Hochverfügbare Zugangsnetze auf Basis von Ethernet

Dietmar Tölle • Markus Zeller • Rudi Knorr

Ethernet gewinnt als kommende Transportplattform für Zugangsnetze an Bedeutung, da es einfach, schnell und kostengünstig zu realisieren ist. Deshalb wird Ethernet in den nächsten Jahren sicherlich ATM als bevorzugte Technik in den Zugangsnetzen ablösen. Doch vorher müssen noch einige Herausforderungen bewältigt werden.

Für eine umfassende Anwendung der Ethernet-Technik für Zugangsnetze – auch Access Aggregation Networks (AAN) genannt – sind noch einige Herausforderungen zu bewältigen. Der Datenverkehr mehrerer Kunden wird in den Zugangsknoten (Access Nodes) des Dienstanbieters zusammengeführt (aggregiert), z. B. in DSLAM für DSL oder drahtlosen Zugangspunkten für WLAN oder WiMAX. Über deren Uplink-Verbindungen werden die Daten an ein ausschließlich Ethernet-gestütztes Zugangsnetz weitergeleitet. Der Vorteil in der Nutzung von Ethernet liegt in der Bandbreiteneffizienz, dem großen Bandbreitenbereich von Ethernet-Verbindungen und der geringen Komplexität in Bezug auf Betrieb, Management und Wartung. Allerdings weist Ethernet in der jetzigen Version einige Mängel auf, die die Nutzung als Provider-Technik noch nicht zulassen. Zu Lösen sind die Bereitstellung bestimmter Dienstgüten (Quality of Service, QoS), die Zuverlässigkeit der Netzknoten und Verbindungen (Links) sowie die Fähigkeit der Fehlererkennung und Fehlerbehebung. Eine schnelle Fehlerbehebung erreicht man z. B. mit einem fehlertoleranten Netzdesign, also die Nutzung redundanter Netzressourcen zur Sicherstellung des Betriebs trotz auftretender Fehler.

Fehlertolerantes Netzdesign für Ethernet

Eine hundertprozentige Verfügbarkeit bedeutet, die Funktionsfähigkeit eines Netzes zu jeder Zeit zu gewährleisten und damit die beauftragten Netzverbindungen und Dienste auch tatsächlich ohne Unterbrechung zu realisieren. Auf Grund der beschränkten Lebensdauer der Netzelemente und Verbindungen lässt sich allerdings eine Verfügbarkeit von 100 % nicht erreichen (siehe **Kasten**). Mit Hochverfügbarkeit wird daher die Nähe zur hundertprozentigen Verfügbarkeit bezeichnet. Das Ziel ist eine Verfügbarkeit von 99,999 % (das entspricht einem Ausfall von rd. 5 min pro Jahr); realistische Werte liegen zwischen 99 % und 99,99 %.

Auf einen Blick

Die Fraunhofer-Einrichtung für Systeme der Kommunikationstechnik ESK hat die gängigen Lösungen für die Absicherung der Hochverfügbarkeit bei Zugangsnetzen mit Ethernet untersucht. Die Umschaltzeiten im Fehlerfall erreichen – so die Messungen – noch nicht die von SDH gewohnten Werte von unter 50 ms.

Um die Hochverfügbarkeit von Zugangsnetzen sicherzustellen, gibt es drei Ebenen des Netzdesigns, die weitgehend unabhängig voneinander betrachtet und optimiert werden können: Geräteebene, Verbindungsebene und Netzebene.

Hochverfügbarkeit auf Geräteebene

Eine sehr nahe liegende Lösung zur Sicherstellung der Hochverfügbarkeit ist die Vermeidung von Fehlern und Ausfällen in den Netzknoten. Auch innerhalb dieser Ebene gibt es wiederum mehrere Optionen: Zum einen lassen sich hochwertigere, aber damit auch teurere Einzelkomponenten nutzen. Zudem lässt sich die Betriebssicherheit einzelner Komponenten durch die geeignete Wahl der Betriebsbedingungen verbessern. Bei Netzkomponenten mit einem hohen Anteil von Halbleitern spielen der thermische Arbeitspunkt und die elektromagnetische Verträglichkeit (EMV) eine große Rolle für die Betriebssicherheit. Bei modularen Systemen ist auch die Zuverlässigkeit der mechanischen Verbindungen entscheidend. Zudem können innerhalb eines Geräts betriebskritische Einzelkomponenten redundant ausgelegt werden.

Typische redundant ausgelegte Systemkomponenten sind die Stromversorgung, der Steuerrechner oder auch die Schaltmatrix. Hiervon sind häufig zwei Komponenten gleichzeitig im Gerät, wobei nur eine davon aktiv ist (Master), die andere ist im Standby-Modus. Die Standby-Komponente überwacht die Funktionsfähigkeit des Masters und übernimmt, falls dieser ausfällt.

Einige Komponenten sind nur einfach vorhanden, etwa die Backplane oder die Hauptplatine. Diese müssen entsprechend ausfallsicherer sein, was aber auf Grund ihrer Funktion und der Betriebsbedingungen häufig gegeben ist.

Alle diese Maßnahmen zielen auf die Erhöhung der Verfügbarkeit des Geräts, was die Kosten zwar erhöht, aber nicht verdoppelt. Allerdings erhält man mit diesen Maßnahmen nur eine explizite Redundanz, das heißt, die redundanten Komponenten können nur eine ganz bestimmte Komponente absichern.

Eine Vielzahl von Standards, Richtlinien und Vorschriften gewährleisten die Betriebssicherheit von Geräten. Die

Mathematische Beschreibung der Verfügbarkeit (Availabilty, *A*)

Die Systemkomponente i habe die Fehlerrate λ_i (durch Stichprobentests ermittelt).

Die Wahrscheinlichkeit P für Fehler der Komponente i im Zeitintervall T (Lebensdauer $L \le T$) beträgt dann: $P(L \le T) = 1 - e^{-\lambda T}$.

Die Zuverlässigkeit (Reliability, R; Komponente i überlebt Zeitintervall T) beträgt: $R = e^{-\lambda T}$.

Für die mittlere Zeit, bis ein Fehler auftritt (Mean Time to Failure, durchschnittlich), gilt: $MTTF_i = 1/\lambda_i$.

Für reparierbare Systeme gelten zusätzlich folgende Angaben:

MTTR (Mean Time to Repair) bzw. *MDT* (Mean Down Time) ist die Zeit zur Fehlererkennung, Reparatur und Wiederinbetriebnahme.

Die mittlere Zeit zwischen zwei Fehlern (Mean Time between Failure) beträgt: $MTBF = MTTF + MTTR$.

Für die Verfügbarkeit (Availability) gilt:
$A = MTTF/MTBF = MTTF/(MTTF + MTTR)$.

Für die Nichtverfügbarkeit (Unavailability) gilt:
$E = 1 - A = MTTR/(MTTF + MTTR)$.

Sind mehrere Komponenten logisch in Reihe geschaltet, d. h. führt der Ausfall einer Teilkomponente zum Ausfall des Gesamtsystems (z. B. Kabelsegmente), so gilt:

$A_{ges} = A_1 \cdot A_2 \cdot \ldots \cdot A_n \Rightarrow E_{ges} = 1 - A_{ges}$.

Sind mehrere Komponenten logisch parallel geschaltet, d. h. führt ein gleichzeitiger Ausfall von $n-1$ Teilkomponenten nicht zum Ausfall des Gesamtsystems (z. B. bei redundanten Kabeln), so gilt:

$E_{ges} = E_1 \cdot E_2 \cdot \ldots \cdot E_n \Rightarrow A_{ges} = 1 - E_{ges}$ und

$A_{ges} = 1 - (1-A_1) \cdot (1-A_2) \cdot \ldots \cdot (1-A_n)$.

Bei vielen Systemen ist die Berechnung von A_{ges} bzw. E_{ges} komplexer, da es sich um beliebig komplexe Kombinationen aus Reihen- und Parallelschaltungen handelt.

Beispiele:

Eine ungeschützte Netzverbindung habe eine Verfügbarkeit von $A = 98{,}5\,\%$; dann haben zwei identische, komplett parallele, ungeschützte Netzverbindungen eine Gesamtverfügbarkeit von $A_{ges} = 99{,}9775\,\%$ und drei identische, komplett parallele, ungeschützte Netzverbindungen eine Gesamtverfügbarkeit von $A_{ges} = 99{,}9996625\,\%$.

Soll eine Gesamtverfügbarkeit von $A_{ges} = 99{,}999\,\%$ erreicht werden, so sind entweder zwei parallele Netzverbindungen mit jeweils einer Einzelverfügbarkeit von $A = 99{,}684\,\%$, drei parallele Netzverbindungen mit jeweils einer Einzelverfügbarkeit von $A = 97{,}846\,\%$ oder vier parallele Netzverbindungen mit jeweils einer Einzelverfügbarkeit von $A = 94{,}38\,\%$ notwendig.

wichtigsten sind hier die NEBS- (Network Equipment Building System) bzw. die ETSI-Konformität, insbesondere EN60950 für Sicherheit, EN55022/4, ETSI 300 386 und NEBS GR-1089 für EMV, NEBS GR-63 für physikalische Sicherheit und ETSI 300 019 für Umweltbedingungen.

Hochverfügbarkeit auf Verbindungsebene

Die Verfügbarkeit muss auch auf der Verbindungsebene, also den Netzkanten, möglichst hoch sein. Die Netzkante bezeichnet in diesem Fall die Punkt-zu-Punkt-Verbindung zwischen zwei Netzelementen. Auch auf dieser Ebene gibt es zwei Möglichkeiten, die Verfügbarkeit zu erhöhen: über die Sicherheit einer Einzelverbindung oder durch eine redundante Auslegung der Netzkante.

Die Verfügbarkeit einer Netzkante lässt sich oft nur ungenau beziffern, mathematisch entspricht sie aber der Reihenschaltung einer Vielzahl von Einzelabschnitten (siehe Kasten). Die Einzelabschnitte haben eine unterschiedliche Verfügbarkeit, die meist sehr nahe an 100 % liegt. Bei der Reihenschaltung wird die Gesamtverfügbarkeit aber maßgeblich durch die geringste Verfügbarkeit einer Einzelkomponente bestimmt. Zudem ergibt sich ein Gesamtausfall schon bei einem beliebigen Einzelausfall.

Bei der Planung der Trasse für eine physikalische Verbindung ist es also wichtig, Einzelkomponenten mit einem hohen Gefährdungspotenzial zu vermeiden. In der Praxis wird die Verfügbarkeit von Netzkanten lediglich geschätzt.

Bei der redundanten Auslegung einer Netzkante werden zwischen zwei zu verbindenden Netzelementen Einzelverbindungen parallel geschaltet. Fällt eine der Einzelverbindungen aus, übernehmen die verbleibenden Verbindungen den Verkehr. Diese Methode wird in Kombination mit der Geräteredundanz benutzt, da nicht nur die Verbindung an sich gedoppelt wird, sondern auch die Schnittstellen an den Geräten. Für die physikalischen Verbindungen ist noch zu beachten, dass diese nicht den gleichen Weg nutzen, da sonst im Fehlerfall alle Verbindungen auf Grund derselben Ursache – und damit gleichzeitig – ausfallen. Werden die Verbindungen doch parallel geführt, so ist diese redundante Auslegung nur zum Schutz gegen den Ausfall einer Schnittstelle verwendbar.

Eine tatsächlich getrennte Wegeführung der beiden Verbindungen ist meist sehr kostenintensiv. Eine Möglichkeit zur Kombination dieser beiden Methoden liefern spezielle bauliche Maßnahmen. In vielen Fällen können diese bei typischen Fehlerfällen – Bagger trifft Kabel – einen doppelten Ausfall verhindern, aber bei massiver mechanischer Gewalteinwirkung helfen diese Maßnahmen auch nicht mehr. Außerdem existiert noch eine weitere kritische Fehlerquelle (single point of failure): Sollte eines der Geräte doch ausfallen, so nützt die redundante Ausführung der Verbindung nichts.

Ethernet bietet einige Verfahren zur redundanten Auslegung von Verbindungen zwischen zwei Netzelementen. Einfache Verfahren beruhen auf Port-Spiegelung, d. h. der Verkehr eines Ports wird senderseitig komplett und ohne jegliche Veränderung auf einen zweiten – den redundanten – Port gespiegelt. Auf der Empfangsseite wird daher der Verkehr doppelt empfangen. Bei Ethernet ist es sehr wichtig, dass dieser doppelte Verkehr nicht weitergeleitet wird, weshalb bei problemlosem Empfang ein Satz der Daten vom Netzelement verworfen wird. Bei diesem Verfahren spricht man auch von 1 + 1-Redundanz.

Ein anderes Verfahren sendet den Verkehr nur über eine der Verbindungen, die andere ist inaktiv. Fällt der aktive Link aus, übernimmt der inaktive den Betrieb. Hier sind die Fehlererkennung und die darauf folgende Umschaltung wichtig. Die Verzögerung liegt hier zwischen 0,05 s und 1 s. Man spricht auch von 1 : 1-Redundanz.

Eine weitere Möglichkeit ist die Aggregation mehrerer physikalischer Verbindungen zu einer logischen Verbindung. Dabei wird der Verkehr des logischen Links sendeseitig auf die physikalischen Verbindungen aufgeteilt. Die Auslastung der einzelnen physikalischen Verbindungen wird reduziert, so dass bei einem einzelnen Ausfall der Verkehr der ausgefallenen Verbindung auf die verbleibenden aufgeteilt werden kann. Bekannte Verfahren sind hier Load Sharing und das Link Aggregation Control Protocol (LACP, IEEE 802.3ad). Man spricht hier von einer *N*:1-Redundanz, wobei *N* auch 1 sein kann.

Dietmar Tölle, Markus Zeller und Rudi Knorr sind an der Fraunhofer-Einrichtung für Systeme der Kommunikationstechnik ESK in München tätig.

Maßnahmen auf der Netzebene

Die beiden oben angesprochenen Ebenen haben den deutlichen Nachteil, nur eine explizite Redundanz bereitzustellen. Die installierten redundanten Netzressourcen dienen zum Schutz einer ganz bestimmten Netzressource und keiner anderen. Dies erhöht die Kosten zum Teil erheblich, der Netzbetreiber kann aber hierbei sehr gut planen und steuern, ob und welche Redundanzen er nutzt.

Wesentlich flexibler ist die Redundanz auf Netzebene. Dieses Konzept sieht die Nutzung von ungeschützten Netzkomponenten vor, die Ausfallsicherheit wird durch die gegenseitige Absicherung erhöht. Bei diesen Verfahren sind daher die Fehlererkennung und die Steuerung der Netzressourcen essenziell für die Hochverfügbarkeit. Die Fehlererkennung und Netzsteuerung wird entweder zentral von einem Netzmanagementsystem übernommen oder autonom durch die Netzelemente mittels eines geeigneten Protokolls realisiert.

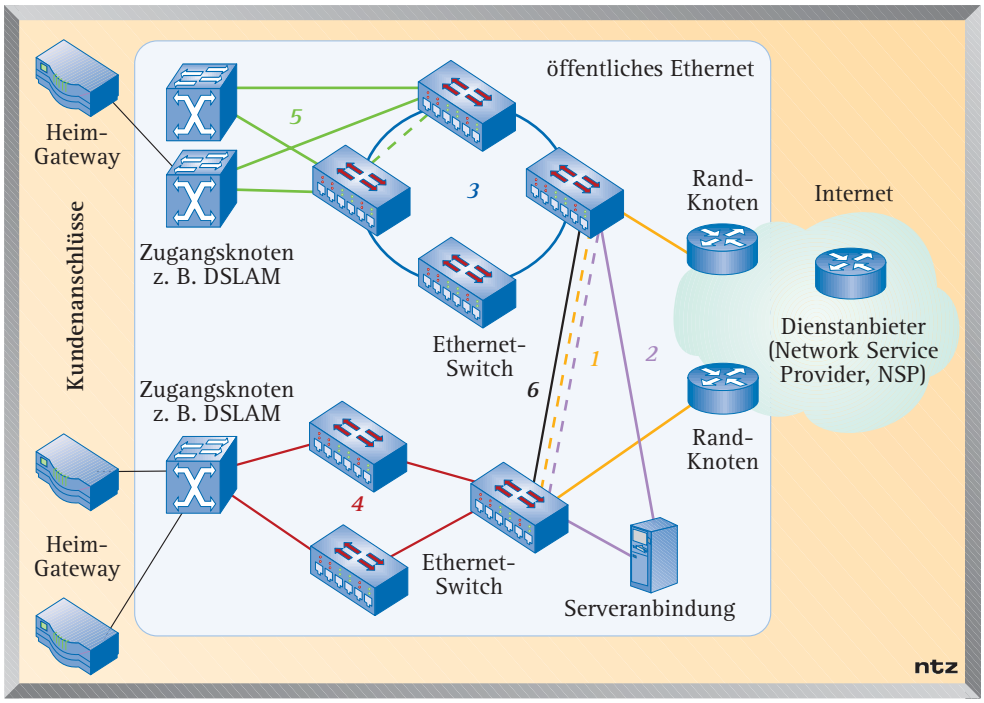
Ein Grund für die hohe Kosteneffizienz von Ethernet ist die Fähigkeit, einfach und effizient Redundanz auf der Netzebene bereitzustellen. Ethernet lässt sich prinzipiell zu jeder beliebigen Topologie zusammenschalten, allerdings ist es sehr anfällig für Schleifenbildungen, die bei redundanten Netzen immer auftreten. Die Netzelemente unterstützen aber in der Regel zahlreiche Protokolle zur Schleifenvermeidung.

Die Funktionsweise dieser Protokolle beruht auf demselben Prinzip. Das Netz wird durch Austausch von Topologie-Informationen in eine logische Struktur, beispielsweise einen Baum oder einen aufgetrennten Ring, abgebildet. Dabei werden einige Netzkanten logisch blockiert, um Schleifen zu vermeiden. Fällt jetzt eine Netzkomponente innerhalb der aktiven logischen Struktur aus, werden die bisher blockierten Netzkanten wieder freigegeben bzw. die logische Struktur wird neu gebildet und kann jetzt auch die bisher blockierten Ressourcen enthalten. Bei Reaktivierung der ausgefallenen Netzkomponente wird die logische Struktur wiederum neu berechnet, und eine bestimmte Netzkante wird wieder blockiert.

Zudem kann man mit Ethernet das physikalische Netz in mehrere virtuelle Netze aufteilen, die sich alle wie eigenständige Ethernet-Netze verhalten. So gibt es eine Vielzahl von Kombinationsmöglichkeiten für die Aufteilung von aktiven und logisch blockierten Netzressourcen, und der Grad der Redundanz ist sehr fein einstellbar. Da diese Protokolle einen vergleichsweise geringen Konfigurationsaufwand bei hoher Kapazität und Flexibilität aufweisen, ist Ethernet sehr gut geeignet für die Zugangsnetze (Access Aggregation Networks).

Im **Bild** sind einige Protokollbeispiele zu sehen, deren Reaktionszeiten auf Fehler sind in der **Tabelle** aufgeführt. Es ist unüblich, dass die Netzelemente zum Kundenanschluss gedoppelt sind, da dies zu kostspielig wäre. Aber häufig beginnt das fehlertolerante Design mit der zweiseitigen Anbindung der Zugangsknoten (active und protecting uplink) an verschiedene Aggregations-Netzelemente (*5*). Die Aggregation mehrerer DSLAM läuft über einen Ring (*3*) oder Aggregationskaskaden (*4*). Die Anbindung an die IP-Router (*1*) bzw. netzinterne Server (*2*) ist sehr wichtig und daher unbedingt redundant auszuführen. Die einzelnen Netzabschnitte können über eine Verbindungsleitung (*6*) miteinander verbunden und gegebenenfalls über das IP-Netz gesichert werden, wobei hier eine komplexere Netzsteuerung notwendig ist.

Im Fehlerfall sollte die Umschaltung vom „working“ auf den „protecting“ Link innerhalb der für SDH typischen 50 ms möglich sein. Ethernet erfüllt diese Anforderung nach momentanem Stand bei weitem noch nicht. Die Fraunhofer-Einrichtung für Systeme der Kommunikationstechnik ESK hat die gängigen Lösungen untersucht; die gemessenen Umschaltzeiten zeigt die Tabelle.



Zugangsnetz (Access Aggregation Network, AAN) mit verschiedenen Redundanzverfahren

Protokoll	Nutzbar für …	Fehlerreaktion	Initialisierung
STP (Spanning Tree Protocol)	Baumstrukturen (*4*, ggf. *5*)	15 000 ms … 50 000 ms	10 000 ms … 30 000 ms
RSTP, MSTP (Rapid, Multiple STP)	Baumstrukturen (*4*, ggf. *5*, *3*)	ca. 600 ms	ca. 400 ms
EAPS (Ethernet Automatic Protection Switching)	Ringstrukturen (*3*, *4*, ggf. *5*)	100 ms … 250 ms	100 ms
VRRP (Virtual Router Redundancy Protocol)	logische Verbindung zweier Netzelemente (*1*)	2 500 ms … 3 500 ms	100 ms
SLB (Server Load Balancing)	doppelte Anbindung eines Netzelements (*2*)	500 ms … 2 000 ms	100 ms

Vergleich der Redundanzverfahren

Die Zeiten weichen zum Teil deutlich von den geforderten 50 ms ab. Am geeignetsten erscheinen die Ring-Protection-Varianten. Allerdings sind diese Varianten meist proprietär und zu den standardisierten Verfahren wie den Spanning Trees (STP) inkompatibel. Die schnellen standardisierten Verfahren brauchen heute etwa zehnmal so lang wie die Standard-SDH-Verfahren. Die von der IEEE bevorzugten STP müssen also weiter entwickelt werden. Aktuelle Arbeiten befassen sich unter anderem mit der Integration von Ringen, der Beschleunigung der Protokolle und der Erweiterung des Radius der Protokolle, d. h. der Anzahl der Netzelemente pro Baumstruktur.

Strategien zum fehlertoleranten Netzdesign

Redundanz auf der Geräte- oder Verbindungsebene kann sehr kostspielig sein, da es sich um explizite Redundanz handelt. Die Systemhersteller sehen für viele ihrer Netzelemente Redundanz vor, der Betreiber kann den Grad der verwendeten Redundanz auch sehr gut steuern. Bei der Verbindungsredundanz ist es sehr schwierig und kostenintensiv, ein hohes Maß an Redundanz bereitzustellen. Außerdem verhindern die Randbedingungen häufig die gezielte Bereitstellung von Redundanz. Hier hat der Betreiber oft viel weniger Entscheidungsmöglichkeiten als er gerne hätte. Die Möglichkeiten von Ethernet zur Bereitstellung von Redundanz auf Netzebene ist eine ideale Ergänzungsmöglichkeit zum Erreichen der Hochverfügbarkeit.

Dabei zeigt sich allerdings, dass die aktuellen Varianten die Anforderungen der Netzbetreiber noch nicht erfüllen können. Außerdem ist ein kombinierter Schutz von Ring- und Baumstrukturen oft noch nicht möglich. Die Fraunhofer ESK arbeitet an Lösungen zur Optimierung dieser Verfahren. Eine dieser Lösungen – Diesel (Distributed Self-Learning) – beruht auf der Verbreitung von Topologie-Informationen und mehrfacher Weiterleitungsmetriken. Somit kann im Fehlerfall sofort reagiert werden. Auch ein Verfahren für die Beschleunigung des RSTP-Verfahrens (Rapid Spanning Tree Protocol) wurde entwickelt. ■