storage-magazin.de

Eine Publikation von speicherguide.de

Storage für den Mittelstand

Bild: speicherguide.de via DALL-E

Lokale Datenspeicherung als Grundlage digitaler Selbstbestimmung

2025
03

Editorial

ANHÄNGIGKEITEN REDUZIEREN

Karl Fröhlich
Chefredakteur
speicherguide.de

Liebe Leserinnen und Leser,

Anfang des Jahres hatte ich schon angemerkt, dass die 2020er Jahre unter keinem Glücksstern stehen. Die hat echt der Teufel gesehen. Zurest hatten wir kein Glück, nun kommt auch noch Pech hinzu…

Nach Pandemie, russischem Angriffskrieg gegen die Ukraine und dem Broadcom/VMware-Debakel droht nun ausgerechnet die US-Regierung, die IT-Welt weiter ins Chaos zu stürzen. Die Androhung von Zöllen und Einfuhrstopps ist das eine – doch viel gravierender ist die strukturelle Erpressbarkeit unserer Behörden, unserer Wirtschaft und unserer Regierung durch US-amerikanische Tech-Konzerne.

Der Hintergrund: US-Hersteller unterliegen der US-Gesetzgebung – einschließlich Gesetzen wie dem US-Cloud-Act, der Behörden Zugriff auf gespeicherte Daten auch außerhalb der USA ermöglicht (**siehe auch hier**). Der US-Cloud Act ist nicht neu, doch bislang haben wir seine Reichweite meist stillschweigend hingenommen. Vor Trump konnte man den Amerikanern zumindest ein gewisses Maß an Kooperationsbereitschaft unterstellen – heute ist das nicht mehr selbstverständlich.

Was passiert, wenn es nicht nach dem Willen des US-Präsidenten geht haben wir bereits gesehen: Microsoft hat beispielsweise das Mail-Konto des Chefanklägers des Internationalen Gerichtshofs nach Trump-Sanktionen gesperrt. Die Eingriffsmöglichkeiten sind also real und nicht hypothetisch.

Deshalb ist der Ruf nach einer digitalen Souveränität Europas nicht bloß ein weiterer Trend, der durch die Medien geistert. US-Anbieter sind selbstverständlich nicht plötzlich »die Bösen«, ganz im Gegenteil: Viele von ihnen liefern hervorragende Technologien. Aber: Sie verfügen über keinen verlässlichen Schutzmechanismus gegenüber politischem Druck aus Washington.

Die Verantwortung liegt also bei uns – bei den europäischen IT-Verantwortlichen, Strategen und Entscheiderinnen. Vor dem Einsatz US-amerikanischer IT-Produkte sollte künftig stets eine gründliche Risikoanalyse stehen, idealerweise mit klaren Eskalationsszenarien.

Und hier kommt oft der Einwand: »Aber europäische Alternativen sind doch gar nicht so gut, die können einfach nicht mithalten.«

Wirklich? Das mag in Einzelfällen zutreffen – aber in vielen Bereichen ist diese Einschätzung veraltet oder schlicht falsch. Auch in Europa gibt es Anbieter mit ausgereiften, skalierbaren und leistungsfähigen Lösungen – gerade im Bereich Storage, Security und Infrastruktur-Management. Wer nie über den Atlantik hinausblickt, verpasst Innovationen vor der eigenen Haustür. Was fehlt, ist nicht die Qualität der Alternativen, sondern oft der Wille, diese ernsthaft zu prüfen.

Das ist keine Situation, die wir einfach aussitzen können. Bitter bleibt, dass erneut äußere Einflüsse unsere IT-Strategien diktieren.

Gerade deshalb setzen wir in diesem eMagazine einen inhaltlichen Schwerpunkt auf digitale Souveränität. Wir werfen einen Blick auf europäische Alternativen, auf neue Denkansätze – und zeigen, dass abseits der eingefahrenen Pfade erstaunlich viel Leben steckt. Es muss nicht immer ein US-Produkt sein.

Ihr Karl Fröhlich,
Chefredakteur speicherguide.de

Bild: speicherguide.de via DALL-E

IT-Strategie und Investitionsklima

WAS DEN MITTELSTAND IN DER IT WIRKLICH BESCHÄFTIGT

Knappe Budgets, steigende Bedrohungslage, neue Regulierungen: Die IT im Mittelstand steht unter wachsendem Druck. Gleichzeitig wächst der Anspruch an Souveränität, Sicherheit und Nachhaltigkeit. Was Unternehmen jetzt wissen und beachten sollten.

Eigene Daten, eigene Regeln:
Das Comeback der lokalen Speicherung

ON-PREMISES: BEWÄHRTE GRUNDLAGE MODERNER DATENSPEICHERUNG

Obwohl die »Cloud-Ecke« vermeintlich lauter ist, bleibt die lokale Datenspeicherung in Rechenzentren und IT-Abteilungen der zentrale Baustein moderner Datenarchitekturen. On-Premises bedeutet dabei weder Rückstand oder technologische Starrheit, sondern steht für gezielte Kontrolle, angepasste Performance und regulatorische Sicherheit.

Bild: speicherguide.de via DALL-E

IT-Infrastruktur auslagern, ohne die Verantwortung abzugeben

AUSGELAGERTE RECHENZENTREN: IT-BETRIEB IM MIETMODELL

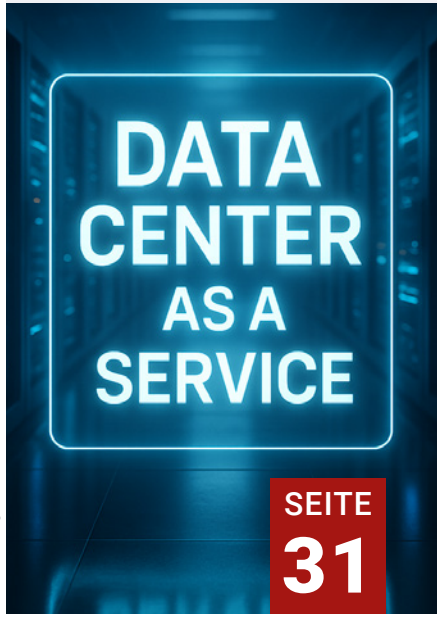
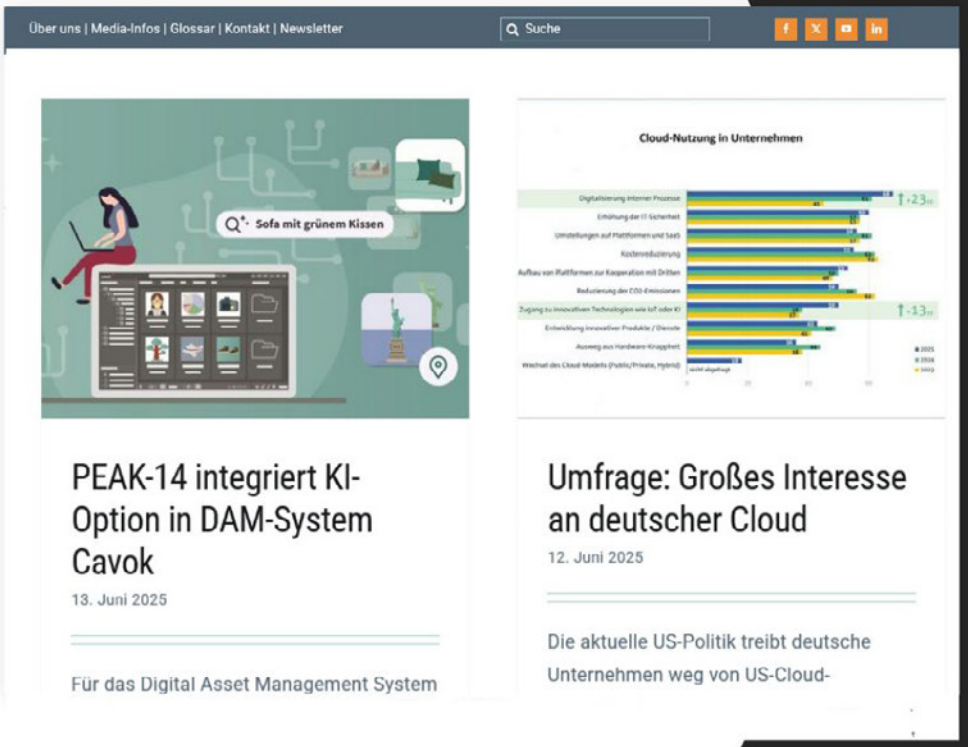


Bild: speicherguide.de via DALL-E

Übersicht
Storage-Anbieter

ECM GUIDE



Digitalisierung verstehen

www.ECMGUIDE.de

Bild: speicherguide.de via DALL-E

Karl Fröhlich
speicherguide.de

IT-Strategie und Investitionsklima

WAS DEN MITTELSTAND IN DER IT WIRKLICH BESCHÄFTIGT

Knappe Budgets, steigende Bedrohungslage, neue Regulierungen: Die IT im Mittelstand steht unter wachsendem Druck. Gleichzeitig wächst der Anspruch an Souveränität, Sicherheit und Nachhaltigkeit. Was Unternehmen jetzt wissen und beachten sollten.

IT-Abteilungen im Mittelstand stehen seit jeher unter Druck: begrenzte Ressourcen, wachsender technischer Anspruch und eine Vielzahl parallel zu bewältigender Aufgaben. Viele dieser Herausforderungen sind struktureller Natur und begleiten Unternehmen dauerhaft. So sind die Teams in mittelständischen Unternehmen häufig klein und bestehen oft genug aus Generalisten. Spezifisches Fachwissen, etwa in den Bereichen Cybersecurity oder Cloud-Integration, muss meist extern eingekauft werden.

Historisch gewachsene IT-Landschaften mit vielen Einzellösungen erschweren nicht nur den täglichen Betrieb, sondern blockieren oftmals auch dringend benötigte Modernisierungen. Besonders außerhalb der Ballungsräume gestaltet sich zudem die Suche nach qualifizierten Fachkräften schwierig. Viele IT-Spezialisten zieht es in größere Unternehmen, die attraktivere Karrierepfade und Entwicklungsmöglichkeiten bieten. Hinzu kommt, dass in vielen Organisationen übergreifende Strategien zur Datenhaltung und -analyse fehlen. Datensilos und unstrukturierte Datenhaltung verhindern effiziente Prozesse und hemmen die Digitalisierung.

Angriff aus dem Nichts: Sicherheitslage wird zum Risiko-Faktor

Die IT-Sicherheit bleibt dabei eine Dauerbaustelle. Laut **Bundesamt für Sicherheit in der Informationstechnik** (BSI) ist die Bedrohungslage so hoch wie nie. Dennoch sind viele mittelständische Unternehmen nicht ausreichend vorbereitet. **Hannes Heckel**, Head of Marketing bei **FAST LTA**, bringt es auf den Punkt: »Jeder gesparte Euro kann nach einem Angriff mehrfach zu Buche schlagen.« Angesichts immer ausgefeilterer Angriffsmethoden, teils unterstützt durch KI, steigen die Risiken kontinuierlich. Gleichzeitig ist das Sicherheitsniveau vieler Unternehmen unzureichend – auch, weil Know-how und Ressourcen fehlen.

Bild: Fast LTA

Hannes Heckel
Fast LTA

»Jeder gesparte Euro kann nach einem Angriff richtig teuer werden.«

Laut **Veeam**-Studie werden 69 Prozent der Ransomware-Opfer mehrfach angegriffen – ein deutlicher Beleg für Nachholbedarf in der IT-Sicherheit. Die Folgen erfolgreicher Attacken sind für kleine und mittlere Unternehmen oft existenzbedrohend: Neben Produktionsausfällen und Reputationsverlust drohen Bußgelder bei Verstößen gegen die DSGVO sowie finanzielle Schäden durch Lösegeldzahlungen oder Datenverlust. In vielen Fällen dauert es Wochen, bis der Betrieb vollständig wiederhergestellt ist – sofern überhaupt ein funktionsfähiges Backup existiert.

Als Minimalausstattung empfiehlt sich für kleine Firmen ein Sicherheitskonzept bestehend aus: Multi-Faktor-Authentifizierung (MFA), Endpoint-Protection, E-Mail-Filterung, regelmäßigen Patch-Zyklen, Offline-Sicherungen sowie einer Backup-Strategie mit Immutability-Funktion. Zusätzlich sinnvoll sind Netzwerksegmentierung, ein Wiederherstellungstestplan, Awareness-Schulungen für Mitarbeitende und ein Notfallhandbuch für den Ernstfall. Wer externe IT-Partner nutzt, sollte zudem auf klar definierte Verantwortlichkeiten und SLA-Vorgaben achten.

Neue Baustellen dank Digitalisierung, NIS-2 & KI

Der Druck zur Digitalisierung ist hoch – und betrifft längst nicht mehr nur die Einführung neuer Tools, sondern die grundlegende Anpassung bestehender Geschäftsprozesse. **Patricia Hillebrand**, International Channel Manager bei **RNT Rausch**, verweist auf die doppelte Dimension: »Der Schutz vor Cyberattacken wie Ransomware und die digitale Transformation – nicht nur in punkto Tools, sondern auch in Bezug auf Prozesse – sind zentrale Themen.«

Bild: RNT Rausch

Patricia Hillebrand
RNT Rausch

»Prozesse sind genauso wichtig wie Tools, wenn es um digitale Transformation geht.«

Auch **Albrecht Hestermann**, Head of Sales & Marketing bei **actidata**, sieht die Bedrohungslage als treibende Kraft: »Ganz klar – Cyber Security. Vielleicht darf hier auch die sogenannte ›German Angst‹ mit erwähnt werden, jedoch – getrieben durch globale Krisen und Veränderungen – sehen wir das auch in der kompletten EU.«

Parallel zwingt die NIS-2-Richtlinie viele Unternehmen, ihre Sicherheits- und Datenstrategie zu überdenken. **Bernd Widmaier**, Director Sales bei **Starline**, beobachtet: »Firmen wollen verstärkt die Hoheit über ihre Daten und Dienste wiedererlangen und verlagern diese in eigene IT-Strukturen oder zu lokalen Partnern.« Dieser Trend geht häufig einher mit einer gewissen Skepsis gegenüber Cloud-Diensten, insbesondere wenn diese aus den USA stammen.

Bild: Actidata

Albrecht Hestermann
Actidata

»Open-Source ist kein Allheilmittel, aber ein erster Schritt in die Eigenverantwortung.«

Bild: Starline

Bernd Widmaier
Starline

»Die Cloud ist kein Allheilmittel – viele holen sich ihre Daten wieder zurück ins Haus.«

Bernhard Seibold, Vice President Product Management bei **Thomas-Krenn**, sieht die Ursachen tiefer: »Für den Mittelstand ist die schiere Masse der Herausforderungen inzwischen wirklich immens. Ob IT-Sicherheit, Digitalisierungsprojekte, KI oder Datenwachstum und Archivierung – hier müssen KMUs mit einem kleineren Team all das erledigen, was auch Konzerne betrifft. Hinzu kommen Themen wie komplexe Branchen-Zertifizierungen, wie sie etwa große OEMs von ihren Zulieferern fordern. In dieser Gemengelage ist es nicht so leicht, die besten Lösungen zu finden, da richtet sich der Blick oft zwangsläufig auf die bekannteste oder kurzfristig einfachste.«

Politik, Monopole, Misstrauen: Die Rückkehr der digitalen Souveränität

Auch die politische Lage wirkt sich auf das Investitionsverhalten aus. »Die Unberechenbarkeit der US-Regierung hat Auswirkungen auf die globale Preisgestaltung. Viele Unternehmen zögern mit Investitionen«, betont Fast-LTA-Manager Heckel. Gleichzeitig sehen viele Unternehmen die Notwendigkeit, gerade jetzt in IT-Sicherheit und Unabhängigkeit zu investieren – nicht zuletzt, um sich Wettbewerbsvorteile zu sichern. Für viele Betriebe wird die Herkunft der Dienste und Produkte zum Auswahlkriterium. Die Debatte um digitale Souveränität gewinnt dadurch an Schärfe.

Hillebrand warnt vor Monopolisten: »Da im Weißen Haus eine einzelne Person willkürlich und mit einer einzigen Unterschrift E-Mail-Konten sperren oder gar den Stecker ziehen kann, ist es unumgänglich, dass Europa endlich Gas gibt.« Auch Widmaier bestätigt: »Die IT-Verantwortlichen schauen plötzlich genau hin, wo ihre Cloud-Dienste wirklich gehostet werden.«

Open-Source statt Lock-in: Unabhängigkeit ist mehr als Idealismus

Der Open-Source-Markt könnte von dieser Entwicklung profitieren. Viele Unternehmen suchen nach Alternativen zu proprietären Systemen, auch weil diese mit Vendor-Lock-ins und späteren Kostensteigerungen verbunden sind. »Viele Hersteller ködern mit Freebies und schlagen dann durch die Hintertür zu«, kritisiert Hillebrand. Starline setzt daher gezielt auf Open-Source-Lösungen wie *Proxmox* oder *TrueNAS* – Technologien, die gerade im Mittelstand den Einstieg in Hochverfügbarkeit und Datensouveränität erleichtern.

Auch Actidata-Manager Hestermann betont: »Hand aufs Herz: Hardware ohne US-Know-how und damit ohne zukünftige Updates wird es wohl in naher Zukunft kaum geben. Insofern hilft Open-Source auch nur bedingt. Aber, was bitte nicht zu unterschätzen ist, eine Open-Source-Lösung bietet eine sehr große Hardware-Kompatibilität, so dass eigene Lösungen hier auf vielen Plattformen eingesetzt werden können – auch wenn diese vielleicht ein wenig älter sind.« Für ihn ist Open-Source ein erster Schritt zur Eigenverantwortung: »Der Einsatz von Open-Source-Lösungen sollte immer in Betracht gezogen werden – es ist ein wichtiger erster Schritt, unabhängiger zu werden.«

KI kommt – aber wie und zu welchem Preis?

Auch der technologische Fortschritt bringt neue Herausforderungen mit sich. Der Einsatz generativer KI wird zunehmend ernsthaft diskutiert, doch viele Unternehmen sind unsicher, wie sie diese Technologie sinnvoll einsetzen können. Gleichzeitig steigen die Anforderungen an die IT-Infrastruktur – sowohl hinsichtlich Performance als auch in Bezug auf Datensicherheit. Die Preisgestaltung vieler KI- und SaaS-Dienste bleibt intransparent, was zu zusätzlicher Zurückhaltung führt.

Bild: Thomas-Krenn

Bernhard Seibold
Thomas-Krenn

»KMUs müssen mit einem kleineren Team all das erledigen, was auch Konzerne betrifft.«

Für mittelständische Unternehmen stellt sich dabei vor allem die Frage, ob ihre bestehende Infrastruktur überhaupt für KI-Anwendungen geeignet ist. Denn KI benötigt je nach Anwendungsfall hohe IOPS-Raten, schnellen Zugriff auf große Datenmengen, geringe Latenz und ausreichend Bandbreite – Anforderungen, die klassische Storage-Systeme oft nicht erfüllen. Gerade Trainingsdaten und Modelle müssen häufig bewegt, zwischengespeichert oder mehrfach abgerufen werden.

Es gilt zu prüfen, ob bestehende NAS- oder SAN-Systeme dafür ausreichen oder ob der Einsatz von NVMe-Storage, GPU-optimierten Servern oder All-Flash-Systemen nötig ist. Für erste Analysen oder Inferencing-Aufgaben lassen sich auch vorhandene Systeme mit Software-Lösungen wie *TensorFlow Lite*, KI-Module in *Microsoft 365* oder Open-Source-Tools ergänzen. Wer strukturiert plant, kann mit Bordmitteln starten – braucht für skalierbare KI-Projekte jedoch meist zusätzliche Investitionen. Kleinere bis mittlere KI-Vorhaben im KMU-Umfeld – zum Beispiel Chatbots, Log-Analysen oder Bildklassifikation – bewegen sich häufig in einem Budget-Rahmen von 10.000 bis 50.000 Euro, inklusive externer Beratung, initialer Hardware-Anpassung und Betriebsaufwand.

Zwischen Spardruck und Notwendigkeit: Das IT-Budget als Zankapfel

Trotz dieser Unsicherheiten wird weiter investiert – wenn auch mit Bedacht. Laut **Bitkom** wird der ITK-Markt in Deutschland im Jahr 2025 ein Volumen von 232,8 Milliarden Euro erreichen, ein Plus von 4,6 Prozent gegenüber dem Vorjahr. Besonders dynamisch entwickeln sich die Segmente Software und IT-Services. Dennoch bleibt das Budget im Mittelstand begrenzt. Laut der **VOICE/metrics**-Studie »IT-Agenda 2025« liegt das durchschnittliche IT-Budget bei 4,75 Prozent des Unternehmensumsatzes. Für 2025 wird nur noch ein moderates Wachstum von 2,4 Prozent erwartet.

Bild: Datadobi

Sascha Hempe
Datadobi

»Datenmanagement ist der Schlüssel – nicht mehr Speicherplatz.«

RNT-Managerin Hillebrand hält dagegen: »Inhaber sollten rund fünf Prozent des Jahresumsatzes in ihre IT-Infrastruktur investieren – auch wenn sie nur wenige TByte an Daten halten.« Im internationalen Vergleich liegt Deutschland im Mittelfeld. In Ländern wie den USA oder Skandinavien ist der Anteil am Umsatz, der in IT fließt, zum Teil deutlich höher. Die Ursachen sind vielfältig: längere Investitionszyklen, konservative Entscheidungsprozesse, aber auch ein geringeres Verständnis für den strategischen Wert von IT.

Budget: Sicherheit trifft Realität

Was den Mittelstand aktuell besonders umtreibt, ist die Frage, wie sich Sicherheit, Hochverfügbarkeit und wirtschaftlicher Betrieb miteinander vereinbaren lassen. Widmaier fasst es so zusammen: »Unsere Kunden wünschen sich Hochverfügbarkeit für schmales Geld – plus ein automatisiertes Backup, das im Zweifel schnell wiederherstellbar ist.« Die Ansprüche an Redundanz und Schutz vor Angriffen steigen – ohne dass die Budgets im gleichen Maße wachsen.

Auch das Thema Datenorganisation gewinnt an Relevanz: »Neben den steigenden Kosten ist für KMUs wichtig, wie sie ihre Daten schützen und ihre Compliance sicherstellen können«, erklärt **Sascha Hempe**, Regional Sales Manager DACH & Nordics bei **Datadobi**. »Auch künstliche Intelligenz wird in dem Zusammenhang ein immer bedeutenderes Thema, denn Unternehmen müssen sicherstellen, dass ihre Daten gut organisiert und für neue KI-Anwendungen bereit sind.«

Datenwert statt Datenmenge: TByte nicht weniger wert als PByte

Kleinere Unternehmen erkennen zunehmend den Wert ihrer Daten. Hillebrand mahnt: »Auch TByte sind oft geschäftskritisch – und müssen wie PByte behandelt werden.« Der Markt bietet inzwischen eine Vielzahl an Lösungen mit Enterprise-Funktionalität, die auch für mittelständische Budgets geeignet sind. Entscheidend ist, dass IT-Entscheider sich mit dem Wert ihrer Daten aktiv auseinandersetzen – und ihre Investitionen entsprechend priorisieren.

Auch Datadobi-Manager Hempe verweist auf zentrale Fragen des Datenmanagements: »Viele Unternehmen versuchen, die Datenflut zu bewältigen, indem sie mehr und mehr Speicherplatz bereitstellen. Doch das ist keine nachhaltige Strategie. In erster Linie sollten sie sicherstellen, dass sie einen transparenten Überblick über ihre Datenlandschaft und Richtlinien für die Verwaltung ihrer Daten haben – also zum Beispiel, was aufbewahrt werden muss, was archiviert werden sollte, was gelöscht werden kann.«

Hestermann sieht allerdings Nachholbedarf beim Know-how im Mittelstand: »Die Administratoren konzentrieren sich ganz besonders auf Zertifizierungen von Microsoft, *Cisco* etc., aber leider weniger auf Hardware und Open-Source. Vielleicht liegt es daran, dass Weiterbildung hier doch ein schweres Kapitel ist und eben wohlklingende Namen eher ein Garant für den nächsten Karriereschritt sind.«.

Fazit: Ohne Strategie keine Sicherheit – und kein Wachstum

Der Mittelstand steht vor einem Paradigmenwechsel. IT wird nicht länger nur als Kostenfaktor gesehen, sondern als strategische Ressource. Wer den Wandel aktiv gestaltet, kann sich Wettbewerbsvorteile sichern – durch Sicherheit, Souveränität und eine zukunftsfähige IT-Strategie.

Die eingangs genannten strukturellen Probleme – zu kleine IT-Teams, fragmentierte Systemlandschaften, Fachkräftemangel – sind weiterhin ungelöst. Viele mittelständische Unternehmen haben in der Corona-Pandemie in Digitalisierung investiert, etwa in Home-Office-Arbeitsplätze, VPN-Zugänge, Collaboration-Tools oder erste Cloud-Dienste. Diese Infrastrukturen wurden vielfach pragmatisch aufgebaut und laufen inzwischen im Dauerbetrieb – aber häufig ohne ein strategisches Update oder umfassende Integration in die IT-Gesamtsystematik.

Heute stehen viele IT-Abteilungen an einem Wendepunkt: Die »Corona-IT« hat ihre Funktion erfüllt, muss aber nun konsolidiert, modernisiert oder teilweise sogar bereits ersetzt werden. Typische nächste Schritte sind: Die Vereinheitlichung der Backup- und Sicherheitskonzepte, eine kritische Überprüfung der Cloud-Nutzung (Kosten, Compliance, Souveränität), der Aufbau eines aktiven Datenmanagements und – zunehmend – die Integration von KI-Komponenten.

Kurz gesagt: Es geht jetzt nicht mehr um kurzfristige Notfalllösungen, sondern um belastbare und nachhaltige IT-Strukturen. Wer diese Phase aktiv gestaltet, schafft Experten zufolge die Basis für künftige Wettbewerbsfähigkeit – technologisch wie organisatorisch. So kann etwa eine konsolidierte, automatisierte Backup-Umgebung nicht nur die IT entlasten, sondern im Ernstfall auch den Geschäftsbetrieb deutlich schneller wiederherstellen – ein Vorteil, der im Wettbewerb den Unterschied machen kann. ■

Bild: speicherguide.de via DALL-E

Karl Fröhlich
speicherguide.de

Eigene Daten, eigene Regeln:
Das Comeback der lokalen Speicherung

ON-PREMISES: BEWÄHRTE GRUNDLAGE MODERNER DATENSPEICHERUNG

Obwohl die »Cloud-Ecke« vermeintlich lauter ist, bleibt die lokale Datenspeicherung in Rechenzentren und IT-Abteilungen der zentrale Baustein moderner Datenarchitekturen. On-Premises bedeutet dabei weder Rückstand oder technologische Starrheit, sondern steht für gezielte Kontrolle, angepasste Performance und regulatorische Sicherheit.

Die lokale Datenspeicherung ist der Standard und war dies bereits lange bevor die ersten Hyperscaler ihre virtuellen Lagerhallen für Daten errichteten. On-Premises gilt als bewährte Methode, Daten sicher, performant und unter eigener Kontrolle zu speichern. Sie hat sich in Jahrzehnten technischer Entwicklung etabliert und stetig weiterentwickelt – von einfachen Server-Räumen bis hin zu hochmodularen, automatisierten IT-Architekturen im Rechenzentrum.

Der oft zitierte »Cloud-first«-Ansatz hat zwar vielerorts für Bewegung gesorgt, doch an der grundlegenden Bedeutung lokal betriebener Speicherlösungen hat das wenig geändert. Im Gegenteil: Gerade durch die rasante Entwicklung von Cloud-Angeboten und damit verbundenen Fragen rund um Datenschutz, Datenkontrolle und Betriebsrisiken rückt die Stärke von On-Premises erneut ins Bewusstsein vieler IT-Verantwortlicher. Die Frage lautet heute nicht mehr »entweder oder«, sondern: Welche Daten gehören wohin.

»Manche Daten sind gut in der Cloud aufgehoben und andere nicht«, erklärt **Patricia Hillebrand**, International Channel Manager bei **RNT Rausch**. »Hybride Umgebungen, die die Vorteile aus beiden Welten bieten, sind die Zukunft.«

Auch **Sascha Hempe**, Regional Sales Manager DACH & Nordics bei **Datadobi**, betont die Bedeutung strategischer Datenverteilung: »Firmen setzen oft ganz bewusst auf hybride Lösungen, um von deren Vorteilen profitieren zu können. Daher ist eine Strategie für das Management unstrukturierter Daten besonders wichtig: zu verstehen, welche Daten man hat, und diese kontrolliert auf die richtigen Plattformen zu verteilen – ob Cloud oder On-Premises.«

Cloud ist gut – aber nicht gut genug für alles

Trotz der allgegenwärtigen Präsenz von Cloud-Technologien behauptet die lokale Datenhaltung ihren festen Platz – und gewinnt in bestimmten Szenarien sogar an Relevanz. Dazu zählen etwa Backup- und Recovery-Umgebungen mit kurzen Wiederherstellungszeiten (RTO), Anwendungen mit besonders hohen I/O-Anforderungen wie Echtzeitverarbeitung in der Produktion, gesetzlich regulierte Branchen mit strikten Datenschutzauflagen sowie KI- und CAD-Workloads, die eine lokale GPU-Verarbeitung erfordern. Auch bei sensiblen Forschungsdaten oder in Unternehmen mit stark dezentralen Standorten bietet On-Premises oftmals die notwendige Sicherheit und Flexibilität.

Der »Cloud-only«-Ansatz, der noch vor wenigen Jahren als Zukunftsmodell propagiert wurde, hat sich in der Breite nicht durchgesetzt. Stattdessen beobachten viele IT-Verantwortliche eine zunehmende Rückbesinnung auf hybride Infrastrukturen mit einer starken On-Premises-Komponente. Gründe dafür sind unter anderem steigende Anforderungen an Datenschutz, Kostentransparenz, Performance sowie die Kontrolle über unternehmenskritische Daten.

Vor allem regulatorische Vorgaben wie die DSGVO, das IT-Sicherheitsgesetz oder branchenspezifische Compliance-Richtlinien haben dazu geführt, dass Datenhaltung auf eigenem Terrain wieder als strategischer Vorteil wahrgenommen wird. In Bereichen mit sensiblen Informationen – etwa im Gesundheitswesen, in der Finanzbranche oder in der öffentlichen Verwaltung – stellt On-Premises den einzig praktikablen Weg dar, um alle regulatorischen Anforderungen zu erfüllen.

Wer seine Daten wirklich im Griff haben will, baut selbst

Mit dem Begriff On-Premises ist der Betrieb von IT-Infrastruktur im eigenen Rechenzentrum oder an einem kontrollierten Unternehmensstandort gemeint. Server, Speicher, Netzwerke und Sicherungssysteme werden nicht als Dienstleistung eines externen Anbieters genutzt, sondern unter eigener Verantwortung bereitgestellt und betrieben. Dabei reicht die Spannweite von klassischen 19-Zoll-Racks in klimatisierten Technikräumen bis hin zu hochmodularen Edge-Systemen in der Fertigungshalle – je nach Anwendungsfall.

On-Premises ist vor allem dann sinnvoll, wenn Unternehmen den vollständigen physischen und logischen Zugriff auf ihre Daten benötigen, wenn Latenzen kritisch sind oder wenn ein kontinuierlicher Betrieb ohne Internetanbindung sichergestellt werden muss. Beispiele sind Echtzeit-Analyse in der Industrieproduktion, Backup- und Recovery-Anwendungen mit kurzem RTO, hochsensible personenbezogene Daten oder proprietäre Forschungsergebnisse in der Produktentwicklung.

Ein On-Premises-Szenario besteht im Wesentlichen aus folgenden Infrastrukturelementen:

- **Server**: Als Rechenknoten für Applikationen, Datenbanken und Virtualisierung – meist in Form von Rack- oder Blade-Systemen.
- **Storage**: Zentrale oder verteilte Speicherlösungen, von klassischen SAN/NAS-Systemen über All-Flash-Arrays bis hin zu software-definierten Plattformen.
- **Netzwerke**: Strukturierte Verkabelung, Switche, Firewalls und segmentierte Netze zur sicheren und performanten Kommunikation.
- **Backup und Recovery**: Lokale Sicherungslösungen, oft ergänzt durch Bandlaufwerke, Snapshots oder dedizierte Backup-Appliances – teilweise auch mit Replikation in ein zweites Rechenzentrum oder in die Cloud.

Diese Infrastruktur wird durch Monitoring- und Management-Tools ergänzt, um Verfügbarkeit, Performance und Sicherheit zu gewährleisten.

Primär- und Sekundärdaten: Zwei Klassen, zwei Anforderungen

Ein zentraler Aspekt bei der Planung und Bewertung lokaler Speicherstrategien ist die Unterscheidung zwischen Primär- und Sekundärdaten. Beide Datentypen stellen sehr unterschiedliche Anforderungen an Infrastruktur, Performance, Verfügbarkeit und Wirtschaftlichkeit – und genau darin liegt eine der Stärken von On-Premises-Umgebungen: Sie lassen sich gezielt und differenziert auf diese Anforderungen hin auslegen.

Primärdaten benötigen PS – keine Performance-Kompromisse

Primärdaten sind produktive, geschäftskritische Informationen, die direkt im operativen Betrieb anfallen. Dazu zählen Datenbanken, Transaktionen, Sensorwerte aus der Produktion oder virtualisierte Workloads – also all jene Inhalte, die in Echtzeit verfügbar sein sollen. Entsprechend hoch sind die Anforderungen: niedrige Latenzen, hohe IOPS, dauerhafte Verfügbarkeit und zuverlässige Absicherung gegen Ausfälle.

On-Premises-Systeme bieten hier einen entscheidenden Vorteil, da sie direkt im Unternehmensnetz betrieben werden, kurze Datenwege ermöglichen und sich individuell auf die jeweilige Anwendung abstimmen lassen. Technisch kommen in der Regel All-Flash-Arrays, NVMe-over-Fabrics-Architekturen oder hyperkonvergente Systeme zum Einsatz, die sich durch hohe Performance, Redundanz und tiefe Integration in die IT-Landschaft auszeichnen.

Sekundärdaten: effizient und langfristig sicher

Im Gegensatz dazu stehen Sekundärdaten, also Daten, die zwar wichtig, aber nicht permanent produktiv im Einsatz sind. Dazu gehören Backups, Archivdaten, Logs, Replikate, Testumgebungen oder E-Mails mit Aufbewahrungsfristen. Die Anforderungen sind hier ganz andere: Kosteneffizienz, Skalierbarkeit, Integrität und oft auch regulatorische Nachweisbarkeit.

Auf diese sogenannten »Cold Data« wird meist selten zugegriffen. In solchen Fällen bieten On-Premises-Ansätze wie Objektspeicher mit S3-Kompatibilität, Bandlösungen mit WORM-Funktion oder klassische NAS-Systeme eine passende Antwort. Auch Air-Gap-Strategien oder Offline-Medien lassen sich lokal einfacher umsetzen als in reinen Cloud-Umgebungen – ein wichtiger Schlüsselfaktor für Datenschutz und Ransomware-Abwehr.

Die Technik muss passen – nicht alles gehört in denselben Topf

Typische Anwendungsszenarien verdeutlichen diesen Unterschied: Während eine produktive Datenbankumgebung auf hochperformanter Primärspeicher-Infrastruktur mit Redundanz und Snapshots betrieben wird, genügt für das wöchentliche Backup oder die Langzeitarchivierung von Logdateien eine kosteneffiziente Plattform mit starker Deduplizierung oder Bandspeicher. Gleiches gilt für KI-Workloads: Trainingsdaten und Modelle benötigen schnellen Zugriff und Rechenleistung – die Ergebnisse oder historischen Trainingsläufe hingegen können auf Sekundärspeicher ausgelagert werden.

Diese differenzierte Speicherarchitektur lässt sich im On-Premises-Betrieb gezielt abbilden. Gerade weil Unternehmen hier über das Design, die Zugriffspfade und die Schutzmechanismen selbst bestimmen, lässt sich ein Höchstmaß an Effizienz, Sicherheit und Datenhoheit erzielen – maßgeschneidert für jede Datenklasse.

Abgrenzung zu Cloud-, Hybrid- und Edge-Modellen

Im Gegensatz zu **Public-Cloud**-Lösungen – bei denen Infrastruktur, Speicher und Services vollständig bei externen Anbietern betrieben und nach Verbrauch abgerechnet werden – behält On-Premises die vollständige Kontrolle in der Hand des Unternehmens. Dafür entfällt allerdings auch die Flexibilität der dynamischen Skalierung und die Entlastung bei Betrieb und Wartung.

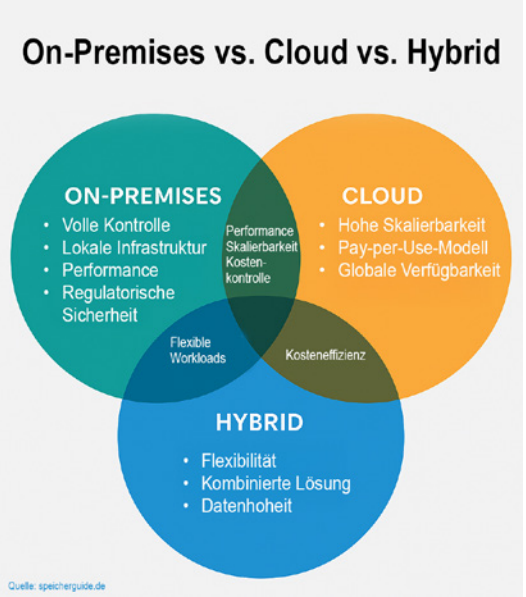
Hybrid-Cloud-Modelle kombinieren beide Welten: Lokale Systeme bilden stabile und performante Basiskomponenten, während nicht-kritische Daten oder Workloads in die Cloud ausgelagert werden – etwa für Archivierung, Analyse oder internationale Zusammenarbeit.

Edge-Computing hingegen ergänzt die On-Prem-Logik um dezentrale Systeme in unmittelbarer Nähe zur Datenquelle. Typisch sind Industrieanlagen, Logistikzentren oder Außenstellen, in denen Daten lokal vorverarbeitet und anschließend an zentrale On-Prem- oder Cloud-Systeme weitergegeben werden.

Die Wahl zwischen diesen Modellen ist keine reine Technologiefrage, sondern hängt maßgeblich vom Anwendungskontext, den regulatorischen Rahmenbedingungen und der strategischen Ausrichtung des Unternehmens ab.

Moderne Storage-Infrastrukturen sind hybrid

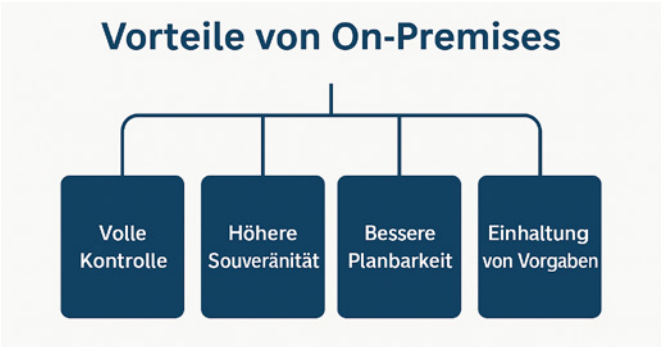
Die Backup-Infrastruktur in die Cloud zu verlagern, ist in IT-Abteilungen mittelständischer Unternehmen ein gängiges Thema. Oft genug handelt es sich um eine pragmatische Entscheidung: Platzmangel im Server-Raum, steigende Wartungskosten und die Aussicht auf »unbegrenzten Speicherplatz« lassen die Angebote der Cloud-Anbieter und Hyperscalers mehr als attraktiv erscheinen. Nachgelagert treten unerwartete Probleme auf, oft schon nach wenigen Monaten: Latenzen bei der Wiederherstellung, gestiegene Kosten durch Datenrückführung und Unsicherheiten bei der Auditierung der Datenspeicherung. Viele Firmen setzen deswegen auf hybride Lösungen mit starker On-Premises-Komponente.



Vorteile im Vergleich: On-Premises, Cloud und hybride Speicherstrategien –überlappende Stärken und exklusive Merkmale.

Laut **Albrecht Hestermann**, Head of Sales & Marketing bei **actidata**, sind dies keine Einzelfälle: »Onsite-Lösungen (On-Premises) sind Installationen, die unabhängig machen. Daten in der Cloud sind vom Wesen her dann nicht mehr ausschließlich im alleinigen Zugriff der Besitzer. Klar, ein deutscher Provider hat sein Rechenzentrum vielleicht in Deutschland; doch wohin repliziert er? Zur USA? Oder nach China? Wer weiß das schon.«

Auch **Bernhard Seibold**, Vice President Product Management bei **Thomas-Krenn**, betont: »Die Argumente für und gegen On-Prem sind die gleichen wie vor einem Jahr und wie vor fünf Jahren. Die Frage ist, wie gewichte ich welche Argumente. Das ist letztlich schon eine individuelle Entscheidung, die von mehreren technischen



Grafik: speicherguide.de

Volle Datenkontrolle, höchste Performance, maximale Sicherheit – lokal gespeicherte Daten bieten entscheidende Stärken für Unternehmen.

Faktoren abhängt.« Zugleich verweist er auf die Problematik internationaler Rechtslagen: »Egal, wo die Server stehen – der US-Cloud-Act besagt, dass ein Zugriff auf Daten möglich ist, wenn das anbietende Unternehmen aus den USA stammt.«

Ähnlich argumentiert **Hannes Heckel**, Head of Marketing bei **FAST LTA**: »Sensible Daten gehören niemals in Public-Cloud-Umgebungen, schon gar nicht, wenn unerlaubter oder unerwünschter Zugriff nicht ausgeschlossen werden kann. Im Hinblick auf schnelles, vollständiges Recovery nach einer Cyber-Attacke ist es unter Umständen überlebensnotwendig, einen bestimmten Anteil von Backup-Daten vor Ort zu halten.«

Bernd Widmaier, Director Sales bei **Starline**, unterstreicht die finanzielle Planbarkeit: »Bei On-Premises hat man die volle Hoheit über seine Daten und ist gleichzeitig von willkürlichen Preissteigerungen gefeit.« Für ihn ist auch eine unternehmenseigene Cloud-Infrastruktur denkbar – etwa für lokale CAD- oder KI-Workloads, die nur temporär hohe Rechenleistungen benötigen.

Technologische Vielfalt: Mehr als nur klassische Server

Technologisch betrachtet umfasst On-Premises-Storage heute weit mehr als klassische Rack-Server mit internen Festplatten. Unternehmen setzen auf hochperformante All-Flash-Systeme, NVMe-over-Fabrics, hyperkonvergente Infrastrukturen oder software-definierte Speicherlösungen, die sich nahtlos in moderne DevOps-Umgebungen integrieren lassen. Diese Technologien ermöglichen nicht nur eine bessere Ressourcenauslastung, sondern auch eine höhere Automatisierung und Orchestrierung lokaler Ressourcen. Gerade bei Anwendungen mit hohen I/O-Anforderungen oder sensiblen Daten bleibt der Betrieb vor Ort konkurrenzfähig.

Datenhoheit und Compliance als Schlüsselfaktoren

Ein zentrales Argument für On-Premises ist und bleibt die Datenhoheit. In regulierten Branchen wie dem Finanzwesen, der öffentlichen Verwaltung oder dem Gesundheitswesen ist die Nachvollziehbarkeit der Datenflüsse essenziell. Die Einhaltung von Vorschriften wie der DSGVO, ISO 27001 oder branchenspezifischer Auflagen ist lokal oft einfacher und transparenter umzusetzen als in global verteilten Cloud-Umgebungen. Zugriffskontrollen, Logging und physische Trennung von Netzwerken können auf On-Prem-Infrastruktur unternehmensspezifisch konfiguriert werden. Das reduziert nicht nur das Risiko von Compliance-Verstößen, sondern schafft auch Vertrauen bei Kunden und Partnern.

Wirtschaftlichkeit: TCO realistisch betrachten

Wirtschaftlich betrachtet ist die Gleichung deutlich komplexer, als viele Anbieter sie darstellen. Zwar locken Public-Cloud-Angebote mit niedrigen Einstiegskosten, Pay-per-Use-Modellen und scheinbar grenzenloser Skalierbarkeit – vor allem für Startups oder Projekte mit unklarer Datenentwicklung kann das kurzfristig attraktiv wirken. Doch sobald Daten in größerem Umfang entstehen und dauerhaft gespeichert werden müssen, kippt die Rechnung.

Ein häufig unterschätzter Kostenfaktor in der Cloud sind sogenannte Egress-Fees – also Gebühren für das Abrufen oder Verschieben von Daten aus der Cloud heraus. Je nach Anbieter und Volumen können sich diese Summen schnell im fünf- oder gar sechsstelligen Bereich pro Jahr bewegen. Wer beispielsweise regelmäßig Backup-Daten zur Wiederherstellung aus der Cloud ziehen muss, zahlt für die eigentlichen Storage-Kosten meist weniger als für den Traffic zurück ins Rechenzentrum.

Cloud rechnet sich nicht automatisch

Hinzu kommt, dass sich Cloud-Nutzungskosten nur schwer langfristig planen lassen. Preise ändern sich, Leistungen werden paketweise neu strukturiert, und nicht selten wachsen auch die Abhängigkeiten vom jeweiligen Anbieter – ein Umstand, der mit dem Begriff Vendor Lock-in umschrieben wird. In der Praxis bedeutet das: Wer einmal große Datenmengen in ein proprietäres Cloud-Ökosystem geladen hat, muss für Migrationen oder Wechsel nicht nur tief in die Tasche greifen, sondern riskiert zudem längere Ausfallzeiten und organisatorische Hürden.

Dem gegenüber steht eine sorgfältig geplante On-Premises-Infrastruktur mit klar kalkulierbaren Kosten: Investitionen in Hardware, Software und Betrieb lassen sich über mehrere Jahre abschreiben. Die Anschaffung erfolgt einmalig, Wartungskosten und Energieverbrauch sind kalkulierbar, Lizenzen planbar – und Erweiterungen lassen sich durch modulare Systeme gezielt steuern. Besonders in Umgebungen mit konstant hohen oder wachsenden Datenvolumen – etwa bei industriellen IoT-Anwendungen, Archivsystemen in Forschungseinrichtungen oder im Medienbereich – spielt diese Planbarkeit ihre Stärken aus.

On-Prem: Kein Kostenfresser per se

Ein Beispiel aus der Praxis: Ein mittelständisches Unternehmen im Automotive-Zulieferbereich betreibt ein eigenes Rechenzentrum mit 500 TByte Speicherkapazität. Die einmaligen Investitionskosten lagen bei rund 250.000 Euro, hinzu kamen jährlich rund 35.000 Euro für Betrieb, Wartung und Strom. Im Gegenzug erhält das Unternehmen vollständige Kontrolle über seine Daten, kann Recovery-Tests regelmäßig lokal durchführen und zahlt keinen Cent für Egress oder Datenrückführung. Ein ähnliches Setup in der Public-Cloud hätte in fünf Jahren laut interner Berechnungen fast das Doppelte gekostet – und wäre

bei gleichen Latenzanforderungen schlicht nicht umsetzbar gewesen.

Kurz gesagt: On-Premises ist kein Kostenfresser per se – im Gegenteil. Bei hoher Auslastung, vorhersehbarem Bedarf und langfristiger Nutzung kann sie die wirtschaftlich nachhaltigere Wahl sein. Entscheidend ist nicht das Preismodell, sondern die Passung zum Workload und zur strategischen Ausrichtung der IT.

Unabhängigkeit und Know-how als strategische Faktoren

Nicht zuletzt sprechen auch strategische Überlegungen für On-Premises: Der Aufbau eigener Kompetenzen, die Unabhängigkeit von Drittanbietern und die Fähigkeit, kritische IT-Ressourcen autonom betreiben zu können, wird in vielen Unternehmen wieder als wichtiger Faktor wahrgenommen. Das gilt insbesondere in geopolitisch instabilen Zeiten oder bei Lieferkettenproblemen.

Allerdings darf man die Herausforderungen nicht verschweigen. On-Premises erfordert Investitionen in Hardware, Räume, Energie und Fachpersonal. Ohne ein durchdachtes Lifecycle-Management, klare IT-Prozesse und passende Monitoring-Tools drohen schnell Komplexität und Betriebskosten aus dem Ruder zu laufen. Zudem fehlt in vielen IT-Abteilungen das spezialisierte Know-how für moderne Speicherarchitekturen, was den Trend zu Managed-Services und On-Premises-as-a-Service-Angeboten (OPaaS) verstärkt.

OPaaS: Kontrolle behalten, Betrieb auslagern

OPaaS ist ein Modell, das als Brücke zwischen klassischem Eigenbetrieb und Public-Cloud-Nutzung gilt. Die Idee: Unternehmen nutzen lokale IT-Infrastruktur – wie Server, Speicher, Netzwerke oder ganze Racks – an ihrem eigenen Standort oder in einem Colocation-Rechenzentrum. Der Unterschied zum traditionellen On-Premises-Betrieb liegt jedoch im Betriebsmodell: Die Hardware wird nicht gekauft, sondern vom Anbieter bereitgestellt und meist vollständig gemanagt – inklusive Monitoring, Wartung, Lifecycle-Management und Support.

Unternehmen sollen dabei die Datenhoheit und die Kontrolle über die physischen Standorte behalten, während sie gleichzeitig Aufwand, Komplexität und Risiko des Betriebs outsourcen. Zudem wird das Modell verbrauchsbasiert abgerechnet – je nach Anbieter als Miet- oder Pay-per-Use-Modell, teilweise mit Mindestlaufzeiten oder Service-Levels.

Zu den bekanntesten Angeboten gehören unter anderem **HPE** *GreenLake* und **Pure Storage** *Evergreen// Forever*. Auch **Bechtle** bietet Infrastruktur im OPaaS-Modell, auf Basis von Partnertechnologie wie **Dell**, **NetApp** oder **Lenovo**. Betrieb, Support und Lifecycle liegen beim Systemhaus – das Rechenzentrum steht beim Kunden.

Aus Sicht der Datensouveränität lassen sich diese Angebote heute nicht mehr guten Gewissens befürworten, zumindest nicht mit Storage-Systemen von US-Anbietern. Gemietete Speicher verbleiben im Besitz des Herstellers und damit greift der US-Cloud-Act.

Als Alternative wäre **Fujitsu** *uSCALE* zu nennen. Der japanische Anbieter richtet sich an Unternehmen, die IT-Infrastruktur wie ein Cloud-Service vor Ort betreiben möchten. Fujitsu übernimmt die Bereitstellung, Wartung und das Monitoring der Systeme. Abgerechnet wird nach tatsächlicher Nutzung, ergänzt um eine Grundpauschale. Auch regionale für KMU-fokussierte OPaaS-Angebote machen Schule: Beim in Friedberg (bei Augsburg) ansässigen Systemhaus **NCS** können Unternehmen komplette RZ-Infrastruktur im ortsansässigen *LEW Green Data Center* beziehen – inklusive Storage, Security und Backup – als Managed-Service.

Die Kosten hängen stark vom Leistungsumfang, der Vertragslaufzeit, dem Service-Level und der eingesetzten Hardware ab. Grundsätzlich lassen sich OPaaS-Angebote in drei grobe Preismodelle einordnen: pauschale Miete pro Monat, Pay-per-Use und gemischte Modelle mit einer festen Basiskapazität plus flexible Nutzung bei Spitzenlasten.

Hybride Modelle als Zukunftsszenario

Die Zukunft der Datenspeicherung liegt daher nicht im Entweder-oder, sondern in hybriden Szenarien: Lokale Systeme, die geschäftskritische Workloads performant, sicher und unter voller Datenhoheit verarbeiten, kombiniert mit Cloud-Komponenten, die für zusätzliche Flexibilität sorgen – etwa bei der Skalierung, bei der Archivierung großer Datenmengen oder für die standortübergreifende Zusammenarbeit in global agierenden Teams.

Entscheidend ist dabei nicht die eingesetzte Technologie im Einzelnen, sondern die passgenaue Integration aller Komponenten in eine übergeordnete, unternehmensspezifische IT-Strategie. Eine durchdacht integrierte hybride Infrastruktur erlaubt es, die jeweiligen Stärken beider Welten optimal zu nutzen. »Wie die Performance, Kontrolle und Planbarkeit von On-Premises mit der Elastizität und globalen Erreichbarkeit der Cloud«, sagt Fast-LTA-Manager Heckel. Dabei lassen sich Datenflüsse, Speicherorte und Zugriffspfade gezielt steuern, priorisieren und absichern – technisch beispielsweise durch Tiered-Storage, intelligente Data-Mover, Policy-Engines oder Cloud-Gateways. »Wobei wir die Cloud in den meisten Fällen nur als dritte Instanz sehen«, schränkt Heckel ein.

Die gezielte Kopplung schafft nicht nur Effizienz, sondern soll auch für eine bessere Kostenkontrolle, Compliance-Abbildung und betriebliche Resilienz sorgen. »Unternehmen speichern sensiblen Daten lokal, während weniger kritische Inhalte automatisiert in Cloud-Archive ausgelagert werden«, ergänzt Thomas-Krenn-Manager Seibold. »Backup-Strategien lassen sich mit cloud-basierten Offsite-Kopien ergänzen, ohne dass der laufende Betrieb auf lokale Wiederherstellungszeiten verzichten muss.«

Auch für die zunehmende Nutzung von KI und Analytics bieten hybride Speicherarchitekturen einen klaren Vorteil: Trainingsdaten lassen sich lokal effizient vorbereiten und analysieren, während rechenintensive oder kollaborative Prozesse in die Cloud ausgelagert werden können – ohne den Kontrollverlust über Datenhaltung und Governance zu riskieren.

Kurz gesagt: Die richtige Mischung macht den Unterschied – aber nur dann, wenn sie nicht aus Zufall oder reiner Kostenerwägung entsteht, sondern aus einer klaren Zielsetzung heraus. Auch wenn die Cloud-Fraktion es anders sieht, On-Premises bleibt als fester Bestandteil einer ausgewogenen Infrastruktur unverzichtbar. ■

FAST LTA

www.fast-lta.de

Wir sind die Spezialisten für Sekundärspeicher, für Archivierung und Backup.

Unsere Produkte und Services helfen mittelständischen Anwendern, Datensicherung und Datenmigration zu vereinfachen, rechtliche und regulatorische Risiken zu minimieren, und das langfristige Risiko, Daten zu verlieren, nachhaltig zu verringern.

Sitz der Gesellschaft:
München

Jahr der Gründung:
1999

Zielgruppe:
KMUs, VARs und Industriekunden

Actidata

actidata.com

Die actidata Storage Systems GmbH mit Sitz in Dortmund ist ein innovativer IT-Hersteller mit Schwerpunkten im Bereich Backup, Storage und Archivierung. Das Unternehmen konzentriert sich mit einem Netzwerk professioneller Systemhäuser auf das Industrie- und Geschäftskundensegment mit dem Ziel, professionelle Speicherlösungen zu platzieren.

Sitz der Gesellschaft:
Dortmund

Jahr der Gründung:
2009

Zielgruppe:
Systemhäuser, VARs und Industriekunden

Thomas-Krenn

www.thomas-krenn.com

Die Thomas-Krenn.AG ist einer der größten Hersteller individueller Server- und Storage-Systeme in Deutschland. Derzeit produzieren rund 230 Mitarbeiterinnen und Mitarbeiter alle Systeme am Standort Freyung nach dem Build-to-Order-Prinzip. Thomas-Krenn steht für höchste Servicequalität bei hardwarenaher Entwicklung, Auftragsfertigung, Lösungsentwicklung, Produktveredlung und Logistik – als zuverlässiger Partner für Industrie, Systemhäuser, Dienstleister und mittelständische Endkunden aller Branchen.

Sitz der Gesellschaft:
Freyung

Jahr der Gründung:
2002

Zielgruppe:
KMUs, Behörden, Bildungseinrichtungen, Systemhäuser, VARs und Industriekunden

EDP

www.edp-germany.de

Die EDP Vertriebs GmbH ist Spezialanbieter für IT-Lösungen mit Fokus auf die Bereiche: Storage-Management, Netzwerkmanagement und Infrastrukturmanagement. Qualifizierte und engagierte Mitarbeiter stellen sicher, dass über die kompetente Beratung die beste wirtschaftliche und technologische Lösung zum Einsatz kommt. Der Name EDP steht für Kompetenz, Innovation und kundenorientierte Lösungen. Und das alles aus einer Hand.

Sitz der Gesellschaft:
Walldorf

Jahr der Gründung:
2014

Zielgruppe:
KMUs, Behörden, Bildungseinrichtungen, Systemhäuser, VARs und Industriekunden

Mit einem NAS aus Deutschland unabhängiger werden – geht das?

BACKUP UND NAS AUF OPEN-SOURCE-BASIS

Hand aufs Herz: Hardware ganz ohne US-Know-how – und damit ohne zukünftige Firmware- oder Sicherheits-Updates – wird es auf absehbare Zeit kaum geben. Auch Open-Source hilft hier nur bedingt weiter. Doch was häufig unterschätzt wird: Eine Open-Source-Lösung bietet eine hohe Hardware-Kompatibilität. Eigene Systeme lassen sich dadurch auf unterschiedlichsten Plattformen betreiben.

Albrecht Hestermann
actidata Storage Systems

Datenschutz, Datensicherheit, Datenverfügbarkeit, Datensicherung, Wiederherstellung, Archivierung, Auslagerung, Integrität – all das sind aktuell hochrelevante Themen, insbesondere für Admins im Mittelstand. Meist geht es vorrangig darum, Angriffe über das Netzwerk – etwa durch Ransomware – und unautorisierten Datenzugriff abzuwehren.

Backup wird dabei oft vernachlässigt oder nur halbherzig umgesetzt. Doch ein fehlerhaftes Backup verhindert im Ernstfall die Wiederherstellung produktiver Systeme oder virtueller Maschinen. Eine definierte Datensicherungsstrategie ist deshalb zentraler Bestandteil eines wirksamen Cybersicherheit-Konzepts.

actidata RMx-DX6 – Kombiniertes NAS- & Backup-System mit actidata-Betriebssystem.

Aktueller Trend: Cloud-Daten lokal sichern

Es ist bemerkenswert, dass Unternehmen zunehmend eine zusätzliche lokale Sicherung ihrer Cloud-Daten umsetzen – beispielsweise für *Microsoft 365*. Ob das Vertrauen in Cloud-Provider schwindet oder der Wunsch nach Redundanz unternehmenskritischer Daten ausschlaggebend ist – wahrscheinlich beides. Der Bedarf ist jedenfalls klar vorhanden und wird teils sogar als existenziell eingeschätzt.

Open-Source – ein Gebot der Stunde?

Open-Source klingt zunächst vielversprechend als Grundlage für kostengünstige Storage-Lösungen: ein offener Standard, allgemein zugänglich, entwickelt von einer engagierten Community. Ziel ist ein Speichersystem, das unabhängig von großen Herstellern betrieben werden kann – geschützt vor unautorisierten Zugriffen durch Unternehmen, Gruppen oder Behörden.

Ganz so einfach ist es allerdings nicht: Eine Lösung besteht immer aus Hard- und Software. Und bei der Hardware ist das Know-how in Deutschland und Europa eher begrenzt. Umso wichtiger ist ein Systemdesign, das möglichst breit hardwarekompatibel ist. Gerade Open-Source-Software ist durch ihre Abwärtskompatibilität oft auch auf älterer Hardware einsetzbar. Die Kombination aus Flexibilität und

Bild: via DALL-E

Kurzüberblick: actidata RMx-DX6

Das *actidata RMx-DX6* ist ein Hochverfügbarkeits-Backup-System für Mittelstand und Enterprise, das NAS-Leistung, Off-Site-Wechselplatten und offene Hardware kombiniert. Es bietet eine robuste, skalierbare und administrative-fit Lösung – mit der nötigen Performance und Sicherheit für zentrale Datensicherungsszenarien.

- **Modulares 2U Rack-NAS mit DX6-Wechselplatten**
  Das RMx-DX6 kombiniert ein leistungsstarkes Linux-NAS mit einem integrierten DX6-Wechsellaufwerk für zuverlässige Offsite-Backups.
- **Leistung & Konnektivität**
  Ausgestattet mit 6 Enterprise-HDDs im RAID (RAID 0, 1, 5, 6 inkl. Hotspare), Dual-10/25 GbE SFP28 und vier weiteren 1 GbE-Ports. Host-IPMI für professionelles Monitoring und Management.
- **Offline-Medium DX6**
  DX6-Wechselmedien (4-20 TByte) basierend auf CRU-Technik bieten hohe Transferleistung (~250 MByte/s) und sind offenes Format–SATA-Festplatten ohne proprietäre Restriktionen
- **Backup-Workflow B2D2DX6 (Backup-to-Disk-to-DX6)**
  Das integrierte actidata Backup-Tool überträgt zunächst auf das interne NAS-RAID, anschließend automatisiert auf die eingelegte DX6-Wechselplatte – ideal für die 3-2-1-Backupregel
- **Betriebssystem & Flexibilität**
  Linux-NAS-OS mit M.2 NVMe-Cache, Installation alternativer Backup-Software (z. B. *IPERIUS*, *Windows Server IoT for Storage*) möglich.
- **Service & Support**
  36 Monate *Fast-Exchange*-Garantie mit Vorabaustausch kaputter Komponenten, kostenloser deutschsprachiger Support – optional bis 60 Monate

What's next: Business-NAS der Calderas-Serie

Die Roadmap für 2025 ist bereits gesetzt. Noch in diesem Jahr will actidata die neue CALDERAS-Produktlinie einführen. Diese Business-NAS-Systeme basieren ebenfalls auf dem firmeneigenen Linux-Betriebssystem und setzen auf robuste Server-Hardware inklusive professionellem Mainboard-Management.

Die Calderas-Systeme sind als Backup-Ziel konzipiert und sollen sich nahtlos in unternehmensweite Datensicherungsstrategien integrieren. Das Motto: einfach zu bedienen, leistungsoptimiert, kosteneffizient – und mit Support direkt aus Deutschland. ■

Kompatibilität ist daher essenziell für den Aufbau möglichst unabhängiger Systeme.

actidata setzt auf Open-Source – made in Germany

Ein Open-Source-basierter Ansatz sollte grundsätzlich in Betracht gezogen werden, wenn es um mehr Unabhängigkeit geht. **actidata** verfolgt genau diesen Weg und bietet kombinierte NAS- und Backup-Systeme, die auf *DX6*-Medien zur externen (Offsite-)Lagerung der Backup-Sets setzen. Grundlage ist bewährte Hardware in Verbindung mit dem von actidata in Deutschland entwickelten Linux-basierten Betriebssystem.

Ein Alleinstellungsmerkmal ist das integrierte Backup-Tool, das Sicherungssätze automatisch und zeitgesteuert auf ein eingelegtes DX6-Medium überträgt. Dabei handelt es sich um ein offenes Speichermedium auf Basis handelsüblicher Enterprise-Festplatten – ohne proprietäre Zugriffsbeschränkungen. Und trotzdem wird Sicherheit großgeschrieben: Zwei-Faktor-Authentifizierung gehört bei den actidata-Systemen zur Standardausstattung.

Weitere Informationen:

actidata Storage Systems GmbH
Wulfshofstr. 16
44149 Dortmund
Tel. 0231/96 36 32-0
E-Mail: info@actidata.com
www.actidata.com

Controller X und Silent Brick Pro, Max und SE

MODERNES BACKUP OHNE TAPE UND OHNE CLOUD

Hannes Heckel
FAST LTA

Ein vollständiges und sicheres Backup sowie ein zuverlässiges und schnelles Restore sind die Grundlagen zum Schutz vor Datenverlust und den Folgen einer Cyber-Attacke. Durch die ständige Bedrohung durch Cyber-Attacken verschiebt sich der Fokus vom Backup zum Recovery. RTO und RPO definieren die unterschiedlichen eingesetzten Technologien, Immutability schützt vor Datenmanipulation und Datenverlust.

Zwei in der Datensicherung eingesetzte Technologien stehen derzeit bei vielen Unternehmen und Behörden auf dem Prüfstand. Tape schneidet mit der Einführung von LTO-10 die Verbindungen zu Vorgängergenerationen komplett ab, was eine vollständige Migration aller Systeme und Daten notwendig macht. Und die Public-Cloud entpuppt sich in vielen Fällen als wesentlich teurer als angenommen und leidet zudem unter rechtlichen Bedenken durch US-CLOUD-Act und die neue US-Regierung. Viele Mittelständler überlegen deshalb, ihre Daten zurückzuholen und sehen sich nach lokalen Alternativen für die Datensicherung um.

Das erneuerte **FAST LTA** *Silent Brick System* mit *Controller X* und den Speichermodule *Silent Brick Pro*, *Max* und *SE* bietet für den Mittelstand hohe Flexibilität, beste Performance sowie höchste Sicherheit und maximale Datensouveränität.

Datensicherung für schnelles Recovery

Die ersten beiden Instanzen eines modernen Backups kümmern sich um den schnellen und vollständigen Restore. Egal, ob Daten versehentlich gelöscht, durch falsche Synchronisierung ungewollt überschrieben, oder durch einen Angriff verschlüsselt oder gelöscht wurden: Eine schnelle und vollständige Wiederherstellung erfordert hohe Performance und Sicherheit für Tage, Wochen oder Monate an Daten.

Im Primary-Target liegt der Fokus auf Geschwindigkeit. Hier werden Daten der letzten Wochen abgelegt – deren Verlust würde den Betrieb unmittelbar beeinträchtigen. Die Recovery-Time-Objective (RTO, Zielsetzung für die Zeit zur Wiederherstellung) legt fest, wie lange dies im Ernstfall dauern darf. In den meisten Fällen werden hierfür Flash- bzw. NVMe-basierte Speicher eingesetzt, die für einen schnellen Restore, insbesondere beim Wiederherstellen aus beliebigen Incrementals oder Differentials, ausreichende Performance bereitstellen.

Der neue Silent Brick Controller X bietet dafür die ideale Grundlage. Die dazu passenden Silent Brick Pro Speichereinheiten enthalten je zwölf NVMe-Module, die zusammen 48 oder 96 TByte (brutto) bereitstellen und je nach Konfiguration bis zu 6 GB/s kontinuierlichen Datendurchsatz erlauben. Im Controller *X6080* mit 8 Slots sind so bis zu 768 TByte auf zwei Höheneinheiten realisierbar.

Die zweite Instanz, das Secondary-Target, ist die eigentlich Basis-Sicherung, die meist einige Monate zurückreicht. Für definierte Wiederherstellungspunkte sorgen Full Backups (synthetisch oder aktiv) in regelmäßigen Abständen. Die großen Daten-

mengen liegen auf HDD-Arrays, die durch Redundanz vor Datenverlust durch Ausfall einzelner Festplatten schützen. Besonders wichtig ist es inzwischen, die Backups selbst vor den Folgen eines Angriffs zu schützen. Immutability sorgt hier für Schutz vor Datenmanipulation (Verschlüsselung) und Verlust (Löschen) durch Cyber-Attacken.

Der Controller X lässt sich extern mit bis zu 26 Silent Brick Max erweitern. Die Speichereinheiten enthalten jeweils zwölf robuste SATA-Festplatten und bieten Kapazitäten von bis zu 240 TByte (brutto) auf einer Höheneinheit. Konfigurationen mit bis über 6 PByte pro Controller sind dadurch realisierbar.

Auslagerung für den Extremfall

Cyber-Attacken zielen inzwischen primär auf Backups, um eine erfolgreiche Wiederherstellung nach einem Angriff zu verhindern. Aber auch ein Komplettausfall eines Speichersystems kann den Verlust kompletter Backups zur Folge haben. Deshalb ist es wichtig – und in vielen Umgebungen auch vorgeschrieben – weitere Maßnahmen zum Schutz vor Datenverlust umzusetzen.

Die Auslagerung von Daten (Geo-Redundanz) gehört hier zu den bewährten Methoden. Dabei werden in regelmäßigen Abständen und über Monate oder Jahre hinweg vollständige, eigenständige Backups angelegt. Auch diese Backups müssen per Immutability vor den Folgen eines möglichen Angriffs geschützt sein.

Physisch entnehmbare Speichermedien trennen die Daten vom Storage-System und erzeugen ein sogenanntes Air-Gap.

Eine Art der Auslagerung ist das Überführen von Full Backups in eine separate Speicherinstanz, die auf einer anderen Technologie basiert. Hier haben sich S3-kompatible Objektspeicher als Quasi-Standard durchgesetzt. Ursprünglich für Cloud-Anwendungen entwickelt, bieten lokale S3-kompatible Speicher maximale Datensouveränität ohne die zusätzlichen Kosten, die bei Cloud-Speichern durch Datenübertragungen entstehen.

Im Silent Brick System kann ein S3-kompatibler Objektspeicher mit Immutability durch Object-Locking im selben System parallel betrieben werden. Auf den schnellen Silent Brick Pro lassen sich bei Bedarf Geschwindigkeiten bis zu 1 GB/s erreichen.

Physisch entnehmbare Speichermedien ermöglichen den sogenannten Air-Gap. Da keine Verbindung mehr zum eigentlichen Speichersystem besteht, besteht so höchste Sicherheit vor Missbrauch. Das Medium selbst sollte ebenfalls vor Datenverlust, etwa durch Beschädigung, geschützt sein.

Mit dem neuen Silent Brick SE bietet das Silent Brick System ein speziell für Air-Gap optimiertes, mit gut 500g besonders leichtes und kompaktes Speichermedium an. Dank physischer Entnahme inklusive automatischem Auswurf ist echtes Air-Gap möglich. 12 oder 24 TByte (brutto) stehen je Silent Brick SE zur Verfügung. Die zwölf NVMe-Module sind durch Redundanz-Codierung vor Datenverlust durch Ausfall geschützt, das stabile Aluminiumprofil und die Schutzhülle halten physische Strapazen fern.

Optimierungen zur Kostenkontrolle

Die oben beschriebenen Methoden schützen Daten und Backups vor Manipulation und Verlust. In heutigen Umgebungen entstehen beim Backup jedoch häufig so große Datenmengen, dass vorgesehene Backup-Fenster oft nicht mehr eingehalten werden können. Noch dramatischer wirken sich steigende Datenmengen auf das Recovery aus. Im Falle eines Angriffs stellt die Ausfallzeit den größten Kostenfaktor dar. Ein schnelleres Recovery reduziert die durch eine Cyber-Attacke anfallenden Kosten also erheblich.

Speziell für den Einsatz mit *Veeam Backup & Recovery* hat das Silent Brick System dafür eine besondere Technologie, die Backups sehr viel kleiner und Backup & Restore erheblich schneller macht: Fast Clone Support. Dabei werden im Primary und Secondary Target Vervielfachungen von Daten weitgehend vermieden. Bereits geschriebene Datenblöcke werden bei weiteren Incrementals und Synthetic Fulls nur noch referenziert, was fast keinen Speicherplatz und minimale Schreib-/Lese-Vorgänge erfordert. Je nach Anwendung können so Kapazität und Dauer für Backup und Wiederherstellung um 50 Prozent und mehr verringert werden.

Noch effektiver für die Reduzierung von Kapazitätsbedarf und Dauer ist es, die Menge der zu sichernden Quelldaten zu reduzieren. Da in vielen Unternehmen große Datenmengen auf File Servern liegen und sich nach der Nutzung gar nicht mehr oder nur noch selten verändern, würden diese Daten immer wieder unverändert gesichert werden und so das Backup unnötig vergrößern.

Eine separate Sicherung von File Servern auf einen sicheren Speicher bietet dabei gleich mehrere Vorteile. Da die Sicherung besonders einfach ist – Replizierung der Inhalte auf den sicheren Speicher – lässt sich die Komplexität der Datensicherung deutlich reduzieren. Der Backup-Speicher kann schlanker dimensioniert oder zum gleichen Preis mit höherem Anteil an Performance-Komponenten betrieben werden. Der vielleicht wichtigste Aspekt ist jedoch, dass im Falle einer Cyber-Attacke Daten überhaupt nicht mehr wiederhergestellt werden müssen, um wieder nutzbar zu sein.

Das erfordert eine transparente Sicherung (1:1-Kopie) und einen hochsicheren Speicher. Mit den Silent Cubes hat FAST LTA ein Storage im Portfolio, das Daten mittels Hardware-WORM versiegelt und so zu 100 Prozent vor Manipulation und Löschen schützt. ■

Weitere Informationen:

FAST LTA GmbH

Rüdesheimer Str. 11,
80686 München
Tel. 089/89 047-0
E-Mail: info@fast-lta.de
www.fast-lta.de

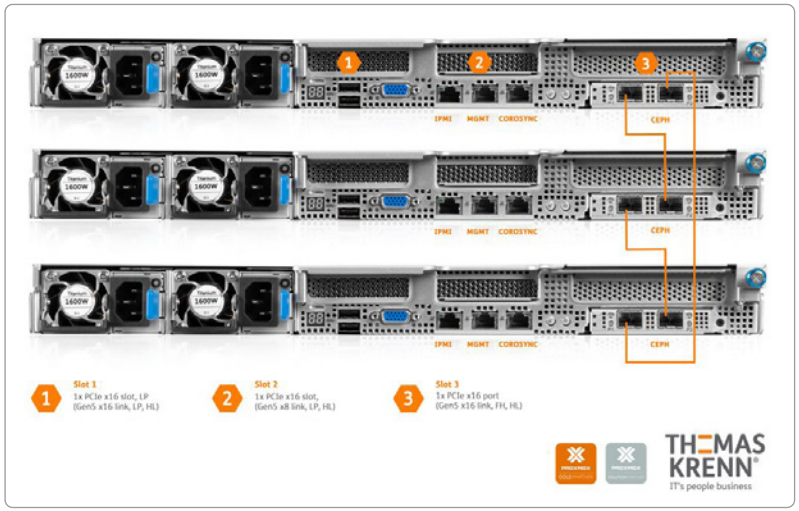
Überblick ZFS und Ceph im Proxmox-VE-Kontext

ENTSCHEIDUNGSHILFE: DAS RICHTIGE STORAGE FÜR PROXMOX VE

Jonas Sterr
Thomas-Krenn

Proxmox VE ist eine Open-Source-basierte Virtualisierungs-Software, die seit etwa 2023 rasant an Beliebtheit gewonnen hat. Der Einstieg gestaltet sich dank eines intuitiven Web-UI leicht. Dabei gibt es viele Möglichkeiten, Storage-Lösungen anzubinden oder Software-Defined Storage anzulegen und zu verwenden. Wir zeigen, welche Storage-Lösung zu welchem Anwendungsfall passt.

Wer in die Welt von *Proxmox VE* (PVE) eintauchen möchte, wird sich auch Gedanken zur Storage-Auswahl machen. Zwischen den verfügbaren Storages wird wie folgt grundsätzlich unterschieden: Welche Architekturmodelle sind nutzbar, also file- oder blockbasiert, oder beides; kann das Storage für Cluster, für mehrere Hosts bereitgestellt werden, ist es geteilt oder repliziert verfügbar (Shared-Storage); können virtuelle Ressourcen nach Ablage auf dieses Storage gesnapshottet werden, oder werden Snapshots eingeschränkt / gar nicht unterstützt?



Full-Meshed-Network für Ceph – mögliche Beispielkonfiguration mit Thomas-Krenn-Servern.

Bild: Thomas-Krenn

Daraus resultieren zwei weitere Fragen: Möchte ich den Host später als Cluster-Member verwenden, oder wird der Single-Host immer ein Einzel-System bleiben? Und falls die Entscheidung zu einem Cluster-System oder einem zukünftigen Einsatz innerhalb eines Clusters getroffen wurde: Muss/möchte ich ein Shared-Storage verwenden, das zum Beispiel auf iSCSI, NFS, Fibre-Channel (FC) oder NVMe over TCP setzt? Oder wäre *Ceph* im Zusammenspiel mit einer Hyper-Converged-Infrastructure (HCI) die bessere Alternative?

Cluster-Perspektive & Storage-Strategie: Weichenstellung bei der Hardware-Wahl

Zu Frage 1: Relevant in dieser Hinsicht ist vor allem die Auswahl der Hardware – wird beispielsweise ein Server-System mit RAID-Controller verwendet, ist für dieses System ein HCI-Setup mit ZFS (Single-Node) oder Ceph (ab 3-Node-Cluster) ausgeschlossen: Beide Lösungen unterstützen keine RAID-Controller bzw. verbieten diese ausdrücklich. So ist bei der Auswahl der ersten Nodes schon mehr oder weniger gesetzt, ob dieser jemals Teil eines HA-Clusters sein wird oder nicht. Bei Unklarheit: Eine Entscheidung zu ZFS mit Einsatz eines Host-Bus-Adapters (HBA) für den ersten Single-Host gibt die Flexibilität, alle Möglichkeiten für künftige Upgrade-Wege offen zu halten.

Tipp: Viele RAID-Controller kann man im Nachhinein in einen HBA-Modus versetzen.

Zu Frage 2: Mit iSCSI sowie allen anderen Storage-Konzepten mit Logical-Unit-Numbers (LUN) – also FC, NVMe-Over-TCP – lassen sich im PVE-Kontext keine Snapshots erstellen. NFS bietet im Zweifel unter Verwendung eines anderen Disk-Formats mit dem Namen »qcow2« die Möglichkeit, Snapshots von solchen Disks zu erstellen – derzeit allerdings nicht über Windows-Server mit virtuellem TPM-Modul. Letztlich fällt Shared-Storage realistisch meist aus dem Rennen, da Snapshotting oft ein zentraler Bestandteil von IT-Prozessen ist und der Wechsel auf Alternativen, etwa Backups statt Snapshots vor einem Upgrade, unerwünscht und/oder zu umständlich ist.

ZFS

Ein häufiges PVE-Storage ist daher das Zettabyte File System (ZFS). Denn es kann sowohl bei PVE-Single-Hosts als auch bei einem ZFS-2-Node-Cluster mit asynchroner VM- oder CT-Replikation verwendet werden. Empfohlen für einen produktiven Enterprise-Betrieb ist der Einsatz von Flash-Speicher in Kombination mit RAID-1 oder RAID-10; der Einsatz von RAID-5 und RAID-6 ist aufgrund des negativen Impacts durch die Paritäts-Bildung auf die Storage-Performance nicht die beste Auswahl. Ist die Entscheidung für ein RAID-Level gefallen, sollte in ZFS pro Terabyte Bruttospeicher ein Gigabyte RAM für das SDS eingerechnet werden – zusätzlich zu den geplanten virtuellen Ressourcen wie VMs und Container. Extra CPU-Ressourcen muss man bei modernen Enterprise-Servern nicht miteinrechnen.

ZFS passt auch gut zu 2-Node-Clustern mit einem externen Witness-Host. Bei gleichnamigen Storage-Pools auf Node-1 und Node-2 kann pro virtuelle Ressource ein Replikations-Job angelegt werden, welcher asynchron in definierten Intervallen – z. B. alle fünf Minuten – eine Replikation der Storage-Disks durchführt. Einmal eingerichtet, können VMs und Container auch ins HA-Programm von PVE hinzugefügt werden. Bei Ausfall eines Knotens starten VMs und Container so automatisch auf dem zweiten Server neu; der Storage-Stand entspricht dabei dem letzten stattgefundenen Replikations-Job. Ein möglicher Storage-Verlust hängt somit davon ab, wie oft asynchron über Replikations-Jobs repliziert wird.

RAID-Flexibilität und Performance mit ZFS

Zusätzlich lassen sich im ZFS-Kontext sehr komplexe RAID-Level bauen; sie sind flexibler/skalierbarer und sehr performant. Denn neben den herkömmlichen RAID-Leveln können auch solche mit mehreren virtuellen Devices (VDEV) gebildet werden. Ein Beispiel: Ein RAID-1 benötigt ja immer zwei Festplatten zur Datenspiegelung. Das limitiert die Performance, da auf einen Datenträger geschrieben und auf dem zweiten gespiegelt werden kann. Die einzelne Disk ist somit ein Flaschenhals. Mit ZFS lassen sich in einem Pool mehrere RAID-1-Gruppen in Form von VDEVs betreiben, die zusammen den Pool als eine Einheit bilden. So lässt sich zum Beispiel ein großer Storage-Pool aus fünf RAID-1-Gruppen erstellen. Die Vorteile: erhöhte Performance, da auf jede RAID-Gruppe parallel schreibend/lesend zugegriffen werden kann; Flexibilität bei der Skalierung, da Erweiterungen des Pools einfach über zusätzliche RAID-Gruppen möglich sind.

Ceph

Ceph ist eine Open-Source-Storage-Lösung, die in Proxmox VE eine umfangreiche und nutzerfreundliche Integration erfahren hat. Es kann über einen Installations-Wizard der Web-UI vollständig installiert, konfiguriert und über ein eigenes Dashboard administriert werden. Und es ist die richtige Storage-Auswahl für alle, die ein PVE-Cluster mit einem Storage betreiben möchten, das sowohl hyperconverged als auch mit synchronen Storage-Writes arbeitet – ein Muss für ein Enterprise-Storage!

Doch welche Hardware-Anforderungen gibt es zu beachten? Der Regelfall wird der Einsatz einer hyperkonvergenten Infrastruktur sein. Ergänzend zu den Ressourcen für VMs und Container müssen Arbeitsspeicher, CPU, Netzwerk und natürlich Storage-Speicher für Ceph einberechnet werden. Außerdem gibt es eine Mindestanzahl an Nodes. Die Mindestanforderungen in der Übersicht:

- **Node-Anzahl:** Für die Ablage von mindestens drei Storage-Kopien ist ein Minimum von drei Nodes erforderlich.
- **Datenträger:** in der Regel NVMe; wenn nicht möglich SATA/SAS-SSDs. HDDs sind nicht performant genug für ein Enterprise-Setup. Pro Node ist es empfehlenswert, mindestens mit vier Datenträgern zu kalkulieren, bei drei Nodes also mindestens zwölf NVMe-SSDs im Gesamtsystem.
- **CPU:** Ein CPU-Thread pro Datenträger wird empfohlen – bei zum Beispiel vier Datenträgern je Node also vier Threads pro Node.
- **Netzwerk:** Die Kommunikation wird bei Ceph über das Netzwerk abgewickelt, daher bedeutet ein schnelleres Netzwerk auch ein schnelleres Storage. Deswegen sind Netzwerkkarten mit 25 Gbit/s das empfohlene Minimum, das Optimum sind 100 Gbit/s.

Wer diese Sizing-Rules berücksichtigt hat, kann Ceph nach erfolgreichem Clustering über die Web-UI installieren und im Anschluss dort die erforderlichen Dienste erstellen sowie administrieren. Zuletzt wird ein Storage-Pool angelegt – bei Ceph ist hierbei einer für eine dreifache Datenspeicherung erforderlich (Pool-Parameter: SIZE). Diese drei Kopien werden auf unterschiedliche Nodes geschrieben, bevor der Storage-Client (PVE-Hypervisor) ein Write-Acknowledgement bekommt. Geschriebene Daten sind also sicher und korrekt dreifach abgelegt. Über den Pool-Parameter MIN-SIZE wird außerdem die Anzahl an mindestens vorhandenen Replikaten definiert, damit der Ceph-Pool erreichbar ist. Haben Sie all dies berücksichtigt, steht der Hochverfügbarkeit Ihrer Ressourcen in Proxmox VE nichts mehr im Weg!

Projekte profitieren von erfahrenen Partnern

Generell gilt noch: Hilfreich ist bei entsprechenden Projekten ein erfahrener Partner, der hochwertige Server-Hardware und umfassendes Proxmox-Know-how vereint. Ein solches Unternehmen ist die **Thomas-Krenn.AG**, die seit vielen Jahren Proxmox Gold sowie Solution Partner ist und für **Proxmox VE maßgeschneiderte Server-Lösungen** entwickelt und liefert – von der Hardware-Auswahl über die Proxmox-Vorinstallation bis hin zu kompletten Cluster-Setups. Die Server sind speziell für den Einsatz mit Proxmox VE konfiguriert und getestet, für maximale Leistung und Stabilität. ■

Weitere Informationen:

Thomas-Krenn.AG

Speltenbach-Steinäcker 1,
94078 Freyung
Tel. 08551/91500
Proxmox optimierte Thomas-Krenn-Systeme
www.thomas-krenn.com/speicherguide25

MagStor schließ Support-Lücke für ehemalige Overland-Tandberg-Kunden

NEUER HERSTELLER FÜR LTO-TAPE-LÖSUNGEN

MagStor positioniert sich als zuverlässiger Partner für Unternehmen, die auf LTO-Bandtechnologie setzen. Ein besonderer Fokus liegt dabei auf Kunden, die Produkte von Overland-Tandberg im Einsatz haben. Neben kompatibler Hardware bietet der US-Hersteller weltweiten Service, Support und Garantieoptionen für bestehende Installationen – in Deutschland exklusiv über EDP.

Hannes Sieverling
EDP

Mit dem Rückzug oder der eingeschränkten Verfügbarkeit von Overland-Tandberg-Produkten und -Services steht für viele IT-Abteilungen die Frage im Raum: Wie lassen sich bestehende Bandarchitekturen weiter betreiben, erweitern oder sinnvoll ersetzen? **MagStor** hat sich genau auf diese Lücke spezialisiert und bietet eine breite Palette an Lösungen, die sowohl bestehende Overland-Tandberg-Infrastrukturen unterstützen als auch durch neue Optionen ergänzen.

Wichtig für frühere Kunden von Overland-Tandberg: Seit der Betrieb des Herstellers im Februar 2025 eingestellt wurde, sind sämtliche bestehenden Garantie- und Support-Verträge erloschen – entsprechende Hinweise wurden von Overland-Tandberg an die Kundschaft kommuniziert.

Service und Support für Bestandskunden

Ein zentrales Element des MagStor-Angebots ist der direkte technische Support für Anwender, die bisher mit Overland-Tandberg-Systemen gearbeitet haben. Über das firmeneigene Support-Portal steht ein kostenloser technischer Support zur Verfügung – inklusive Fehlerdiagnose, Hilfestellung bei der Fehlerbehebung und Firmware-Updates. Die Services richten sich dabei sowohl an Endkunden als auch an IT-Dienstleister, die im Auftrag ihrer Kunden bestehende Bandlösungen betreuen.

MagStor bietet darüber hinaus weltweite Garantie- und Vor-Ort-Service-Verträge für bestehende Installationen. Diese richten sich explizit an Overland-Tandberg-Kunden, deren Systeme weiter betrieben werden sollen, jedoch auf herstellerseitige Unterstützung verzichten müssen. Derartige Service-Angebote können helfen, vorhandene Investitionen abzusichern und die Nutzungsdauer der Infrastruktur zu verlängern.

Kompatible Ersatzlaufwerke und Komponenten

Neben der Service-Unterstützung liefert MagStor auch passende Ersatz- und Upgrade-Komponenten für bestehende Bandumgebungen. Dazu zählen neue, mit Overland-Tandberg-Systemen kompatible LTO-7-, LTO-8- und LTO-9-Bandlaufwerke. Diese werden mit voller MagStor-Garantie ausgeliefert und ermöglichen sowohl Reparaturen als auch gezielte Upgrades innerhalb bestehender Tape-Architekturen. Ebenfalls im Portfolio sind Ersatzteile für Overland-Tandberg-Bandbibliotheken. Dazu zählen unter anderem Komponenten wie Netzteile, Roboterarme oder Magazinkassetten. Diese Ersatzteile ermöglichen es, defekte Systeme instand zu setzen, ohne vollständig auf neue Hardware umsteigen zu müssen.

MagStor Bandbibliotheken und LTO-Lösungen

MagStor beschränkt sich jedoch nicht auf die Kompatibilität zu Overland-Tandberg, sondern bietet ein breites Portfolio eigener LTO-Produkte an – darunter Bandlaufwerke und Tape Librarys für unterschiedlichste Einsatzszenarien.

Im Bereich der Laufwerke stehen sowohl Desktop- als auch Rackmount-Varianten zur Verfügung. Die Produktreihe reicht von kompakten 1U-Einstiegsmodellen bis hin zu skalierbaren Enterprise-Lösungen.

Die *M1000*-Serie richtet sich an kleinere IT-Umgebungen und bietet in einem 1U-Gehäuse acht Slots mit einer unkomprimierten Kapazität von bis zu 144 TByte bei LTO-9. Diese Modelle sind mit SAS- oder Fibre-Channel-Schnittstellen ausgestattet und liefern eine Datenübertragungsrate von bis zu 1,1 TByte/h. Für mittelgroße Anforderungen bietet die M2000-Serie in einem 2U-Gehäuse 24 Slots und unterstützt bis zu zwei Laufwerke, was eine unkomprimierte Kapazität von bis zu 288 TByte ermöglicht.

Die *M3000E*-Serie ist ausgelegt für große IT-Umgebungen: Sie beginnt mit 40 Slots in einem 3U-Gehäuse und lässt sich durch Erweiterungseinheiten auf bis zu 640 Slots und 48 Laufwerke skalieren. Dadurch ist mit LTO-9 eine unkomprimierte Gesamtkapazität von bis zu 11,5 PByte möglich. Diese Modelle unterstützen sowohl SAS- als auch Fibre-Channel-Schnittstellen und bieten eine maximale Datenübertragungsrate von 23,1 TByte/h.

Die Tape-Librarys von MagStor zeichnen sich durch ihre modulare Architektur und breite Schnittstellenunterstützung aus, darunter SAS, Fibre-Channel und Thunderbolt sowie 10GbE. Sie sind kompatibel mit LTO-7-, LTO-8- und LTO-9-Medien und unterstützen Funktionen wie LTFS, AES-256-Verschlüsselung und WORM. Alle Modelle bieten eine dreijährige Garantie ohne zusätzliche Lizenzkosten für die Nutzung aller Slots. Dank standardisierter Treiber sind sie mit den meisten Backup- und Archivierungs-Programmen kompatibel, was eine nahtlose Integration in bestehende IT-Infrastrukturen ermöglicht.

MagStor Thunderbolt Slim LTO-9: Desktop-Tape-Drive mit Highspeed-Anschluss

Das *MagStor LTO-9 Slim Desktop Tape Drive* gehört zu den meistverkauften Modellen – insbesondere im Mac-Umfeld. Es kombiniert die LTO-9-Technologie mit einem kompakten Formfaktor und einem Thunderbolt 3/USB4-Anschluss, der besonders in Mac-Umgebungen und kreativen Workflows beliebt ist. Mit einer nativen Kapazität von 18 TByte pro Band und eine native Transferrate von bis zu 400 MByte/s eignet sich das Laufwerk hervorragend für Backup, Archivierung und Offload großer Datenmengen – etwa aus Video- oder CAD-Produktionen.

Das Besondere: Die Thunderbolt-Schnittstelle ermöglicht eine direkte, schnelle Anbindung an Workstations – ganz ohne SAS-Host-Adapter oder komplexe Infrastruktur. Damit wird Tape »plug-and-play«-fähig für den Schreibtisch-Einsatz. Die Lösung ist TAA-konform, läuft ohne proprietäre Software und unterstützt LTFS für den dateibasierten Zugriff auf Bänder wie auf ein externes Laufwerk. Das Gerät eignet sich insbesondere für Einzelanwender, kleine Studios und Teams mit professionellem Anspruch an Sicherheit, Mobilität und Performance.

Herstellerhintergrund und Zielgruppen

MagStor ist eine Marke der US-amerikanischen Firma **Magnext Inc.**, die sich seit fast 20 Jahren auf Speicherlösungen im Enterprise-Segment spezialisiert. Der Fokus liegt dabei klar auf LTO-Tape-Technologie, insbesondere für professionelle Medien-Workflows, Rechenzentren und Archivierungsanwendungen. Mit einem Firmensitz in Ohio und internationaler Vertriebsstruktur positioniert sich der Hersteller als unabhängiger Anbieter, der sowohl auf eigene Entwicklungen als auch auf OEM-Beziehungen mit etablierten Hardwareherstellern zurückgreift. Die Produkte sind »Made in USA« und richten sich bewusst an professionelle Anwender, die Wert auf langlebige und wartbare Infrastruktur legen.

Die adressierte Zielgruppe umfasst unter anderem Medienunternehmen, Forschungsinstitute, Behörden sowie kleine und mittelständische Unternehmen bis hin zu großen Rechenzentrumsbetreibern. Besonders für Organisationen, die LTO-Bänder als Langzeitarchiv nutzen oder mit großen Datenmengen in offline-fähigen Speicherlösungen arbeiten, bietet MagStor eine attraktive Kombination aus Service, Verfügbarkeit und Produktvielfalt.

EDP mit Sitz in Walldorf hat für MagStor die Exklusiv-Distribution in Deutschland und Europa übernommen. Ursprünglich als Spezialist für Barcode-Labels gestartet, hat sich EDP zu einem der größten Magnetband-Distributoren Europas entwickelt. Das Unternehmen liefert zuverlässig an Systemhäuser, Integratoren und Fachhändler – mit einem beachtlichen Tempo: Während der üblichen Bürozeiten geht im Schnitt alle 60 Sekunden ein Band über den virtuellen Ladentisch. Neben Standard-Medien umfasst das Sortiment auch Library-Zubehör, Etikettierungslösungen und maßgeschneiderte Services rund um Tape-Infrastrukturen. Zudem vertreibt EDP auch Quantum-Laufwerke und Tape-Librarys.

Tape-Technologie: Langlebig, sicher und kosteneffizient

Und den Bandtechnologien gehört weiterhin die Zukunft: Trotz wachsender Cloud-Nutzung und steigender SSD-Verbreitung bleibt Tape ein relevanter Baustein in der Datenspeicherstrategie vieler Unternehmen. LTO-Bänder punkten durch eine besonders hohe Datensicherheit im Offline-Betrieb (Air-Gap), niedrige Kosten pro TByte und eine bewährte Lebensdauer von bis zu 30 Jahren. Zudem lässt sich die Kapazität durch neue LTO-Generationen kontinuierlich steigern. Gerade für Archivierung, Backup und Compliance-relevante Daten sind Tapes auch in hybriden Infrastrukturen oft die robusteste Lösung. MagStor greift diesen Bedarf auf und kombiniert klassische Tape-Vorteile mit modernen Schnittstellen und flexiblem Service. ■

Weitere Informationen:

MagStor
7100 Huntley Rd,
Columbus OH 43229
Tel. 1-844-624-7867
https://magstor.com/

Exklusiv-Distributor:

EDP Vertriebs GmbH
Otto-Hahn-Str. 1c, 69190 Walldorf
Tel.: +49 (0) 62 27 / 8287 – 0
E-Mail: sales@edp-germany.de
www.edp-germany.de

Deutsche Firmen müssen umdenken

US-CLOUD-ACT: DAS UNTERSCHÄTZTE RISIKO

Der US-Cloud-Act existiert seit 2018 – und doch rückt er gerade jetzt ins Zentrum regulatorischer Diskussionen. Hintergrund: Veränderungen auf US-Seite gefährden die Datensicherheit europäischer Unternehmen. In diesem Artikel beleuchten wir, was der Act regelt, welche Risiken für deutsche Firmen bestehen – und warum gerade jetzt Handeln gefragt ist.

Karl Fröhlich
speicherguide.de

Der US-»Clarifying Lawful Overseas Use of Data« (Cloud Act), verabschiedet am 23. März 2018, erweitert den bereits bestehenden Stored-Communications-Act: US-Behörden dürfen über Gerichtsbeschluss oder Vorladung (»warrant or subpoena«) Daten von US-Firmen verlangen – unabhängig davon, wo die Daten physisch gespeichert sind. Das Gesetz richtet sich ausdrücklich an Unternehmen mit Sitz in den USA – unabhängig davon, ob Tochtergesellschaften im Ausland tätig sind. Diese extraterritoriale Reichweite stellt deutsche Unternehmen vor ernsthafte Probleme.

Konsequenzen für deutsche Firmen

Bei Nutzung von US-Cloud-Diensten (z. B. *AWS, Microsoft Azure, Google Cloud*) riskieren Unternehmen, dass personenbezogene Daten deutscher Bürger an US-Behörden übermittelt werden – im Widerspruch zur EU-Datenschutz-Grundverordnung (DSGVO). Diese verlangt beim Datentransfer in Drittstaaten besondere Schutzmechanismen – ein direkter Konflikt. 2019 erklärten EU-Aufsichtsbehörden, der Cloud-Act könne alleine keine rechtskonforme Datenübermittlung begründen.

Warum ist der Cloud-Act heute wieder problematisch?

Seit Jahresbeginn 2025 hat sich die Situation verschärft: US-Präsident *Trump* entließ im Februar drei demokratische Mitglieder des *Privacy and Civil Liberties Oversight Board* (PCLOB), das den Datenschutz bei Geheimdiensten überwacht – ein Kernbestandteil des neuen EU-US-Rahmenwerks, dem *Transatlantic Data Privacy Framework* (TADPF). Das US-Justizministerium kann nun innerhalb von 45 Tagen beginnen, Bidens Richtlinien zu nationaler Sicherheit zu ändern – und das könnte die Schutzmechanismen aus dem TADPF aufheben.

Aktuelle Auswirkungen & Gegenmaßnahmen

Unternehmen reagieren bereits: Zahlreiche EU-Rechenzentrumsbetreiber (*OVHcloud*, *Deutsche Telekom*, *Exoscale*) verzeichnen steigende Anfragen. Hyperscaler wie Microsoft, Google und AWS bieten inzwischen »souveräne Cloud-Instanzen« innerhalb Europas an – mit »data boundary«-Versprechen und juristischer Kontrolle. Das funktioniert aber nicht (siehe Seite 24).

Abgrenzung zu Schrems II und Privacy Shield

Der US-Cloud-Act ist nicht identisch mit dem viel diskutierten »Schrems II«-Urteil des EuGH von 2020 oder dem gescheiterten Abkommen »Privacy Shield«. Während sich Schrems II auf die Unzulässigkeit von Datenübermittlungen in die USA ohne adäquaten Schutz bezieht – insbesondere wegen des Zugriffsrisikos durch US-Geheimdienste –, definiert der Cloud-Act die rechtlichen Befugnisse für genau diesen Zugriff. Anders gesagt: Schrems II beschreibt das Problem aus europäischer Sicht. Der Cloud-Act ist das Problem aus US-amerikanischer Gesetzgebungsperspektive.

Fazit: Handlungsbedarf für deutsche Firmen

Deutsche IT-Verantwortliche müssen jetzt handeln: Eine »Plan B«-Strategie umfasst:

1. Prüfung datenschutzrechtlicher Risiken beim Einsatz von US-Clouds,
2. Einsatz europäischer Anbieter,
3. Verträge mit EU-Standardklauseln (SCC) und umfassende Transfer-Impact-Assessments.

Nur so lässt sich das Risiko legaler Konflikte und von Reputationsverlusten vermeiden. Auch wenn der Cloud-Act schon 2018 in Kraft trat – die jüngsten politischen Entscheidungen in Washington machen ihn aus europäischer Perspektive brandaktuell. ■

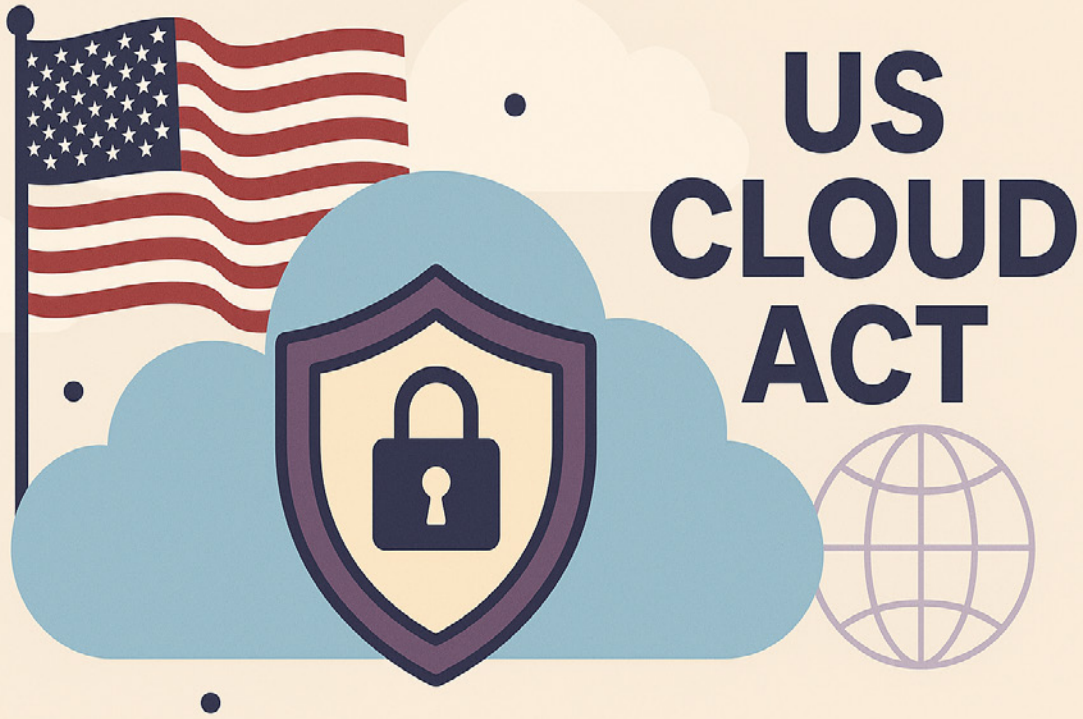


Bild: speicherguide.de via DALL-E

Karl Fröhlich
speicherguide.de

US-Cloud-Act: Relevantes Kriterium ist nicht der Speicherort

SOUVERÄNE CLOUD-ANGEBOTE VON US-ANBIETERN – **ETIKETTENSCHWINDEL MIT ANSAGE**

US-Cloud-Anbieter werben mit »souveränen« Cloud-Lösungen für europäische Kunden. Doch die Datensouveränität entpuppt sich bei genauer Betrachtung als juristische Fiktion. Denn wo US-Recht greift, hilft auch ein europäischer Rechenzentrumsstandort wenig.

Ob *Microsoft Azure European Cloud*, *AWS European Sovereign Cloud* oder *Google Cloud Sovereign Solutions* – alle großen US-Anbieter versprechen ihren europäischen Kunden, Daten ausschließlich innerhalb der EU zu verarbeiten, abgesichert durch lokale Schlüsselverwaltung, europäisches Personal und organisatorische Trennung. Klingt gut. Ist aber in wesentlichen Punkten irreführend.

Problem Nr. 1: Der Cloud-Act kennt keine Grenzen

Laut US-Cloud-Act sind amerikanische Unternehmen verpflichtet, auf Anordnung von US-Behörden Daten herauszugeben – auch dann, wenn diese physisch in Europa gespeichert und von einer europäischen Tochtergesellschaft verwaltet werden.

Relevantes Kriterium ist nicht der Speicherort, sondern die Unternehmensstruktur: Solange die Muttergesellschaft Zugriff auf Systeme oder Personal hat, ist der Zugriff legal möglich – und unterliegt sogar in vielen Fällen einer Geheimhaltungspflicht (»Gag Order«).

Problem Nr. 2: Governance bleibt in US-Hand

Auch wenn Datenzugriffe technisch abgesichert scheinen, bleibt die Verwaltung – also das »Control Plane« – oft zentralisiert. Vertrauenswürdigkeit und Compliance-Versprechen der Anbieter helfen nur begrenzt, wenn nicht gleichzeitig eine konsequente rechtliche Entkopplung erfolgt – was bisher in keinem Fall vollständig gelungen ist.

Problem Nr. 3: Missverständlicher Begriff der Souveränität

Der Begriff »souverän« suggeriert eine Autonomie, die faktisch nicht existiert. Eine Cloud ist erst dann wirklich souverän, wenn sie vollständig unter europäischer Kontrolle steht – technisch, personell und juristisch. Genau das ist bei den Angeboten von US-Hyperscalern nicht gegeben.

Alternative: EU-Anbieter mit Sitz in Europa

Anbieter wie *IONOS*, *OVHcloud*, *Wolkenkraft*, *Univention*, *Exoscale* oder *Citycloud* betreiben ihre Dienste vollständig innerhalb der EU und unterliegen ausschließlich europäischem Recht. Für Branchen mit hohen Datenschutzanforderungen – etwa im Gesundheitswesen, der öffentlichen Verwaltung oder KRITIS-Infrastrukturen – sind sie derzeit die einzige echte Option.

Fazit: Vertrauen ist keine Strategie

Die Anbieter argumentieren häufig mit technischen Schutzmaßnahmen: Georedundanz, EU-Schlüsselverwaltung und Zero-Trust. Das schützt zwar vor Hackern – nicht aber vor legalem Zugriff durch US-Behörden.

Was fehlt, ist eine eindeutige Trennung der Verwaltungsrechte – sowohl auf technischer als auch auf juristischer Ebene. Diese ist nur schwer umzusetzen, wenn zentrale Verwaltungsstrukturen weiter aus den USA heraus gesteuert werden.

Selbst wenn sich US-Anbieter vertraglich verpflichten, Behördenanfragen offenzulegen oder Daten nicht herauszugeben, steht das im Konflikt mit US-Recht. Sie sind per Gesetz gezwungen, bestimmte Anfragen geheim zu halten. Das bedeutet: Kunden würden es nicht einmal erfahren, wenn auf ihre Daten zugegriffen wird.

Die souveräne Cloud von US-Anbietern ist bestenfalls ein taktisches Zugeständnis – keine tragfähige Strategie für langfristige digitale Unabhängigkeit. Unternehmen, die Datenschutz und regulatorische Sicherheit ernst nehmen, sollten sich nicht auf Marketingversprechen verlassen – sondern auf Rechtslage und Kontrolle. Denn, Vertrauen ist keine Strategie. ■

Bild: speicherguide.de via DALL-E

US-Cloud-Act versus europäische Datensouveränität: *Die Grafik zeigt drei zentrale Kriterien, an denen Unternehmen prüfen sollten, ob ein Cloud-Angebot echte digitale Unabhängigkeit ermöglicht – oder nur scheinbare Sicherheit bietet. Entscheidend ist nicht der Standort der Daten, sondern wer rechtlich Zugriff darauf hat.*

US-Cloud-Act (18 U.S.C. § 2713): Zugriffspflicht auch auf Daten außerhalb der USA

Der zentrale Passus des US-Cloud-Act verpflichtet US-Anbieter zur Herausgabe von Daten unabhängig vom Speicherort:

»A provider of electronic communication service or remote computing service shall comply with the requirements of this chapter to preserve, backup, or disclose the contents of electronic communications, ***regardless of whether such communication, record, or other information is located within or outside of the United States.****« (18 U.S.C. § 2713, Stored Communications Act, geändert durch CLOUD Act 2018)*

Das bedeutet: Sobald ein Unternehmen der US-Gerichtsbarkeit unterliegt – etwa durch Hauptsitz oder Konzernstruktur – kann ein US-Gericht die Herausgabe anordnen, selbst wenn die Daten in einem europäischen Rechenzentrum gespeichert sind und von einem europäischen Tochterunternehmen verwaltet werden. Die Verpflichtung kann zudem mit einer Schweigeverpflichtung (»Gag Order«, § 2705 SCA) verbunden sein.

Cloud- und Hosting-Provider: Wenn die Abhängigkeit zum Risiko wird

CLOUD GONE? WARUM EXITSTRATEGIEN UNVERZICHTBAR SIND

Die Auslagerung von IT-Ressourcen bringt Skalierbarkeit – aber auch Abhängigkeit. Kommt es zur Insolvenz oder Geschäftsaufgabe eines Anbieters, stehen Unternehmen oft unter Zeitdruck. Prominente Beispiele zeigen: Eine Exitstrategie muss Teil jeder Cloud-Strategie sein – von Anfang an.

Karl Fröhlich
speicherguide.de

Die Cloud verspricht Flexibilität, Effizienz und Skalierbarkeit. Doch was passiert, wenn der Dienstleister plötzlich ausfällt – sei es durch Insolvenz, Geschäftsaufgabe oder Anbieterwechsel? Genau dann zeigt sich, ob Unternehmen vorgesorgt haben. Denn mit jeder Auslagerung geht immer auch eine technische und wirtschaftliche Abhängigkeit einher.

Dass dieses Risiko real ist, zeigen mehrere namhafte Beispiele:

2013 ging der US-Anbieter **Nirvanix**, spezialisiert auf Cloud-Storage, überraschend in die Insolvenz. Kunden blieben nur wenige Wochen, um teils mehrere PByte an Daten zu migrieren. Der Schweizer Dienst **Wuala** stellte seinen Betrieb 2015 ein – mit immerhin rund drei Monaten Vorlauf für die Datenrückführung. In Großbritannien traf es Behörden und Unternehmen, als der auf den Public-Sektor spezialisierte Anbieter **UKCloud** 2022 seinen Betrieb einstellte. Auch große Namen wie der Colocation-Dienstleister **Cyxtera** (Insolvenzverfahren 2023) zeigen: Größe allein schützt nicht vor dem Aus.

Selbst 2024 kam es erneut zu einem prominenten Fall: Die Berliner **Cortado Mobile Solutions** stellte die Plattform *Teamplace* ein – zu aufwändig im Betrieb, wie das Unternehmen erklärte. Trotz dieser Risiken verzichten viele Unternehmen auf eine vorgedachte Exitstrategie. Sie sollte daher von Beginn an Teil jeder Cloud-Architektur sein – vertraglich geregelt und regelmäßig getestet.

Eine solide Exitstrategie umfasst:

- vertraglich garantierte Exportmechanismen für alle Daten,
- abgestimmte Migrationspfade mit definierten Formaten, APIs und Fristen,
- Fallback-Optionen im Fall eines abrupten Dienstendes sowie
- klare Notfallprozesse für operative Übergangsszenarien.

Auch der Zugriff auf technische Dokumentationen, Schlüssel oder Plattform-Instanzen sollte im Ernstfall gewährleistet sein. Andernfalls drohen Datenverluste, Compliance-Verstöße oder Systemausfälle.

Was tun, wenn der Cloud-Anbieter ausfällt?

Vier Punkte, die eine Exitstrategie enthalten sollte

Datenrückführung (Data Export)	Migrationspfade (Migration Readiness)	Vertragliche Vorkehrungen (Legal Safeguards)	Nofallprozesse ((Fallback & Business Continuity))
• Format und Umfang regeln • APIs und Tools vorab definieren • Fristen für Datenverfügbarkeit festlegen	• Szenarien für Wechsel durchspielen • Plattformunabhängige Formate wählen • Testmigration regelmäßig durchführen	• Kündigung und Insolvenzregelungen • Zugriff auf technische Unterlagen sichern • ggf. Escrow-Vereinbarung nutzen	• Ersatzbetrieb (On-Prem/ Zweitanbieter vorbereiten • IT und Fachabteilungen in Exit-Pläne einbinden • Wiederanlaufzeiten definieren und üben

Grafik: speicherguide.de

Fazit: Auf Ausfälle und Probleme vorbereitet sein

Die Frage ist nicht, ob der Cloud-Provider ausfällt – sondern ob man vorbereitet ist, wenn es passiert. Das kann auch durch Katastrophen geschehen, wie 2021, als nach einem nicht selbst verschuldeten Brand der französische Cloud-Anbieter **OVHcloud** erst nach drei Wochen mit dem Wiederanlauf beginnen konnte. ■

Bild: speicherguide.de via DALL-E

Karl Fröhlich
speicherguide.de

Cloud oder eigene Cloud? Mittelstand zwischen Flexibilität und Kontrolle

CLOUD-STORAGE IM MITTELSTAND – DIGITALE FREIHEIT MIT VERANTWORTUNG

Cloud-Storage bietet Skalierbarkeit, Kosteneffizienz und Agilität – doch IT-Verantwortliche in KMU müssen Datenkontrolle, Datenschutz und versteckte Kosten im Blick behalten. In Zeiten wachsender Abhängigkeit von US-Hyperscalern überlegen viele, On-Premises oder eigene Private Clouds zu priorisieren.

Mittelständische Unternehmen sehen sich mit rasant wachsenden Datenmengen, flexiblen Arbeitsformen und digitalen Transformationsprojekten konfrontiert. On-Premises-Infrastrukturen erreichen hierbei oft technische und finanzielle Grenzen. Cloud-Storage verspricht hier Abhilfe – mit Skalierbarkeit, automatisierten Prozessen, geografischer Redundanz und modernen APIs für Backup, Zusammenarbeit und Analyse. Gleichzeitig steigt der Druck, sensible Kundendaten sicher und DSGVO-konform zu verwalten.

Cloud-Storage: Vorteile für kleine und mittlere Unternehmen

Cloud-Storage bietet eine Reihe Vorteile – insbesondere dort, wo Flexibilität, Skalierbarkeit und schnelle Verfügbarkeit gefragt sind. Einer der größten Pluspunkte ist die **bedarfsgerechte** Nutzung: Speicher-Ressourcen lassen sich flexibel anpassen, ohne dass teure Überkapazitäten vorgehalten werden müssen. Das sogenannte OPEX-Modell – also laufende Betriebsausgaben statt hoher Anfangsinvestitionen – soll den Anbietern zufolge die Budgets schonen und eine genauere Kostenplanung ermöglichen – in der Praxis ist aber oft das Gegenteil der Fall.

Ein weiterer Pluspunkt liegt in der Verfügbarkeit und Ausfallsicherheit: Cloud-Anbieter stellen Daten redundant über mehrere Standorte hinweg bereit. Für KMU bedeutet das eine deutliche Verbesserung der Resilienz, etwa im Fall eines lokalen Hardware- oder Stromausfalls im eigenen Rechenzentrum. Zudem profitieren Unternehmen von automatisierten Prozessen, etwa bei Backups, Replikation oder Lifecycle-Management – alles lässt sich über APIs und Management-Tools effizient steuern.

Bild: Luckycloud

Luc Mader
Luckycloud

»Souveränität beginnt nicht beim Datenstandort, sondern bei der Kontrolle über den Betreiber – solange ein US-Konzern Infrastruktur und Updates steuert, bleibt die Cloud angreifbar.«

Auch in der Entwicklung und Innovation kann die Cloud punkten. Test- und Entwicklungsumgebungen lassen sich bei Bedarf schnell bereitstellen und wieder löschen, was agile Projekte vereinfachen kann. Wer mit KI oder Datenanalysen experimentiert, findet in der Cloud die notwendige Infrastruktur, ohne in dedizierte Hardware investieren zu müssen.

Für KMU, die verteiltes Arbeiten ermöglichen wollen, bieten sich zudem Collaboration-Plattformen an: Über Cloud-Storage lassen sich Dateien standortübergreifend synchronisieren und sicher bereitstellen – ein wichtiges Argument in Zeiten von Home-Office und Remote-Teams.

Kurz gesagt: Überall dort, wo Skalierbarkeit, Agilität und Geschwindigkeit entscheidend sind, kann Cloud-Storage auch für Mittelständler eine sinnvolle Ergänzung oder Alternative zur klassischen IT-Infrastruktur sein.

Gängige Einsatzszenarien:

- **Backup und Archiv:** Cloud ergänzt lokale Sicherungen – Restore-Tests zeigen die Wiederherstellbarkeit.
- **Dateizugriff und Teamarbeit:** DSGVO-konforme Cloud-Lösungen unterstützen Home-Office und standortübergreifende Zusammenarbeit.
- **Test- und Entwicklungsumgebungen:** Schnell bereitgestellt und nach Projektende zurückgezogen.
- **KI und Analytics:** Cloud-Lösungen liefern die Rechenleistung, die für ersten KI-Einsatz benötigt wird.

Risiken & Stolpersteine – Kritische Betrachtung

Trotz aller Vorteile birgt Cloud-Storage – insbesondere bei den großen US-Hyperscalern – erhebliche Risiken, die KMUs nicht unterschätzen sollten. Ein zentrales Problem ist die fehlende digitale Souveränität: Zwar werben Anbieter wie **AWS**, **Microsoft** oder **Google** mit europäischen Regionen und Datenschutzversprechen, doch faktisch bleibt die Kontrolle über Infrastruktur, Updates und Zugriffe beim Anbieter – und damit oft außerhalb europäischer Rechtsräume.

Luc Mader, CEO des Berliner Anbieters **luckycloud**, bringt es auf den Punkt: »Souveränität beginnt nicht beim Datenstandort, sondern bei der Kontrolle über den Betreiber. Solange ein US-Konzern Herr über Infrastruktur und Updates bleibt, sind technische Sicherheitsversprechen nur so viel wert wie die nächste Executive-Order in Washington.« Selbst wenn Daten physisch in der EU liegen, können US-Behörden über den sogenannten US-Cloud-Act Zugriff erzwingen – auch ohne dass betroffene Unternehmen dies nachvollziehen oder verhindern könnten. »Ein US-Richter, ein interner Befehl, ein nicht nachvollziehbares Update – und das war's mit der Autonomie«, warnt Mader.

Bild: Nextcloud

Frank Karlitschek
Nextcloud

»Digitale Souveränität heißt: volle Kontrolle über Daten und Infrastruktur – proprietäre, intransparente Clouds liefern nur Souveränitäts-Washing statt echter Unabhängigkeit.«

»Digitale Souveränität bedeutet, dass Organisationen die volle Kontrolle über ihre Daten, ihre Infrastruktur und die zugrundeliegende Technologie haben«, ergänzt **Nextcloud**-CEO **Frank Karlitschek**. »Wenn ein US-Konzern verspricht, europäische Souveränität zu garantieren, bleibt das ein Widerspruch: Die rechtlichen sowie techno-

Ionos Object Storage

IONOS Object Storage ist eine S3-kompatible, ISO-27001- und DSGVO-konforme Speicherlösung, die beliebig skalierbar und nutzungsbasiert abrechenbar ist. Sie bietet TLS-/AES-verschlüsselte Übertragung, Server-Seiten-Verschlüsselung, Versioning, WORM-Lock, Lifecycle-Regeln und georedundante Architektur. Die Cloud soll sich in Anwendungen, Backups und Pipelines integrieren – unterstützt durch Block- sowie NFS-Storage, Kubernetes und Big-Data-Systeme.

Der Objektspeicher bietet redundante Speicherung auf mehreren Storage-Knoten – ohne zusätzliche Kosten. Die Daten werden laut Anbieter redundant gesichert und lassen sich mit Replikation jederzeit wiederherstellen. Für 1 GByte berechnet Ionos 0,007 Euro pro 30 Tage. Ein Data-Scan kostet pro GByte 0,0014 Euro und Data-Return 0,0005 Euro pro GByte.

Mehr unter:
www.cloud.ionos.de/storage/

Luckycloud

Luckycloud ist eine modulare Cloud-Plattform des gleichnamigen Unternehmens aus Berlin, die für datenschutzkritische Anwendungen entwickelt wurde. Der Dienst bietet sicheren Speicher (3 GByte bis 500 TByte), S3-kompatiblen Zugriff, Datei-Versionierung, Ende-zu-Ende-Verschlüsselung (optional), Zero-Knowledge-Prinzip sowie kollaborative Funktionen wie Mail, Kalender und Online-Meetings.

Die Infrastruktur ist vollständig in ISO-27001-zertifizierten Rechenzentren in Deutschland gehostet. Die Preisgestaltung ist transparent und flexibel: zum Beispiel Teams-Paket ab 3,50 Euro/Monat netto, Business ab 10 Euro, Enterprise ab 45,02 Euro und S3-Speicher ab 3,00 Euro/100 GByte. Luckycloud richtet sich an KMU, Behörden und Organisationen mit hohen Anforderungen an Datenschutz, Skalierbarkeit und Compliance.

Mehr unter:
www.luckycloud.de

Plusserver

PlusServer bietet ein umfassendes Storage-Portfolio »Made in Germany« für Unternehmen mit Ansprüchen an Skalierbarkeit, Sicherheit und Compliance. Der S3-kompatible Objektspeicher richtet sich an datenintensive Anwendungen wie Backup, Archivierung oder Data-Lakes und ist ab 0,022 Euro pro GByte nutzbar.

Ergänzend stehen ein flexibel buchbares Network-Storage (iSCSI, NFS, SMB) sowie ein agentenbasiertes, verschlüsseltes Backup-as-a-Service zur Verfügung. Alle Services laufen in ISO 27001- und BSI C5-zertifizierten Rechenzentren in Deutschland und sollen DSGVO-Vorgaben erfüllen. Plusserver richtet sich mit seinem nutzungsbasierten Modell an mittelständische und große Unternehmen mit verlässlichem Support, klarer Abgrenzung zu Hyperscalern und technischer Integrationstiefe.

Mehr unter:
www.plusserver.com/produkt

Securecloud

SecureCloud ist eine in Deutschland gehostete Cloud-Plattform für sichere Zusammenarbeit, Dokumentenmanagement und rechtssichere Workflows. Die Lösung wurde für Unternehmen konzipiert, die Wert auf Datenschutz, Integrationsfähigkeit und ein hohes Maß an Benutzerkontrolle legen. Securecloud ist vollständig auf eine browserbasierte Nutzung ausgelegt – ohne lokale Software-Installation. Zentrale Module umfassen SecureShare, SecureWork und SecureSign – ergänzt um Versionierung, Volltextsuche und Audit-Trails.

Speicher pro Nutzer: 50 GByte (Business), 100 GByte (Advanced/Enterprise); Preise auf Anfrage. Funktionen sind in allen Paketen enthalten, die Skalierung erfolgt über Nutzeranzahl. Der Betrieb erfolgt in ISO-27001-zertifizierten Rechenzentren.

Mehr unter:
www.securecloud.de

Telekom Object Storage Service

Der Telekom Object Storage Service (OBS) bietet Unternehmen ein S3-kompatibles, skalierbares und deutsches Cloud-Speichermodell mit 99,95 prozentiger Verfügbarkeit und einer garantierten Haltbarkeit von bis zu 99,9999999999 Prozent (12 Neunen). Mit einem dreistufigem Speichermodell (Standard/Warm/Cold), AES-256-Verschlüsselung, detailliertem IAM und Lifecycle-Regeln deckt OBS Archivierung, Backup, Medien-Hosting, Data-Lakes und DevOps-Szenarien ab.

Ab zirka 0,024 Euro pro GByte nutzungsabhängig, mit gestaffeltem Traffic und API-Request-Kosten. Zudem lässt sich OBS in die Open Telekom Cloud integrieren und komplett absichern. Die Telekom spricht damit vor allem Unternehmen mit Compliance-, Datenschutz- und Infrastruktur-Anforderungen aus Deutschland an.

Mehr unter:
open-telekom-cloud.com/produkte-services

logischen Abhängigkeiten, die proprietäre Architektur und die fehlende Transparenz bleiben bestehen. Eine Cloud kann nur dann souverän sein, wenn sie unabhängig betrieben und kontrolliert wird – und das geht nur mit Open-Source. Alles andere ist reines Souveränitäts-Washing.«

Doch nicht nur geopolitische Risiken machen Hyperscaler problematisch für den Mittelstand. Auch technische und betriebswirtschaftliche Stolpersteine sind häufig: Die Komplexität der Preismodelle – etwa durch zusätzliche Gebühren für Datenabflüsse (Egress), API-Aufrufe oder Restore-Vorgänge – kann zu erheblichen Mehrkosten führen, insbesondere wenn Cloud-Storage intensiv genutzt wird. Hinzu kommen Lock-in-Effekte durch proprietäre Schnittstellen oder spezielle Storage-Formate, die einen späteren Anbieterwechsel erschweren oder verteuern.

Auch aus Sicht der Betriebssicherheit gibt es kritische Punkte: Der Cloud-Kunde hat keinen direkten Einfluss auf Software-Updates, Patch-Zyklen oder Änderungen an der Infrastruktur. Diese Abhängigkeit kann im Ernstfall – etwa bei Änderungen im Support oder plötzlichen Ausfällen – zu erheblichen Problemen führen, insbesondere wenn Service-Level-Agreements (SLAs) vage formuliert oder einseitig anpassbar sind.

Für viele mittelständische Unternehmen stellt sich daher nicht die Frage, ob Cloud genutzt werden soll – sondern *wie* sie kontrolliert, souverän

und wirtschaftlich sinnvoll eingesetzt werden kann.

Kostenfalle: Günstiger Einstieg, teurer Betrieb

Ein oft genannter Vorteil von Cloud-Storage ist die vermeintliche Kosteneffizienz – doch dieser Eindruck relativiert sich schnell bei genauer Betrachtung. Zwar sind die Einstiegskosten gering, insbesondere für Speicherplatz selbst: Hyperscaler wie *AWS*, *Microsoft Azure* oder *Google Cloud* locken mit günstigen Preisen pro Gigabyte und minimalem Aufwand für die Inbetriebnahme. Doch die tatsächlichen Betriebskosten hängen stark von der Nutzung ab – und genau hier wird es für viele mittelständische Unternehmen teuer.

Ein wesentlicher Kostenfaktor ist der **Datenabfluss**: Während das Hochladen von Daten in die Cloud in der Regel kostenfrei ist, wird das Herunterladen (Egress) meist hoch bepreist. Schon alltägliche Vorgänge wie Backups, Datenabgleiche oder Wiederherstellungen können unerwartet hohe Gebühren verursachen. Auch **API-Zugriffe**, etwa bei der Nutzung von S3-kompatiblen Tools wie *Veeam* oder Archivsystemen, werden einzeln abgerechnet und können sich bei intensiver Nutzung summieren. Hinzu kommt, dass **hochverfügbare Konfigurationen** – etwa Multi-Zone-Replikation – zusätzliche Kosten durch redundante Datenhaltung und Datenverkehr verursachen.

Besonders problematisch wird es, wenn Unternehmen hybride Szenarien mit Cloud-Anbindung betreiben oder regelmäßig größere Datenmengen zwischen Cloud und lokalem Rechenzentrum bewegen. Auch wenn Speicherklassen wie *Glacier* auf den ersten Blick günstig wirken, werden zusätzliche Gebühren für den Zugriff auf archivierte Daten fällig – inklusive Wartezeiten.

Für volatile Daten, Testumgebungen oder selten genutzte Archive kann Cloud-Storage wirtschaftlich sein. Doch für Anwendungen mit hohem Zugriff, festem Datenvolumen oder strengen Restore-Zeiten ist eine On-Prem- oder Private-Cloud-Lösung langfristig oft die kostengünstigere und besser planbare Alternative. Unternehmen sollten deshalb nicht nur die reinen Speicherkosten vergleichen, sondern die komplette Nutzungskette analysieren – inklusive Zugriffsmuster, Bandbreite und möglichen Rückführungen.

Typische Auswahlkriterien für KMU:

- **Kontrolle & Souveränität:** Wer kontrolliert Updates, Infrastruktur, Zugriffe?
- **Datenschutz & Compliance:** DSGVO-Konformität, Verschlüsselung, Logs, Auftragsverarbeitung.
- **Kostenstruktur:** Transparente Gebühren (Speicher, Transfer, API-Aufrufe).
- **Lock-in:** API-Kompatibilität und Wechselbarkeit.
- **Performance:** Latenz-Einschätzung je nach Anwendungsfall.
- **Support & Management:** Eigenbetrieb vs. Managed Services.
- **Hybridfähigkeit:** Nahtlose Integration zwischen Cloud und lokalem Storage.

Infrastrukturen: Hybriden Umgebungen gehört die Zukunft

Cloud-Storage bietet KMU Vorteile bei Flexibilität, Agilität und Innovation – ideal für nicht-kritische Workloads, agile Projekte oder KI-Piloten. Doch zentrale Unternehmensdaten, sensible Kundendaten und regulierte Prozesse verdienen höchste Kontrolle – On-Premises oder Private-Cloud sind hier oft effizienter und sicherer. Grundsätzlich gelten Hybrid-Modelle als Standard: Wichtige Workloads und Daten bleiben lokal – Cloud ergänzt in ausgewählten Szenarien.

Mit deutschen Anbietern wie zum Beispiel **Hetzner**, **IONOS**, Luckycloud, **PlusServer**, **SpaceNet** oder **Telekom** lassen sich souveräne, DSGVO-konforme Cloud-Strategien umsetzen. In Kombination mit On-Prem-Systemen und souveränen Private-Cloud-Infrastrukturen im eigenen Rechenzentrum kann ein robustes, wirtschaftliches und zukunftsfähiges IT-Setup entstehen, auch für den Mittelstand. ■

Karl Fröhlich
speicherguide.de

IT-Infrastruktur auslagern, ohne die Verantwortung abzugeben

AUSGELAGERTE RECHENZENTREN: IT-BETRIEB IM MIETMODELL

Ein Rechenzentrum-as-a-Service verspricht Unternehmen eine flexible, sichere und kosteneffiziente IT-Infrastruktur. Besonders für kleine und mittlere Unternehmen kann ein RZ ohne eigene Hardware oder Personalaufwand eine echte Entlastung darstellen. Das Leistungsspektrum reicht von virtualisierten Servern über Datensicherung bis hin zu vollständig gemanagten IT-Umgebungen.

Die digitale Transformation ist längst kein Projekt mehr – sie ist Dauerzustand. Für viele Unternehmen bedeutet das: Die Anforderungen an die eigene IT-Infrastruktur steigen kontinuierlich, während Ressourcen, Fachpersonal und Budgets stagnieren. Die Folge sind überlastete IT-Abteilungen, Sicherheitslücken, ineffiziente Insellösungen – und strategischer Stillstand. Genau hier setzt das Konzept des Rechenzentrum-as-a-Service (RZaaS) an.

Statt eigene Server, Storage-Systeme, Firewalls und Backups im Keller oder Server-Raum zu betreiben, mieten Unternehmen beim RZaaS-Modell eine komplette IT-Infrastruktur im Rechenzentrum eines Dienstleisters – inklusive Betrieb, Wartung, Skalierung und Support. Der entscheidende Vorteil: Die Verantwortung für Verfügbarkeit, Sicherheit und Skalierbarkeit wird ausgelagert, die Kontrolle über Daten und Prozesse bleibt erhalten.

Mietmodell kann Ressourcen sparen

Besonders kleine und mittlere Unternehmen (KMUs) sollen von einem ausgelagerten Rechenzentrum profitieren. Oft genug fehlt hier das spezialisierte Personal für den Betrieb kritischer IT-Systeme, während gleichzeitig hohe Anforderungen an Datenschutz, Verfügbarkeit und Performance bestehen. Auch Firmen mit verteilten Standorten oder hybriden Arbeitsmodellen sehen sich zunehmend mit komplexen IT-Strukturen konfrontiert, die intern kaum mehr zu bewältigen sind.

Für diese Zielgruppen kann RZaaS ein hohes Maß an Entlastung bieten, bei gleichzeitigem Zugewinn an Sicherheit und Flexibilität. Aber auch IT-Abteilungen in wachstumsstarken Unternehmen oder solche, die Legacy-Infrastrukturen ablösen wollen, greifen verstärkt auf das Modell zurück.

Was genau wird eingekauft?

Unternehmen mieten im Rahmen von RZaaS eine passgenaue Kombination aus Compute-Ressourcen (virtuelle Server, Containerumgebungen, ggf. Bare-Metal-Systeme), Speicherlösungen (File, Block, Objekt), Netzwerkkomponenten (inkl. Firewalls, VPN, VLANs) sowie Backup- und Sicherheitsdienste. Der Betrieb erfolgt in einem professionellen Rechenzentrum, meist mit ISO-Zertifizierungen, DSGVO-konformer Datenhaltung und redundanter Infrastruktur.

Abgerechnet wird auf Miet- oder Pay-per-Use-Basis, oft mit festen monatlichen Paketen oder nutzungsabhängigen Modellen – je nach Anbieter.

Welche Einsatzszenarien sind typisch?

Die Anwendungsmöglichkeiten reichen von der vollständigen Auslagerung der IT bis hin zu ergänzenden Szenarien wie:

- Hosting von ERP-, DMS- oder CAD-Systemen
- Bereitstellung von Test- und Entwicklungsumgebungen
- Backup-Standort oder Disaster-Recovery-Lösungen
- Unterstützung von mobilen Arbeitsplätzen und Home-Office
- Entlastung von Produktions-IT oder Shop-Filialinfrastrukturen

Dabei lassen sich Ressourcen flexibel skalieren – ideal für saisonale Schwankungen oder Wachstumsszenarien.

RZaaS: stabiler Betrieb zu kalkulierbaren Kosten

Für IT-Verantwortliche kann der Schritt zum RZaaS ein Befreiungsschlag sein. Keine ständige Sorge mehr vor Systemausfällen, Sicherheitslücken oder nächtlichen Alarmen – der Betrieb läuft stabil und planbar im Hintergrund.

Richtig kalkuliert sollten Firmen von Kostentransparenz und berechenbaren Betriebskosten profitieren. Zudem spricht der Wegfall von Investitionen und reduzierter interner Aufwand für Wartung und Betrieb für das Modell. Gerade KMUs sparen damit häufig nicht nur Geld, sondern auch an personellen Kapazitäten.

Was viele nicht erwarten: Ein ausgelagertes Rechenzentrum kann schneller auf veränderte Anforderungen reagieren als eine eigene Umgebung – etwa beim Onboarding neuer Standorte, bei Lastspitzen oder bei der Anbindung externer Systeme. IT wird dadurch nicht nur stabiler, sondern auch agiler.

Grafik: speicherguide.de

Rechenzentrum-as-a-Service bietet zahlreiche Vorteile – von Entlastung der IT bis zu kalkulierbaren Kosten. Doch auch Aspekte wie Anbieterbindung und eingeschränkte Individualisierung sollten bedacht werden.

Typische Einwände gegen RZaaS

Trotz vieler Vorteile gibt es auch typische Einwände. Einige Unternehmen fürchten den Kontrollverlust über ihre Daten – doch moderne RZaaS-Angebote setzen auf vollständige Transparenz und DSGVO-konforme Strukturen. Auch der Sicherheitsaspekt wird häufig kritisch gesehen, dabei bieten professionelle Rechenzentren meist höhere Schutzmechanismen als lokal betriebene Server-Räume.

Ein weiterer Vorbehalt betrifft die vermeintlich höheren Kosten. Dabei wird oft vergessen, dass viele interne IT-Kosten in Personal, Ausfallzeiten oder Wartung nicht direkt sichtbar sind. Ein neutraler Kostenvergleich über drei bis fünf Jahre zeigt meist: Das Betriebsmodell kann auch finan-

ziell sinnvoll sein – besonders bei wachsendem IT-Bedarf.

Konkrete Nachteile eines RZaaS-Modells

So überzeugend das Konzept eines ausgelagerten Rechenzentrums auch sein mag – RZaaS ist kein Modell ohne Einschränkungen. Je nach Anforderungen und Rahmenbedingungen können bestimmte Nachteile relevant werden, die Unternehmen bei der Planung mitbedenken sollten.

Ein zentraler Punkt ist die Abhängigkeit vom Dienstleister. Wer den Betrieb seiner Infrastruktur vollständig auslagert, überträgt nicht nur technische Aufgaben, sondern auch Verantwortung. Das erfordert Vertrauen und belastbare Service-Level-Vereinbarungen (SLAs). Fehlt es an Transparenz, Zuverlässigkeit oder Kommunikation, kann diese Abhängigkeit zur Schwachstelle werden – vor allem bei Störungen oder Änderungen im laufenden Betrieb.

RZaaS: Viel Standard, wenig Individualisierbarkeit

Ein weiterer möglicher Nachteil ist die eingeschränkte Individualisierbarkeit. Viele RZaaS-Angebote basieren auf standardisierten Umgebungen, was zwar Effizienz bringt, aber auch Anpassungsgrenzen setzt. Spezielle Anforderungen, etwa für Legacy-Anwendungen, exotische Software oder proprietäre Schnittstellen, lassen sich nicht immer problemlos abbilden. Hier ist eine frühzeitige technische Abstimmung mit dem Anbieter wichtig.

Grafik: speicherguide.de via DALL-E

Von ERP-Hosting bis zum Backup-Standort: RZaaS eignet sich für vielfältige Szenarien im Mittelstand und schafft Flexibilität bei stabiler Betriebsbasis.

Netzwerkengpässe bei der Anbindung

Auch die Netzwerkanbindung kann zum Engpass werden – insbesondere bei datenintensiven Anwendungen oder Echtzeitszenarien. Anders als bei lokaler Infrastruktur hängt die Performance im RZaaS-Modell von der Qualität der Verbindung zum Rechenzentrum ab. Hohe Latenzzeiten oder instabile Leitungen können den Nutzen deutlich einschränken, etwa im CAD-Umfeld oder bei Remote-Desktops.

Ein weiterer Aspekt betrifft die Kostenstruktur. Zwar verspricht RZaaS in vielen Fällen eine bessere Planbarkeit und geringere Investitionen, doch bei unkontrollierter Nutzung oder fehlender Ressourcen-Optimierung können die monatlichen Betriebskosten ansteigen. Nicht genutzte virtuelle Maschinen, überdimensionierte Speicherkapazitäten oder unklare Bereitstellungsmodelle führen schnell zu unerwarteten Belastungen.

Schließlich bleibt bei manchen Unternehmen ein gewisses Unbehagen hinsichtlich der Datenhoheit. Selbst wenn der Anbieter in Deutschland sitzt und DSGVO-konform arbeitet, empfinden manche IT-Verantwortliche die physische Distanz zur eigenen Infrastruktur als Kontrollverlust. Hier hilft Transparenz – etwa durch dokumentierte Sicherheitskonzepte, Zugriffsprotokolle oder sogar Rechenzentrumsführungen.

RZaaS vs. Cloud-Service eines Hyperscalers

Während Cloud-Services von Hyperscalern wie *AWS*, *Microsoft Azure* oder *Google Cloud* meist mit unbegrenzter Skalierbarkeit, modernsten Entwicklungswerkzeugen und globaler Infrastruktur werben, verfolgt ein Rechenzentrum-as-a-Service (RZaaS) einen anderen Ansatz – näher am klassischen IT-Betrieb und deutlich betreuungsorientierter.

Beim RZaaS-Modell mieten Unternehmen nicht nur technische Ressourcen wie virtuelle Server, Speicher oder Netzwerk-Komponenten, sondern erhalten eine vollständig gemanagte IT-Umgebung – inklusive Betriebssystem, Backups, Sicherheitsmaßnahmen, Monitoring und persönlichem Support. Das Ziel: Entlastung interner IT-Teams, klare Verantwortlichkeiten und eine stabile, kontrollierbare Infrastruktur.

Im Gegensatz dazu versteht sich ein Hyperscaler in erster Linie als technische Plattform: Kunden stellen sich aus einem umfangreichen Baukasten die gewünschten Services selbst zusammen – vom virtuellen Rechner bis zur KI-Plattform. Der Betrieb, die Konfiguration und der Schutz dieser Umgebung liegen weitgehend in der Verantwortung des Kunden. Wer hier erfolgreich arbeiten will, benötigt in der Regel ein erfahrenes DevOps-Team, Kenntnisse in Cloud-Architektur und ein sauberes Sicherheitskonzept.

Hyperscaler: Datenschutz- & Sicherheitsbedenken

Auch in Bezug auf Standort und Datenschutz gibt es Unterschiede. Während Hyperscaler weltweit operieren und die genaue Datenlokation oft nur eingeschränkt steuerbar ist, setzen RZaaS-Anbieter auf Rechenzentren mit klar benannten Standorten – in Deutschland bzw. Europa – und erfüllen gezielt Anforderungen an DSGVO, ISO-Zertifizierungen oder branchenspezifische Compliance-Vorgaben. Besonders für mittelständische Unternehmen mit sensiblen Daten ist das ein entscheidender Vertrauensfaktor. Ein weiterer kritischer Punkt ist die digitale Souveränität: Die Hyperscaler sind allesamt US-Anbieter, die dem US-Cloud-Act unterliegen.

RZaaS-Anbieter mit besserem Service-Verständnis

Ein weiterer Unterschied liegt im Service-Verständnis: RZaaS-Anbieter sollten persönliche Ansprechpartner bereitstellen, feste SLAs und Beratung auch über den reinen Technikbetrieb hinaus – zum Beispiel bei Migration,

Grafik: speicherguide.de via DALL-E

Skalierbarkeit, Sicherheit, Fokus auf das Kerngeschäft: Die Auslagerung der IT-Infrastruktur kann Unternehmen strategisch wie operativ deutlich entlasten.

Modernisierung oder Anwendungsintegration. Bei Hyperscalern findet der Support in der Regel automatisiert oder über Self-Service-Portale statt und ist oft an zusätzliche Kosten oder Enterprise-Verträge gebunden. Mit menschlichen Service-Technikern zu telefonieren, ist nicht vorgesehen. Kurz gesagt: RZaaS ist eine betreute Infrastruktur für Unternehmen, die Stabilität, Datenschutz und persönlichen Service suchen. Hyperscaler-Clouds sind hochskalierbare Entwicklungs- und Betriebsplattformen für IT-Teams mit starkem Innovations- und Eigenbetriebsfokus.

Beide Modelle haben ihre Berechtigung – die Wahl hängt davon ab, welche Anforderungen, Kompetenzen und Prioritäten ein Unternehmen hat. Wobei Hyperscaler-Clouds weder für personenbezogene noch für unternehmenskritische Daten verwendet werden sollten.

RZaaS ist kein IT-Ersatz

RZaaS bietet viele Vorteile, ist aber kein Selbstläufer. Wer realistische Erwartungen hat, das eigene Nutzungsszenario gut kennt und auf einen passenden Partner setzt, kann die möglichen Nachteile aktiv adressieren – und das volle Potenzial des Modells nutzen.

Wichtig zu beachten ist: RZaaS ist kein Ersatz für eine IT-Abteilung. Das ausgelagerte RZ sorgt für eine erhebliche Entlastung – aber es ersetzt die IT-Abteilung nicht vollständig. Denn: Während das operative Tagesgeschäft rund um Server-Betrieb, Backups, Netzwerkverfügbarkeit oder System-Updates an den Dienstleister übergeben wird, bleiben viele strategische, organisatorische und fachliche Aufgaben weiterhin intern relevant.

Die technische Umsetzung kann ausgelagert werden – aber nicht die Verantwortung für Digitalisierung und IT-Entscheidungen. Eine interne Rolle mit technischer oder digitaler Kompetenz bleibt unverzichtbar – gerade als Ansprechpartner für Dienstleister, Fachabteilungen und Geschäftsführung. Diese Rolle kann je nach Unternehmensgröße schlank besetzt sein, darf aber nicht entfallen. Die Geschäftsleitung muss diese Themen aktiv begleiten, auch wenn sie diese nicht selbst umsetzt.

Fazit: RZaaS nutzen – Verantwortung behalten

RZaaS bietet insbesondere für mittelständische Unternehmen eine praxistaugliche Möglichkeit, die eigene IT-Infrastruktur zu modernisieren und zu entlasten. Auf die Kontrolle und Sicherheit von Daten und Ressourcen muss die IT-Abteilung dabei nicht verzichten. Entscheidend ist die Wahl eines zuverlässigen Partners, ein klares Erwartungsmanagement und eine IT-Strategie, die Betrieb und Verantwortung sinnvoll austariert.

Was kostet ein ausgelagertes Rechenzentrum?

Die Kosten für ein Rechenzentrum-as-a-Service hängen stark vom gewählten Leistungsumfang ab. Für kleinere bis mittlere Umgebungen lässt sich jedoch ein realistischer Preisrahmen skizzieren:

Beispielhafte monatliche Kosten (netto) für ein mittelständisches Unternehmen mit zirka 100 Mitarbeitenden:

Leistungskomponente	Beispielausstattung	Monatliche Kosten*
Compute (8–10 virtuelle Maschinen)	inkl. Windows-Server, Datenbank, DMS	ca. 900–1.200 Euro
Storage (5–10 TB inkl. Snapshot-Funktion)	File- und Datenbankdaten, gesichert	ca. 300–500 Euro
Backup & Recovery	tägliche Sicherung, 30 Tage Vorhaltezeit	ca. 150–300 Euro
Netzwerk & Security	Firewall, VPN, VLAN, Monitoring	ca. 200–400 Euro
Betrieb & Support (Managed Service)	inkl. Updates, Monitoring, Störungsbehebung	ca. 300–500 Euro

Gesamt: ca. 1.800 bis 2.900 € pro Monat (zzgl. individueller Sonderanforderungen)

** Die Preise dienen der Orientierung und variieren je nach Anforderungen, SLA-Stufe und Skalierung. Individuelle Angebote sind üblich und empfehlenswert.*

Rechenzentrum nutzen – ohne eigene Cloud

COLOCATION: KONTROLLE BEHALTEN, INFRASTRUKTUR AUSLAGERN

Karl Fröhlich
speicherguide.de

Eine Colocation bietet Unternehmen die Möglichkeit, ihre eigene IT-Hardware in einem professionellen Rechenzentrum zu betreiben. Die Verantwortung für Strom, Klimatisierung und Sicherheit liegt beim Anbieter – die Kontrolle über Server und Anwendungen bleibt beim Kunden. Eine flexible Alternative zum Eigenbetrieb – besonders geeignet für Georedundanz, Backup und Disaster-Recovery.

Nicht alle Unternehmen sind bereit, ihre IT vollständig in fremde Hände zu geben – und oft lassen sich bestehende Infrastrukturen gar nicht ohne Weiteres in die Cloud oder ein RZaaS-Modell (Rechenzentrum-as-a-Service) überführen. Vor allem Unternehmen mit Spezial-Hardware, sensiblen Anwendungen oder branchenspezifischen Anforderungen suchen nach Alternativen zur Eigenhaltung – ohne dabei die Kontrolle über ihre Systeme zu verlieren. Colocation ist für viele dieser Szenarien ein pragmatischer Mittelweg.

Beim Colocation-Modell stellen Unternehmen ihre eigene IT-Hardware – Server, Storage, Netzwerk-Komponenten – in einem Rechenzentrum eines Dienstleisters auf. Der Anbieter sorgt für Strom, Kühlung, physische Sicherheit, Zutrittskontrollen, Brandschutz und oft auch für redundante Anbindungen. Die Hardware bleibt im Besitz des Kunden, ebenso wie die Verantwortung für Betriebssysteme, Applikationen und Wartung.

Vorteile für anspruchsvolle IT-Umgebungen

Colocation eignet sich vor allem für Organisationen, die aus regulatorischen, technischen oder wirtschaftlichen Gründen nicht auf gemanagte IT-Strukturen setzen wollen oder können. Auch dann, wenn bereits Investitionen in leistungsfähige Systeme erfolgt sind, kann die Verlagerung der Infrastruktur in ein professionelles Umfeld sinnvoll sein.

Typische Einsatzszenarien sind:

- Betrieb von Legacy-Systemen mit besonderen Anforderungen
- Einsatz von Spezialhardware, etwa GPU-Servern oder High-Density-Umgebungen
- Anbindung an zentrale Dienste mit hohen Verfügbarkeitsanforderungen
- Ersatz für unsichere oder überlastete Serverräume am Standort

Der große Vorteil: Unternehmen behalten die volle Hoheit über ihre IT, profitieren aber gleichzeitig von professionellen Rahmenbedingungen – etwa durch ISO-zertifizierte Rechenzentren, Notstromversorgung, Brandschutz und Zugangskontrollen.

Grenzen und Herausforderungen

Bei RZaaS mieten Unternehmen »funktionierende IT« – bei Colocation mietet man »Platz für seine eigene IT«. Wer seine Systeme im Rechenzentrum unterbringt, bleibt für deren Betrieb, Updates, Monitoring und Störungsbeseitigung selbst verantwortlich. Auch die physische Betreuung – etwa beim Austausch von Festplatten oder bei Hardware-Defekten – erfordert eine Remote- oder Vor-Ort-Lösung, sofern der Anbieter keine Smart-Hands-Services anbietet.

Hinzu kommt: Skalierung ist bei Colocation nicht ohne Weiteres möglich. Neue Hardware muss beschafft, aufgebaut und in Betrieb genommen werden – was Zeit und Personal bindet. Wer maximale Flexibilität sucht, ist mit einem RZaaS-Modell oft besser bedient.

Fazit: Colocation als stabile Brücke zwischen Eigenbetrieb und Outsourcing

Colocation bietet professionelles Hosting für eigene Systeme – und damit einen zuverlässigen Mittelweg zwischen Eigenbetrieb und vollständigem IT-Outsourcing. Wer besondere Anforderungen hat oder den Schritt zum vollständigen Outsourcing (noch) nicht gehen will, findet hier ein stabiles Betriebsmodell mit klaren Verantwortlichkeiten.

Eine Colocation ist auch eine Option für Unternehmen, die eine zweite Betriebsumgebung benötigen, ohne doppelt in Infrastruktur zu investieren. Beispielsweise um einen Backup- und Disaster-Recovery-Standort aufzubauen oder um eine Georedundanz zu schaffen bzw. um Lastverteilung und Hochverfügbarkeit sicherzustellen. Sie profitieren von professionellen Rahmenbedingungen, ohne die Kontrolle über ihre Hardware aufzugeben – gerade bei Szenarien, bei denen RZaaS oder Cloud-Angebote zu starr oder zu wenig anpassbar wären. ■

Kerstin Mende-Stief
speicherguide.de

Unterschiedliche Anforderungen von Trainings- und Inferenzphasen

LEISTUNGSSTARKE STORAGE-INFRASTRUKTUR FÜR KI-WORKLOADS

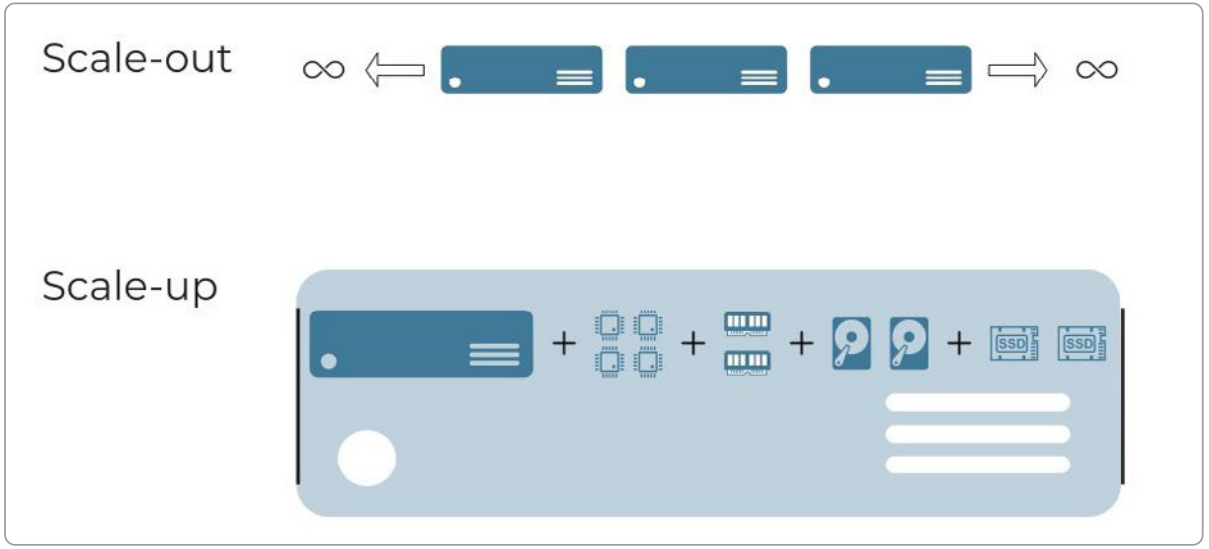
Eine optimale Storage-Infrastruktur ist entscheidend für den Erfolg von KI-Anwendungen. Der Leitfaden beleuchtet die Unterschiede zwischen Trainings- und Inferenzphasen und deren spezifische Anforderungen an Speicherlösungen. Wir erklären, warum NVMe-SSDs, skalierbare Architekturen und schnelle Netzwerkverbindungen erfolgskritisch sind und zeigen Strategien zur Leistungssteigerung und Kosteneffizienz auf.

Die Storage-Infrastruktur ist entscheidend für den Erfolg von KI-Anwendungen. Dabei müssen verschiedene Aspekte berücksichtigt werden: Große Kapazität, niedrige Latenz, hoher Durchsatz, Skalierbarkeit, Datenbeständigkeit oder die Integration mit den gängigen KI-Werkzeugen und Frameworks sind ebenso wichtig wie die wirtschaftliche Tragfähigkeit einer Lösung. Bei komplexen maschinellen Lernprozessen (Machine-Learning, ML) spielen zudem erweiterte Funktionen wie Metadaten-Management oder Versionierung eine wesentliche Rolle. Die Vielzahl dieser Anforderungen unterscheidet Storage für KI-Workloads von Speicherprodukten für klassische Unternehmens-Anwendungen wie ERP oder CRM. Auch bei den KI-Workloads selbst gibt es Unterschiede: Das Training beispielsweise eines großen Sprachmodells (Large Language Model, LLM) stellt andere Anforderungen als dessen praktische Anwendung. In diesem Artikel gehen wir auf Bedürfnisse an die Leistung ein, erklären warum diese wichtig sind, und stellen geeignete Storage-Architekturen bzw. Lösungen vor.

Prämisse: Unterschied zwischen Training und Inferenz berücksichtigen

Bei der Betrachtung der Speicheranforderungen für KI-Anwendungen (Künstliche Intelligenz) ist es wichtig, zwischen den beiden Hauptphasen Training und Inferenz zu unterscheiden. Jede der beiden Phasen stellt eigene Anforderungen an Speichersysteme. Das Verständnis dieser Unterschiede ist entscheidend für die Auswahl und Implementierung geeigneter Speicherlösungen, die sowohl die Trainings- als auch die Inferenz-Phasen von KI-Anwendungen effektiv unterstützen und so eine optimale Leistung und Effizienz im gesamten KI-Workflow gewährleisten. Das Training erfordert erhebliche Investitionen in skalierbare, leistungsstarke Speicherplattformen. Bei der Inferenz kommt es vor allem auf Geschwindigkeit und Zuverlässigkeit an.

Im Unternehmenskontext müssen geeignete Speicherlösungen beide Phasen optimal bedienen können. Selbst vortrainierte KI muss auf ihre spezielle Einsatzumgebung abgestimmt werden und kontinuierlich weiterlernen. Das ist notwendig, um sich an unternehmensspezifische und neue Daten, Muster oder Anforderungen anzupassen. Dieser Prozess wird als maschinelles Lernen oder Machine-Learning (ML) bezeichnet. Je mehr Daten ein KI-System verarbeitet, desto besser kann es Muster erkennen, Vorhersagen treffen und Entscheidungen treffen. Ohne kontinuierliches Lernen würden die Ergebnisse mit der Zeit weniger präzise oder die KI gänzlich obsolet werden. Die Leistungsfähigkeit wiederum wirkt sich im Betrieb unmittelbar auf das Nutzererlebnis aus und kann in bestimmten Einsatzszenarien sogar sicherheitskritisch werden.



Grafik: Kerstin Mende-Stief

Bei einer Scale-out-Architektur wird die Kapazität durch Hinzufügen neuer Geräte erweitert. Beim Scale-up werden einem Gerät einzelne Komponenten wie Arbeitsspeicher, zusätzliche Prozessoren oder Festplattenkapazität hinzugefügt.

Leistung und Kapazität

KI-Workloads erfordern eine konstant hohe Leistung. Für das Training greifen KI-Anwendungen häufig sequentiell auf Daten zu. Hohe IOPS (Input/Output Operations Per Second) und konstant schnelle Lese- und Schreibvorgänge gewährleisten eine reibungslose Leistung auch bei Spitzenbelastungen. Skalierbare Speicherlösungen ermöglichen es Unternehmen, wachsende Datenmengen ohne Leistungseinbußen zu bewältigen. Mit Load-Balancing lässt sich die Belastung auf verschiedene Speichergeräte verteilen. So kann die Gesamtleistung erhöht und Engpässe vermieden werden.

Bei der Inferenz kommt es vor allem auf die effiziente Ausführung der vorab trainierten Modelle an. Je nach Anwendung entscheiden Modellgröße (Anzahl der Parameter) und Kontextlänge (maximale Größe der kombinierten Eingabe und Ausgabe) über die Relevanz der Ergebnisse. Je mehr Informationen zur Verfügung stehen und in Kontext gesetzt werden können, desto genauer ist das Resultat. Ein Modell mit 70B Parametern liefert bessere Antworten als ein Modell mit 7B Parametern. Größere Modelle und Sequenzlängen haben jedoch andere Ansprüche an die Rechenleistung und Speicherbandbreite.

Skalierbare, verteilte Speicherlösungen (zum Beispiel Ceph, HDFS oder Lustre) ermöglichen die effiziente Verwaltung großer Datenmengen. RAID bzw. Erasure-Coding sorgen für eine hohe Verfügbarkeit der Daten. Eine Herausforderung kann die Verteilung der Daten über unterschiedliche Plattformen und Standorte hinweg sein. Eine effiziente Metadatenverwaltung ist entscheidend für die Organisation der Datenstruktur und

beschleunigt die Suche nach bestimmten Datensätzen.

Scale-out vs. Scale-up

Vor allem für das Training künstlicher Intelligenz wird eine Vielzahl unterschiedlicher Informationen benötigt. Die Speicherlösungen müssen über eine große Kapazität verfügen und mit großen Datensätzen umgehen können. Je mehr Daten zur Verfügung stehen, desto genauer oder relevanter kann das Ergebnis ausfallen. Das gilt sowohl für das Training als auch für die Anwendung, beispielsweise in Analysen (Analytic AI) oder Vorhersagen (Predictive AI). Eine Plattform sollte also nahtlos skalierbar sein. Scale-out-Architekturen sind einfacher im Betrieb und daher besser geeignet als ein reiner Scale-up-Ansatz.

In einer Scale-out-Architektur kann die Kapazität im laufenden Betrieb beliebig erweitert werden. Beim Scale-up-Ansatz setzt die Hardware der maximal möglichen Erweiterung eine physische Grenze. Das Clustern einzelner Appliances zu einem logischen Verbund ist komplexer und erfordert oft spezielles Fachwissen. Sowohl die Installation als auch der Betrieb können komplex und aufwändig sein. Selbstverständlich können auch einzelne Geräte in einer Scale-out-Architektur mit einzelnen Komponenten wie Arbeitsspeicher oder HDDs bzw. SSDs nachgerüstet werden. Vor allem Speicherkapazität lässt sich bei den gängigen Herstellern problemlos im laufenden Betrieb ergänzen. Gegebenenfalls müssen IT-Manager die RAID-Konfiguration manuell anpassen.

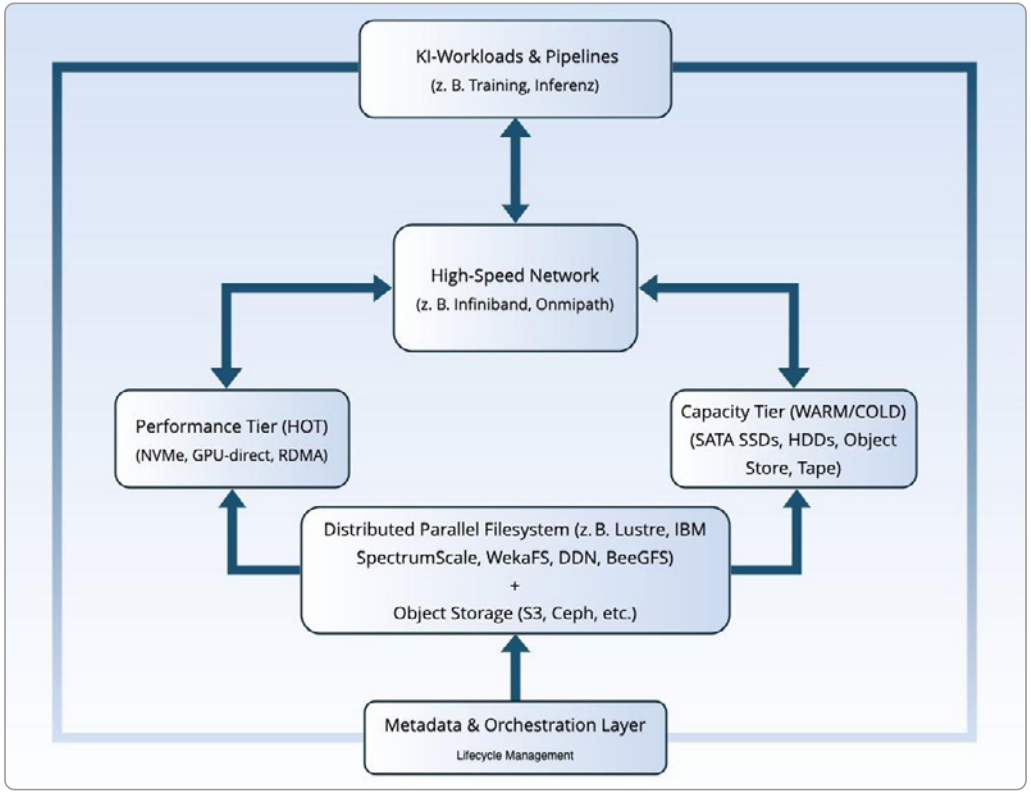
Netzwerk und Kommunikation

Kritisch sind bei KI-Workloads insbesondere Laufzeiten bei Dateizugriffen (Latenz) und die Geschwindigkeit, mit der Daten gelesen bzw. geschrieben werden können. Physische Limitationen wie Bandbreite oder Einschränkungen aufgrund veralteter Technologie können zu Ineffizienz führen. Jeder Weg, den die Daten zurücklegen, jede Schicht, die sie durchlaufen, jeder Kontextwechsel und jede Protokollübersetzung verursachen Verzögerungen. Netzwerk-Protokolle wie TCP belasten den gesamten Übertragungsweg und erhöhen die Paketumlaufzeiten besonders zwischen entfernten Standorten.

Kommunikation mit Speichermedien

Eine wichtige Rolle im Hardware-Stack spielen unter anderem die der Übertragung zugrunde liegenden Protokolle. Das speziell für SSDs entwickelte NVMe-Protokoll basiert auf dem PCIe-Standard und erlaubt es Speichermedien, direkt mit der CPU zu kommunizieren. Zusätzliche Controller oder die für SAS/SATA benötigten AHCI- bzw. SCSI-Adapter sind nicht nötig. Der Vorteil wird beim Vergleich der Latenzen deutlich, die sich von mehr als 100 Millisekunden (ms) bei SATA über 100 Mikrosekunden (µs) mit SAS auf bis zu 10 µs bei PCIe reduzieren lassen.

Features wie ein Full-Duplex-Mode sind bei gleichzeitigen Schreib- und Lesezugriffen nützlich. Zwar ist auch SAS full-duplex-fähig. Allerdings sind die Bandbreiten bei PCIe – und damit NVMe – um ein Vielfaches höher. PCI-Express-Verbindungen unterstützen bis zu 16 Lanes, die gebündelt werden können. Die aktuell genutzte PCIe Gen 5 ermöglicht eine Datenübertragungsrate von knapp 4 GByte/s pro Lane. Mit jeder neuen Generation verdoppelte sich bisher die Bandbreite. Die für 2025 angekündigte Gen 7 soll knapp 16 GByte/s pro Lane erlauben. SAS hängt seit über zehn Jahren bei 12 GByte/s fest, für den schnelleren 24G-Standard (SAS-4) gibt es kaum Geräte. Die Anbindung schneller Speichermedien wie SSDs sollte daher prinzipiell über NVMe erfolgen. Die Hardware muss dafür passende Anschlüsse für NVMe-Geräte bieten (M.2, U.2/U.3 oder EDSFF).



Grafik: Kerstin Mende-Stief

Im Zentrum der Architektur stehen das Highspeed-Netzwerk und eine funktionale Trennung zwischen einem Performance-Tier für hochfrequent genutzte Daten und einem Kapazitäts-Tier für skalierbare, kosteneffiziente Speicherung.

Auch bei den IOPS gibt es große Unterschiede. Hochleistungs-SSDs schaffen über 1 Million IOPS. Das hat unter anderem mit der unterstützten Anzahl Warteschlangen (Queues) und Befehlen je Warteschlange (Kommandos oder engl. Commands) zu tun. PCIe unterstützt 64.000 (64k) Command-Queues mit jeweils 64k Befehlen pro Queue. Zum Vergleich: eine SATA-Schnittstelle unterstützt gerade mal eine einzelne Queue mit maximal 32 Befehlen. Das dem SAS zugrunde liegende SCSI-Protokoll erlaubt nur 256 Kommandos in der (einzigen) Command-Queue.

Netzwerk-Infrastruktur

Bei KI-Anwendungen beeinflusst die Entfernung der Daten zu den Rechen-Ressourcen sowohl die Latenz als auch die Leistung. Schnelle Netzwerkverbindungen wie Infiniband (unter anderem **Nvidia**) oder das inzwischen quelloffene Omni-Path (**Cornelis Networks**) beschleunigen die Übertragung von Daten zwischen Storage und Rechenknoten. Beide sind Implementationen von RDMA (Remote Direct Memory Access), das den direkten Zugriff auf entfernte Speicher ermöglicht, ohne CPUs und Hauptspeicher der jeweiligen Rechner zu belasten. Latenz wird quasi auf die Signallaufzeit der zugrunde liegenden physischen Leitung (Kupfer, Glas) reduziert (wire speed).

Datenmanagement

KI arbeitet sowohl mit strukturierten als auch mit unstrukturierten Daten wie Datenbanken, Bildern und Textdateien. Das Speichersystem muss diese verschiedenen Datentypen effizient verwalten und abrufen können.

Die Bereinigung der Daten führt zu relevanteren Ergebnissen und entlastet die Infrastruktur (Festplatten-I/O, Speicherplatzbedarf, CPU-/GPU-Last, Netzwerkverkehr). Beim so genannten Data-Pruning werden irrelevante Daten ausgeschlossen und redundante Informationen entfernt (Deduplizierung).

Komprimierung der Daten vor der Ablage kann die Übertragungsmenge reduzieren. Allerdings ist zu beachten, dass bereits komprimierte Formate wie JPEG, MP3, H.265 oder MPEG sich nicht wesentlich weiter komprimieren lassen. Aus dem HPC stammende Formate wie HDF5 oder zarr sind für den Umgang mit großen Datasets optimiert. Auch lassen sich Daten für die Übertragung serialisieren. Dafür werden die Daten in ein für den Transport geeignetes Format konvertiert.

Von besonderer Bedeutung für KI sind Metadaten. Es wird zwischen drei Haupttypen unterschieden: administrativ, strukturell und informativ. Die Informationen ermöglichen es, Daten zu strukturieren, zu organisieren, zu kategorisieren, zu klassifizieren und vor allem, sie auch wiederzufinden. Informative Metadaten liefern zudem einen wichtigen Kontext für das Training und den Betrieb von KI-Algorithmen. Metadaten müssen nicht nur aktuell, sondern vor allem vollständig sein. Moderne Storage-Plattformen sind in der Lage, Metadaten automatisiert zu erfassen und zu organisieren. Dennoch sollten menschliche Mitarbeiter sowohl Vollständigkeit als auch Richtigkeit zumindest stichprobenartig prüfen.

Richtlinien helfen bei der Ablage von Daten. Vor allem in verteilten Systemen ist es wichtig, bereits in der Planungsphase die Dateistruktur und Ablageorte für bestimmte Daten festzulegen.

KI-Workloads – insbesondere im Training und bei der Inferenz großer Modelle – erzeugen massive Datenvolumina und erfordern hohe Bandbreite, geringe Latenz und flexible Skalierbarkeit. Eine moderne, auf diese Anforderungen zugeschnittene Storage-Architektur bildet das Rückgrat einer performanten und zukunftssicheren KI-Infrastruktur.

Das Performance-Tier ist auf maximale Leistung ausgelegt: NVMe-SSDs bieten in Kombination mit Technologien wie *NVIDIA GPUDirect Storage (GDS)* und *Remote Direct Memory Access (RDMA)* einen extrem schnellen Datenzugriff – ideal für Deep-Learning-Training auf GPU-Clustern. Typische Einsatzszenarien sind etwa Datenvorverarbeitung und das kontinuierliche Nachladen von Batch-Daten beim Modelltraining.

Anbieter von Data-Management-Lösungen erfassen Daten aus verschiedenen Quellen und helfen bei der Organisation, Verwaltung und Analyse. Hier sind einige Beispiele: **Alation** arbeitet mit Gamification als Motivator für die Pflege der Metadaten. **Data Dynamics** ist spezialisiert auf die Verwaltung unstrukturierter Daten und hat eine eigene KI entwickelt, mit der Daten unter anderem automatisiert klassifiziert werden können. **Komprise** legt großen Wert auf Informationssicherheit und schützt mit seinen umfangreichen Analysewerkzeugen sensible Informationen zuverlässig.

Der effiziente und unterbrechungsfreie Betrieb einer solchen Speicherarchitektur erfordert ein durchgängiges Monitoring- und Optimierungskonzept. Im Fokus stehen dabei Kennzahlen wie Bandbreite, Latenz, Zugriffsmuster, IOPS und die Auslastung einzelner Komponenten oder Tiers.

Mit Hilfe spezialisierter Tools wie *Prometheus, Grafana, NetApp Active IQ* oder *DDN Insight* können Unternehmen die Performance in Echtzeit überwachen, Engpässe frühzeitig identifizieren und auf Basis historischer Daten Kapazitätsplanungen durchführen. Darüber hinaus lassen sich Optimierungspotenziale – etwa beim Tiering-Verhalten oder in der Verteilung von Hot- und Cold-Data – gezielt aufdecken und automatisiert umsetzen.

Ein transparenter Einblick in die Storage-Nutzung ist nicht nur für technische Effizienz wichtig, sondern auch zur Kostenkontrolle und Einhaltung regulatorischer Vorgaben.

Fazit: Die perfekte Architektur besteht aus einer Kombination

Die perfekte Storage-Architektur für das Training von KI ist eine Kombination aus leistungsfähigen, skalierbaren und zuverlässigen Speicherlösungen, die hohe Geschwindigkeit, Redundanz und Sicherheit bieten. Durch die Integration von modernen Technologien wie NVMe-SSDs, verteilten Dateisystemen und Hochgeschwindigkeitsnetzwerken können die Anforderungen des KI-Trainings effektiv gemeistert werden. Darüber hinaus ist die Implementierung von Metadaten-Management, Sicherheitsfunktionen und Überwachungs-Tools entscheidend, um die Gesamteffizienz und Zuverlässigkeit der Storage-Architektur zu gewährleisten. ■

Weiterführende Links:

Lesen Sie eine ausführliche Fassung auf ➔ **speicherguide.de**

storage-magazin.de
Eine Publikation von speicherguide.de

Newsletter-Abonnenten erhalten die neue Ausgabe jeweils »linkfrisch« an ihren Mail-Account.Registrieren Sie sich bitte hier. Beachten Sie auch unser Archiv im Download-Bereich.

Unser Team

Karl Fröhlich
Chefredakteur
speicherguide.de

Michael Baumann
Redaktion
speicherguide.de

Peter Marwan
Redaktion
speicherguide.de

Kerstin Mende-Stief
Redaktion
speicherguide.de

Bettina Röber
Mediaberatung
speicherguide.de

storage-magazin.de
eine Publikation von speicherguide.de
Karl Fröhlich
Ginsterweg 12, 81377 München
Tel. +49 (0) 89-740 03 99
E-Mail: redaktion@speicherguide.de

Chefredaktion, Konzept:
Karl Fröhlich *(verantwortlich für den redaktionellen Inhalt)*
Tel. 089-740 03 99
E-Mail: redaktion@speicherguide.de

Redaktion:
Karl Fröhlich, Uli Neiss, Kerstin Mende-Stief

Schlussredaktion:
Brigitte Scholz, Michael Baumann

Titelbild:
speicherguide.de/Dall-E

Layout/Grafik:
Uwe Klenner, Layout und Gestaltung,
Rittsteiger Str. 104, 94036 Passau,
Tel. 08 51-9 86 24 15
www.layout-und-gestaltung.de

Mediaberatung:
Bettina Röber
Tel. +49 177 8487001
E-Mail: broeber@speicherguide.de