

Netzwerkkritische physische Infrastruktur: Optimierung des Geschäftswerts

Wendy Torell

White Paper Nr. 117

APC[®]
Legendary Reliability[®]

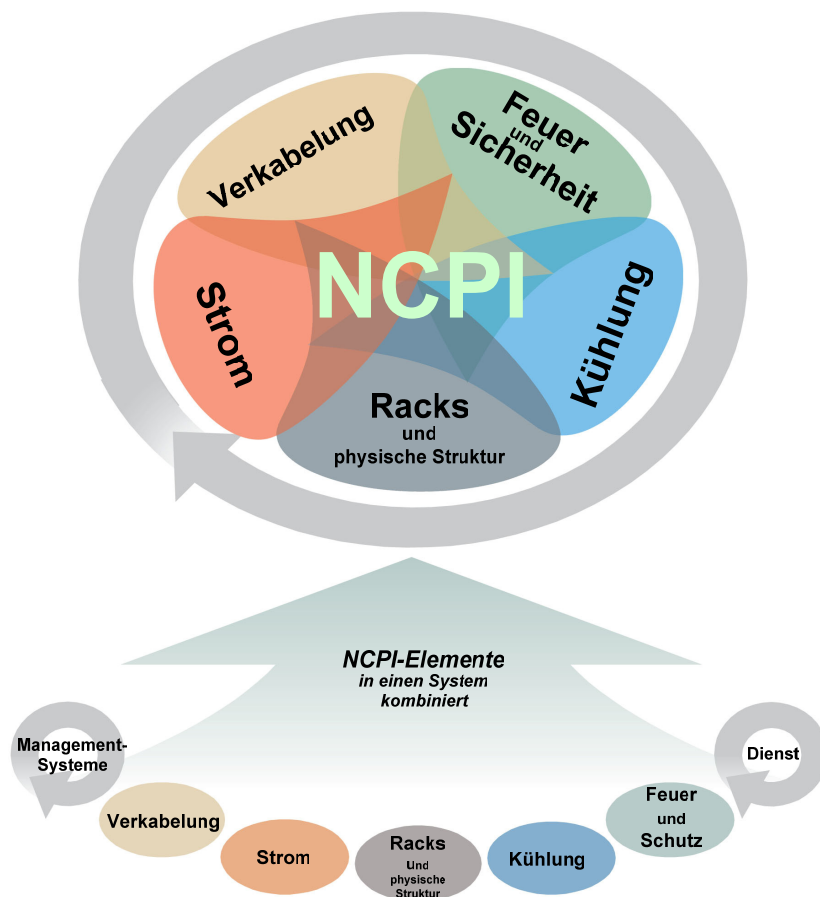
Zusammenfassung

Um mit den rasanten Veränderungen in der Geschäftswelt Schritt zu halten, müssen Unternehmen den Wert ihrer Investitionen in puncto NCPI (Netzwerk-kritische physische Infrastruktur) neu überdenken. Die Beurteilung von Verfügbarkeit und Up-Front-Kosten genügt nicht mehr, um geeignete Geschäftsentscheidungen zu treffen. Agilität oder Geschäftsflexibilität sowie geringe Gesamtbetriebskosten sind genauso wichtig für Unternehmen, die auf einem sich verändernden globalen Markt erfolgreich sein möchten.

Einleitung

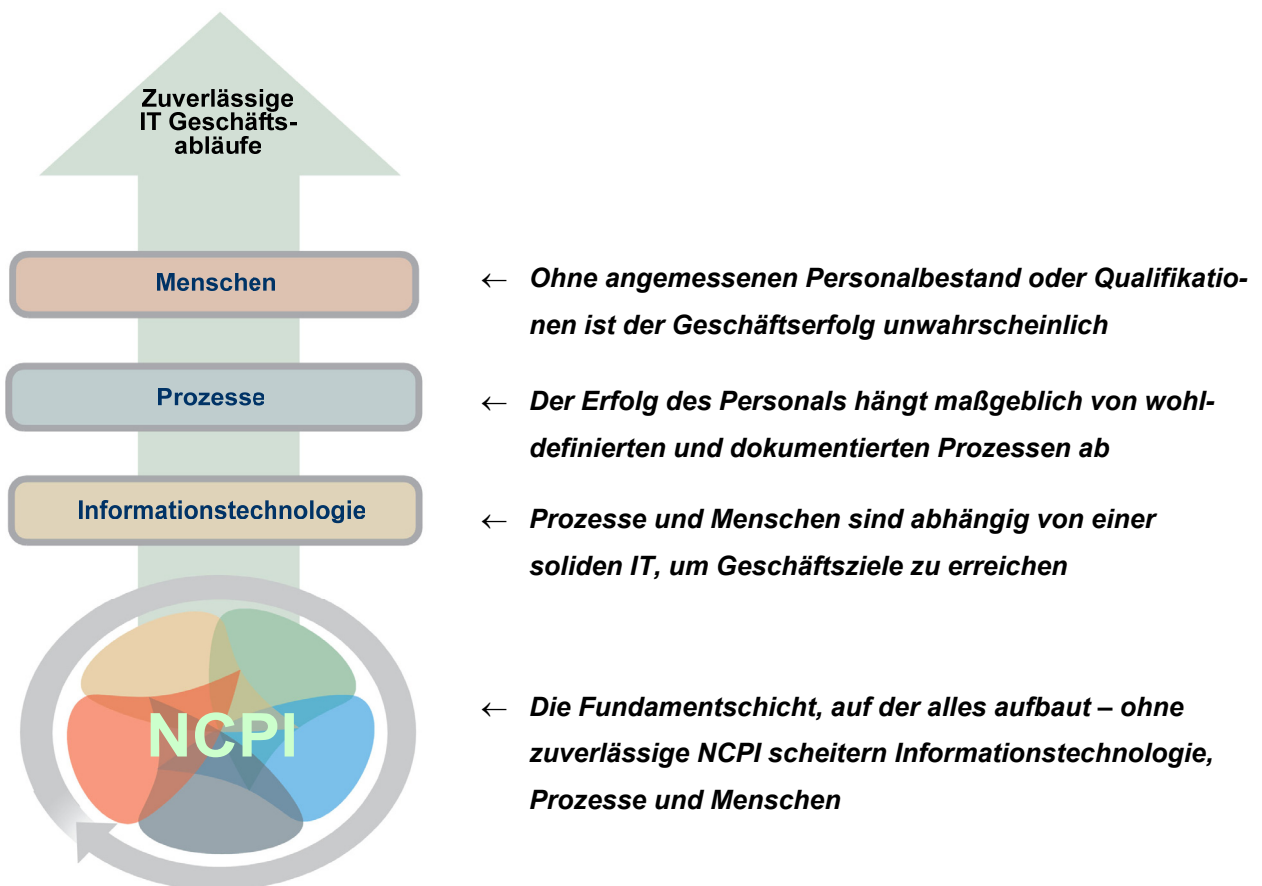
Die netzwerkkritische Infrastruktur (NCPI) ist das Fundament für die Informationstechnologie (IT) und Telekommunikationsnetzwerke. Es handelt sich dabei um das Rückgrat des Unternehmens, da sich hier Elemente wie Kühlung, Gehäusegestaltung, Sicherheit, Brandschutz und Verkabelung befinden, die eine funktionierende Informationstechnologie ermöglichen. **Abbildung 1** zeigt diese kritischen NCPI-Elemente und deren Einbindung in ein nahtloses End-to-End-System, das durch Verwaltungssysteme und Dienste unterstützt wird. Sieht man die NCPI nicht als einzelne Bausteine, sondern als Ganzes, ist es notwendig, ein integriertes und verständliches System zu entwickeln und einzusetzen, das wie erwartet funktioniert. Werden die einzelnen Elemente getrennt von anderen NCPI-Elementen angeschafft, ist das Endresultat meist ein komplexes und schwer einschätzbare NCPI-System, das aus vielen Produkten verschiedener Hersteller besteht, die nicht für die Zusammenarbeit untereinander konzipiert wurden. Da eine Fülle von Verwaltungssystemen benötigt wird, um dem gesamten System Transparenz zu verleihen, wird die Verwaltung immer komplexer. Zudem wird das Abschließen vieler Serviceverträge unumgänglich. **Anhang A** dieses White Papers liefert detailliertere Informationen zu jedem dieser NCPI-Elemente.

Abbildung 1 – Ein integriertes NCPI-System



Die meisten erfolgreichen Unternehmen von heute sind abhängig von einer soliden Informationstechnologieplattform. Um diese IT-Geschäftsabläufe verwalten und koordinieren zu können, müssen vier Schichten oder Bausteine vorhanden sein. **Abbildung 2** illustriert dieses Schichtenmodell und die Bedeutung eines integrierten NCPI-Systems als Fundament für die Aufrechterhaltung von Geschäftsfunktionen. Zusätzlich zur NCPI gehören hierzu die *Informationstechnologie*, *Prozesse* und *Menschen*, die die Funktion dieser Systeme unterstützen. Unter Informationstechnologie versteht man die Datenverarbeitung und -sicherung sowie Kommunikationssysteme, und zwar in Form von Hard- und Software. Ohne richtige Planung und Konzeption dieser Technologie können das Netzwerk und folglich auch das Unternehmen nicht funktionieren. Alle *Prozesse* für die Verarbeitung im Datacenter oder in einer IT-Umgebung müssen auf einfache und verständliche Weise klar definiert, gut dokumentiert und standardisiert sein. Ohne effektive Implementierung derartiger Prozesse sind Probleme beim Betrieb und bei der Wartung der Systeme unvermeidbar. Unerwartete Ausfallzeiten sind die zwangsläufige Folge. Außerdem müssen diese Abläufe von *Menschen* unterstützt werden. Dies wiederum erfordert qualifizierte Mitarbeiter mit geeignetem Ausbildungsstand. Ohne die richtige Planung in Sachen Personal und Qualifikation sind menschliche Fehler unvermeidbar.

Abbildung 2 – NCPI ist eine wichtige Schicht für zuverlässige IT-Geschäftsabläufe



Dieses White Paper erörtert eine notwendige Veränderung in der Denkweise der IT-Planer: Bei NCPI-Anlageentscheidungen muss der Geschäftswert konkreter einbezogen werden. Die steuernden Faktoren dieser neuen Wertkriterien werden hier vorgestellt.

Optimierung der NCPI für einen höheren Geschäftswert

Die Beurteilung des Geschäftswerts orientiert sich aus Sicht des Unternehmens im Allgemeinen an drei Kernzielen:

- Maximierung der Erträge
- Minimierung der Kosten
- Bessere Nutzung der Ressourcen

Ungeachtet der Branche führen diese drei Ziele letztendlich zu höheren Gewinnen und einem besseren Cashflow. NCPI-Investitionen werden getätigt, weil sie diese drei Geschäftsziele sowohl direkt als auch indirekt beeinflussen. Manager beschaffen Generatoren, Klimageräte, Sicherheits- und USV-Systeme, um diese als „Versicherungspolicen“ einzusetzen. Jedes Netzwerk oder Datacenter birgt Risiken von Ausfallzeiten durch Strom- und Temperaturprobleme in sich. Die Investition in NCPI minimiert sowohl diese Risiken als auch andere. Wie wirkt sich das auf die drei oben genannten Kerngeschäftsziele (Gewinne, Kosten und Ressourcen) aus? Bei Systemausfällen werden Ertragsströme gebremst oder gar unterbrochen, Geschäftsausgaben erhöht und die Ressourcen ineffektiv genutzt. Deshalb gilt: Je effektiver die NCPI etwaige Ausfallzeiten reduziert, desto größer der Wert für das Unternehmen bei der Erreichung aller drei Ziele.

Bisher stützte sich die Beurteilung des Wertes der NCPI für Unternehmen auf zwei Kernkriterien: Verfügbarkeit und Up-Front-Kosten. Eine verbesserte **Verfügbarkeit** (verfügbare Betriebszeit) des NCPI-Systems und somit bessere Geschäftsabläufe ermöglichen weiterhin fortlaufende Erträge und eine effektivere Nutzung (oder Produktivität) der Ressourcen. In einem Unternehmen beispielsweise, das Kreditkartenabrechnungen durchführt und dessen Systeme nicht funktionieren, können Einkäufe per Kreditkarte nicht mehr getätigt werden, wodurch der Ertragsstrom für die gesamte Ausfallzeit unterbrochen ist. Zusätzlich können Angestellte, deren Systeme nicht online sind, nicht produktiv arbeiten. Die Minimierung der **Up-Front-Kosten** der NCPI führt zu höheren Kapitalerträgen. Sind die NCPI-Kosten gering und die Risiken/Kosten der Ausfallzeit hoch, ist der Geschäftsfall einfacher zu justieren.

Während dies immer noch gilt, erfordern die sich schnell verändernden IT-Umgebungen zwei zusätzliche Kriterien für die Beurteilung des Wertes einer NCPI für ein Unternehmen. Zunächst müssen Geschäftspläne flexibel auf sich verändernde Marktsituationen, Chancen und Umweltfaktoren reagieren können. Investitionen, durch die Ressourcen gebunden werden, beschränken die Fähigkeit flexibel zu reagieren. Ohne diese Flexibilität oder **Agilität** werden viele Chancen verspielt.

Ein weiteres, zusätzliches und sehr wichtiges Wertkriterium für Unternehmen sind die **Gesamtbetriebskosten (TCO)**. Obwohl die Up-Front-Kosten immer noch ein sehr wichtiger Faktor sind, geben sie nicht alle Informationen wieder. Sie lassen den Entscheidungsträger, was die Langzeitkosten einer Lösung, inklusive Betriebs- und Wartungskosten angeht, im Dunkeln. Die Up-Front-Kosten sind besonders dann ein Kriterium von unverhältnismäßig hohem Stellenwert, wenn bei der Anschaffung von Kapitalgütern sehr konventionell ausgerichtete Projektkonzepte verfolgt werden. Investitionskosten wurden bislang in der Buchführung für ein Projekt aus Steuer- und Abschreibungsgründen oft von laufenden Kosten getrennt. Obwohl Up-Front-Kosten nur einen Bruchteil der TCO darstellen, reichten sie aus, um ein Projekt und damit verbundene Anschaffungen bewilligt zu bekommen. Dabei waren Dinge wie Stromrechnungen nie Teil der Entscheidungsfindung. Diese Art von Ausgaben wurden oft als „übliche“ Grundaussgaben abgetan, die einfach nur ein notwendiges Übel darstellten – sie waren einfach eine Begleiterscheinung des Projekts und wurden nicht von Projektgeldern gedeckt, sondern mit Betriebsmitteln bezahlt. Nun erkennen Entscheidungsträger bei der Abwägung unternehmenskritischer Entscheidungen die Bedeutung dieser „sonstigen“ Kosten. Eine durchdachte Investitionsentscheidung berücksichtigt sowohl Anlaufkosten als auch laufende Betriebsausgaben.

Abbildung 3 zeigt den Sinneswandel hinsichtlich der Kriterien für die Einschätzung des Wertes der NCPI für Unternehmen ausgehend von sich langsam verändernden bis hin zu den sich schnell verändernden Unternehmen von heute. Diese Gleichung liefert keine „quantitativen Werte“. Sie unterstreicht vielmehr die Faktoren, die für die Erreichung eines hohen Geschäftswerts wichtig sind. Verfügbarkeit und Agilität, die den Zähler der Gleichung bilden, müssen maximiert werden, um den Geschäftswert zu erhöhen. Die Gesamtbetriebskosten, die hier den Nenner oder den unteren Teil der Gleichung wiedergeben, müssen verringert werden, um den Geschäftswert zu erhöhen.

Dieses Dokument soll Betreiber von Datacenter dazu veranlassen, die Begriffe Verfügbarkeit und Kosten neu zu bewerten und einen neuen Performancevektor, nämlich die Agilität, zu berücksichtigen, der in der Geschäftswelt von heute nicht mehr fehlen darf. Sämtliche Performancevektoren werden letztlich in Geldwerte übertragen. Die Überlegung ist daher notwendig, wie die Berechnung „(Verfügbarkeit x Agilität) / Gesamtbetriebskosten“ zu optimieren ist.

Abbildung 3 – Sinneswandel bei der Bewertung der NCPI für Unternehmen

$$\text{Wert} = \frac{\text{Verfügbarkeit}}{\text{Vorab-Kosten}}$$

Der Sinneswandel

$$\text{Wert} = \frac{\text{Verfügbarkeit} \times \text{Flexibilität}}{\text{Gesamtbetriebskosten (TCO)}}$$

Bei der Betrachtung der neuen „Geschäftswertgleichung“ für NCPI sind für jedes der drei Kriterien zahlreiche Parameter zu berücksichtigen. Die folgenden Abschnitte behandeln die Parameter, die den größten Einfluss auf den gesamten Geschäftswert haben.

VERFÜGBARKEIT – Performancevektor Nr. 1

Wie bereits erwähnt, ist die Verfügbarkeit einer der Schlüsselfaktoren für den NCPI-Geschäftswert. Dies ist der Hauptgrund, dass in NCPI investiert wird. Wenn Systeme nicht verfügbar sind, sind wichtige Geschäftsziele gefährdet.

$$\text{Wert} = \frac{\text{Verfügbarkeit} \times \text{Agilität}}{\text{TCO}}$$

Der Begriff Verfügbarkeit wird von verschiedenen Menschen verschieden interpretiert. Manche verbinden damit einfach nur die Zuverlässigkeit der Systemkomponenten. Aber es geht hier um viel mehr. Die technische Definition von Verfügbarkeit lautet:

Verfügbarkeit ist der Grad, in dem ein System oder eine Komponente funktionsfähig und betrieblich nutzbar ist, wenn es bzw. sie verwendet werden soll [IEEE 90].

Die Zuverlässigkeit ist sicherlich eine wichtige Größe für die Betriebsbereitschaft der Systeme und Komponenten, jedoch spielen andere Faktoren wie die mittlere Wiederherstellungszeit (Mean Time to Recover, MTTR) und menschliche Faktoren ebenfalls eine wichtige Rolle. Systeme fallen aus verschiedenen Gründen aus, sowohl geplant als auch ungeplant. Ein Beispiel für eine geplante Ausfallzeit ist die regelmäßige Wartung eines Systems. Unter den Ursachen einer ungeplanten Ausfallzeit findet man menschliches Versagen und Naturkatastrophen. Überraschenderweise sind die Ausfallzeiten aufgrund von Naturkatastrophen wesentlich kürzer als die, die im täglichen Betrieb durch menschliches Versagen verursacht werden. Nach Einschätzung von Branchenexperten machen die Ausfallzeiten, die durch menschliches Versagen verursacht werden, 40 - 60 % aller Ausfallzeiten aus.

Der Bedarf an Verfügbarkeit wird auf verschiedene Arten ausgedrückt, dabei sind manche mehr quantitativ orientiert als andere. Gemäß der technischen Definition wird die Verfügbarkeit als Prozentanteil der Zeit gemessen, während der die Systeme betriebsbereit sind, z. B. 99,99 %. Viele Betreiber von Datacentern streben die „fünf Neuner“ an. Andere klassifizieren die Verfügbarkeit über die Ausfallzeit. Zum Beispiel steht 99,999 % für fünf Minuten Ausfallzeit pro Jahr. Eine dritte Möglichkeit, die Verfügbarkeit zu beschreiben, ist die Gliederung in Ebenen (engl. Tiers), wobei auf jeder Ebene eigene Bedingungen für die Redundanz und die benötigten Lösungen gelten. Das Uptime Institute ist beispielsweise eine Organisation, die einen Satz solcher Verfügbarkeitsebenen definiert hat.

Wenn die Verfügbarkeit der Hauptgrund für NCPI-Investitionen ist, wie entscheidet man, welche Verfügbarkeit gefordert ist? Natürlich wünschen sich alle Betreiber eines Datacenters die höchstmögliche Verfügbarkeit. In der Praxis bestimmen jedoch die Kosten der Lösung, wo die angestrebte Marke gesetzt wird. Unternehmen arbeiten oft an konkurrierenden Projekten und mit begrenzten Budgets. Durch eine NCPI-Lösung, die für eine hohe **Zuverlässigkeit** entwickelt wurde, **menschliches Versagen** verhindert und kürzeste **Reparaturzeiten** zulässt, kann höchstmögliche Verfügbarkeit erreicht werden. Diese drei Schlüsselfaktoren für eine hochverfügbare NCPI (Zuverlässigkeit der Systemkomponenten, Reparaturzeit, menschliches Versagen) werden im Folgenden näher erörtert.

Zuverlässigkeit der Systemkomponenten

Unter Zuverlässigkeit versteht man die Wahrscheinlichkeit, mit der ein Gerät, ein System oder Prozess seine zugewiesene Funktion fehlerlos über eine bestimmte Zeit hinweg ausübt. „Zeit“ ist in dieser Definition eine wichtige Komponente. Zuverlässigkeit wird oft damit verwechselt, dass „ein Produkt X zu 98,5 % zuverlässig“ ist. Ohne einen Zeitrahmen anzugeben, in dem diese Zuverlässigkeit gilt, ist eine solche Aussage bedeutungslos. Die richtige Aussage müsste lauten: „Produkt X wird mit einer Wahrscheinlichkeit von 98,5 % drei Jahre lang fehlerlos arbeiten“. Manchmal wird die Zuverlässigkeit auch als Unzuverlässigkeit oder in Form der Fehlerwahrscheinlichkeit wiedergegeben – im Prinzip die gegenteilige Variante. Das Unzuverlässigkeitsäquivalent zu der oben genannten Aussage müsste dann lauten: „Die Fehlerwahrscheinlichkeit für drei Jahre liegt bei 1,5 %.“

Die physische Infrastruktur von Datacentern umfasst viele Komponenten. Damit ein System als Ganzes zuverlässig ist, müssen alle Teile dieses Systems zuverlässig sein. Betreiber von Datacentern wollen nicht mit Ausfällen von Komponenten konfrontiert werden, ganz gleich wie schnell diese wiederhergestellt werden können. Ausfälle sind auch in puncto Mannstunden und Geld eine kostspielige Angelegenheit. Je zuverlässiger die Komponenten sind, desto unwahrscheinlicher ist das Auftreten derartiger Probleme.

Mittlere Reparaturdauer (Mean Time to Recover, MTTR)

Die mittlere Reparaturdauer spielt bei der Verfügbarkeit eines wiederherstellbaren Systems eine wichtige Rolle. Datacenter müssen so konzipiert sein, dass sie hinsichtlich ihrer Systemkomponenten so zuverlässig sind wie möglich. Bei Datacentern jedoch, die bereits 10 bis 15 Jahre in Betrieb sind, ist ein Ausfall von Systemen oder Subsystemen in diesem Zeitrahmen unvermeidbar. Je nach Redundanz des Systems müssen diese Ausfälle nicht unbedingt Ausfallzeiten generieren. Ungeachtet dessen, ist es wichtig, bei einem Ausfall so wenig Zeit wie möglich in die Reparatur zu investieren. Das bedeutet: Die Systemdiagnose benötigt nur wenig Zeit, Ersatzkomponenten stehen bereit und das System kann einfach repariert oder ersetzt werden. Viele haben von Horrorgeschichten gehört, in denen die Reparatur ausgefallener Datacenter Tage gedauert hat, oder solche Vorfälle bereits miterlebt. Die Folgen für ein Unternehmen sind verheerend.

Menschliches Versagen

Wenn es um die Verfügbarkeit von NCPI geht, spielt menschliches Versagen eine große Rolle. Wie vorhin schon erwähnt, zeigen verschiedene Studien, wie menschliches Versagen in 60 % der Fälle der Hauptgrund unvorhergesehener Ausfallzeiten ist. Während der gesamten Lebensdauer eines Datacenters wird das IT-System meistens viermal oder öfter ausgetauscht. Dabei ändern sich oft die Anforderungen an Stromversorgung, Kühlung und Sicherheit. Beispielsweise kann ein neuer Server mit einer anderen Anschlussbuchse ausgestattet sein, sodass Änderungen am Stromversorgungskabel erforderlich sind. Diese instabile Umgebung, die regelmäßigen Änderungen unterliegt, stiftet Chaos und ist sehr anfällig für menschliches Versagen.

Im Allgemeinen bieten Hersteller ihren Kunden oder potenziellen Kunden Informationen über Verfügbarkeit und Zuverlässigkeit, vernachlässigen jedoch den Faktor Mensch. Diese Fakten sind irreführend, wenn die Fehler in den eigentlichen Systemkomponenten nur einen derart geringen Teil aller Ausfälle bedingen und der Anteil an menschlichem Versagen so groß ist. Sinnvollere Angaben über die Verfügbarkeit eines NCPI-Systems enthielten auch diese menschlichen Fehler. Durch die Berücksichtigung dieser Fehler in den Angaben müssten Hersteller ihre Systeme weniger komplex gestalten, einfache sowie intuitive Schnittstellen einbinden und die Möglichkeit vorsehen, mühelos auf regelmäßige Veränderungen zu reagieren. Letztendlich würde dadurch der Geschäftswert gesteigert.

Checkliste für Verfügbarkeitsaspekte

Tabelle 1 zeigt oft übersehene Verfügbarkeitsaspekte, die einen großen Einfluss auf den Geschäftswert besitzen. In der Vergangenheit waren viele davon einfach nicht realisierbar. Heute gibt es jedoch keinen Grund mehr, diese Fragen mit „Nein“ zu beantworten.

Tabelle 1 – Liste der Verfügbarkeitsaspekte

Verfügbarkeitsaspekte
Wurden die Komponenten jedes NCPI-Bestandteils in hoher Stückzahl hergestellt, um die Zuverlässigkeit des Equipments zu verbessern?
Wurde in der NCPI auf Redundanz geachtet, um die Wahrscheinlichkeit eines Ausfalls des Datacenters / IT-Systems zu verringern?
Kann ein Fehler im Ernstfall in weniger als einer Stunde behoben werden?
Wurden die Elemente des NCPI-Systems für nahtlose Interoperabilität konzipiert?
Wurde das System mit Blick auf geringste Komplexität entwickelt?
Bietet das System intuitive und einfache Schnittstellen sowie proaktives Management?
Hat das System die Fähigkeit, mühelos auf regelmäßige Veränderungen zu reagieren, ohne menschliches Versagen zu fördern?

AGILITÄT – Performancevektor Nr. 2

Der Begriff Agilität steht für die Fähigkeit eines Systems, sich an eine Veränderung anzupassen. Sich an Veränderungen anpassen kann vieles bedeuten – die termingemäße Installation, besonders bei zeitlich begrenzten Projekten (**Einsatzgeschwindigkeit**). Ebenso kann die **Skalierbarkeit** gemeint sein, mit der auf veränderte Unternehmensgrößen reagiert und eine Überdimensionierung möglichst ausgeschlossen wird, oder die Fähigkeit, sich schnell auf die immer größer werdende Zahl an Veränderungen der Unternehmens-IT anzupassen (**Rekonfigurierbarkeit**). Klassische Beispiele für Unternehmen mit mangelnder Agilität und daraus folgendem Scheitern sind unter den Co-Location-Unternehmen zu finden. Diese Unternehmen hatten hohe Investitionen getätigt, um solide, hoch sichere Infrastrukturen zu entwickeln. Man *glaubte*, damit den Bedarf potenzieller Kunden für das Hosting kritischer IT-Komponenten zu treffen. Da sich die Systeme nicht an veränderte Geschäftsanforderungen anpassen konnten, setzte man für den „worse case“ auf Kapazität. Konkrete Planung wurde durch Bauchgefühl ersetzt. Ein großer Überhang an nicht verwendeter Infrastruktur und leere Auftragsbücher waren die Folge.

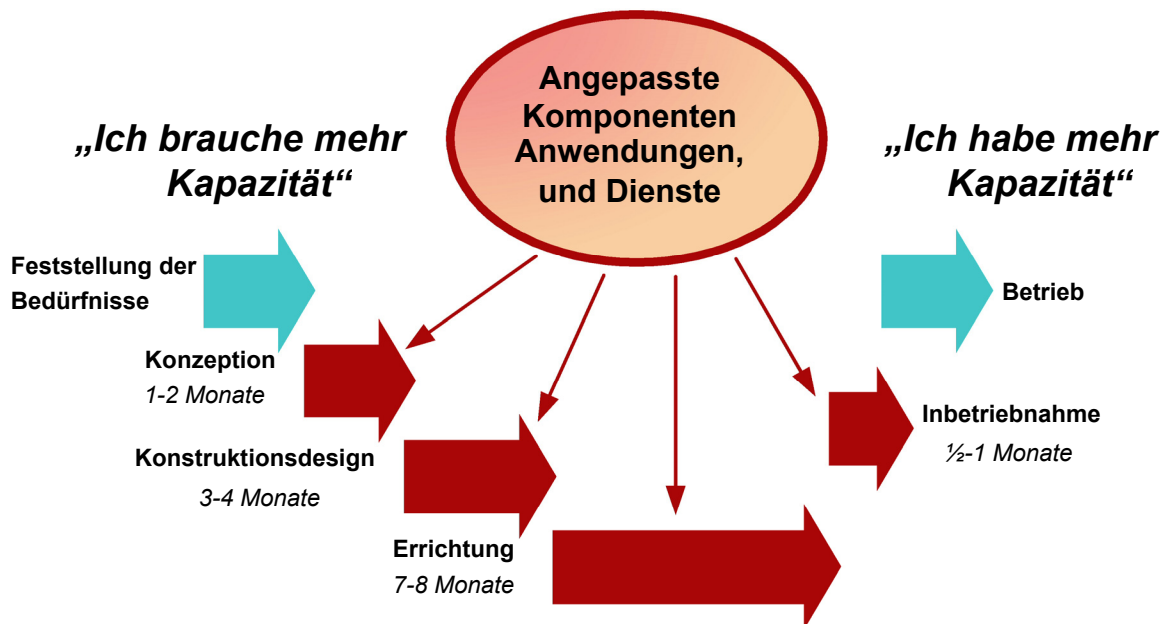
$$\text{Wert} = \frac{\text{Verfügbarkeit} \times \text{Agilität}}{\text{TCO}}$$

Agilität ist ein entscheidendes Teil in dem Geschäftswertpuzzle – hinsichtlich der endgültigen Investitionen genauso wichtig wie alle anderen Bestandteile des Geschäftswerts. Wenn die NCPI-Investition für die momentane Situation getätigt werden kann und agil genug ist, zukünftige Bedürfnisse abzudecken (anstatt die nächsten zehn Jahre vorherzusagen), entfällt das Risiko falscher Vorhersagen und die Kapitalerträge werden optimiert.

Schnelligkeit der Bereitstellung

Die Schnelligkeit der Bereitstellung gibt an, wie schnell das NCPI-System konzipiert, entwickelt, installiert und in Betrieb genommen werden kann. In der Vergangenheit handelte es sich hier um Jahre. **Abbildung 4** zeigt die typische Zeitleiste solcher Projekte. Die sehr lange Zeitleiste ist auf die Tatsache zurückzuführen, dass konventionelle Ansätze ein hohes Maß an Anpassung mit sich brachten. Um mehr Kapazität zu erreichen, benötigte man größere Schritte der Anpassung – den einmaligen Entwurf, die einmalige Konstruktion, den eigentlichen maßgeschneiderten Konstruktionsprozess sowie die Inbetriebnahme, die darauf abzielte, viele Systemkomponenten zu integrieren, die gar nicht dafür ausgelegt waren.

Abbildung 4 – Typische Zeitleiste von 400 Tagen für den Einsatz von NCPI-Legacy-Systemen



Mit den rasanten Unternehmensanforderungen von heute ist diese starre Zeitleiste bestenfalls unzureichend, schlimmstenfalls jedoch katastrophal. Unternehmen benötigen Datacenter, die innerhalb weniger Monate errichtet werden.

Betreiber von Datacenter kennen weitere damit verbundene Anekdoten, und zwar die Projektverzögerungen durch wiederholte Versuche, die Zukunft vorauszusagen. Projektzeiten können verkürzt werden, indem die Zeiten und Ressourcen für die Kapazitäts- und Dichteplanung entfernt werden, bevor ein derartiger Bedarf aufkommt.

Skalierbarkeit

Wenn Betriebsmittel nicht frei fließen (was normalerweise der Fall ist), ist die Skalierbarkeit der Schlüssel für die Agilität. Ein durchschnittliches Datacenter ist deutlich überdimensioniert. Untersuchungen zeigen, dass typische Datacenter von heute weniger als 50 % ihrer Infrastruktur nutzen. Gemäß „Data Center Power Requirements: Measurements from Silicon Valley“, von J.D. Mitchell-Jackson, u. a., werden Datacenter in der Tat *nur zu einem Drittel* genutzt. 66 % sind keine geringe Verschwendung. Die Kapitalrendite für dieses Projekt könnte dreimal höher sein. Zwei Drittel der Investitionen könnten anderen Unternehmensprojekten mit größerem Renditepotenzial zugewiesen werden.

Betreiber von Datacentern haben nicht überbaut, weil sie Ressourcen vergeuden wollten. Sie haben überbaut, weil sie, auf der Analyse zukünftiger Geschäftsanforderungen bauend, vom „worse case“ in Sachen Kapazität ausgegangen sind, ohne mehr als einen Zeitraum von drei bis vier Jahren überblicken zu können. Bis vor kurzem war die Alternative dazu sogar negativer – zu unterbauen, ohne die Möglichkeit zu besitzen, die Kapazität unterbrechungsfrei und mit geringen Kosten zu erhöhen. APC White Paper Nr. 37, „Überdimensionierte Datacenter und Serverraum-Installationen: Kostenvermeidung“ behandelt die Bedeutung des Rightsizing im Detail.

Rekonfigurierbarkeit

Co-Location-Unternehmen, die 50 % ihrer Einrichtung mit Gleichstrom und 50 % mit Wechselstrom ausgebaut haben, erfahren später, dass 99 % ihrer Kunden Wechselstrom benötigen. Die Gleichstrominvestition war aufgrund der mangelnden Agilität nahezu ein Totalverlust. Was wäre, wenn dasselbe Co-Location-Unternehmen nur das bereitgestellt hätte, was seine Kunden benötigen, und nur im Bedarfsfall für verschiedene Stromdichten, Redundanzebenen (gezielte Verfügbarkeit) und verschiedene Spannungen sowie Anschlussstypen gesorgt hätten, ohne das Datacenter abreißen zu müssen?

Man stelle sich ein anderes Unternehmen vor, das ein Datacenter gebaut hat, nur um nach der Inbetriebnahme festzustellen, dass diese Einrichtung an einen anderen Ort verlegt werden muss. Die Mehrheit der eingesetzten Ressourcen kann bei Legacy-Systemen nicht verlagert werden, was letztendlich einen großen finanziellen Einfluss auf das Unternehmen hat, da eine beträchtliche Zweitinvestition für den neuen Standort getätigt werden muss. Portable, umsetzbare Systeme können den Gewinn bedeutend maximieren.

Checkliste für Agilitätsaspekte

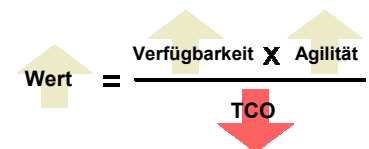
Tabelle 2 zeigt oft übersehene Agilitätsaspekte, die einen großen Einfluss auf den Geschäftswert besitzen. Wie bei den Verfügbarkeitsaspekten, können diese Punkte in einem erfolgreichen NCPI-System erwartet werden. Bei den rasanten Veränderungen der Geschäftsbedingungen von heute geht der Geschäftserfolg mit der dezidierten Verfolgung dieser Agilitätsaspekte einher.

Tabelle 2 – Liste der Agilitätsaspekte

Agilitätsaspekte
Benötigt der Bau eines neuen Datacenters Tage anstatt Monate und Jahre?
Kann die Kapazität des bestehenden NCPI-Systems schnell erhöht werden? Kann sie im nächsten Monat verdoppelt werden?
Kann die NCPI des Datacenters mit einem minimalen Bauaufwand installiert werden?
Kann im Bedarfsfall ein großer Teil der Infrastruktur an einen anderen Ort verlagert werden?
Ist es möglich, das Datacenter innerhalb von Monaten zu verlagern?
Kann ein Teil des Datacenters mit redundanter NCPI (gezielte Verfügbarkeit) ausgestattet werden?
Kann ein Austausch der Anschlussstypen innerhalb von Minuten während einer Modernisierung der IT erfolgen?
Kann je nach Bedarf die Backup-Laufzeit skaliert werden?

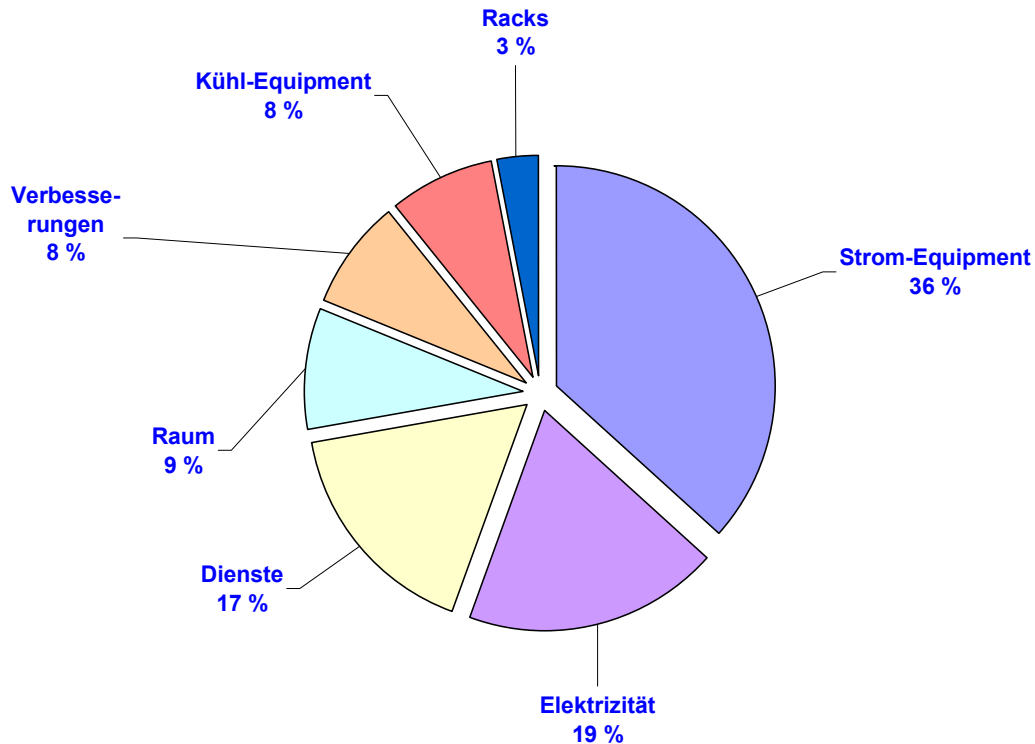
Gesamtbetriebskosten – Performancevektor Nr. 3

Die Gesamtbetriebskosten (TCO) sind für die meisten Manager klar – es handelt sich hier um real ausgegebenes Geld. Die Gesamtbetriebskosten eines Datacenters hängen von der Größe der Einrichtung ab. Es *ist* jedoch möglich, die gesamten Kosten auf eine für alle Größen von Datacenters brauchbare Weise auszudrücken. Ein guter Weg, die Anforderungen an die Gesamtbetriebskosten darzustellen, ist die Angabe der Gesamtkosten je Rack über die Lebensdauer des Datacenters. Ein typischer Wert für NCPI-Kosten je Rack über die Gesamtlebensdauer ist 100.000 USD. Das ist eine beträchtliche Investition (vergleichbar mit den Investitionskosten für IT-Equipment) und deshalb ist es für den Betreiber unabdingbar zu verstehen, welchen Wert er für die Ausgaben erhält.



Die Gesamtbetriebskosten werden nicht einfach nur durch die Anfangsinvestitionen für NCPI gebildet, sondern enthalten alle anderen Kosten für den Betrieb und die Wartung dieser Investition für die gesamte Lebensdauer (normalerweise 10 - 15 Jahre). Es sollte klar sein, dass in der Geschäftswelt von heute die Gesamtbetriebskosten für den Geschäftswert ausschlaggebend sind und nicht nur alleine die Up-Front-Kosten. **Abbildung 5** zeigt eine typische Aufgliederung der Kosten für Datacenter mit herkömmlichen Systemen.

Abbildung 5 – Aufgliederung der Gesamtbetriebskosten für herkömmliche Systeme



Ungefähr 50 % dieser Kosten sind Betriebsausgaben, 50 % sind Kapitalkosten. Von diesem Gesamtwert wird ein großer Anteil (grob 30 %) aufgrund mangelhafter Planung und Entscheidungen vergeudet.

Das „Oversizing“ ist der Hauptgrund für das Sprengen der Gesamtbetriebskosten. Daraus ergibt sich eine beträchtliche Vergeudung, u. a. von **Kapitalkosten**, **Betriebskosten** und besonders von **Energiekosten**. Durchschnittlich werden über 50 % der eingesetzten NCPI-Kapazität vergeudet. Während diese Vergeudung in erster Linie durch Unsicherheit entsteht, kann die Antwort darauf nicht einfach das Verkleinern von Konzepten sein – es gibt sicherlich Fälle, in welchen die wirkliche Last der erwarteten gleichkommt oder diese übersteigt, wobei die Datacenter eine NCPI-Lösung einsetzen, die skalierbar ist, um somit zu jeder Zeit den Anforderungen gerecht zu werden. Niemand möchte für ein Datacenter mit einer Kapazität von einem Megawatt verantwortlich sein, nur um zu erfahren, dass die wirkliche Auslastung die 300-kW-Marke nie überschritten hat. Das APC White Paper Nr. 6, „Ermitteln der Total Cost of Ownership (TCO) für Datacenter- und Serverraum-Infrastrukturen“, beschreibt eine Methode für die Messung der Gesamtbetriebskosten und schlägt Strategien vor, um diese Kosten zu minimieren.

Kapitalkosten

Der Begriff Kapitalkosten, auch CAPEX genannt, bezeichnet das Geld, das von einem Unternehmen für den Erwerb oder die Modernisierung von Immobilien, Fertigungsanlagen und Systemkomponenten ausgegeben wird. Damit verbindet sich die Erwartung, dass das Unternehmen über einen langen Zeitraum hinweg (über mindestens ein Jahr) einen Profit erzielt. Diese Kosten werden über die gesamte Lebensdauer der Anlagegüter abgeschrieben und unterliegen besonderen steuerrechtlichen Behandlungen. Die Kapitalkosten für ein NCPI-Projekt sind sehr reale und messbare Ausgaben. Darunter versteht man sowohl die Kosten für das NCPI-Equipment als auch die Kosten, die mit der Konzeption und dem Einsatz dieses Equipments verbunden sind. Die beste Möglichkeit, Kapitalkosten zu optimieren (reduzieren) ist die Skalierbarkeit der Infrastruktur. Die Anpassung der Kapazität an die Last verhindert eine erhebliche Überinvestition, die im Bereich bis zu 300 % liegen kann. Eine andere Möglichkeit, die Kapitalkosten zu senken, ist die Verringerung der benötigten technischen Arbeit. Die Implementierung einer Lösung, die Arbeiten vor Ort während der Installation minimiert, drückt diese vorausgehenden Arbeitskosten.

Betriebskosten

Unter Betriebskosten, auch OPEX genannt, versteht man sämtliche mit dem laufenden Betrieb eines Unternehmens verbundene Kosten. Die Betriebskosten für eine NCPI-Einrichtung umfassen das NCPI-Betriebspersonal, Schulungsausgaben sowie Wartungs- und Reparaturkosten. Energiekosten sind ebenfalls Betriebskosten. Diese müssen jedoch gesondert dargestellt werden, um den bedeutenden Einfluss auf den Geschäftswert darzustellen.

Es ist wichtig, den Grad der prädiktiven Wartung in diesem System zu beachten, da dadurch der Bedarf an programmierter Wartung reduziert werden kann, was nicht nur die Kosten senkt, sondern auch das Fehlerisiko minimiert. Der Umfang der benötigten Dienste hängt auch zum Großteil von der Komplexität und der Anpassung des Systems ab. Umfang und Art der Reparaturen werden gleichzeitig durch die Beschaffenheit des Systems bestimmt. Modulare Systeme ermöglichen eine bessere Selbstwartung und den Austausch fehlerhafter Komponenten, während angepasste Legacy-Systeme mehr Reparaturingriffe vor Ort erfordern.

Je komplexer und kundenspezifischer ein System ist, desto umfangreicher die Arbeiten für Konstruktion und Standortvorbereitung sowie letztendlich der Bedarf an technischem Personal für die Systembetreuung. Das bedeutet im Allgemeinen höhere Ausgaben für die Schulungen der Angestellten, damit die Systeme erfolgreich betrieben werden können.

Energiekosten

Das größte Potenzial für die Einsparung von Gesamtbetriebskosten findet man in Form von Effizienzkosten. Ein überdimensioniertes USV-System ist weniger effizient und benötigt wesentlich mehr Energie für den Betrieb als ein genau auf die Last abgestimmtes System. Grund dafür ist, dass der Wirkungsgrad steigt, sobald sich die Ist-Last der projektierten Last nähert.

In redundanten USV-Systemen hat auch die Größe der redundanten Module einen großen Einfluss auf den Wirkungsgrad. Dies ist darauf zurückzuführen, dass bei einer typischen N+1 Konfiguration die Last auf alle Module verteilt wird. Betrachten wir ein Datacenter mit einer Last von 80 kW. Bei einem USV-System, das sich aus vier 20-kW-Systemen für die Last und einem weiteren 20-kW-System für die Redundanz zusammensetzt, würden die 80 kW auf alle fünf Systeme verteilt werden. Anders ausgedrückt würde jede USV 16 kW oder 80 % der Kapazität tragen. Dieses System könnte auch mit zwei 80-kW-Systemen arbeiten, wobei jedes mit 50 % der Kapazität laufen würde. In diesem Beispiel hat das erste Konzept einen höheren Wirkungsgrad und ermöglicht somit geringere Energiekosten.

Die gewählten NCPI-Technologien können einen großen Einfluss auf die Energiekosten ausüben. Herkömmliche USV-Systeme haben zum Beispiel nur einen Wirkungsgrad von 85 % bei 50 % Last. Neuere Online-Technologien, wie die USV-Systeme mit Deltawandlung, erzielen bei diesen geringen Belastungswerten einen Wirkungsgrad von 93 % oder mehr. Mit Stromkosten von gut 0,10 USD/kWh in den meisten Gegenden der USA trägt dieser hohe Wirkungsgrad in hohem Maß zur Verringerung der Gesamtbetriebskosten bei.

Checkliste für Aspekte der Gesamtbetriebskosten

Tabelle 3 zeigt oft übersehene Aspekte der Gesamtbetriebskosten, die einen großen Einfluss auf den Geschäftswert ausüben. Wie bei den Verfügbarkeits- und Agilitätsaspekten sind diese Punkte für richtige Investitionsentscheidungen in Sachen NCPI essenziell.

Tabelle 3 – Aspekte der Gesamtbetriebskosten

Aspekte der Gesamtbetriebskosten
Wird dieses Projekt durch Gesamtbetriebskosten oder nur durch Up-Front-Kosten bewertet?
Besitzt die NCPI die richtige Kapazität, um die Kapitalinvestitionen zu optimieren?
Ist der elektrische Wirkungsgrad für das Datacenter durch passende Skalierung optimiert?
Sind im Konzept hoch effiziente NCPI-Komponenten berücksichtigt?
Sind Langzeit-Wartungskosten für das Datacenter minimiert?

Strategie für die Optimierung des Geschäftswerts

In der heutigen Geschäftswelt müssen bestimmte zeitgemäße Kriterien erfüllt werden, um den Geschäftswert zu optimieren. **Abbildung 6** zeigt, worum es sich dabei handelt. Diese Kriterien drücken sich wieder in realen Kosten aus, die das Unternehmen in verschiedenster Weise betreffen – nicht nur rein finanziell, sondern auch im Hinblick auf Strategie und Planung.

Abbildung 6 – Die neue Geschäftswertgleichung

$$\text{Wert} = \frac{\text{Verfügbarkeit} \times \text{Agilität}}{\text{TCO}}$$

Mit NCPI-Legacy-Systemen – mit ihrer Einmal-Konzeption sowie ihrem unflexiblen, üblicherweise überdimensionierten Design – konnte kein System aufgebaut werden, das alle drei Performancevektoren dieser Gleichung optimiert. Es war nicht machbar, ein qualitativ hochwertiges System (Verfügbarkeit) innerhalb einer vernünftigen Zeitspanne (Agilität) zu vernünftigen TCO-Kosten (Gesamtbetriebskosten) zu erhalten. Im Allgemeinen musste man immer eine dieser Variablen vernachlässigen, um die anderen beiden zu optimieren. Wollte man es schnell und günstig, so opferte man die Qualität. Wollte man es qualitativ hochwertig und schnell, konnte man sich das sehr wahrscheinlich nicht leisten. Dieses Paradigma gilt glücklicherweise heute nicht mehr.

Es gibt eine bewährte Strategie für die Optimierung aller drei Geschäftsanforderungen – durch Standardisierung, oder noch spezieller, durch *modulare Standardisierung*. Bei Anwendung derselben Ansätze wie in der Automobilentwicklung zum Beispiel kann der Geschäftswert von NCPI erhöht werden. Betrachtet man die Server der Vergangenheit, passte jeder Server in ein anderes Rack, was zu Inkompatibilitäten geführt hat. Seit die Server standardisiert wurden, passen die Server in nahezu jedes Rack. Das ist nur eines von vielen Beispielen dafür, in welche Richtung die NCPI-Industrie geht. Das APC White Paper Nr. 116, „Standardisierung und Modularität in der netzwerkkritischen physischen Infrastruktur“, beschreibt die breit gefächerten Vorteile für eine NCPI-Investition durch Standardisierung und Modularität.

Ergebnisse

Die NCPI ist die Basis eines Datacenters. Die Veränderungen in der heutigen Geschäftswelt, besonders in den verkürzten Modernisierungszyklen der IT, wecken das Bedürfnis nach einer veränderten Denkweise hinsichtlich des Geschäftswerts der NCPI. Die in diesem Dokument beschriebene neue NCPI-Geschäftswertgleichung ist hier zusammengefasst. Die Verfügbarkeit muss neu begriffen werden, auch unter Berücksichtigung von Schlüsselfaktoren wie menschliches Versagen. Der Begriff „Kosten“ muss ebenfalls neu definiert werden: Up-Front-Kosten alleine sind nicht länger ausreichend für fundierte Geschäftsentscheidungen – was wirklich den Wert bestimmt, ist die Gesamtheit der Kosten über die Lebensdauer der Infrastruktur. Agilität ist zudem ein Performancevektor, der beachtet werden muss, weil er direkt die Fähigkeit der NCPI misst, auf unvorhersehbare Anforderungen und Möglichkeiten zu reagieren.

Der Schlüssel zu einem hohen Geschäftswert unter diesem neuen Paradigma sind Standardisierung, Modularität und Skalierbarkeit. Die NCPI-Branche muss sich von ihrer tradierten Haltung lösen, das zwei verschiedene Einrichtungen zwei vollkommen verschiedene Konzeptionen besitzen und mit zwei verschiedenen Arten von Problemen konfrontiert sind. Systeme müssen modular und skalierbar sein, damit sie den sich ändernden Anforderungen gerecht werden, die Zuverlässigkeit und Verfügbarkeit der Infrastruktur erhöhen und die Gesamtbetriebskosten optimieren.

Die Autorin:

Wendy Torell ist Verfügbarkeitsingenieurin bei APC in W. Kingston, RI. Sie erläutert Kunden die Vorgehensweise der Verfügbarkeitslehre und Konstruktionsverfahren, um die Verfügbarkeit der Datacenterumgebungen der Kunden zu optimieren. Ihren Bachelor-Abschluss in Maschinenbau hat sie am Union College in Schenectady, NY, gemacht. Frau Torell ist eine ASQ-zertifizierte Zuverlässigkeitsingenieurin.

Anhang: Elemente einer NCPI

Eine NCPI umfasst Stromversorgung, Kühlung, Racks und physische Struktur, Sicherheit und Feuerschutz, Verkabelung sowie die Verwaltung und Wartung dieser Komponenten. Die NCPI ist die Basis für die Zuverlässigkeit von IT-Abläufen in einem Unternehmen. Als Basis muss sie zuverlässig und nach dem Motto „Aus den Augen, aus dem Sinn“ funktionieren. Nachfolgend sind Kurzbeschreibungen eines jeden NCPI-Elements zu finden.

Stromversorgung

Es gibt viele Elemente, die eine solide Stromversorgungsinfrastruktur ausmachen, inklusive aller Systeme vom Eingang der Stromeinspeisung des Gebäudes bis hin zu den Lastwerten des Datacenters oder Netzwerks. Das umfasst auch den Hauptstromverteiler, die Generatoren, die USV-Systeme und Batterien, die Überspannungsschutzgeräte, Transformatoren, Verteileranlagen und Netztrennschalter.

Das APC White Paper Nr. 4 „Essenzielle Stromversorgungsanforderungen für Datacenter der nächsten Generation“ behandelt die Probleme und Herausforderungen bei herkömmlichen Stromversorgungssystemen und zeigt eine immer größer werdende Fülle von Anforderungen.

Kühlung

Kühlsysteme werden benötigt, um Abwärme aus dem Datacenter zu leiten. Darunter findet man Klimageräte für Computerräume, sämtliche damit verbundene Subsysteme, die für den Betrieb dieser Klimageräte benötigt werden, wie Kühler, Kühltürme, Kondensatoren, Pumpen, Rohrleitungen und alle Verteilungskomponenten auf Rack-Ebene.

Das APC White Paper Nr. 5, „Essential Cooling System Requirements for Next Generation Data Centers“, behandelt die Probleme und Herausforderungen bei herkömmlichen Kühlsystemen und zeigt eine immer größer werdende Fülle von Anforderungen.

Racks und physische Struktur

Es gibt viele Systeme, die als Teil der physischen Struktur eines Datacenters betrachtet werden können. Die kritischsten Elemente sind die IT-Racks, die das IT-Equipment beherbergen und die physischen Raumelemente wie doppelte Decken und Böden (sowohl doppelte Böden als auch Betonplattenböden).

Das APC White Paper Nr. 7, „Essential Rack System Requirements for Next Generation Data Centers“, behandelt die Probleme und Herausforderungen bei herkömmlichen Racksystemen und zeigt eine immer größer werdende Fülle von Anforderungen. Das APC White Paper Nr. 19, „Eignung von Doppelböden für Datacenter“, betrachtet die Gründe für den Einsatz von Doppelböden in der Vergangenheit und legt dar, dass der weit verbreitete Gebrauch in Datacentern nicht länger gerechtfertigt ist.

Sicherheit und Feuerschutz

Sicherheits- und Feuerschutzsysteme sind für den Erhalt der Integrität, Sicherheit und Verfügbarkeit eines Datencenters essenziell. Die enthaltenen Subsysteme sind hier Geräte für die physische Sicherheit auf Raum- und Rack-Ebene sowie Feuermelde- und Löschsyste-me. Unter den Geräten für die physische Sicherheit findet man beispielsweise biometrische Sicherheit, Schlüssel, Codes und Steckkarten. Intelligente Rauchmelder, Feuerlöschsysteme mit Alternativgas und lineare Wärmemelder sind Beispiele für Feuermelde- und Löschsyste-me.

Das APC White Paper Nr. 82, „Physical Security in Mission Critical Facilities“, behandelt die Elemente für die physische Sicherheit im Detail und bietet Möglichkeit für eine optimierte Sicherheit der Einrichtung. Das APC White Paper Nr.83, „Mitigating Fire Risks in Mission Critical Facilities“, beschreibt im Detail Erkennung, Bekämpfung und Verhinderung von Bränden und zeigt die besten Methoden für die Minimierung von Brandrisiken.

Verkabelung

Die Kabelinfrastruktur umfasst alle Datenkabel, die Teil des Datencenters sind, sowie alle Stromversorgungskabel für die Speisung aller Lasten. Kabelpritschen und Geräte für das Kabelmanagement sind genauso wichtig für die IT-Infrastruktur, weil sie die Wahrscheinlichkeit eines Ausfalls aufgrund menschlichen Versagens und Überhitzung verringern können.

Management

Der Begriff „Management“ ist ein Element von NCPI, der alle oben beschriebenen Elemente umfasst. Um eine zuverlässige NCPI zu erhalten, ist es wichtig, dass man alle Komponenten der physischen Infrastruktur unter Beobachtung hat. Unter Management versteht man Systeme wie Building Management Systeme (BMS) Netzwerk Management Systeme (NMS), Komponenten-Manager (wie APC InfraStruXure) und andere Überwachungs-Hard- und Software.

Das APC White Paper Nr. 14 „Essential NCPI Management Requirements for Next Generation Data Centers“ behandelt die Probleme und Herausforderungen bei herkömmlichen NCPI-Managementsystemen und zeigt eine immer größer werdende Fülle von Anforderungen.

Dienste

Es gibt eine Fülle von notwendigen Diensten, um NCPI-Systeme über die gesamte Lebensdauer zu erhalten. Diese Dienste können in fünf Kategorien unterteilt werden: (1) Beratungs- und Konzeptionsdienste; (2) Installationsdienste; (3) Wartungs- und Reparaturdienste; (4) Überwachungsdienste und (5) Inbetriebnahmedienste.

Das APC White Paper Nr. 12 „Zeitgemäße Servicekonzepte für NCPI-Systeme in Datencentern“ behandelt die Probleme und Herausforderungen bei herkömmlichen NCPI-Dienstangeboten und zeigt eine immer größer werdende Fülle von Anforderungen.