzung ist nicht zulässig. Nicht erlaubt dagegen sind (nicht abschliessend) das Anbieten der ISO-Dateien gegen Entgelt (insb. auch Unkostenbeitrag), das Einbinden der ArchivistaBox-CD in eine andere Distribution, das Anbieten von kommerzieller Schulung und Support sowie das Verwenden der Logos in irgendeiner Form.

Diese Genauigkeit ist uns deshalb wichtig, weil wir nicht möchten, dass unter dem Namen Archivista CDs in den Umlauf gelangen, die z.B. ungefragt Festplatten auch dann formatieren, wenn diese bereits Daten enthalten. In einem solchen Falle kann die Festplatte mit der originären ArchivistaBox nicht automatisch gelöscht bzw. neu formatiert werden; bei einer modifizierten CD ist dies möglich (es mag sogar Fälle geben, wo es z.B. in Testumgebungen sinnvoll ist). Und daher gibt es diese Einschränkungen. Wo Archivista draufsteht, ist auch Archivista drinnen.

3.3.1 Einsatz Community-Version

Die Community-Versionen der ArchivistaBox sind so ausgelegt, dass im privaten Umfeld sehr effizient gearbeitet werden kann. Der gewerbliche Einsatz ist nicht gestattet. Es ist klar, dass diese Grenze fließend ist.

Open Source Software, die von Firmen stammt, wird oftmals für den privaten Gebrauch kostenfrei zur Verfügung gestellt. Dies bedeutet, dass der Einsatz der ArchivistaBox im privaten Umfeld klar zulässig, innerhalb einer Firma dagegen explizit ausgeschlossen ist. Selbst wenn ich als Privatperson gewerblich mit einem Produkt arbeite, ist dies lizentechnisch nicht durch die Private-Use-Regelung abgedeckt.

Die Business-Editionen bieten weiter einen auf Firmen ausgerichteten Umfang. Die Unterschiede bei ArchivistaVM als Community-Version zur Business-Edition sind ausführlich beschrieben unter:

<http://www.archivista.ch/de/pages/aktuell-blog/archivistabox-2012i.php>

3.3.2 Community-Version

- Zielgruppe: Private ohne Support
- Hardware: Ohne Hardware
- Anwendung: Ausschliesslich privater Einsatz
- Handbuch: Mini-Handbuch (ca. 100 Seiten)
- Stabilität: Tägliche Builds (bei Neuerungen)
- Festplatten: 2 Platten (Raid1)
- Virtualisierung: Unterstützung 1 Socket
- Cluster-Modus: DRBD ohne Software-RAID

3.3.3 Business-Edtion VM

- Zielgruppe: Unternehmen/Private mit Support
- Hardware: Mit oder ohne Hardware (ArchivistaVM Light)
- Anwendung: Kommerzieller Einsatz
- Handbuch: PDF-Datei (ca. 300 Seiten) und Online
- Stabilität: Build-On-Request (eigene CD pro Kunde)

- Festplatten: bis 24 Platten (Raid1, Raid10, Raid0, RaidX)
- Virtualisierung: Beliebige Anzahl von Sockets
- Closter-Modus: DRDB mit Software-RAID

3.3.4 Abschliessende Bemerkungen

Natürlich würde der Autor gut gerne den ganzen Tag neue Features implementieren, welche in allen Versionen für alle Benutzer/innen zur Verfügung stehen. Das ginge solange gut, wie die Firma Archivista GmbH das Gehalt am Monatsende überweisen könnte bzw. würde. Allerdings liegt hier doch das Problem, dass die Entwicklung letztlich auch durch die Community getragen werden müsste. Dies wird bei Open Source (zumindest ist der Autor davon überzeugt) im Grundsatz umso schwieriger, je spezifischer ein Produkt auf bestimmte Aufgaben ausgerichtet ist. Dies deshalb, weil die Community bei stärker spezifizierten Projekten eher klein(er) sein wird.

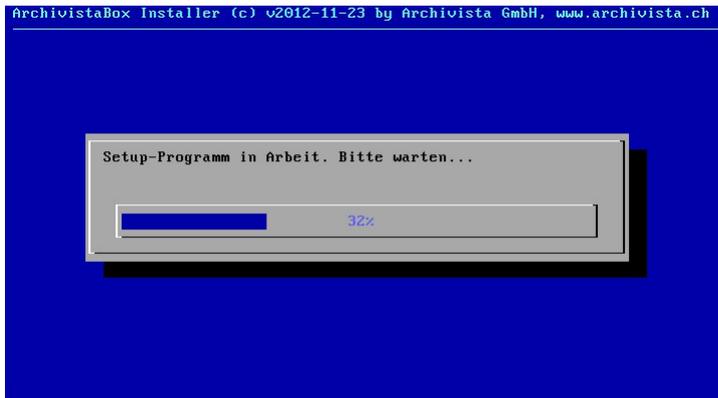
Bei einem DMS-Produkt ist ein reines Community-Projekt daher eher schwierig, bei der Virtualisierung wäre die Bandbreite der Einsatzzwecke zwar höher, aber auch hier gilt, praktisch alle Projekte werden durch Firmen gepflegt. Oder anders herum gesagt, hätte es damals auf dem Markt bereits ein gutes webbasiertes Frontend mit einem schlanken OS-Unterbau gegeben, so wäre ArchivistaVM wohl nicht entstanden. Mag sein, dass hier Desktop-Virtualisierung eine Alternative darstellen könnte, aber Desktop-Virtualisierung lässt sich kaum automatisieren.

Sagen wir es so, eine Firma, die hardwarenahe Software entwickelt, wird nur schon bei der Automatisierung der Testumgebungen sehr schnell an Grenzen stossen. Eine Firma wie Archivista GmbH könnte weder ArchivistaDMS noch ArchivistaVM ohne ArchivistaVM entwickeln. Gäbe es Alternativen? Es gibt immer Alternativen! Produkte wie Ganeti sind aber deutlich aufwändiger zu implementieren. Für grosse bis extrem grosse Firmenstrukturen (Ganeti ist ein Google-Projekt) mag der Einarbeitungsaufwand von einigen Personentagen nicht entscheidend sein. Der Aufwand an Personal wird im tiefsten Promille-Bereich (bzw. einem Bruchteil davon) liegen, in einer KMU sind es schnell ein paar Jahres-Prozente der gesamten Lohnsumme, als mitnichten nichts.

ArchivistaVM wurde und wird daher für KMU-Umgebungen entwickelt, um Server-Virtualisierung mit einfachsten Mitteln zu betreiben. Mag sein, das nicht jedes Feature der Business-Edition in der Community-Version zur Verfügung steht. Solange aber ein kleiner DRBD-Cluster als Bausatz im Bereich von 500 Euro machbar ist, solange dürfte die Community-Version anderen Lösungen, bei der bereits der Hardware-Kontroller um die 500 Euro 'wegfrisst', eine gute Wahl darstellen.

3.4 Hochfahren eines Clusters

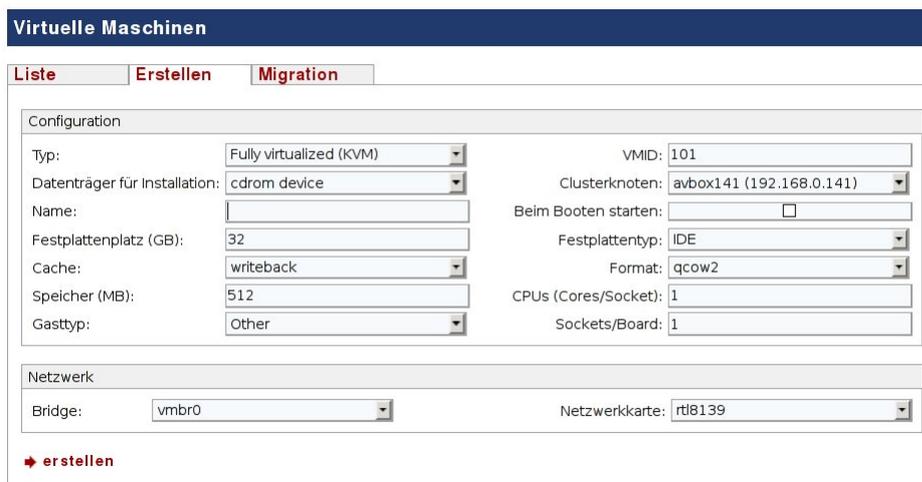
Um den Cluster einzurichten (die Platten müssen leer sein), können Sie die vorbereiteten Sticks einschieben und auf den Start-Knopf aller Rechner drücken. Nach etwa 1 Minute bleibt das Setup bei 92 Prozent stehen. An dieser Stelle wird der Cluster eingerichtet. Nach etwa einer weiteren Minute sollte das Initialisieren des Clusters beendet sein; Sie gelangen dabei zum Login-Screen von ArchivistaVM.



Die Installation wird auf jedem Rechner in der Datei `'/ram.log'` protokolliert. Sofern der Cluster nicht korrekt aufgebaut wurde, sollte diese Datei begutachtet werden.

3.5 Installation von Gästen

Gäste können in einem Cluster auf jedem verfügbaren Knoten eingerichtet werden. Alle Gäste werden aber über den Master verwaltet. Meistens werden dazu zunächst ISO-Dateien bereitgestellt. Diese wiederum können (sofern sie unter 2 GByte gross sind) über Hochladen direkt ins Verzeichnis `'/var/lib/vz/template/iso'` des Master-Knotens geschrieben werden.

The image shows a web-based configuration interface for virtual machines. The title is "Virtuelle Maschinen". There are three tabs: "Liste", "Erstellen", and "Migration", with "Erstellen" being the active tab. The configuration is divided into two sections: "Configuration" and "Netzwerk".
In the "Configuration" section, the following fields are visible:
- Typ: Fully virtualized (KVM) (dropdown)
- Datenträger für Installation: cdrom device (dropdown)
- Name: (empty text field)
- Festplattenplatz (GB): 32 (text field)
- Cache: writeback (dropdown)
- Speicher (MB): 512 (text field)
- Gasttyp: Other (dropdown)
- VMID: 101 (text field)
- Clusterknoten: avbox141 (192.168.0.141) (dropdown)
- Beim Booten starten: (checkbox, unchecked)
- Festplattentyp: IDE (dropdown)
- Format: qcow2 (dropdown)
- CPUs (Cores/Socket): 1 (text field)
- Sockets/Board: 1 (text field)
In the "Netzwerk" section:
- Bridge: vmbro (dropdown)
- Netzwerkkarte: rtl8139 (dropdown)
At the bottom left of the configuration area, there is a red button labeled "erstellen".

Wichtig zu wissen ist dabei, dass die ISO-Dateien automatisch auf alle Knoten verteilt werden. Dieser Vorgang kann einige Sekunden Zeit in Anspruch nehmen.

3.6 Migration von Gästen

Im Unterschied von früheren Versionen können beim Release 2012/XI erstmalig RAW-Instanzen mit 'sparse' migriert werden. Dabei geht es darum, dass beim RAW-Format nur der tatsächlich verwendete Platz der virtuellen Festplatte in der sogenannten Sparse-Datei auf dem VM-Host verwendet wird. Wurde nun eine solche Instanz auf einen anderen Knoten gezügelt, so verlor die Datei diese Sparse-Funktionalität.

Dies hatte den unschönen Nebeneffekt, dass eine zuvor als 32 GByte formatierte Platte, die effektiv z.B. 4 GByte Platz benötigte, danach die gesamten 32 GByte in Anspruch nahm. Dies war nicht nur unschön, es bestand auch die Gefahr, dass bei einer späteren Rückmigration urplötzlich der notwendige Platz fehlte (4 GByte wären allemal noch da, 32 GByte hingegen nicht, diese werden nunmehr aber benötigt).

3.7 Datensicherung

Bei ArchivistaVM werden im Grundsatz immer die gesamten Gäste gesichert. Selbstverständlich kann die Datensicherung aber auch von innen heraus mit einer eigenen Backup-Lösung des Betriebssystems durchgeführt werden. Dabei können aber meistens nur die Daten, nicht aber der gesamte Gast, gesichert werden. Dies ist bei den nachfolgenden Sicherungsarten anders, hier werden immer die gesamten Gäste gesichert.

Datensicherung

Jobs

Neuen Job erstellen

Clusterknoten: Mail senden an:

Zielverzeichnis: Alle VMs einschließen:

Externes Gerät: Speichere alte Backups (1-x):

Wochentag: Mon Die Mit Don Fri Sam Son

Startzeit: :

Modus:

Dateien komprimieren:

VMID	Status	Name	VM-Typ	Festplattengröße	
<input checked="" type="checkbox"/>	101	stopped	dms	qemu	32.00 GB

➔ **speichern**

3.7.1 Stop-Modus

Bei der Datensicherung im Stop-Modus werden die Instanzen heruntergefahren. Danach erfolgt die Datensicherung auf die externe Platte oder aber in ein Verzeichnis auf des lokalen Knotens. Nach dem vollständigen Backup werden die Instanzen wieder hochgefahren. Als Nachteil des Stop-Modus resultiert, dass während der Sicherung die Gäste nicht zur Verfügung stehen.

3.7.2 Cluster-Modus

Die Datensicherung im Cluster-Modus arbeitet effizienter als jene im Stop-Modus, der Cluster-Modus bedingt aber (wie der Name sagt) einen ArchivistaVM-Cluster. Dabei werden die Instanzen ebenfalls kurz heruntergefahren. Sobald dies der Fall ist, wird die Datensicherung auf dem alternativen Rechner (secondary) gestartet. Dabei wird die Platte auf dem zweiten Rechner aus dem DRDB-Verbund ausgeklinkt. Sobald dies der Fall ist, kann der Gast auf dem ersten Rechner wieder hochgefahren werden. Auf dem zweiten Rechner erfolgt nun die Datensicherung. Ist diese beendet, werden die Differenzen vom ersten auf den zweiten Rechner nachgeführt.

3.7.3 Nachtrag zum 'Patzer' anlässlich des Vortrages

Anlässlich des Vortrages funktionierte die cluster-Datensicherung in dem Sinne nicht, als dass der gezeigte gesicherte Gast von der Sicherungsplatte nicht hochgefahren werden konnte. Zunächst sei hier angemerkt, dass das Überprüfen der Sicherung in ca. einer Minute nur daher möglich war, dass es 'Restore-on-the-fly' in ArchivistaVM überhaupt gibt. Weiter sollten im produktiven Umfeld die Optionen 'unsave' und 'qcow2' mit einer gewissen Vorsicht zum Einsatz kommen; wobei dies nicht die Ursache war. Vielmehr war es so, dass die verwendete USB-Platte (älteres Modell) den Strom über den USB-Port bezog und nicht sauber horizontal (mangels längerem USB-Kabel der Platte sowie den engen Platzverhältnissen) ausgerichtet war. Eine horizontal flach liegende Platte verbraucht weniger Strom als eine irgendwie schräg liegende. Dies wiederum ist dann zentral, wenn dieser direkt über den USB-Port bezogen werden muss, und der 'Saft' hier an sich bereits bei einer flach liegenden Platte knapp ist.

Natürlich sind Probleme bei der USB-Übertragung in den Log-Dateien protokolliert, aber dies ändert nichts daran, dass eine Datensicherung, die nicht sauber durchgeführt wird, unbrauchbar ist. Die Erfahrung des Autors ist, dass Log-Dateien zu wenig überprüft werden. Insofern war der Autor anlässlich des Vortrages kein gutes Vorbild, weil er die Log-Datei auch nicht überprüfte – die Quittung erhielt er anlässlich des Vortrages keine Minute später. Nur, im Alltag erfolgt die Quittung leider meist erst dann, wenn es zu spät ist, d.h. wenn im Konfliktfall die Datensicherung nicht wiederhergestellt werden kann. In diesem Sinne würde der Autor den Rat abgeben, die Log-Dateien nach jeder Sicherung anzusehen. Zudem sollten nur USB-Festplatten verwendet werden, die einen externen Stromanschluss haben. Der Patzer hätte sich damit vermeiden lassen. Der Autor möchte sich für den Patzer an dieser Stelle nochmals entschuldigen und versichert, das Kabel in Zukunft dabei zu haben.

3.8 Neustart und Updates

ArchivistaVM kann jederzeit neu gestartet werden. Dazu muss (ab Release 2012/XI) der USB-Stick nicht (mehr) eingelegt sein. Der Inhalt des Sticks wurde bereits beim Hochfahren auf die internen Festplatten (erste Partition) geschrieben. Selbstverständlich werden beim Start die richtigen Einstellungen wieder aktiviert.

Um eine ArchivistaBox später in eine neue Version zu überführen, kann ein Stick mit einem neuen Release gebootet werden. Dabei wird das neue Release wieder auf die Festplatte geschrieben.

3.9 Restore-on-the-fly

Damit 'Restore-on-the-fly' bei der Datensicherung aktiviert ist, muss die ArchivistaBox CD mit dem Namen 'vmbbackup.iso' (bei ArchivistaDMS wäre es dmsbackup.iso) ins Verzeichnis '/var/lib/vz/template/iso' kopiert werden. Danach wird der Inhalt dieser CD jeweils bei der Datensicherung mit auf die externe Festplatte geschrieben.

Um nun ein Restore-on-the-fly zu starten, kann die so erstellte Datensicherung an einem 64-Bit-Computer (Atoms gehen zwar, sind aber mangels Chipsätze für die Virtualisierung sehr langsam) angehängt werden. Der Computer muss dabei neu gestartet werden, wobei beim Starten die Festplatte der Datensicherung zum Booten auszuwählen ist.

Danach wird direkt ArchivistaVM hochgefahren, allerdings erfolgt der Start in einem speziellen Backup-Modus. Es werden immer zwingend die IP-Kenndaten erfragt, dagegen wird kein Auswahlmenü für die Festplatte/n angezeigt, die Instanzen werden direkt ab der Sicherungsplatte (immer die erste Partition) so eingebunden, dass mit diesen gearbeitet werden kann.

Achtung: Es empfiehlt sich, darauf zu achten, dass der Rechner entweder nicht mit dem Netzkabel nach aussen verbunden ist (Kabel ausstecken!), oder es gilt sicherzustellen, dass die Instanzen, die ab der Sicherungsplatte gestartet werden sollen, nicht bereits an anderer Stelle gestartet sind, da ansonsten sowohl Konflikte bei den IP- als auch Mac-Adressen auftreten.

Wichtig zu wissen ist weiter, dass im Backup-Modus die gemachten Einstellungen (insbesondere die IP-Kenndaten) beim Runterfahren nicht zurück auf die Festplatte geschrieben werden. Damit wird verhindert, dass diese Einstellungen permanent (bei einem weiteren Start im Backup-Modus) zur Anwendung kommen.

4 RAM-Modus und Ausblick

Der Autor erlebt es oft, dass der RAM-Modus entweder nicht oder gar falsch verstanden wird. Zuweilen wird er sogar als Störefried betrachtet. Daher sei dem RAM-Modus ein eigenes Kapitel gewidmet. Etwas breiter gefasst erfolgt in diesem Kapitel ein Rückblick, eine Momentaufnahme und auch ein Blick in die Zukunft, selbstverständlich immer auch im Hinblick auf den RAM-Modus, aber nicht nur.

4.1 Historisch gesehen

Als 2010 die neue ArchivistaBox mit 64 Bit entstanden ist, wurden einige Linux-Systeme genauer unter die Lupe genommen. Dabei hat uns TinyCoreLinux gut gefallen, eine schlanke Distribution (ca. 10 MByte), die in ca. 20 Sekunden komplett im RAM aufgebaut werden kann. Allerdings war/ist TinyCoreLinux auf 32 Bit fokussiert. Dies hinderte uns aber nicht daran, das Konzept beim Installer zu verwenden. Folglich ist dabei die RAM basierte Installation entstanden.

Da von TinyCoreLinux bekannt war, dass es möglich sein muss, einen X-Server direkt im RAM zu starten, und sich der Autor letztes Jahr im Urlaub über die lange Installationszeit eines Windows-Netbooks ärgerte, entstand die Idee, es wäre toll, wenn die ArchivistaBox auf einem beliebigen Rechner gestartet werden könnte, ohne das darunterliegende Betriebssystem zu zerstören (es ging dabei eher ums Prinzip als um Liebe zu Windows).

4.2 Rückmeldungen vom Support

Parallel dazu stellte sich im Support heraus, dass das Booten von Festplatten mit Grub1 mitunter seine Tücken haben kann. Gerade bei mehreren Platten musste nach der Installation sichergestellt werden, dass ab der richtigen Platte gebootet wird, weil z.B. Bios und Betriebssystem die Festplatten in geänderter Reihenfolge abbilden.

Die Installation der ArchivistaBox auf sämtliche Festplatten (jeweils erste Partition) hätte ab ca. 2 bis 4 Platten deutlich zu lange gedauert. Und war im BIOS die Bootreihenfolge klar geregelt, so konnte es je nach Hauptplatine passieren, dass bei neu eingehängtem USB-Stick die Reihenfolge im BIOS zerstört wurde. Weiter gab es Probleme mit einem Debian-Skript (mdadm), welches die Festplattenverbünde hochfahren sollte. Plötzlich fehlte da z.B. `'/dev/md1'`, dafür erschien plötzlich `'/dev/md120'`, auch dies führte nicht unbedingt zu erfreulichen Resultaten.

Diese Phänomene waren selten, nur sehr wenige Kunden haben im produktiven Umfeld den Patzer überhaupt gesehen, aber sie waren/sind in der Entwicklung und überhaupt als Gefühl unschön. Interessanterweise gab es diese Probleme bei der Installation im RAM nicht. Alle Platten wurden korrekt erkannt, die Reihenfolge der Platten stimmte immer. Und ja, der RAM-Modus war natürlich sehr schnell, um Faktoren schneller als jede aktuelle SSD-Festplatte.

4.3 RAM für alle Fälle

Daraus ist irgendwann im März dieses Jahres die Idee entstanden, anstelle der aufwendigen Arbeiten, um Grub2 mit allen Möglichkeiten und in gesamtem Umfang zu unterstützen, direkt und in Zukunft komplett auf den RAM-Modus umzusteigen.

Bereits im Mai 2012 wurde ArchivistaVM als Cluster live im RAM-Modus anlässlich des Linuxtages 2012 in Berlin gezeigt. Mag sein, die Kommunikation, dass der RAM-Modus für alle ArchivistaBox-Systeme kommen würde, nicht genügend klar erfolgte, aber offen und ehrlich gesagt ist es wohl eher so, dass wir heute Informationen zwar irgendwie wahrnehmen, aber keinesfalls handeln, bevor es uns betrifft.

Nur so ist es zu erklären, dass mit der Einführung der ArchivistaBox 2012/IX einige Community User förmlich überrascht wurden, weil es die konventionelle ArchivistaBox-ISO-Dateien nicht mehr gab (die Übergangsfrist betrug immerhin fast ein Jahr). Dabei ist ein Umstieg von der konventionellen zur RAM-basierten ArchivistaBox keine Hexerei, die gewünschte Daten-Partition kann direkt beim Hochfahren mit `'datadisk=/dev/sda4'` problemlos und ultimativ angesprochen werden. Auch wenn etwas Handarbeit benötigt wird, so ist der Umstieg deutlich einfacher als die früheren Migrationen sowohl bei ArchivistaDMS wie ArchivistaVM.

4.4 Mehr, nicht weniger Flexibilität

Der Autor stellt weiter fest, dass eine ISO von ca. 330 MByte bei ArchivistaDMS (Verkleinerung auf unter die Hälfte) bzw. 100 MByte bei ArchivistaVM (Verkleinerung auf unter einen Drittel) kaum wohlwollend zur Kenntnis genommen werden. Dass dabei die Boot-Zeiten um den Faktor 1:2 oder gar 1:3 schneller sind, dass der Speicherbedarf in einem Jahr von 4 auf ca. 1.7 GByte bei ArchivistaDMS bzw. von 2 auf ca. 700 MByte bei ArchivistaVM gesenkt wurde, dies alles erforderte aber viele harte Stunden, Wochen, ja Monate an Entwicklungszeit.

Dass mit dem RAM-Modus sämtliche selbst zerschossenen Installationen ganz einfach durch einen Neustart wieder gerade gerichtet werden, wird ebenso kaum erkannt, vielmehr wird bemängelt, es könnten keine Pakete mehr nachinstalliert werden. Einmal abgesehen davon, dass sehr wohl Pakete nachinstalliert werden können (am Vortrag wird epdfview in einigen

Sekunden nachinstalliert), bietet doch gerade der RAM-Modus die Möglichkeit, Pakete nach Lust und Laune aufzuspielen, ohne Gefahr zu laufen, dass etwas beim nächsten Neustart nicht mehr funktionieren wird. Es kostet einen einigermaßen geübten User vielleicht ein paar Minuten, um die wirklich gewünschten Pakete mit einem Skript automatisiert nach dem Hochfahren wieder zu aktivieren, ist das nun wirklich so kompliziert? Und dann gab es da erst kürzlich die Supportmeldung, die Platte könne nicht platt gemacht werden, auch wenn dies dank RAM-Modus mit dd ein Kinderspiel ist.

4.5 Mehr Tempo mit weniger Hardware

Der Autor sieht bis heute keine Nachteile beim RAM-Modus. Dagegen ermöglicht der RAM-Modus in der Entwicklung ein Tempo einzuschlagen, das ansonsten nur mit extrem teuren (und mässig stabilen) SSD-Platten machbar wäre. Die ArchivistaBox läuft im RAM-Modus selbst auf leistungsschwachen Rechnern (C- und E-Serie von AMD, bei ArchivistaDMS reicht gar ein Atom) erstaunlich schnell und flott, wie es mit keiner Festplatte so machbar wäre.



Zur Erinnerung, wie sonst wäre anlässlich dieses Vortrages machbar, den Aufbau eines kompletten Clusters mit 18 Watt Rechnern in weniger als drei Minuten und komplett automatisiert zu zeigen, wenn nicht im RAM-Modus? Hätte jemand vor einem Jahr gefragt, ob es sinnvoll ist, die ArchivistaBox (insbesondere ArchivistaDMS) auf einem solchen Mini-Rechner zu betreiben, so hätte der Autor eher davon abgeraten. Heute ist dies möglich. Offen gestanden hätte der Autor es vor einem Jahr als unmöglich erachtet, mit Debian die gesamte ArchivistaVM-Lösung in unter 100 MByte zu packen. Dank dem RAM-Modus konnte dieses Feature (und viele andere ebenfalls) erst realisiert werden, in deutlich kürzerer Zeit und weit effizienter.

4.6 Blick in die Zukunft

Die Entwicklung der ArchivistaBox wird wohl nie abgeschlossen sein. Ein System, das sich nicht weiterentwickelt, wird früher oder später immer Gefahr laufen. nicht mehr à jour zu

sein. Dank dem RAM-Modus besteht nun eine weit leistungsfähigere Entwicklungsumgebung. Daher kann die ArchivistaBox heute wie morgen mehr neue Features realisiert werden. Dabei gibt es keinen fixen Menüplan, die ArchivistaBox wird vielmehr in agiler Art und Weise (d.h. ohne starre Zyklen) ständig dort erweitert, wo Bedarf besteht.

Einige Beispiele, wohin die Reise gehen könnte, seien hier (ohne Anspruch auf eine Realisierung) angeführt: Einfacheres Erweitern eines Clusters, Stromsparfunktionen (z.B. Abschalten einzelner Cores) sowie das Durchreichen von USB-Geräten direkt im Web-Browser (dies erfordert heute noch ganz wenig Handarbeit im Browser). Die tollsten Ideen entstehen aber in der täglichen Arbeit mit ArchivistaVM, an dieser Stelle sei allen gedankt, die mit wertvollen Feedbacks dazu beigetragen haben, dass ArchivistaVM heute mit diesem Umfang zur Verfügung steht. Herzlichen Dank für die Aufmerksamkeit.