

Thomas-Krenn Open-E RA1112 Metro Cluster



Spieglein, Spieglein im RZ

von Jürgen Heyer

Mit dem Open-E RA1112 (All-Flash) Metro Cluster hat Thomas-Krenn eine sehr kompakte Speicher-Appliance aufgelegt, die auf dem robusten ZFS-Dateisystem und dem bekannten Betriebssystem Open-E JovianDSS basiert. Dabei bilden zwei identische Server einen Metro Cluster, um durch einen synchronen Spiegelbetrieb eine hochverfügbare Datenhaltung zu gewährleisten. Wir haben uns diesen Server-Doppelpack im Labor genauer angesehen.



Thomas-Krenn aus dem niederbayerischen Freyung bietet seit 20 Jahren Hardware nach dem Build-to-Order-Prinzip mit Fokus auf individuelle Server- und Speichersysteme an. Für den Bau der Systeme setzt der Hersteller auf renommierte Hard- und Softwarekomponenten, was sich auch in dem von uns getesteten Metro Cluster widerspiegelt. Mit nur einer Höheneinheit im 19-Zoll-Rack zeichnet sich der Server durch eine sehr kompakte Bauform aus, wobei sich die benötigte Nutzkapazität durch unterschiedliche Plattenbestückungen individuell konfigurieren lässt.

Der Betrieb des Open-E RA1112 ist auch als Einzelsystem möglich, laut Thomas-Krenn werden diese Appliances aber mittlerweile überwiegend als Paare für einen Betrieb im Failovercluster mit synchroner Datenspiegelung verkauft. Idealerweise laufen diese an zwei Standorten, um so im Katastrophenfall und dem Verlust eines Standorts keine Daten zu verlieren und weiter Zugriff zu haben. Hinsichtlich der Entfernung sind Kabellängen von bis zu 100 Meter Länge Standard, mit speziellen und teureren Transceiver-Kabeln sind bis zu 80 Kilometer möglich.

Praktischerweise konnte uns Thomas-Krenn in Zusammenarbeit mit Open-E und Intel trotz aktueller Lieferprobleme rechtzeitig ein topaktuelles Serverpaar

zur Verfügung stellen, das bei Erscheinen dieses Tests allgemein erhältlich sein soll. Anpassungen an die jeweilige Umgebung und Netzwerkinfrastruktur sind durch das Baukastensystem ganz individuell je nach Lieferfähigkeit im Onlineshop konfigurierbar.

Bewährte Hardware von Asus und Intel

Basis des RA1112 ist der ASUS-Server RS500A-E11-RS12U, der mit vollem Remotemanagement mit KVM over LAN und IPMI 2.0 über einen eigenen Netzwerk-Port ausgestattet ist. Serienmäßig verbaut sind zwei weitere 1-GBit/s-Netzwerkanschlüsse auf RJ45-Basis, von denen in der eingerichteten Konfiguration einer zur Verbindung zum Managementnetz für das Open-E-Betriebssystem zum Einsatz kam.

Im Testgerät steckten zusätzlich zwei weitere Dual-Port-Netzwerkkarten aus dem Hause Broadcom, eine BCM57416 mit 10-GBE-Base-T-Anschlüssen, passend für die Anbindung an unser Datennetzwerk, sowie eine BCM57414 mit 25-GBE-SFP28-Anschlüssen als doppelte Direktverbindung zwischen den beiden Servern für die Datenspiegelung. In das Gerät lassen sich für die Datenspeicherung von vorn in 2,5-Zoll-Hot-Plug-Rahmen zwölf SAS- oder SATA-Festplatten oder -SSDs (NVMe) einschieben.

In unserem Testgerät waren zwölf 6,4-TByte-SSDs vom Typ Intel Solidigm NVMe D7-P5620 eingebaut. Für das Betriebssystem sind im Inneren zwei M.2-Steckplätze vorgesehen, in denen im Testgerät zwei NVMe-SSDs mit je 480 GByte Kapazität steckten, die im Spiegelbetrieb liefen. Aufgrund dieser All-Flash-Ausstattung waren keine zusätzlichen Datenträger für Write-Logs oder als Lese-Cache erforderlich, wie wir es von früheren Teststellungen mit ZFS-Dateisystem und Open-E JovianDSS-Betriebssystem kannten.

Das Einschalten der Appliance erfolgt über einen kleinen Taster, der auf der Front im schmalen Rahmen unter den Hot-Plug-Einschieben zu finden ist. Zwei weitere Taster ermöglichen einen Reset und das Schalten der Ortungs-LED. Einige LEDs im Rahmen zeigen Plattenzugriffe sowie Netzwerkverkehr an. Identische Taster samt LEDs sind auf der Rückseite zu finden. USB-Port gibt es in der Front keinen, auf der Rückseite sind es zwei. Dazu kommt ein VGA-Anschluss für einen Monitor. So ist es möglich, bei Bedarf direkt am Gerät beispielsweise die IP-Adressen anzupassen.

Die Stromversorgung erfolgt über zwei redundante Hot-Swap-Netzteile mit je 800 Watt, was einen Tausch im laufenden Betrieb erlaubt. Zum Öffnen des Gehäusedeckels dient eine fixierte Rändel-

The screenshot shows the Open-E JovianDSS web interface. The top navigation bar includes 'About', 'Help', a trial warning, and 'Logout'. The left sidebar contains menu items: Storage, User Management, Failover Settings, Storage Settings, Backup & Recovery, System Settings, and Diagnostics. The main content area is titled 'System Settings' and has tabs for Administration, Network, MPIO, System, Settings management, Task scheduler, and Update. The 'Network' tab is active, showing an 'Interfaces' section with a table of network configurations. Buttons for '+ Add VLAN' and '+ Create bond interface' are visible above the table.

Name	IP	DHCP	Vendor	Negotiated speed	Cable	Status	Options
bond0	192.168.7.220	No	Ethernet Bonding Driver	25000 Mbps	cable	Active	Options
eth0	192.168.0.220	No	Broadcom Inc. and subsidiaries BCM57416 NetXtreme-E Du...	10000 Mbps	cable	Active	Options
eth1	192.168.6.220	No	Broadcom Inc. and subsidiaries BCM57416 NetXtreme-E Du...	10000 Mbps	cable	Active	Options
eth2 (bond0)	N/A	No	Broadcom Inc. and subsidiaries BCM57414 NetXtreme-E 10...	25000 Mbps	cable	Bond slave	Options
eth3 (bond0)	N/A	No	Broadcom Inc. and subsidiaries BCM57414 NetXtreme-E 10...	25000 Mbps	cable	Bond slave	Options
eth4	192.168.1.220	No	Intel Corporation I350 Gigabit Network Connection (rev 01)	1000 Mbps	cable	Active	Options
eth5	192.168.5.220	No	Intel Corporation I350 Gigabit Network Connection (rev 01)	N/A	no cable	Active	Options

Bild 1: Die Netzwerkkonfiguration eines Clusters ist komplex und erfordert vorausschauende Planung.

schraube, sodass es keine losen Teile gibt, die bei Arbeiten an der Appliance herunterfallen können.

Trotz der vielen Netzwerkanschlüsse war die Verkabelung der Geräte im Handumdrehen erledigt, nicht zuletzt dank der guten Dokumentation von Open-E. Die beiden 25-GBE-Ports der Server werden über ein SFP28-Kabel direkt miteinander verbunden, die mögliche Leitungslänge bestimmt den maximalen Abstand, in dem sich die beiden Server voneinander entfernt aufstellen lassen. Die beiden 10-GBE-Ports jedes Servers werden idealerweise aus Redundanzgründen kreuzweise auf zwei Switches verteilt, der 1-GBE-Port zum Managementnetz jeweils auf einen der beiden Switches. Die verbleibenden Remote-IPMI-Anschlüsse spielen nur eine untergeordnete Rolle, manche Unternehmen nutzen hier eigene VLANs nur für die Remote-Ports, ansonsten finden Sie ebenfalls für das Managementnetz Verwendung.

IP-Adressvergabe erfordert Sorgfalt

Im Gegensatz zur recht einfachen Verschaltung der Netzwerkanschlüsse erfordert die daran anschließende IP-Konfiguration besondere Aufmerksamkeit und Sorgfalt. Sehr hilfreich ist dazu ein Step-by-Step-Guide von Open-E, der mit vielen Screenshots und dazugehörigen Erklärungen die einzelnen Schritte detailliert schildert. Für die Administration der bei-

den Appliances gibt es eine sehr übersichtliche WebGUI, die jeweils über die Managementadressen erreichbar ist. Während die ersten Schritte auf jedem System getrennt vorzunehmen sind, kann die Administration später für den ganzen Cluster über einen Knoten erfolgen.

Obwohl wir unser Testgerät komplett vor-konfiguriert zur Verfügung gestellt bekamen, löschten wir die Einstellungen nochmals, auch weil die eingetragenen IP-Adressen nicht exakt zu unserem Testnetz passten. Zu beachten ist hierbei, dass es für die Kommunikation zwischen den beiden Serverknoten mehrerer Subnetzbereiche (3 x Bonding, 2 x Speicherzugriff, 1 x Management) bedarf und eine Doppelvergabe auf jeden Fall zu vermeiden ist – hier empfiehlt sich das Führen einer genauen Dokumentation. Die beiden Subnetze für den Speicherzugriff müssen dabei unterschiedlich sein zu den Subnetzen, in denen sich die vorgesehenen Clientserver befinden, die dann den Speicher nutzen sollen. Deren Zugriff erfolgt vielmehr analog zu anderen Clusterkonfigurationen über virtuelle IP-Adressen, die sich im Rahmen der Speicherpools definieren lassen.

Bei unseren Testgeräten mussten wir zudem berücksichtigen, dass die eth-Portnummern vom Betriebssystem nicht in gleicher Reihenfolge vergeben waren, sodass wir uns daran nicht orientieren konnten. Vielmehr mussten wir bei der

Thomas-Krenn Open-E RA1112 (All-Flash) Metro Cluster

Produkt

Speichercluster auf Basis von Open-E JovianDSS Linux und ZFS-Dateisystem.

Hersteller

Thomas-Krenn
www.thomas-krenn.com/de

Preis

Der Preis für das zum Test zur Verfügung gestellte System inklusive zwölf Monaten Support liegt bei 69.270 Euro. Die Ausstattung ist individuell konfigurierbar.

Systemausstattung

- Eine Höheneinheit im 19-Zoll-Rack
- Mainboard Asus KMPA-U16
- 1 x AMD EPYC 72F3 (3,7 GHz, 8-Core)
- 512 GByte ECC-DDR4-3200-RAM (8 x 64 GByte)
- 2 x 480 GByte ATP-N600Sc-M.2-NVMe-SSDs mit PLP für das Betriebssystem
- 12 x 6,4 TByte Solidigm-NVMe-D7-P5620 im Hotplug-Rahmen für die Daten
- 2 x 1-Gbit/s-on-Board-LAN (Intel I350-AM2)
- Remotemanagement (KVM over LAN, IPMI 2)
- 1 x Broadcom-BCM57414-Dual-Port-10/25-Gbit-SFP28
- 1 x Broadcom-BCM57416-Dual-Port-10-Gbit-BASE-T
- 2 x 800-Watt-Hot-Swap-Netzteil

Technische Daten

www.it-administrator.de/downloads/datenblaetter

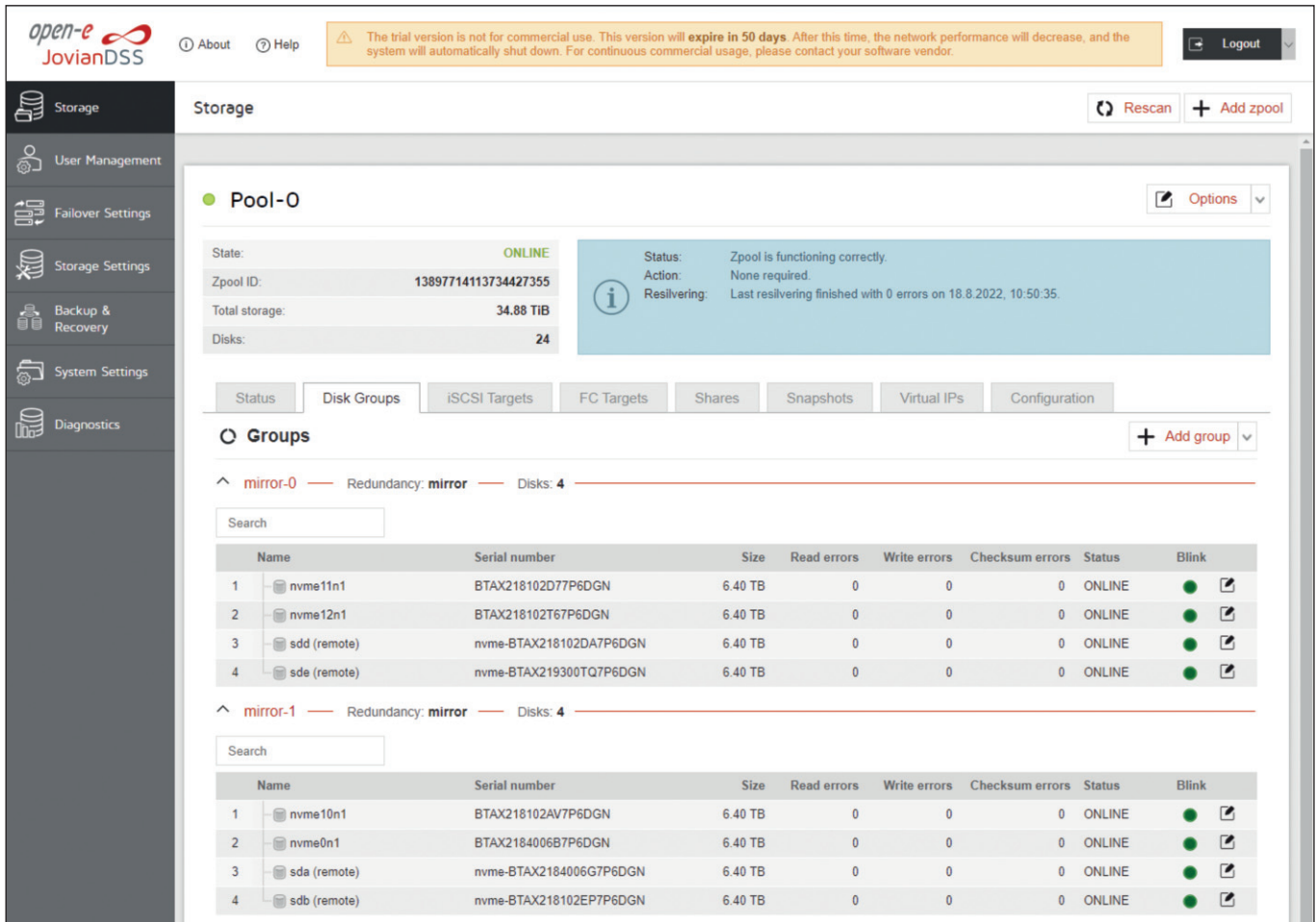


Bild 2: In dieser Ansicht ist gut erkennbar, wie das System für den Cluster-Betrieb mit Spiegelung lokale und entfernte Datenträger zu Disk Groups zusammenfasst.

Konfiguration des Bondings über die 25-GbE-SFP-Ports nochmals eine Verbindung auftrennen, um die Zuordnung zu erkennen. Um nun das Bonding korrekt einzurichten, sind zuerst für die beiden verbundenen Portpaare passende IP-Adressen aus zwei Subnetzen einzutragen. Dann ist auf jedem Server über die beiden Bonding-Ports ein sogenanntes Bonding-Interface zu definieren mit passender IP-Adresse aus einem weiteren Subnetz, um schließlich in einem späteren Schritt darüber einen Spiegelpfad (Mirroring Path) einzurichten.

Die erwähnte Dokumentation der IP-Adressen ist insofern wichtig, als dass nach Anlegen des Bonding-Interfaces in der GUI nicht mehr erkennbar ist, welche IP-Adressen die darunterliegenden Bonding-Ports erhalten haben. Nach dem Erstellen der beiden Bonding-Interfaces ist einem Port noch das Default Gateway zuzuweisen und die Zeitsynchronisation einzurichten, bevorzugt per NTP.

Perfekt geeignet für VMware-Umgebungen

Als Nächstes haben wir die beiden Knoten über das zuvor angelegte Bonding-Interface miteinander verbunden und dann zwei sogenannte Ringe angelegt. Der erste Ring besteht aus den beiden IP-Adressen der Bonding-Interfaces, für den zweiten kann der Administrator ein weiteres Port-Paar wählen, beispielsweise die zwei zueinander passenden Ports für den Datenzugriff durch die Clients.

Danach sind mindestens zwei und maximal sechs Ping Nodes anzugeben, geeignete IP-Adressen in den verschiedenen Subnetzen, mit denen der Cluster später Unterbrechungen im Netzwerk erkennen und eine Split-Brain-Situation vermeiden kann. Anschließend ist der schon erwähnte Spiegelpfad einzurichten, bestehend aus den beiden Bonding Interfaces.

Hilfreich ist, dass die zuletzt beschriebenen Schritte in der WebGUI allesamt auf

einer Seite, genannt "Failover Settings", einzutragen sind und die beiden Knoten stets die Kommunikation prüfen, was bei Erfolg mit einem grünen Hinweis "active" oder "reachable" bestätigt wird. Ist alles grün, kann der Administrator oben auf dieser Seite den Failover-Betrieb starten.

Beim Wechseln auf die "Storage"-Seite konnten wir nun auf zwei Registerblättern sowohl die lokalen Datenträger als auch die entfernten SSDs (remote) des zweiten Knotens sehen. Das ist die Voraussetzung, um einen gespiegelten Datenpool anzulegen, hier Zpool genannt.

Insgesamt lässt sich zur IP-Konfiguration des Clusters feststellen, dass es hier etwas komplizierter klingen mag, als es sich in der Praxis darstellt – gerade wenn der Administrator die grafische Darstellung des Step-by-Step-Guides vor sich und schon etwas Erfahrung mit anderen Cluster-Konfigurationen gesammelt hat. Zu erwähnen ist allerdings noch, dass wir

uns bei der Konfiguration auf die Bereitstellung per iSCSI an eine VMware-Umgebung konzentriert haben. Für die Nutzung als NAS mit NFS oder SMB gibt es in dem Guide etwas abweichende Empfehlungen. Weiterhin existieren eigene Konfigurationshinweise, wenn es beispielsweise nicht möglich ist, das Bonding mittels Direktverbindung zu realisieren.

Failover-Betrieb mit vierfacher Spiegelung

Sofern für den nun folgenden Schritt zum Anlegen eines Zpools noch nicht alle verfügbaren Datenträger angezeigt werden, gibt es die Möglichkeit für einen Rescan. In unseren beiden Knoten standen nun insgesamt 24 Daten-SSDs zur Verfügung, die wir für eine optimale Performance als sechs Spiegelpaare konfigurierten. Dazu wählten wir jeweils zwei NVMEs der beiden Knoten aus, also insgesamt vier SSDs, und bildeten daraus eine Spiegelgruppe.

Dies wiederholten wir sechsmal und hatten am Ende einen Zpool mit rund 34,7 TiByte angelegt. Hot-Spare-Platten standen in diesem Fall nicht zur Verfügung, dazu hätten wir auf eine Spiegelgruppe mit knapp 6 TiByte Kapazität verzichten müssen. In anderen Konfigurationen mit noch mehr Festplatten mögen zusätzliche Hot-Spare-Platten sinnvoll sein, bei unserer Testkonfiguration war aber zu berücksichtigen, dass die Daten durch die Spiegelgruppen jeweils aus vier Platten sowieso viermal identisch vorhanden waren, wodurch die Nutzkapazität auch nur 25 Prozent des Bruttowertes betrug. Selbst beim Ausfall einzelner SSDs besteht in dieser Konfiguration weiterhin Spiegelredundanz, ebenso wenn einer der beiden Knoten ganz ausfällt oder nur neu startet.

In unserem Setup mit einem Zpool über die gesamte Kapazität arbeiteten die beiden Knoten im Active-Passive-Betrieb, weil sich ein Zpool einem Knoten immer nur komplett zuweisen lässt. Legt der Ad-

Bild 3: Beim Anlegen eines neuen Volumes lassen sich Optionen wie Deduplizierung und Kompression auswählen.

ministrator stattdessen zwei oder mehr Zpools an, kann er diese auf die beiden Knoten verteilen und so einen Active-Active-Betrieb fahren.

Wer statt optimaler Performance eine höhere Speichereffizienz anstrebt, kann dies mit Open-E JovianDSS ebenfalls realisieren, muss dazu aber einen Hardware-RAID-Controller mitordern, um mit diesem je Knoten beispielsweise ein RAID-6-Array zu konfigurieren, das dann zwischen den Knoten gespiegelt wird. Das Betriebssystem enthält alle Treiber und Tools für die RAID-Controller von Broadcom (LSI), Microsemi (Adaptec) sowie Areca. Ist der Zpool eingerichtet, muss sich der Admin nochmals mit der IP-Adressvergabe beschäftigen und wie schon erwähnt jedem Zpool eine virtuelle IP-Adresse aus den eigentlichen Speicher-netzen zuordnen. Diese Adressen werden immer von dem aktiven Knoten bedient und können damit bei einem Failover entsprechend wandern.

ZFS-Dateisystem als robuste Basis

Thomas-Krenn hat sich beim Bau des RA1112 dafür entschieden, beim Be-

triebssystem auf JovianDSS von Open-E zu setzen. Herzstück ist das 128-Bit-ZFS-Dateisystem. Es kommt seit vielen Jahren im Linux-Umfeld für Speichersysteme zum Einsatz und ist in der Praxis erprobt. Es gewährleistet nicht nur die Integrität der gespeicherten Daten, sondern unterstützt auch eine fast unbegrenzte Speicherkapazität sowie sehr große Dateien.

Open-E nutzt all diese Vorteile und ergänzt sie um einige nützliche Enterprise-Features. Eine asynchrone Replikation (On- & Off-site Data Protection) lokal oder zu einem entfernten Standort ermöglicht bei einem Totalausfall eine schnelle Wiederherstellung von wichtigen Daten und schützt so zuverlässig vor deren Verlust. Darüber hinaus ist wie bei unserem Test eine Cluster-Konfiguration mit synchroner Spiegelung möglich.

JovianDSS basiert auf Linux und fügt sich dadurch nahtlos in die meisten bestehenden IT-Strukturen ein. Es ist besonders für Nutzer geeignet, die nach einer Unified-NAS- und -SAN-Plattform mit Thin Provisioning, Kompression und Deduplizierung suchen. Trotz der Linux-Basis unterstützt das System neben VMware, Citrix, Linux, RHEL, MacOS, XEN und OpenStack auch Hyper-V. In reinen Microsoft-Umgebungen mag sich allerdings ein auf Windows basierendes NAS-System noch besser integrieren lassen, das hängt aber stets von der individuellen Situation ab.

Bei ZFS handelt es sich um ein 128-Bit-Copy-on-Write-Dateisystem, mit dem sich sehr effizient Snapshots erstellen lassen. Es arbeitet transaktional und wurde speziell für den Serverbereich entwickelt. Die maximale Größe einer Datei kann 16 EiByte (=16*2⁶⁰ Byte) betragen, die Anzahl der verwaltbaren Dateien liegt bei 2⁴⁸. Es besitzt integrierte RAID-Funktionalitäten, ein Volume-Management sowie einen prüfsummenbasierten Schutz vor Dateiübertragungsfehlern bei nur geringem Einfluss auf die Performance.

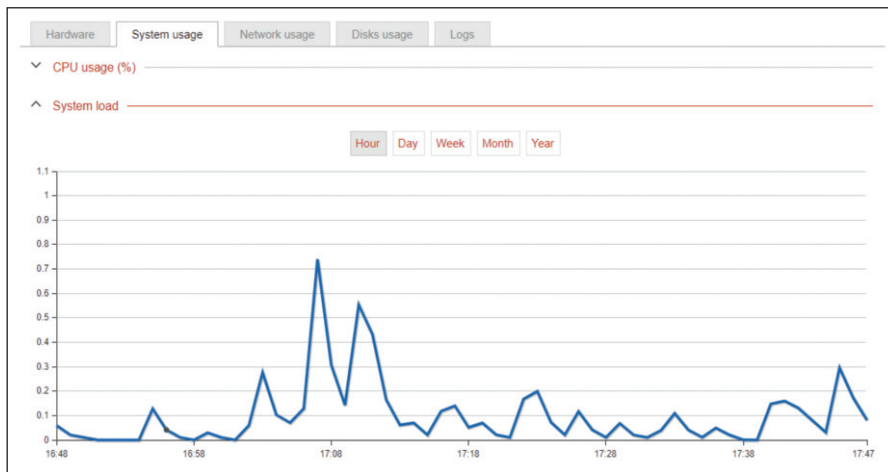


Bild 4: Die Angaben zur aktuellen Lastsituation sind bei JovianDSS recht unübersichtlich auf diverse Menüs verteilt und teils wenig aussagekräftig.

Weiterhin sorgt ZFS dafür, dass das Dateisystem zu jeder Zeit konsistent ist und somit beispielsweise nach einem Stromausfall keine Überprüfung notwendig ist. ZFS unterstützt Deduplizierung, Kompression und Thin Provisioning, was sich in JovianDSS je nach Bedarf aktivieren und deaktivieren lässt.

Übersichtliche WebGUI ohne Dashboard

Auch wenn sich an der Appliance Tastatur und Monitor anschließen lassen, dürfte die Bedienung in der Regel über die WebGUI erfolgen. Die Oberfläche ist mit einem Sieben-Punkte-Menü auf der linken Seite sehr übersichtlich aufgebaut, jedoch vermissen wir ein Dashboard zur Ausgabe der wesentlichen Performance- und Statusdaten.

Zumindest kann der Administrator aber unter dem letzten Menüpunkt "Diagnostics" einen Blick auf einzelne Werte wie die CPU-, System- sowie Speicherauslastung werfen und den Netzwerkverkehr sowie die Plattennutzung getrennt nach Lesen und Schreiben auslesen. Letztere lässt sich dabei nur für jeden Datenträger einzeln betrachten und eine Ausgabe der aktuellen IOPS sowie der Latenz suchten wir vergebens. Hier würden wir uns etwas mehr Übersicht wünschen, denn die tatsächliche Auslastung lässt sich so nur schlecht feststellen.

Ist wie oben beschrieben mindestens ein Zpool angelegt, sind nun je nach Bedarf und Einsatzzweck Volumes oder Datasets

und Freigaben für den eigentlichen externen Zugriff anzulegen. Das RA11112 ist multiprotokollfähig und unterstützt die Blockprotokolle iSCSI und Fibre Channel (FC) sowie die Datei-basierten NFS- und SMB-Freigaben. Die FC-Nutzung setzt voraus, dass sich mindestens ein FC-Hostbusadapter im Gerät befindet, was bei uns nicht der Fall war. Wir entschieden uns daher im Test für iSCSI-Volumes und NFS-Freigaben, um diese in unserer VMware vSphere-Umgebung an die ESXi-Server anzuhängen.

Bei der Anlage eines iSCSI-Targets schlägt der Assistent zuerst einen typischen Standardnamen vor. Anschließend muss der Administrator ein Volume (zvol) anlegen, standardmäßig mit Thin Provisioning, sodass der Speicherplatz nicht reserviert wird. Ein solches Volume lässt sich dann größer definieren als der verfügbare Platz im Pool. Eine Deduplizierung ist möglich, aber per Standard deaktiviert. Zur Kontrolle des iSCSI-Zugriffs unterstützt die Appliance CHAP, Mutual CHAP sowie eine Restriktion anhand von IP-Adressen. Im Test funktionierten die Einrichtung und anschließende Anbindung an unsere drei ESXi-Server problemlos. Ein komplettes Multipathing (MPIO) bis hin zu den Servern konnten wir allerdings mangels einer zweiten 10-GBE-Switches nicht konfigurieren.

Dateibasierte Freigabe möglich

Um statt eines iSCSI-Volumes eine dateibasierte Freigabe zu definieren, ist zuerst ein sogenanntes Dataset anzulegen.

Zur Steuerung der Größe unterstützt JovianDSS sowohl Quotas als auch Reservierungen. Eine Quota beschreibt eine harte Grenze, wie groß ein Dataset werden kann, und beinhaltet dabei eventuelle Snapshots. Die Reservierung sorgt dafür, dass der festgelegte Platz garantiert verfügbar ist. Ist also in einem aktiven Datenpool bereits ein Teil belegt, lässt sich eine Reservierung nicht größer anlegen als der noch freie Platz. Snapshots und Clones gehen in die Reservierung ebenfalls mit ein. Quota und Reservierung lassen sich zudem kombinieren. Innerhalb des Datasets kann der Administrator nun eine Freigabe anlegen und dabei den Zugriff für NFS und/oder SMB erlauben.

Bei einem SMB-Zugriff erlaubt das NAS-System eine Authentifizierung entweder über ein Active Directory oder über LDAP, außerdem über lokale Benutzer und Gruppen. Möglich ist zudem ein offener Gastzugriff ohne Authentifizierung, außerdem eine Beschränkung auf einen lesenden Zugriff. Für eine NFS-Freigabe lässt sich der Zugriff auf IP-Adressen eingrenzen, wahlweise schreibend und lesend oder nur lesend. Im Rahmen unseres Tests integrierten wir den Cluster in unser Active Directory und definierten probeweise einige Freigaben mit unterschiedlichen Zugriffsrechten, was reibungslos funktionierte.

Granularer Remote-Zugriff

Erfreulich granular lässt sich der administrative Zugriff auf die Appliance einstellen. So ist der HTTPS-Zugriff auf die GUI auf bestimmte IP-Adressen einschränkbar. Auch die lokale Konsole konnten wir per Passwort schützen und den REST-Zugriff ein- und ausschalten ebenso wie SNMP und den Zugriff per Remote Konsole (SSH) auf die Kommandozeile (CLI). Logs lassen sich an einen Log-Server weiterleiten und Meldungen per E-Mail versenden.

Um individuell einzelne Datenstände festzuhalten, unterstützt das System manuelle Snapshots für Volumes und Datasets. Für eine Automatisierung lässt sich die Snapshot-Erstellung als Sicherungsaufgabe (Backup Task) einrichten. Hier kann der

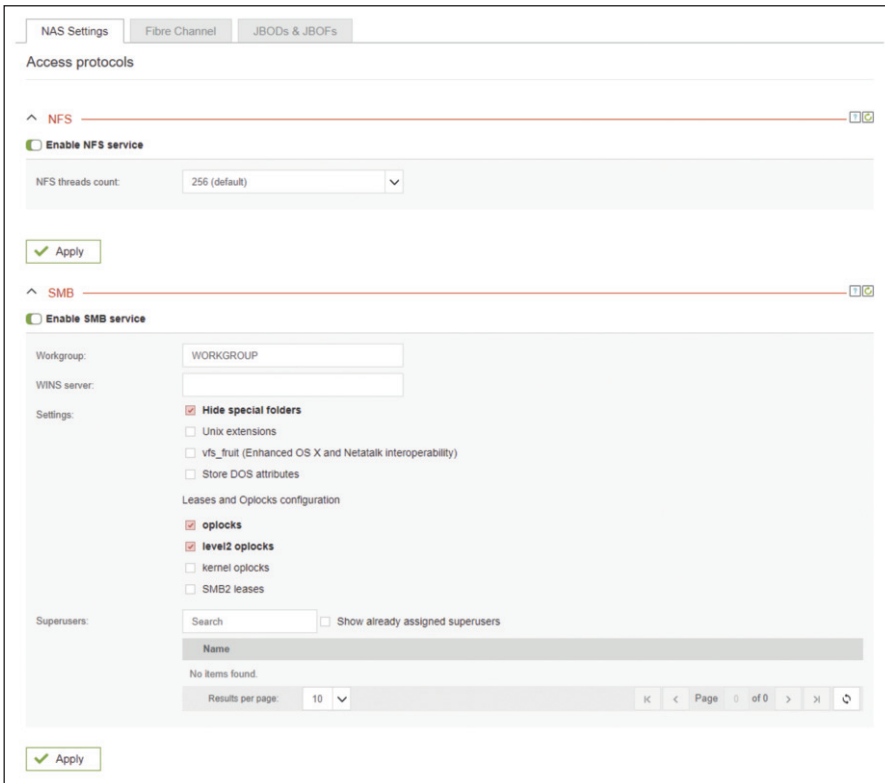


Bild 5: Statt per iSCSI lässt sich auch ein Zugriff per NFS oder SMB erfreulich einfach einrichten.

Administrator eigene Aufbewahrungsregeln definieren, also in welchen Abständen ein Snapshot erzeugt und wie lange dieser dann aufgehoben werden soll.

Überzeugende Spiegelungsperformance

Um unserem Cluster leistungsmäßig etwas auf den Zahn zu fühlen, nutzten wir eine spezielle Linux-VM als Lastgenerator mit verschiedenen Operationen. Auf dieser VM unter CentOS läuft das Unix-Kommando dperf mit verschiedenen Optionen. Wir nutzten hier in erster Linie den IOPS-Test, der einen kontinuierlichen Zufalls-IO erzeugt. Zur Vorbereitung rollten wir diese VM jeweils zweimal auf drei ESXi-Hosts aus, denen wir ein iSCSI-Volume des Clusters zugewiesen hatten, so dass sechs VMs parallel zugegriffen.

Dabei erzielten wir 116.000 IOPS, was ein erfreulich gutes Resultat ist. Zu berücksichtigen ist dabei, dass diese Performance mit synchroner Spiegelung zustande kam.

Im Falle des Betriebs eines einzelnen RA1112 dürfte das Ergebnis sogar noch etwas besser ausfallen. Gemäß Open-E ist das Bonding mit "nur" 25 GBE das

Nadelöhr und Thomas-Krenn überlegt, eventuell zu einem etwas höheren Preis auch eine 100-GBE-Variante anzubieten. Unserer Einschätzung nach diskutieren wir hier allerdings bereits auf einem sehr hohen Leistungsniveau und jeder Interessent sollte im Vorfeld genau die eigenen Anforderungen an Verfügbarkeit und Performance festlegen, um so das optimale Preis-Leistungs-Verhältnis zu erzielen.

Fazit

Der RA1112 Metro Cluster von Thomas-Krenn eignet sich als hochverfügbares Speichersystem für KMUs, die ihre Daten synchron an zwei Standorten halten wollen. Den Haupteinsatzzweck sehen wir vor allem innerhalb von Virtualisierungsumgebungen unter VMware vSphere, Citrix oder KVM. Ein Einsatz in Verbindung mit Microsoft Hyper-V ist ebenfalls möglich und zertifiziert, wobei Microsoft hier Windows-basierende Alternativen anbietet, die sich in manchen Fällen besser in eine Umgebung integrieren lassen dürften.

Im Test präsentierte sich das System mit bis zu 116.000 IOPS bei synchroner Spiegelung überaus performant. Das Linux-

basierte Betriebssystem JovianDSS von Open-E ist einfach und intuitiv konfigurierbar. Allerdings würden wir uns innerhalb der WebGUI mehr konzentrierte Informationen zur aktuellen Auslastung wünschen. Erfreulich übersichtlich ist dagegen die Anzeige des Cluster- und Failover-Status gestaltet. Mit Unterstützung für iSCSI, Fibre Channel, NFS und SMB ist das System vielseitig einsetzbar, sowohl block- als auch dateibasiert.

Die Netzwerkanbindung lässt sich bei Thomas-Krenn individuell konfigurieren, zweimal 1-GBE sowie ein Management-Port für IPMI sind ab Werk an Bord, der Rest wie 10-, 25- oder gar 100-GBE-Karten sind individuell zu ergänzen. Allgemein empfehlen wir beim Sizing derartiger Cluster eine Zusammenarbeit mit dem Hersteller, damit sowohl die Performance passt als auch eine möglicherweise notwendige Erweiterbarkeit gegeben ist und letztlich der Preis stimmt. (In) IT

So urteilt IT-Administrator

Flexibilität beim Hardwaredesign	8
Inbetriebnahme	8
Redundanz	9
Performance	9
Skalierbarkeit	7

Die Details unserer Testmethodik finden Sie unter www.it-administrator.de/testmethodik

Dieses Produkt eignet sich

optimal für den Einsatz in KMUs als hochverfügbarer Speicher-Cluster für Virtualisierungsumgebungen vor allem unter VMware vSphere.

bedingt für die Nutzung in Verbindung mit Hyper-V. Die Kombination ist zwar getestet und bestätigt, eine Speicherplattform auf Windows-Basis dürfte jedoch in manchen Konstellationen eine bessere Integration ermöglichen.

nicht bei geringen Verfügbarkeitsanforderungen, die keinen Cluster-Betrieb verlangen und sich mit einem einfacheren NAS-System abdecken lassen, eventuell auch mit nur einem RA1112-Knoten.