Zuverlässigkeit vor, hinter, unter und über dem Cluster

Benedikt Stockebrand

GUUG UpTimes, September 2005

Wer das Thema „Hochverfügbarkeit“ hört, denkt wohl in den meisten Fällen sofort an einen HA-Cluster. Aber Cluster sind nur ein winziger Teil dessen, was eine zuverlässige Umgebung ausmacht. Das Drumherum wird immer wieder vernachlässigt. So habe ich in der Vergangenheit eine Liste an wesentlichen Punkten gesammelt, an denen die angestrebte Zuverlässigkeit von Systemen immer wieder scheitert.

Eine Anmerkung zur Begrifflichkeit: „Zuverlässigkeit“ ist das Ziel, das man erreichen will. „Redundanz“ ist ein *Mittel, diesem Ziel näherzukommen, und „Hochverfügbarkeit“ ist entweder der dazugehörige Marketing-Begriff oder das, was man durch den systematischen Einsatz von Redundanz erreichen will.*

Geld und Zeit

Zuverlässigkeit ist teuer. Die meisten der folgenden Punkte setzen ein entsprechendes Budget voraus, das deutlich über die doppelte Hardware und die Cluster-Software hinausgeht.

Auch wenn Zeit Geld ist, lässt sich Zeit doch nicht kaufen. Zuverlässigkeit hat viel damit zu tun, sauber zu arbeiten und zu überlegen, um Probleme zu erkennen, bevor sie auftreten. Wer diese Zeit nicht hat, wird zu keinem zuverlässigen System kommen.

Konventionalstrafen

Wer von seinen Zulieferern ernstgenommen werden will, vereinbart mit ihnen Konventionalstrafen. Erfahrungsgemäß sind die in den seltensten Fällen ausreichend hoch, weil ein seriöser Anbieter eine entsprechende Vereinbarung wie eine Versicherung kalkuliert und entsprechend teuer in Rechnung stellt.

Trotzdem ist es essentiell wichtig, die Konventionalstrafen so hoch anzusetzen, dass sie einen entstehenden Schaden tatsächlich decken. Denn einige „seriöse“ Anbieter kalkulieren knallhart, dass sie ein Schaden eben nur diese Konventionalstrafe kostet. Dann werden Zusagen gemacht ohne das geringste Bemühen, sie einzuhalten. Und als Kunde bekommt man zu hören, dass man ja die Konventionalstrafe selbst so angesetzt hat.

Personelles

Zuverlässigkeit fängt bei den Administratoren an – und das nicht nur, wenn man absteigend nach Kosten sortiert. Ohne eine entsprechende Personalstruktur und -organisation ist „Hochverfügbarkeit“ ein Wunschtraum.

Teamgröße

Ein unterbesetztes Team, das ständig damit beschäftigt ist, akute Störungen zurechtzuflicken, kann keinen zuverlässigen Betrieb sicherstellen. Wenn eine redundante Komponente ausfällt, wird die Ersatzkomponente (hoffentlich) den Service übernehmen und einen Ausfall verhindern. Aber ohne Administratoren, die den Ausfall bemerken und die fehlerhafte Komponente austauschen wiederherstellen, gewinnt man hier nur etwas Zeit – bis die Ersatzkomponente auch ausfällt. Dann das gesamte System insgesamt wieder in Betrieb zu bringen kostet gerade bei Clustern erfahrungsgemäß wesentlich mehr Zeit als zwei Ausfälle eines nicht-redundanten Systems. Und wenn das gesamte Administratoren-Team ständig Feuerwehr spielt, wird es immer „wichtigeres“ geben als das ausgefallene redunante Netzteil oder den einen ausgefallenen Cluster-Knoten.

Rufbereitschaft und Schichtbetrieb

Systeme, die zuverlässig sein müssen, brauchen ständige Pflege und Aufsicht. Ohne einen kompetenten Administrator, der bei Problemen sofort eingreifen kann, ist ein Cluster wertlos.

Diese ständige Verfügbarkeit eines Administrators ist in vielen Fällen, vor allem bei kleinen, einzelnen Clustern, der größte Kostenpunkt beim Einsatz einer zuverlässigen Lösung.

Training

Cluster sind komplex und in vielen Situationen anti-intuitiv. Ohne entsprechendes Training kann kein Administrator sie sinnvoll betreuen. „Training“ heißt in diesem Zusammenhang nicht nur, dass *jeder* betroffene Administrator das Schulungsprogramm des Herstellers genossen hat, sondern auch, dass *jeder* betroffene Administrator tatsächlich regelmäßig für den Ernstfall trainiert.

Administrator-Turnover

Wenn Administratoren ständig wechseln, werden sie nie ein Gespür dafür entwickeln, wo eine Umgebung ihre spezifischen Probleme hat oder wie die Kollegen arbeiten. Dementsprechend wird die Umgebung, die man ja auch nicht einfach mal etwas genauer untersuchen kann, weil man damit den

Regelbetrieb gefährden würde, immer weniger verstanden.

Spätestens, wenn ein neuer Administrator einen HA-Cluster „erbt“, ohne seinen Vorgänger noch kennenzulernen, sind Ausfälle vorprogrammiert.

Entwickler

Software muss passend entwickelt werden, um zuverlässig zu sein. Ein Cluster ist dazu nicht zwingend nötig, bietet aber eine Reihe von Mechanismen, um die Entwicklung von hochverfügbarer Software zu erleichtern.

Trotzdem kommt es immer wieder vor, dass Entwickler den Standpunkt vertreten, dass sich „der Cluster um die Hochverfügbarkeit kümmern muss“. Die folgenden Schlammschlachten zwischen Entwicklern und Administratoren, oder auch zwischen Software-Hersteller und Kunde, sind teuer aber nicht konstruktiv.

Auch Entwickler von hochverfügbarer Software müssen also entsprechend qualifiziert sein und durch geeignete Anforderungsdefinitionen dazu verpflichtet werden, ihre Software tatsächlich hochverfügbar zu machen.

Organisatorisches

Auch wenn viele Administratoren den Ruf nach „klar definierten Prozessen“ nicht mehr hören können, weil sich dahinter nur der verklausulierte Vorwurf verbirgt, sie könnten nicht eigenverantwortlich arbeiten, gibt es doch einige organisatorische Voraussetzungen für einen zuverlässigen Betrieb.

Koordination

Administratoren, die sich nicht untereinander koordinieren, arbeiten aneinander vorbei. Gerade wenn Zuverlässigkeit wichtig ist, darf das nicht passieren. Und mit Schichtbetrieb und Rufbereitschaft wird die Koordination untereinander nicht-trivial.

Es ist grundsätzlich egal, ob es zur Koordination ein Logbuch gibt, wo jeder zum Schichtbeginn nachliest, was die Kollegen vor ihm gemacht haben, und wo alle Vorfälle und Aktivitäten dokumentiert werden, oder ob dazu eine Change-Management-Software eingesetzt wird. Wichtig ist aber, dass es eine solche Koordination gibt und sie konsequent gepflegt wird.

Spezifizierte Zuverlässigkeits-Anforderungen

Unterschiedliche Anwendungen verstehen unterschiedliche Dinge unter „Zuverlässigkeit“. Die Steuerung einer Zentralheizung sollte (mindestens im Winter) nicht für mehrere Stunden zu Wartungszwecken ausgeschaltet werden müssen; ein gelegentlicher spontaner Reboot von wenigen Minuten ist aber kein Problem. Die Rechner der Frankfurter Flugsicherung dürfen (von meiner Seite aus gerne jede Nacht) mit entsprechendem Vorlauf notfalls einige Stunden ausgeschaltet werden; ein spontaner Reboot über zwei Minuten dagegen kann fatal sein: “Jungs, bleibt da oben mal kurz wo ihr gerade seid, wir müssen hier unten die Sicherung wieder reindrücken“ ist nicht wirklich eine Option.

Wenn die spezifischen Anforderungen an die Hochverfügbarkeit eines Systems nicht dokumentiert sind, kann weder ein Entwickler noch ein Administrator die nötigen Entscheidungen treffen, um ihnen bei der Entwicklung oder beim Betrieb gerecht zu werden, denn in weniger extremen Fällen ist diese Abwägung nicht so offensichtlich wie in den beiden Beispielen.

Maintenance

Hard- und Software müssen gelegentlich gepflegt werden. Dazu ist es in vielen Fällen nötig, sie außer Betrieb zu nehmen. Passiert das nicht, steigt das Risiko eines ungeplanten Ausfalls kontinuierlich an, bis auch der beste Cluster dieses Risiko nicht mehr kompensieren kann.

Deshalb muss jede auf Zuverlässigkeit ausgerichtete Umgebung systematisch gewartet werden – mit geregelten Wartungsintervallen, die auch tatsächlich zur Pflege des Systems genutzt werden.

Katastrophenplanung

Eine Umgebung, an die Verfügbarkeitsanforderungen gestellt werden, ist immer wichtig genug, dass für einen Ausfall ein Katastrophenplan existiert. Darin ist mindestens geklärt, wer welche Entscheidungskompetenzen hat, mit welcher Priorität betroffene Services wiederhergestellt werden, wer wann worüber informiert werden muss und welche Ziele der Katastrophenplan verfolgt – denn wenn eine juristisch wasserdichte Beweissicherung wichtiger als die schnellstmögliche Wiederinbetriebnahme ist, muss ein Administrator das wissen, um entsprechend zu handeln.

Ein guter Katastrophenplan zeichnet sich darüber hinaus noch durch zwei weitere wichtige Eigenschaften aus: Er ist aktuell, enthält also keine veralteten Informationen, und er ist *kurz*, damit der „Katastrophen-Admin“ ihn in der gegebenen Eile noch lesen kann.

Sicherheit

Auch wenn der eine oder andere Cluster-Anbieter in der Vergangenheit darauf bestanden hat, dass Telnet, Rlogin und FTP ein- und Ssh ausgeschaltet sein müssen, gilt trotzdem: Eine auf Zuverlässigkeit ausgerichtete Umgebung muss auch gegen böswillige Angriffe geschützt werden.

Physikalische Sicherheit

Ein hochverfügbares System muss in einer physikalisch geschützten Umgebung stehen. Ob das eine Lambertz-Zelle in einem hermetisch abgeriegelten Rechenzentrum ist, oder ein abgeschlossener 19”-Schrank in der Besenkammer, spielt dabei keine wesentliche Rolle.

Wichtig ist, dass der Zugang auf Mitarbeiter beschränkt ist, die wissen, was sie tun. Es wäre nicht das erste Mal, dass eine Putzfrau^WRaumpflegerin beim Saugen hinter einem Rechner ein Netzwerkkabel herauszieht (oder ein Systemadministrator es anschliessend mit Tesafilm „fixiert“, weil die Haltezunge abgebrochen ist – aber das ist ein anderes Problem). Und auch dass Produktionsmitarbeiter einen wichtigen Server für eine Runde

Doom missbrauchen soll schon vorgekommen sein.

Schneller physikalischer Zugang

Andererseits ist die Diskussion mit einem Wachmann, wenn ein Admin kurz nach Mitternacht vor der Tür steht und ein ausgefallenes System wiederbeleben will, ausgesprochen teuer: „Um diese Zeit darf niemand mehr in das Gebäude“ – „Doch, sie haben eine Liste in dem Ordner da vorne, wer rund um die Uhr ins Gebäude darf, und da stehe ich drin“ – „Von so einer Liste weiss ich nichts“ – „Die ist in dem zweiten Ordner von rechts, ganz oben drauf“ – „Aber Ihre Chipkarte ist für nachts gesperrt“ – „Nein, die ist genauso freigeschaltet wie ihre“ – „Das kann gar nicht sein“ – „Doch, wenn sie mich reinlassen zeige ich es ihnen“ – „Ich darf sie aber nicht reinlassen“…

Ich habe genau diese Situation schon erlebt, genauso wie ich auch schon erlebt habe, dass ein Wachmann kaum dass er mich nachts erkannt hat schon die anwesende Nachtschicht zusammengetrommelt hat „weil etwas nicht stimmen kann, wenn der um so eine Zeit auftaucht“. Sicherheit und zügiger physikalischer Zugang bei Ausfällen schließen sich nicht gegenseitig aus, setzen aber wieder eine gewisse Organisation und vor allem Routine voraus.

Logische Sicherheit

Ungeschützte „Hochverfügbarkeitslösungen“ sind keine – nicht nur, wenn es um Windowsrechner mit Internetanbindung geht.

Vor dem Aufbau einer zuverlässigen Lösung steht deshalb immer der Aufbau einer Sicherheits-Infrastruktur, die in sich wiederum entsprechend zuverlässig sein muss. Ein Cluster, der hinter einer einzelnen Firewall steht, ist Geldverschwendung.

Gebäudetechnik

Zuverlässige Systeme gehören in entsprechende Räume. Auch dabei werden oft wesentliche Aspekte „übersehen“, weil sie teilweise viel Geld kosten.

Brandschutz

Wo viel Strom fließt, gibt es gelegentlich auch Kurzschlüsse, Funken, Schwelbrände. Wenn dann die Feuerwehr anrücken muss, wird es richtig teuer, denn Löschwasser richtet einige Schäden an.

Mit relativ wenig Aufwand lässt sich ein brauchbarer Brandschutz sicherstellen, wenn mehrere Brandabschnitte eingerichtet werden. Wenn sichergestellt ist, dass nach einem Brand in einem Abschnitt die anderen trotzdem weiterarbeiten können, ist auch bei einem Brand der Betrieb sichergestellt.

Eine geeignete Löschanlage hilft weiter, bei einem Ausfall den Schaden innerhalb des betroffenen Abschnitts und damit Versicherungsprämien auf ein Minimum zu reduzieren.

Andererseits sollte man nicht vergessen, dass die so restriktiven Brandschutzvorschriften in Deutschland ihren historischen Ursprung in einigen mittelalterlichen Stadtbränden haben; man sollte das Thema deshalb nicht überdramatisieren, sondern objektiv durchrechnen.

Wassereinbruch

Einen ernstzunehmenden Brand habe ich noch nie in einem Rechnerraum erlebt, aber einen Wassereinbruch schon. Eine undichte Wasserleitung direkt über einer freistehenden Test- und Entwicklungs-E450 hat mehrere Kubikmeter Wasser abgelassen, bevor jemand das Problem bemerkte. Einzig der durchlässige Doppelboden und die hoch montierten Steckdosen haben einen größeren Schaden verhindert.

Wer ernsthaft vor hat, ein Rechenzentrum im Keller einzurichten, sollte das berücksichtigen. Nicht nur, wenn der Keller in der Nähe von Rhein, Elbe, Nord- oder Ostsee stehen soll oder das Penthouse auf dem Dach einen Pool hat, sondern auch, wenn das Dach notorisch undicht ist.

Strom

Auch wenn wir nicht mit einer so maroden Stromversorgung wie Teile der USA oder einige afrikanische Länder zu kämpfen haben, ist eine zuverlässige Stromversorgung wichtig. Spätestens, wenn die ersten Blade-Systeme im Rechenzentrum einziehen und die Stromrechnung hochtreiben, wird es Zeit, die Dimensionierung der Notstromsysteme zu überprüfen.

Dass Dieselaggregate und Bleiakkus ein Minimum an regelmäßiger Pflege brauchen, sollte jedem Autofahrer einsichtig sein.

Um die Kosten und den Aufwand für Notstromsysteme nicht in beliebige Höhen zu treiben, ist es sinnvoll, essentiell wichtige Komponenten (Server, Storage, …) von nichtessentiellen Komponenten (Monitore, Drucker, …) zu trennen und nicht blind jede Kaffeemaschine an den Notstrom-gestützten Stromkreis anzuschliessen.

Für längerfristige Krisen, wenn allmählich der Dieseltank leer wird, hilft in manchen Fällen noch ein Krisenplan, welche „fast essentiellen“ Systeme notfalls doch noch abgeschaltet werden dürfen, vor allem, wenn die Tankkapazität nur für einige Stunden ausreicht.

Klima

Blades brauchen nicht nur viel Strom, sie erzeugen auch viel Abwärme. Kommt die Klimatisierung nicht mehr gegen die Wärme an, altert zuerst die Hardware schneller als notwendig. Wenn es dann noch wärmer wird, bleiben die ersten Systeme mehr oder weniger kontrolliert stehen.

Eine redundant ausgelegte Klimaanlage, die sich bei 38°C Außentemperatur von selbst abschaltet, hat in einem Rechenzentrum nichts verloren. Trotzdem habe ich so eine Konstruktion an einem Hochsommernachmittag schon ein Rechenzentrum lahmlegen sehen.

Netzverkabelung

Es gibt Verkabelungen, die sehen chaotisch aus, funktionieren aber recht gut. Und es gibt Verkabelungen, wo alle Kabel sauber mit Kabelbindern gebündelt in Kabelwannen liegen, aber kein Mensch mehr nachvollziehen kann, welches Kabel wohin geht und welche Kabel überhaupt noch benutzt werden.

Sicherlich ist es Arbeit, neue Patchkabel so durchzunummerieren, dass sich jedem Ende das dazugehörige Gegenstück zuordnen lässt.

Sicherlich kostet es Geld, neu aufgestellte Schränke gleich so großzügig fest zu verkabeln, dass freifliegende Einzelkabel im Zwischenboden die extreme Ausnahme bleiben.

Sicherlicht ist ein sauberes Kabel-Management nicht unbedingt etwas, was einen Systemadministrator zu freiwilligen Überstunden am Samstagabend motiviert.

Aber wer einmal versehentlich nicht nur einige unidentifizierte, aber offensichtlich tote Kabel im Zwischenboden weggeschnitten hat, sondern dabei auch die Heartbeat-Leitungen eines Clusters erwischt hat, wird diesen zusätzlichen Aufwand zu schätzen wissen.

Externe Netzanbindung

Dass WAN-Strecken anfällig sind, ist allgemein bekannt. Dass man deshalb eine redundante Netzanbindung haben will, auch.

Dass der zweite Anbieter für die redundante Netzanbindung die Leitungen beim ersten Anbieter anmietet, erfährt man gelegentlich erst, wenn das „statistisch Unmögliche“ passiert ist und beide Anbindungen gleichzeitig ausfallen – weil ein Baggerfahrer dreißig Kilometer entfernt eine „eklig dicke Baumwurzel“ aus dem Boden gezogen hat.

Es kommt tatsächlich vor, dass aus einem Rechenzentrum an zwei entgegengesetzten Enden Leitungen zu zwei verschiedenen Vermittlungsstellen führen, aber beide Strecken hinter dem Ortsausgang im gleichen Strang die nächsten fünfzig Kilometer zusammenlaufen, um dann vor dem Ortseingang der Gegenseite wieder getrennt zu werden.

Und was für WAN-Strecken zwischen Rechenzentren geht, funktioniert genauso auch mit der Internet-Anbindung. Immerhin kann man da mit traceroute und ähnlichem vielleicht schon im Vorfeld feststellen, dass da die Redundanz nur auf Papier besteht – in den Verträgen und bei den monatlichen Abbuchungen. Aber dank BGP lässt sich auch hier selten vorhersagen, ob bei allen möglichen Ausfällen die Reserveanbindung die nötige Kapazität hat, um einen geregelten Betrieb aufrechtzuerhalten oder genau dann unter Überlast zusammenbricht.

Gerade in diesem Zusammenhang gelten die Überlegungen zum Thema „Konventionalstrafen“.

Räumliche Trennung und Ausweichsysteme

Um bei Clustern ein Split Brain Syndrome zu vermeiden, stehen Cluster-Knoten selten allzu weit auseinander. Wenn dann eine Natur- oder andere Katastrophe zuschlägt, nutzt auch der beste HA-Cluster nichts mehr. Auch wenn es Bemühungen gibt, Lösungen zu entwickeln, bei denen die Cluster-Knoten entfernt voneinander stehen dürfen, habe ich damit doch einige Probleme.

Mindestens wenn eine Heartbeat-Verbindung zwischen den Cluster-Knoten notwendig ist, wächst deren Anfälligkeit monoton in ihrer Länge. Der Abstand, und damit die maximal mögliche räumliche Trennung, der Cluster-Knoten wird durch dieses Problem und die geforderte Verfügbarkeit nach oben beschränkt.

Wenn Verfügbarkeit auch in einem Katastrophenfall gefordert ist, muss also ein räumlich weit entferntes Ausfallsystem vorhanden sein, das *nicht* auf gängigen Cluster-Techniken basiert. Ob nun eine Standby-Datenbank samt täglichem Taxi-Kurier für die Backup-Bänder, ein asynchroner Plattenspiegel zwischen zwei großen Plattensystem oder eine rsync-basierte Synchronisation: Wichtig ist vor allem, dass bei einem Ausfall der Verbindung die Daten im Ausfallsystem nicht inkonsistent werden und man schnell wieder einen geregelten Betrieb aufnehmen kann.

Hardware

Nach den letzten beiden Abschnitten ist die Frage nach der Hardware schon wieder fast irrelevant – soweit es um die Kosten geht.

Qualität muss sein

Trotzdem muss die Qualität bei der Hardware-Auswahl im Vordergrund stehen. Ein „Doppelfehler“ wegen minderwertiger Hardware ist nicht so unwahrscheinlich, dass man ihn ignorieren könnte. Ob nun das nicht-redundante Billig-Netzteil in einem Arbeitsplatz-PC ausfällt, der als Cluster-Knoten missbraucht wird, oder für einen Cluster-Knoten mit „wirtschaftlichem Totalschaden“ kein baugleicher Ersatz zu bekommen ist: In jedem Fall hat man sein letztes Redundanz-As zeitweise ausgespielt und es steht einige Arbeit an, bis die Hochverfügbarkeit wieder sichergestellt ist.

Gerade bei Plattensystemen hat sich inzwischen übrigens wohl herumgesprochen, dass ein Plattenausfall in einem RAID-System relativ häufig dazu führt, dass durch die hohe Last beim Rekonstruieren der Daten auf einer neuen Platte noch eine weitere Platte ausfällt und die Daten endgültig verloren sind. Spätestens hier sollte klar sein, dass man mit Redundanz keine minderwertige Hardware kompensieren kann.

Das Ersatzteillager

Im gleichen Zusammenhang ist ein ausreichend dimensioniertes eigenes Ersatzteillager wichtig.

In Zeiten, wo man gelegentlich schon um die Hot-Spare-Platte im Software-RAID feilschen muss, ist das oft schwierig zu begründen. Wenn mit dem Lieferant einen “Platinum-mit-Diamanteinlagen”-Wartungsvertrag abgeschlossen wurde und innerhalb von zwei Stunden Ersatz ankommen

muss, hilft vielleicht noch der Hinweis auf die zu schwach dimensionierte Konventionalstrafe.

Aber manchmal sind auch zwei Stunden schon zu lang. Und wer schon einmal einen samstags nachts per Rufbereitschaft mobilisierten Servicetechniker der zweiten Wahl gesehen hat, wie er verschlafen und demotiviert an einem wichtigen System Hardware austauscht, will das vielleicht doch lieber selbst machen.

Test- und Trainingsumgebungen

Wie schon oben gesagt ist es wichtig, dass Systemadministratoren praktisch geübt sind, wenn sie an Produktivsystemen arbeiten.

Außerdem ist es wichtig, Patches vor dem Einspielen auf unerwünschte Nebenwirkungen hin durchzutesten.

Damit ist klar, dass eine Test- und Trainingsumgebung unbedingt notwendig ist. In manchen Fällen kann sie auch gleichzeitig als Ersatzteillager missbraucht werden. Wer viel Geld sparen will und eine marginale Verschlechterung der Verfügbarkeit in Kauf nehmen kann, wird vielleicht die räumlich getrennte Ausfallumgebung mit einem zweiten Satz Festplatten dazu missbrauchen, oder eine einzelne Testumgebung für mehrere produktive Umgebungen benutzen.

Aber ohne jede Testumgebung ist ein zuverlässiger Betrieb unmöglich.

IT-Infrastruktur

Hochverfügbarkeit setzt schließlich eine entsprechende IT-Infrastruktur voraus, wie sie in den meisten Rechenzentren auch existiert. Der Vollständigkeit halber dürfen sie aber hier nicht fehlen.

Monitoring

Ein Cluster, der nicht ständig überwacht wird, ist wertlos, genau wie das redundante Netzteil, dessen Ausfall erst bemerkt wird, wenn auch sein Ersatz ausfällt.

In Rechenzentren ist normalerweise ein mehr oder weniger aufwendiges Monitoring vorhanden. Hochverfügbare Systeme müssen ständig auf Ausfälle einzelner Komponenten überwacht werden – ein „der Service läuft, alles ist ok“ reicht nicht.

Backup

Auch deutschen Richtern ist inzwischen wohl klargeworden, dass Backups im Rahmen der üblichen Sorgfaltspflichten zum Standard gehören. Gelegenheiten, um das zu lernen, hatten sie wohl in der Vergangenheit genug, wenn es um schadensrechtliche Auseinandersetzungen wegen nicht erfolgter oder fehlerhafter Backups ging.

Eine redundante Lösung ist kein Ersatz für ein funktionierendes Backup; spätestens, wenn jemand böswillig ein Cluster-System manipuliert, hilft der Cluster nur, die Manipulation auf allen Knoten einheitlich durchzuführen.

Infrastruktur-Services

Zu guter letzt gibt es noch einige Services, die immer und überall vorhanden sind und erst dann ins allgemeine Bewußtsein rücken, wenn sie ausfallen.

Die wohl entscheidendsten Vertreter dieser Kategorie sind das DNS und die Zeitsynchronisation, üblicherweise per NTP. Aber auch LDAP-Server zur Benutzerauthentisierung, Mail-Gateways zur Alarmierung von Systemadministratoren während der Rufbereitschaft und der Ssh-Zugang per Terminalserver auf die seriellen Konsolen der entsprechenden Server können in diese Kateogorie fallen.

Wenn diese Services nicht selbst auf Hochverfügbarkeit ausgelegt sind, dann kann auch ein dahinter aufgebauter Cluster das nicht kompensieren.

Cluster-taugliche Software

Schließlich muss die Software, die auf einem Cluster eingesetzt wird, den Cluster selbst nutzen. Das ist nicht selbstverständlich; wer es ausprobieren will, darf einmal versuchen, einen Ssh-Cluster zu bauen, der bei einem Failover die laufenden Sessions *nicht* verliert…

Fazit

All diese Voraussetzungen für eine hochverfügbare Umgebung sind bekannt und (hoffentlich) unmittelbar einsichtig. Wirklich Überraschendes dürfte für die wenigsten dabei sein. Trotzdem werden immer wieder wesentliche Punkte übersehen und viel Geld an der falschen Stelle investiert, um die Verfügbarkeit von Systemen zu verbessern.

Wer also versucht ist, das letzte Viertelprozent an Verfügbarkeit herauszuholen, sollte nicht blind auf die Versprechen der Cluster-Anbieter hören, sondern zuerst die übrigen Randbedingungen für den hochverfügbaren Betrieb schaffen. Wenn dann am Ende das Geld für einen Cluster nicht mehr reicht, ist es mit hoher Wahrscheinlichkeit besser investiert worden als in einen Cluster, der unter den gegebenen Umständen keinen praktischen Nutzen bieten kann.