

Implementierungen

Netzwerk-API (=Application Program Interface)

Vom Betriebssystem für das Kommunikationssystem bereitgestellte Schnittstelle. Es liefert die Syntax, mittels der der Protokolldienst abgerufen werden kann.

Socket-API: Interface zwischen einem Anwendungsprogramm und den Kommunikationsprotokollen in einem Betriebssystem. Bietet Operationen zum Erstellen eines Sockets, zum Anbinden des Sockets an das Netzwerk, zum Senden/Empfangen von Nachrichten und zum Schliessen des Sockets.

Das Socket-API ist ein De-facto-Standard.

Schnittstellen zwischen 2 Protokollen:

Prozess-pro-Protokoll Modell: Erfordert teure Prozess-Wechsel (context switches)

Prozess-pro-Nachricht Modell: Erfordert lediglich (billigere) Prozeduraufrufe

Statt send-receive meist send-deliver Semantik

Nachrichtenpuffer

Ineffizienz: Socket Interface verlangt Nachrichtenkopieren zwischen zwei Puffern.

Ähnliche Nachrichtenabstraktion wird von den meisten Betriebssystemen unterstützt:

- Hinzufügen und Entfernen von Headern
- Fragmentierung und Zusammensetzung von Nachrichten
- Logisches Speichern einer Nachricht in einer anderen
- Erzeugen von Nachrichten aus Inhalten von Puffern
- Löschen einer Nachricht, Länge einer Nachricht, Kürzen einer Nachricht

Direktverbindungsnetzwerke

Fünf Aufgaben: Kodierung, Framing, Fehlerüberwachung, zuverlässige Zustellung, Medienzugriffssteuerung

Die Lösungen werden in Netzwerkadaptoren bzw. Netzwerkkarten implementiert:

Bits werden zwischen Netzwerkkarten ausgetauscht, der korrekte Rahmen zwischen Knoten und Netzwerkkarte wird vom Treiber (SW) auf den Knoten gesteuert.

Bitübertragungsschicht

Grundlage aller Netze mit physikalischen Grenzen. Sie betrifft die Übertragung von rohen Bits in einem Übertragungskanal.

Maximale Datenrate:

Für **rauschfreie Kanäle** = $2H \log_2 V$ (bit/Sekunde); H = Bandbreite V= Anzahl diskrete Stufen eines Signals

Nyquist: Wenn ein beliebiges Signal durch einen Tiefpassfilter der Bandbreite H geführt wird, kann das gefilterte Signal vollständig durch Abtastwerte von (genau) 2H Samples pro Sekunde wiederhergestellt werden. Abtastwerte von mehr als 2H pro Sekunde sind nutzlos, da Anteile mit höherer Frequenz, die durch eine höhere Abtastrate entdeckt werden könnten, bereits ausgefiltert wurden.

Bsp. 3-kHz-Kanal kann binäre (zweistufige) Signale nicht mit mehr als 6000 bps übertragen.

Für **verrauschte Kanäle** = $H \log_2(1+S/N)$ (bit/Sekunde); S = Signalstärke, N = Rauschstärke → Rauschabstand = S/N (dB)

Shannon: Ein Kanal mit 3000 Hz Bandbreite und einem Rauschabstand von 30dB (z.B. Telefon) nie mehr als 30000 bps übertragen, gleichgültig, wie viele Signalstufen benutzt werden und wie oft die Abtastwerte genommen werden.

Medien

Terrestrische Übertragungsmedien (verdrehte Kabelpaare, Koaxialkabel, Lichtwellenleiter)

Aerische Übertragungsmedien (Funk, Mikrowelle, Infrarot, Laser)

Kodierung

- Non-Return to Zero (NRZ)
- Non-Return to Zero Inverted (NRZI)
- Manchester Kodierung
- 4B/5B Kodierung

Sicherungsschicht

Die Protokolle dieser Schicht definieren die Organisation von Daten in Rahmen und die Übertragung von Rahmen.

Aufgabe:

Rauschende Leitungen für die Vermittlungsschicht in fehlerfreie Kommunikationskanäle verwandeln. → durch Rahmen. Fehlerüberwachung und Flusststeuerung

Dienstarten:

Unbeständige verbindungslose Dienste: flexibel und optimistisch; Problembearbeitung erfolgt auf höherer Schicht. Unabhängige Rahmen werden an den Empfänger geschickt, ohne Empfangsbestätigung. Kommt zur Anwendung, wenn Fehlerquote sehr niedrig ist und die Datenwiederherstellung höheren Schichten überlassen wird.

Bestätigte verbindungslose Dienste: Es werden immer noch keine Verbindungen benutzt, aber der Empfang jedes abgeschickten Rahmens wird einzeln bestätigt.

Verbindungsorientierte Dienste: Höchste Qualitätsstufe, 3 Phasen der Datenübertragung (Verbindungsaufbau, Übertragung, Abbau). Jeder Rahmen ist nummeriert → kommt sicher an, und nur einmal, in der richtigen Reihenfolge.

Rahmenerstellung

Zeichenzählung: Ein Feld wird im Header des Rahmens eingefügt, das die Anzahl der Zeichen im Rahmen angibt. Sobald die Sicherungsschicht auf der Empfängerseite dieses Feld liest, weiss sie, wie viele Zeichen ankommen werden und wo folglich das Ende des Rahmens sein wird. Problem: Wenn Fehler, kann Empfänger zwar erkennen, dass der Rahmen falsch ist, kann aber den Anfang des nächsten Rahmens nicht erkennen.

Anfangs- und Endzeichensetzung (Zeichenstopfen): Erneute Synchronisation zur Vermeidung des Problems. Rahmen beginnt mit der ASCII-Zeichenfolge DLE STX und endet mit DLE ETX (DLE = Data link escape, STX = Start of TeXt, ETX = End of TeXt)

Anfangs- und Endflags (Bitstopfen): Jeder Rahmen beginnt und endet mit einem speziellen Bitmuster, nämlich 01111110. Sobald die Sicherungsschicht des Senders 5 aufeinanderfolgende Einsen im Datenstrom entdeckt, stopft er automatisch eine Null in den abzusendenden Bitstrom. Sobald der Empfänger eine Folge von fünf Einsen, gefolgt von einer Null, im ankommenden Datenfluss erkennt, nimmt er das Null-Bit automatisch aus dem Datenfluss heraus.

Adaptierte Kodierregeln der Bitübertragung: Kann nur in Netzen angewandt werden, in denen Kodierung im physischen Medium mit Redundanz verbunden ist.

Fehlerüberwachung

Codierung ermöglicht die Korrektur einiger Fehler und das Erkennen weiterer – je besser desto grössere Redundanz notwendig (ohne Redundanz geht nichts, keine 100%ige Fehlererkennung). Nur Fehlererkennung und Neuübertragung (optimistisches Protokoll) vs. Fehlerkorrektur (pessimistisches Protokoll).

Flusststeuerung

Aufgabe: Schnelleren Sender davon abhalten, den langsameren mit Daten zu überschütten. Das Protokoll enthält genau definierte Regeln, wann ein Sender den nächsten Rahmen abschicken darf.

Flusssteuerungs- & Zuverlässigkeitsprotokolle

Einfaches Simplexprotokoll ohne etwas: Voraussetzung ist ein zuverlässiger Kanal

Stop-and-Wait: Timer, Nummerierung, Huckepacktransport (piggybacking)

Schiebefensterprotokoll: Selektive vs. Gesamte Wiederholung, anwendungsspezifische Verwerfung, zusätzliche Aufgabe = Einhaltung der Reihenfolge (durch Restklassennummerierung). Werden nach der Größe des Sende- und Empfangsfensters eingeteilt.

MAC Teilschicht

Protokolle, mit denen bestimmt werden kann, wer wann in einem Mehrfachzugriffskanal an die Reihe kommt, gehören zu einer Teilschicht der Sicherungsschicht, der **Medium Access Control**. Wichtig für LANs.

Media Access Protocol

Aufgabe ist die Aufteilung der Kanäle.

Vorgehen: Optimistische, pessimistische und hybride Ansätze, Entwicklung von „Bei-Bedarf-Zugriff“-Algorithmen anhand des Grundmodells zum Leistungsvergleich.

Protokolle mit Kollisionen (Konkurrenzprotokolle)

- **ALOHA:** Jederzeit nichtsynchrone Übertragung. Absender kann dank Bestätigungsmöglichkeit der Broadcast-Technologie herausfinden, ob sein Rahmen zerstört wurde. Wenn ja, wartet der Absender eine zufällige Zeitspanne und sendet ihn nochmals. Zeitspanne muss zufällig sein, da sonst die gleichen Rahmen wieder kollidieren.
- **Unterteiltes ALOHA:** diskrete Version. Zeit wird in Intervalle eingeteilt, jedes Intervall entspricht einem Rahmen. Der Sender kann nicht einfach senden, sondern muss auf den nächsten Zeitschlitz warten.
- **1-persistentes CSMA** (Carrier sense multiple access): Trägererkennung & zufällige Zeitspanne warten nach Kollision. Wenn eine Station Daten übertragen will, hört sie zuerst den Kanal ab, ob bereits jemand übermittelt. Bei Kollision wird wieder eine zufällige Zeitspanne gewartet. 1-persistent, da die Station mit einer Wahrscheinlichkeit von 1 sendet, wenn der Kanal frei ist.
- **Nicht-persistentes CSMA:** Es wird eine zufällige Zeitspanne gewartet mit Kanalprüfung, falls bei der ersten Prüfung der Kanal besetzt war. → bessere Kanalauslastung und längere Wartezeiten als bei 1-persistentem CSMA.
- **P-persistentes CSMA:** Getaktet, mit Wahrscheinlichkeit q wird bis zum nächsten Zeitschlitz gewartet. Wenn Kanal frei ist, wird mit einer Wahrscheinlichkeit p gesendet oder gewartet... Vorgang wird so lange wiederholt, bis entweder der Rahmen übertragen ist, oder eine andere Station zu senden begonnen hat.
- **CSMA/CD (CD = Collision Detection):** Abbruch der Übertragung bei Kollisionserkennung. Wenn eine Station übertragen will, hört sie das Kabel ab. Wenn zwei oder mehrere Stationen gleichzeitig auf ein freies Kabel zugreifen, erfolgt eine Kollision. Jede dieser Stationen unterbricht dann die Übertragung, wartet eine zufallsgesteuerte Zeitspanne und wiederholt den ganzen Vorgang.

Carrier Sense Protocols = Protokolle, bei denen Stationen einen Träger (Carrier) abhören und dementsprechend handeln können.

Kollisionsfreie Protokolle

Pessimistisches Protokoll (gut bei hoher Auslastung, im Vergleich zu optimistischen Konkurrenzmethoden (gut bei niedriger Auslastung))

- *Bitmusterprotokoll:* Reservierungsprotokoll
- *Binärer Countdown,* mit Adressrotation, 100% Effizienz möglich

Protokolle mit eingeschränkter Konkurrenz (limited contention protocols):

Ziel wäre: bei Niedriglast Konkurrenz, bei Hochlast kollisionsfrei

Idee: Reduktion der konkurrierenden Knoten, indem in Gruppen unterteilt ein Binär-
musterprotokoll durchgespielt wird (z.B. Interpolation zwischen Bitmusterprotokoll
(eine Station pro Gruppe) und ALOHA (alle Stationen in einer Gruppe))

Lösung: Adaptive Tree Walk (z.B. Syphilistestmethode der US-Armee). Traversieren
eines Baumes in Abhängigkeit auftretender Kollisionen (bei Kollisionsfreiheit in die
Breite, sonst in die Tiefe)

WDMA (Wavelength Division Multiple Access)

Jeder Station werden zwei Kanäle zugeordnet. Ein schmaler Kanal wird als Steuerkanal
bereitgestellt, um die Stationen zu signalisieren, während ein breiter Kanal den
Stationen dazu dient, Datenrahmen auszugeben.

Jede Station hat zwei Sender und zwei Empfänger:

- Ein Empfänger mit fester Wellenlänge zum Abhören seines eigenen Steuerkanals
- Einen einstellbaren Sender zum Senden auf dem Steuerkanal der anderen Station
- Einen Sender mit fester Wellenlänge zum Ausgeben von Datenrahmen
- Einen einstellbaren Empfänger zum Auswählen eines abzuhörenden Datensenders

Es gibt 3 Verkehrsklassen:

- Verbindungsorientiert mit konstanter Datenrate (z.B. unkompliziertes Video)
- Verbindungsorientiert mit variabler Datenrate (z.B. Dateitransfer)
- Datagrammverkehr (z.B. UDP Pakete)

Protokolle für drahtlose LANs

Problem der versteckten und der exponierten Stationen

Lösung: **MACA** (Multiple Access with Collision Avoidance). Konzept ist so, dass der
Sender den Empfänger zur Ausgabe eines kurzen Rahmens anregt, so dass nahegelegene
Stationen diese Übertragung erkennen und für die Dauer des bevorstehenden
(grossen) Datenrahmens nichts übertragen. Dies geschieht zuerst mit Senden eines
RTS-Rahmens (request to send) und Antwort mit CTS-Rahmen (clear to send). Dies
ermöglicht dritten Knoten ihre relative Lokalisierung zu bestimmen.

Ethernet (IEEE 802.3)

Im PARC entwickelt mitte der 70er Jahre. Es beruht auf der CSMA/CD Technologie.
Unabhängig von der Topologie erreichen alle Daten alle Knoten, alle Knoten konkurrieren
deshalb um eine Leitung.

Bustopologie mit Repeatern (max 4 Repeater zwischen 2 Hosts, max. Reichweite
2500m) oder Ethernet Hubs, maximal 1024 Hosts, Basisbandübertragung, Manchesterkodierung,
Bruttodatenrate 10Mbit/s, Datenpakete mit maximal 1.5 kByte Nutzdaten. Verschiedene Verkabelungsvarianten:

Thick Ethernet: 10Base5 (10Mbps, bis zu 500m, 100 Knoten/Segment); gelbes gartenschlauchartiges Kabel, das im Abstand von 2.5m Markierungen aufweist, wo sich Abzweigungen befinden. Anschlüsse werden über sogenannte Vampirabzweige hergestellt. Dabei wird ein Stift vorsichtig bis zur Hälfte in den Kern eines Koaxialkabels gedrückt.

Thin Ethernet/Cheapernet: 10Base 2 (bis zu 200m, 30 Knoten/Segment); lässt sich gut biegen. Anschlüsse über BNC-Stecker im Industriestandard in Form von T-Verbindungen → einfacher, günstiger.

Twisted pair: 10Base-T (bis zu 100m, 1024 Knoten/Segment). Verdrehte Kabelpaare; von allen Stationen aus führt ein Kabel zu einem zentralen Hub.

Glasfaser: 10Base-F (2000m, 1024 Knoten/Segment); Teuer, aber ausgezeichneten Widerstand gegen Rauschen. Wird zur Verkabelung von Gebäuden oder weit voneinander entfernten Hubs bevorzugt.

Es gibt keine zentrale Mediumzugriffskontrolle, Kollisionen werden analog entdeckt.

Format eines Ethernet-Pakets

Präambel 7 Byte, Spezialzeichen 1 Byte, Empfängeradresse 6 Byte, Senderadresse 6 Byte, Längenangabe 2 Byte, Daten 46-1500 Byte, Prüfsumme 4 Byte.

Ethernet-Adressen

Jeder Ethernet-Host hat eine eindeutige Adresse (gehört zur Netzwerkkarte). Zusätzlich Broadcast Adresse (aus lauter 1en) und Multicast-Adressen (erstes Bit eine 1, wird vom Host programmiert).

Promiscuous-Mode eines Netzwerkkadapters stellt Host alle Frames durch.

CSMA Sendealgorithmus

1-persistent

CD, d.h. Kollisionserkennung (minimal 96 Bit bei Fehlversuch, maximal 512 Bit = Headergröße)

Exponential Backoff (jeweils Verdoppelung der Wartezeiten) bei konsekutiven Fehlversuchen)

Fast Ethernet

100 Mbit/s Ethernet, 1995 als Erweiterung der alten Norm, identisch mit Ethernet bzgl. Paketformat, MAC-Verfahren und Dienstschnittstelle, neue Verkabelungsnormen, arbeitet nur mit Hubs und Switches.

Bemerkungen

30% Auslastung sehr/zu hoch, meist zusätzlich Ende-zu-Ende-Flusskontrolle auf Ende-zu-Ende-Basis, einfache Verwaltung und Wartung, keine Switches, kein Routing, einfacher Anschluss neuer Hosts, kostengünstig

Begriffe

Repeater = 2-Weg-Verstärker durch Signalregenerierung und Weiterreichen von Kollisionen

Bridge = Zum Filtern von Paketen zur Isolierung des lokalen Datenaustausches

Hub = Repeater mit Sternkonfiguration, der Kollisionen weitergibt

Switch = Schaltzentrale, sternförmig verkabelt, für Segmente und einzelne Stationen. Pakete werden gepuffert und via Bus intern an den richtigen Ausgangsort kopiert.

Token-Technologie (IEEE-802.5)

Bei IBM Rüsclikon entwickelt, verschiedene Topologien

Token-Ring Netz: Ringtopologie mit Round Robin

Bild S. 101 Comer

MSAU (Multi Access Station Unit), Relais zum Umgehen von Knoten, meist mit mehreren zusammen (quasi Sterntopologie)

Manchester-Kodierung, bis zu 250/260 Stationen, bei IBM twisted pair, Netzwerkkarte besteht aus Empfänger, Sender und dazwischenliegendem Datenspeicher, Minimalgröße (um Token aufzunehmen) des Netzes wird durch Monitorstation sichergestellt.

Daten werden an kreisendes Token (spezielle Bitfolge) angehängt, d.h. Token wird durch Frame ersetzt und (erst) am Ende des Kreises wieder herausgenommen.

Standard Token Haltezeit (THT): 10ms; TRT (Token Umlaufzeit) \leq aktive Knoten * THT + Ringlatenz.

Es gibt das Byte „Rahmenstatus“ und enthält die Bits A und C. Wenn ein Rahmen an der Schnittstelle der Station mit der Zieladresse ankommt, setzt die Schnittstelle während des Durchlaufs Bit A. Kopiert die Schnittstelle den Rahmen auf die Station, setzt sie auch Bit C. (Zuverlässige Zustellung durch A-Bit (Nachricht enthalten) und C-Bit (Nachricht kopiert))

Verschiedene Prioritätsstufen mit Reservierung, verzögerte versus frühe Token-Freigabe.

Wartung durch Monitorstation (= Überwachungsstation)

Aufgaben = zusätzliche Verzögerung, Sicherstellung von Token-Existenz, Elimination von verfälschten und von verwaisten Frames, schauen, dass kein Token verloren geht
Ursachen für Token-Verlust: Bitfehler, Stationsabsturz. Lösung: Timer mit Intervall Anzahl Stationen *THT + Ringlatenz (grösstmögliches tokenloses Zeitintervall). Wenn dieser Timer abläuft, leert die Überwachungsstation den Ring und bringt ein neues Token in Umlauf

Monitorbit zum Erkennen verwaister Frames: Verwaiste Rahmen werden von der Überwachungsstation entdeckt, indem bei jedem durchlaufenden Rahmen das Überwachungsbit im Byte Zugriffssteuerung gesetzt wird.

Bei Ausfallverdacht Versendung eines **Beacon-Frames**: Er enthält die Adresse der eventuell ausgefallenen Station. Hat sich der Beacon-Frame so weit wie möglich vorwärts bewegt, ist ersichtlich, wie viele Stationen ausgefallen sind. Sie werden aus dem Ring genommen.

Selbstermittlung des Nachfolgers bei Ausfall eines Monitors

Bei Ausfall des Empfangs der Heartbeats (spezielle Steuerzeichen) → Claim-Frames und „höchste Adresse gewinnt“ Auswahlverfahren.

Frameformat

Bild S. 324 Tannenbaum

FDDI (Fiber Distributed Data Interface)

Doppelring: einer zur Übertragung der Daten, wenn alles korrekt läuft, der andere wird als Ausweichnetz benutzt, falls das Hauptnetz ausfällt.

Bild S. 104 Comer

Maximal 500 Hosts mit max. 2km Abstand zwischen zwei Hosts, Glasfaserkabel (statt Kupferkabel) mit bis zu 100km Umfang (je Strang), Bruttodatenrate: 100Mbit/s (= zehnfache Geschw. Des Ethernets), 4B/5B Kodierung, Token-Sollumlaufzeit (TTRT) minus gemessener Zeit bestimmt die Nutzungslänge für asynchrones Senden, zusätzlich synchrones Senden.

Token-Wartung: Alle Knoten funktionieren als Monitore, Claim-Frames im Fall eines Tokenausfalls mit Gebot für TTRT, Einigung auf benötigte TTRT

Einsatz als Backbone Netzwerk, d.h. als schnelle Drehscheibe für langsamere, lokale Netze, hat sich als LAN-Technologie nicht durchgesetzt.

Drahtlose Netze

IEEE 802.11

Spektrum auf der Bitübertragungsschicht: Frequenzsprungverfahren und Direct Sequence, sowie Standard für Infrarotsignale (bis 10m)

MACA (multiple Access with Collision Avoidance): Protokoll zur Kollisionsvermeidung

Scanning für mobile Systeme:

Aktives Scanning: 1. Probe-Frame 2. Probe-Response-Frame(s) 3. Auswahl und Association-Request-Frame 4. Association-Response-Frame

Passives Scanning: Beacon-Frames des Verteilsystems

Frame-Format: 4 Adressen, um die Übertragung über Verteilsysteme zu lenken

Vermittlungsschicht

Die Protokolle definieren die Zuweisung von Adressen und die Weiterleitung von Paketen von einem Ende des Netzes zu einem anderen. Bsp. Routing

OSI 1 und 2 bieten Kommunikation zwischen direkt verbundenen Geräten, rahmenweise Übertragung, Fehlererkennung und Fehlerkorrektur, Flusskontrolle, eventuell Zuverlässigkeit

Dies bedeutet: Beschränkung auf Direktverbindungsnetze. Nur Rechner mit direkter physikalischer Verbindung sind erreichbar. Kein einheitliches Adressierungsschema, Format hängt von der Netztopologie ab.

Problem: Nicht jeder kann mit jedem direkt verbunden sein.

Lösung: Knoten dazwischen erfüllen Routingfunktion.

Aufgaben: Routing von Paketen (inkl. Finden des optimalen Wegs), transparente Datenübertragung, Bereitstellung eines einheitlichen Adressierungsschemas (für global eindeutige Adressen)

Grundsätzliche (Routing-) Optionen: Verbindungslos mit **Datagrammen**, verbindungsorientiert mit **virtuellen Verbindungen** und Source-Routing.

Netzwerktopologie

Stern

Alle leitenden Daten gehen zum Zentralrechner, der die Verteilung übernimmt.

Vorteil: Einfache Verkabelung, einfaches Routing

Nachteil: Hohe Netzlast beim Zentralrechner, Ausfallproblem

Bus

Alle an einem Kabel über das alle Daten verteilt werden

Vorteil: Direktverbindung aller Rechner, kein Routing, Ausfall einzelner Knoten unkritisch. Nachteil: Leistungsverlust

Maschennetz

Vorteil: Schnelle Verbindung, hohe Ausfallsicherheit

Nachteil: Unregelmässige Struktur, aufwändige Verkabelung

Hierarchischer Stern

Vorteil: Übersichtliche Struktur, einfach zu administrieren, leicht ausbaubar, geringer Verkabelungsaufwand und kurze Wege.

Nachteil: Hohe Last an zentralen Knoten, schlechte Ausfallsicherheit

Datagramme

Jedes Paket enthält ausreichend Information, damit jeder Switch entscheiden kann, wie er es an sein Ziel bringen kann – nämlich die vollständige Zieladresse.

Weiterleitung mittels **Weiterleitungstabelle**, deren Inhalt vom Routing bereitgestellt wird.

Eigenschaften: Pakete jederzeit an jeden Ort verschickbar, keine Information über Zustellbarkeit, Pakete werden unabhängig voneinander weitergeleitet, Ausfall eines Switches führt nicht notwendigerweise zum Scheitern.

Virtuelle Leitung

In Der Verbindungsphase muss ein Verbindungszustand aufgebaut werden.

Permanente virtuelle Leitungen durch Konfiguration durch einen Netzwerkadministrator

Vermittelte virtuelle Leitungen (switched virtual circuit, SVC) durch Signalisierung
Für jeden Verbindungszustand enthält ein Switch einen Eintrag in der VC-Tabelle: Eingangsinterface, virtual circuit identifier (VCI), Ausgangsinterface, VCI für ausgehende Pakete

Nachteil: Aufbaukosten mindestens 1RTT (Ausnahme optimistische Variante), zusätzlich Abbaukosten; virtuelle Verbindungen sind empfindlich: wenn ein Router zusammenbricht, werden alle stehenden virtuellen Verbindungen abgebrochen, auch wenn der Ausfall nur Sekunden dauert.

Vorteil: QoS-Garantien sind bei VCs möglich, Vermeidung von Überlastungen im Teilnetz

Pakete können Verbindungsnummern anstelle der vollen Zieladresse enthalten.

Gegenüberstellung

Kein Verbindungsaufbau

Jedes Paket hat volle Quell-und Zieladr.

Keine Statusinformation im Teilnetz erforderl.

Jedes Paket wird unabhängig befördert

Routerfehler provozieren nur Paketverlust

Schwierige Überlastüberwachung

Verbindungsaufbau

Jedes Paket hat eine kurze Nummer

Tabelleneintrag für jede virt. Verbind.

alle Pakete haben denselben Pfad

Routerausfall beendet virt. Verbind.

Einfache Überlastüberwachung, wenn ausreichend Puffer bereitgestellt werden.

Routing

Finden eines Weges oder Erkennung der Unmöglichkeit

Aufgabe: Über Ausgangsleitung entscheiden

Kriterien: einfach, robust, stabil, fair, effizient

Optimierungsprinzip = Dreiecksungleichung

Statisches Routing vs. dynamisches Routing, Abstandsrouting vs. Flussbasiertes Routing

Statisches Routing

Shortest Path (z.B. mit Dijkstras Algorithmus)

Flooding (z.B. für Snapshots, verteilte DBs zum Updaten). Jedes ankommende Paket wird über mehrere Ausgangsleitungen gesendet, ausser über diejenige, auf der es angekommen ist. Viele Paketduplikate werden erstellt.

Flussbasiert (Warteschlangenmodelle, M/M/1...) Nicht nur die Topologie sondern auch die Last (Verkehrsaufkommen) wird berücksichtigt. Optimaler Weg mit möglichst wenig Wartezeit wird berechnet. *Siehe Operationsresearch!!!*

Lokales statisches Routing

Statisches Routing in Teilnetzen

Hierarchisches Routing

In grösseren Netzen wachsen auch die Routing-Tabellen, die mehr CPU-Zeit brauchen für die Durchsuchung und mehr Bandbreite für die Statusmeldungen. Darum werden die Router in Regionen eingeteilt. Die Router kennen nur die innere Struktur ihrer Region. In sehr grossen Netzen werden die Regionen weiter in Cluster → Zonen → Gruppen etc. unterteilt.

Dynamisches Routing

Distance-Vector-Routing: Jeder Router verwaltet eine Tabelle, auf deren Grundlage er die am besten bekannte Entfernung zu jedem Ziel und die zu benutzende Leitung zu diesem Ziel ermittelt. Die Tabellen werden durch Austausch von Informationen mit den benachbarten Routern aktualisiert.

→ **Count-to-infinity Problem** („negative Nachrichten verbreiten sich langsamer als positive“), Lösungen unhandlich und problemhaft

Link-State-Routing: Löste das Distance-Vector-Routing ab. Verschicken lokaler Informationen und lokale Konstruktion globaler Informationen

1. Jeder Router muss seine Nachbarn entdecken und ihre Netzadressen feststellen. Dies geschieht durch Senden eines HELLO-Pakets auf jede Punkt-zu-Punkt-Leitung. Der Router am anderen Ende muss eine Antwort zurücksenden, durch die er sich zu erkennen gibt.
2. Er muss die Verzögerung oder die Leitungskosten seiner Nachbarn messen. Es wird ein spezielles ECHO-Paket ausgesendet, um die Verzögerung zu ermitteln. Die andere Seite schickt es sofort wieder zurück.
3. Der Router muss ein Paket zusammenstellen, in dem alles steht, was er gelernt hat. Periodische oder ereignisabhängige Verschickung.
4. Er muss dieses Paket an alle anderen Router senden. Mittels Flooding
5. Der Router muss den kürzesten Pfad zu allen anderen Routern berechnen.

Mobile Hosts

Home Agents und Foreign Agents

Der Fremdagent gibt periodisch die Anwesenheit und Adresse bekannt. Nach dem Betreten eines neuen Bereichs erfolgt die Registrierung (primär agenteninitiiert). Danach nimmt der Foreign Agent mit dem Home Agent Verbindung auf. Der Home Agent überprüft die Sicherheitsinformationen (inklusive Zeitstempel) und bestätigt. Der Foreign Agent informiert den mobilen Host.

Pakete werden jeweils an die Heimatadresse adressiert, nachfolgende Pakete werden zur Fremdadresse umgeleitet.

Broadcast Routing

Broadcast, Flooding, Multidestination-Routing, Spanning Tree. Multicasting setzt Gruppenmanagement voraus.

Überlastüberwachung

Überlast: Last (R'bedürfnis) > Ressourcen

Überlastüberwachung/vermeidung

Bezieht sich auf verschiedene-Sender-zu-Ressourcen-Beziehungen (Flusskontrolle auf Sender-zu-Empfänger-Beziehungen)

Unterschiedlich für best-offer und für QoS Systeme

Klassifikation nach unterschiedlichen Alternativen

Router versus host-zentrisch, Reservierung versus Feedback

Fenster- versus ratenbasiert

(meist best-effort mit Feedback und QoS mit Reservierung)

Bewertung der Qualität der Ressourcenzuteilung nach „Netzwerkleistung“ (= Durchsatz/Verzögerung), sowie „Fairness“

Parameter

Anteil verworfener Pakete

Länge der Warteschlangen

Anteil wiederholter Pakete

Paketverzögerung, Standardabweichung, ...

Reaktion

Lastreduktion oder Ressourcenvergrößerung

Massnahmen

Auf Sicherungsschicht und auf Transportschicht ähnlich

Auf Vermittlungsschicht virtuelle Verbindungen (als Voraussetzung für verschiedene Algorithmen), Warteschlangen, Verwerfen von Paketen, Routing, Lebensdauer-Verwaltung

Alternativen aus Sicht der Steuerungstheorie

Offene Schleifen: Keine Berücksichtigung des Netzzustandes

Lösungsansatz: Gutes Design

Geschlossene Schleifen: Feedbackschleife

Monitoring, Dissemination und Korrektur

Ohne Feedback

Traffic Shaping

Regulierung der durchschnittlichen Übertragungsrate

Einfacher für virtuelle Verbindungen als für Datagramme

Vorzugsweise gemeinsame Vereinbarung zwischen Sender, Empfänger und Teilnetz (Flussspezifikation) und Überwachung des Kundenverhaltens (Traffic Policing)

Leaky-Bucket-Algorithmus

Token-Bucket-Algorithmus

Beispiel einer Flussspezifikation

Merkmale der Eingabe

Maximale Paketgrösse (Byte), Token-Bucket Rate (Byte)

Token-Bucket-Grösse (Byte), Maximale Übertragungsrate (Byte/s)

Gewünschte Dienstqualität

Max Verlustrate (Byte/ μ s)

Spitzenverlust (Pakete) = Max. Anzahl konsekutiv verlorener Pakete

Minimal feststellbare Verzögerung (μ s)

Maximale Verzögerungsabweichung, d.h. Jitter (μ s)

Qualitätsgarantie

Mit Feedback

Kontrollmethoden in Teilnetzen mit virtuellen Verbindungen

Regelung des Verbindungsaufbaus (Admission Control)

Stauumleitung beim Aufbau

Vereinbarung zwischen Host und Teilnetz über Ressourcenreservierung

Choke Pakete

Hop-by-Hop Choke-Pakete

(Weighted) Fair Queueing in Routern

Load-Shedding

Priorisierung (durch Anwendungen), frühzeitiger Beginn erforderlich, unterschied...

Jitter-Kontrolle

Information über Verspätung/Verfrühung in Paketen gespeichert

Verbundnetze

Verbindungen

LAN-LAN, LAN-WAN, WAN-WAN, LAN-WAN-LAN, ...

Verbindungs-BB

Schicht 1: Repeater kopieren Bits zwischen Kabelsegmenten

Schicht 2: Bridges zwischenspeichern und weiterleiten Rahmen

Schicht 3: Mehrfachprotokollrouter übertragen Pakete zwischen ungleichen Netzen

Schicht 4: Verarbeitungsgateways ermöglichen die Netzzusammenarbeit oberhalb Schicht 4

Verkettete virtuelle Verbindungen

Vorausreservation von Puffern, garantierte Reihenfolge, kürzere Header, mehr Routineinformation, weniger Fehlertoleranz

Verbindungsloser Netzverbund

Leichter überlastet, bessere Ausgleichsmöglichkeiten, robust

Tunneling

Quelle und Ziel gleicher Netztyp

Vermittlung durch Mehrfachprotokoll-Router

Fragmentierung/Segmentierung

Paketgrenzen durch HW, BS, Protokolle, Standards, Strategien (Fehlerbehandlung, Fairness bei der Kanalbelegung)

Problem der Rekomposition (im Netz oder am Zielhost)

Unterscheidungs/Klassifikationsmerkmale

Garantierte Dienstqualität (ja, nein, welche), Fehlerbehandlung (zuverlässig, geordnet, ungeordnet), Kostenerfassung und Abrechnung, Sicherheit und Vertrauenswürdigkeit....

Flusskontrolle, Protokolle, Verbindungen ja/nein, Paketgröße, Multicasting ja/nein, flache (802) oder hierarchische (IP) Adressierung....

Einschub: ATM

Asynchronous Transfer Mode

Paketvermittletes Netz, verbindungsorientiert

Zellen konstanter Länge (5+48 kB)

Basiert auf asynchronem Zeitmultiplexen, d.h. Zellen zu beliebigem Zeitpunkt einspeisbar

Virtuelle Verbindungen mit garantierter Datenrate

End-zu-End Fehlerbehandlung

ATM Adaption Layer (AAL)

Segmentierung und Rekombination grösserer Protokolldateneinheiten

Anpassung an gewünschten Übertragungsdienst

Verschiedene Dienstklassen

CBR (constant bit rate)

VBR (variable bit rate), mit und ohne Echtzeitanforderungen

ABR (available bit rate)

UBR (unspecified bit rate)

Gesamtarchitektur entspricht verschiedenen Schichten

Physische Schicht (entspricht OSI 1 & 2)

PMD (Physical Medium Dependant): Bitzeitgabe und physikalischer Zugriff

TC (Transmission Control): Zellen-zu-Bitströmen und vice versa

ATM Schicht (entspricht OSI 2 & 3): Verwaltung, Erzeugung und Beförderung der Zellen

I.e. Flusssteuerung, Erzeugung/Extraktion der Zellenheader, Management des virtuellen Pfades, Multiplexen/Demultiplexen der Zellen

AAL Schicht (ARM Adaption Layer, entspricht OSI 3&4)

SAR (Segmentation und Reassembly): De/Segmentierung

CS (Convergence Sublayer): Bereitstellung der Standardschnittstelle

Verschiedene Sichten

Virtueller Ersatz für physikalische Leitung, normales Computernetzwerk, Zugangsprotokoll, Ersatz für Backplane eines Supercomputers

.... aber SDH (Synchronous Digital Hierachy) Hochleistungsbackbone setzt sich durch

Transportschicht

Zuverlässige Ende-zu-Ende Verbindungen

Verbindungsmanagement in unzuverlässigen Netzen

d.h. unabhängig von den Qualitäten darunterliegender Netzwerke & möglichst ökonomisch und mit garantierter Dienstgüte

d.h. z.B. Umgehen mit verzögerten Duplikatpaketen

D.h. kafkaeske Voraussetzungen

Offerierte Dienste

Aufbau, Nutzung, Abbau von Verbindungen (meist von Applikationen genutzt)

Vergleich von Schicht 2 mit Schicht 4

„Gleich“: Dienste und Funktionen

Medium: Flüchtig vs. Netzwerk als Zwischenspeicher

Dynamik: Stabiles Verhalten vs. Unberechenbar

Zustand: Absturz bedeutet Verlust vs. Probleme wegen überlebenden Teilzuständen

Aufgabe

Auf der Basis von Host-zu-Host Kommunikation eine Prozess-zu-Prozess Kommunikation realisieren

Dabei werden Standard-Basisdienste für Applikationen angeboten und die Unzuverlässigkeit des Netzwerks wird maskiert

Anforderungen „von oben“

Garantierte Übertragung von Nachrichten bei Wahrung der Reihenfolge, ohne Auslassungen und Duplikate

Unterstützung beliebig grosser Nachrichten

Unterstützung der Synchronisation zwischen Sender und Empfänger (inkl. Kontrolle des Sendeflusses durch den Empfänger)

Ermöglichung von konkurrenten Applikationen auf jedem Host

Einschränkungen von unten

Best effort Dienste mit Nachrichtenverlust/Duplikation/Umordnung, Grössenbeschränkungen, unvorhersehbar lange Verzögerungen

Transportschicht zwischen TSAP (Transport Service Access Points) und NSAP (Network Service Access Point)

Transportinstanzen verschicken virtuell TPDU's (Transport Protocol Data Unit) und geben Daten von der Verarbeitungs-/ Transportschnittstelle an die Transport-/Vermittlungsschnittstelle weiter

Netzadressen und Transportadresse (erstere meist in letzterer enthalten)

Einfache Dienstoperationen

Listen: - (blockieren, bis ein Prozess versucht, eine Verbindung aufzubauen)

Connect: Connection Request (Aktiver Versuch eines Verbindungsaufbaus)

Send: Data (...)

Receive: - (blockieren bis data TPDU ankommt)

Disconnect: Disconnection Request (...)

Berkeley Sockets

SOCKET, BIND, LISTEN, ACCEPT, CONNECT, SEND, RECEIVE, CLOSE

Dienstqualität (QoS)

Optionsverhandlung über Parameter der Dienstqualität

Dauer des Verbindungsaufbaus (connection establishment delay)

Ausfallwahrscheinlichkeit beim Verbindungsaufbau (connection establishment failure probability)

Durchsatz (throughput)

Übertragungsverzögerung (transit delay)

Restfehlerrate (residual error ratio)

Schutz (protection)

Priorität (priority)

Störausgleichsverhalten (resilience)

Protokolle

Elemente von Transportprotokollen

Adressierung
Verbindungsaufbau
Verbindungsabbau
Flusssteuerung und Zwischenspeicherung
Multiplexen
Systemwiederherstellung

Adressierung

Name: was & Adresse: wo & Route: wie
Ziel: eindeutiges Adressierungsschema für alle Rechner im Netz- Alternativen: Hierarchische versus flache Adressierung
ISO/OSI:
AFI (authority and format indicator): Typ
IDI (Initial Domain Identifier): Domäne
DSP (Domain Specific Part): lokale Adresse

Alternativen für Empfang und Lieferung

Prozessserver als Proxies für andere Server, z.B. in UNIX Initial Connection Protocol für Server, die nach Bedarf erstellt werden
Name Server oder Directory Server als Vermittler (Auflösung von Namen in physikalische Adressen)

Problem: Verzögerte Duplikate

Lösung: Beschränkung der Lebensdauer von Paketen
Teilstreckenzähler, Zeitstempel auf Paketen, ...
Folgenummern (...)

Verbindungsaufbau

Dreiweg-Handshake-Protokoll
Normal: 1. CR (seq=x) 2. ACK (seq=y, ACK=x) 3. DATA (seq=x, ACK=y)
Altes CR Duplikat: 1. CR (seq=x) 2. ACK (seq=y, ACK=x) 3. REJECT (seq=x, ACK=y)
Alte CR und ACK Duplikate: 1. CR (seq=x) 2. ACK (seq=y, ACK=x) 3. DATA (seq=x, ACK=z) 4. REJECT (ACK=y)

Verbindungsaufbau

Asymmetrischer Aufbau führt zu Datenverlust
Aber: Unlösbares Zwei-Armeen-Problem
Anwendung des Drei-Weg-Handshake-Protokolls
DISCONNECTION REQUEST mit N Timeouts
DISCONNECTION ACK mit Timeout

Flusssteuerung und Zwischenspeicherung

Senderpufferung (bei unzuverlässiger Vermittlungsschicht zwingend) versus Empfängerpufferung und dynamische vs. Statische Zuordnung von Pufferplatz
Optimaler Kompromiss zwischen Quell- und Zielpuffer hängt von der Art des Verkehrs ab (schubweiser Verkehr: Sender, hohe Bandbreite: Empfänger)
Senderpufferung und... bei Flaschenhals Netzwerk

Multiplexen

Verschiedene Transportverbindungen auf die selbe Vermittlungsverbindung aufteilen

Eine Transportverbindung auf verschiedene Vermittlungsverbindungen

Systemwiederherstellung

Kein immer funktionierendes Protokoll

Absturz kann zwischen Weitergabe und Bestätigung oder zwischen Bestätigung und Weitergabe passieren

Ausfälle auf Schicht N nur auf Schicht N+1 behandelbar, und nur, wenn ausreichend Statusinformation vorhanden

Einschub

Bezeichnungen

LAN = Local Area Network (Ethernet, Token-Ring, FDDI, LATM)

MAN = Metropolitan Area Network (i.A. 100-200 km Durchmesser)

WAN = Wide Area Network (X.25, Frame-Relay, IP, ATM)

SAN = System Area Network (auch Storage Area Network)

Eigenschaften

Grösse des zugrundeliegenden Netzwerks schränkt verwendbare Technologien ein
Anwendung von SANs im Cluster Computing, d.h. bei der Koppelung von (Arbeitsplatz-) Rechnern zu Funktions-, Daten- und Lastverbänden.

Internetworking

Internetwork

= Sammlungen von zusammengeschalteten autonomen Systemen, die Pakete zwischen Hosts übertragen.

Probleme: Heterogenität und Skalierbarkeit.

TCP/UDP-IP Protokollhierarchie

Anwendungsschicht: TelNet, FTP, SMTP, HTTP, DNS, NFS, ...

Transportschicht: TCP, UDP

Internetschicht (OSI 3): IP

Netzzugangsschicht: Paketorientierung, zuweilen gesicherter, Datentransport zu direkt verbundenen Knoten

Internet Protokoll (IP)

Zur Beförderung von Datagrammen (Header und Daten) von der Quelle zum Ziel. Einheitliche Adressierung (IP-Adressen). Fragmentierung und Zusammensetzung über langer Datagramme. Protokolle zum Unterhalt konsistenter Routinginformationen. Methoden zur Linderung von Überlast. Generierung von Fehlermeldungen.

Keine quantitativen oder qualitativen Garantien, keine Bestätigung, keine Reihenfolge-treue, kein Schutz vor Duplizieren.

Dienstmodell (IP)

Adressierungsschema + „best effort“ Datagramm-Modell

IP Header (20Byte fix + optionaler Teil variabler Länge (max 40 Byte))

Version (4 Bit), d.h. 0100 = Version des Protokolls, zu dem das Datagramm gehört

Header Length (4 Bit): Da die Header Länge nicht konstant ist, wird das Header-Feld IHL bereitgestellt.

Type of Service (TOS, 8 Bit), mehrfach verändert, nicht standardisiert interpretiert. Der Host kann dem Teilnetz den gewünschten Dienst bekannt geben. Das Feld selbst enthält das Feld Precedence (3Bit, ist eine Priorität von 0 bis 7), drei Flags D, T, R (sie ermöglichen dem Host, auf was er am meisten Wert legt (Verzögerung, Durchsatz, Zuverlässigkeit)) und zwei unbenutzte Bit.

IP Packet Length (Länge des Datagramms, max 65535 Byte, 16 Bit). Beinhaltet das komplette Datagramm – Header und Daten.

Fragmentnummer etc. (32):

Packet **Identification** Number (16). Ist erforderlich, damit der Zielhost feststellen kann, zu welchem Datagramm ein neu angekommenes Fragment gehört. Alle Fragmente eines Datagramms erhalten den selben Identification-Wert.

Don't Fragment Bit: das Ziel ist nicht in der Lage, es wieder zusammensetzen.

More Fragment Bit: bei allen Fragmenten ausser dem letzten gesetzt. Es wird benötigt, um festzustellen, wann alle Fragmente eines Datagramms angekommen sind.

Fragment Offset (13, theoret. Max 8192 Fragmente pro Datagramm). Bezeichnet, an welcher Stelle im Datagramm ein Fragment gehört.

Time to live (TTL, 8): Zähler, bei dem die Lebensdauer von Paketen begrenzt werden kann. Max. Lebensdauer 255 sec.

Protokoll (TCP oder UDP, 8): Die Vermittlungsschicht kann erkennen, an welchen Transportprozess das Datagramm weiterzugeben ist.

Prüfsumme (16): Prüft nur den Header. Ist nützlich zum Erkennen von Fehlern, die durch falsche Speicherwörter in einem Router erzeugt wurden.

Absender-IP-Nummer (32), **Ziel-IP-Nummer** (32): Bezeichnen die Netz- und Hostnummer.

Optionen+Füller (32n): Wurde ausgelegt, um späteren Versionen des Protokolls zu ermöglichen, Informationen zu beinhalten, welche im urspr. Design nicht vorhanden sind. Bsp: **Sicherheit** (wie geheim ist das Datagramm), **Strict Routing** (bestimmt kompletten Pfad), **Loose Source Routing** (Gibt eine Liste von Routern aus, die nicht zu verfehlen sind), **Record Routing** (veranlasst jeden Router, seine Adresse anzuhängen), **Time Stamp** (dito + Zeitstempel)

Bild S. 443, Tanenbaum

IP Adressen (32 Bit)

Hierarchisch, d.h. Netzwerkanteil und Rechneranteil

Form a.b.c.d, $-1 < a, b, c, d < 256$

126 Class-A-Netze mit je ca. 16 Mio Rechner-Nr.

16000 Class-B-Netze mit je ca. 65000 Rechner-Nr. ($127 < a < 192$, 16Bit)

2 Mio Class-C-Netze mit je 254 Rechner-Nr. ($192 < a < 224$, 8 Bit Rechneranteil)

Sonderadressen

0.0.0.0 beim Start

0. bzw. 0.0 bzw. 0.0.0 aktuelles Netz

255.255.255.255 Broadcast im lokalen Netz

127. Schleifentest

die Netznummern werden vom NIC zugewiesen, um Konflikte zu vermeiden.

ICMP (Internet Control Message Protocol)

Transport von Status- und Fehlermeldungen über IP.

Bei: Datagrammen an unerreichbare IP-Nummern, zu grosse Pakete, Zeitüberschreitung, nicht ansprechbarem Zielprotokoll, falsch ausgefülltem IP-Header, Umleitungsinformationen, Anfrage ob eine Maschine am Leben ist und antworten darauf...

Nicht bei: Datagrammen an eine Broadcast-Adresse, Transport von ICMP-Paketen, einem Fehler in der IP-Prüfsumme

Anwendungsbeispiele: Ping, Traceroute

Bild S. 450, Tanenbaum

ARP (Address Resolution Protocol)

Zuordnung von IP-Nummern zu Ethernet-Nummern, weil die Hardware auf der Sicherungsschicht Internet-Adressen nicht versteht.

ARP Request als Ethernet-Broadcast an das lokale Netz. ARP-Reply des angesprochenen Rechners

RARP (Reverse Address Resolution Protocol)

Ermitteln der eigenen IP-Nummer. Z. B. beim Starten einer plattenlosen Workstation. Die Ethernetadresse wird gesendet und man fragt, ob jemand die IP-Adresse dazu kennt. Nachteil: Zieladresse, die vollständig aus 1en besteht → RARP-Server ist nötig, da die Nachrichten nicht über Router weitergegeben werden.

Alternative BOOTP, das UDP-Nachrichten benutzt, die über Router verteilt werden.

DHCP (Dynamic Host Configuration Protocol)

Zur Reduktion der Administrationskosten hat man ein Server mit Information über die Zuteilung von IP-Adressen, als C/S über UDP/IP

Beim Aufstarten fragt ein Rechner um seine IP-Adresse an, in jedem Netzwerk mindestens ein Relaisagent, der die Anfrage weiterleitet. Bei nicht fixen IP-Adressen wird eine dynamische IP-Adresse auf Miete zugeordnet.

Internes Gateway Routing (innerhalb eines autonomen Systems)

Grossteils **OSPF Protokoll** (Open Shortest Path First, 1990): unterstützt 3 Arten von Verbindungen: Punkt-zu-Punkt Leitungen zwischen 2 Routern, Mehrfachzugriffsnetze (Mehrere Router sind angeschlossen, jeder kann direkt mit allen anderen kommunizieren) mit und ohne Broadcasting. Urspr. Distanz-Vektoren-Protokoll (count-to-infinity etc. Probleme)

Anforderungen an den OSPF-Entwurf: Offen, Unterstützung für verschiedene Entfernungsparameter, dynamisch und selbstadaptiv, Lastausgleich, Unterstützung für hierarchische Systeme, Sicherheit, Unterstützung für Tunneling.

OSPF-Lösungskonzept: Abstraktion der tatsächlichen Netze, Router und Leitungen in einen gerichteten Graphen mit kostengewichteten Kanten und Berechnung des kürzesten Pfads. Partielle, nichtüberlappende Überdeckung mit transparenten (bb) Bereichen, jedes autonome System hat Backbone-Bereich, an den alle Bereiche angeschlossen sind. Innerhalb eines Bereiches hat jeder Router die gleiche Link-State-Datenbank, welche den gleichen Algorithmus (des kürzesten Pfades) ausführt, wofür mehrere Graphen (Verzögerung, Durchsatz, Zuverlässigkeit) verwendet werden.

4 Routerklassen: interne (die gänzlich zu einem Bereich gehören), Router an Bereichsgrenzen (verbinden 2 oder mehrere Bereiche), Backbone Router (befinden sich

am Backbone), AS-Grenz-Router (vermitteln zwischen mehreren autonomen Systemen).

Hello Nachrichten werden beim Starten vom Router an alle seine Punkt-zu-Punkt Leitungen gesendet. Jeder Router speist gelegentlich Nachrichten vom Typ **Link-State Update** in alle angrenzenden Router. Diese Nachricht gibt den Status des sendenden Routers und die in der Topologiedatenbank benutzten Kosten an. Es sind zuverlässige Nachrichten mit fortlaufenden Nummern. **Link-State Ack** bestätigt eine Link-State Aktualisierung. **Database Description** Nachrichten geben die Folgenummern aller Link-State Einträge aus, die momentan beim Sender vorhanden sind. Mittels **Link State Request** können beide Partner Link-State Informationen voneinander anfordern. So wird geprüft, wer die aktuellen Daten hat.

Externes Gateway mit BGP (Border Gateway Protocol)

Zwischen mehreren autonomen Systemen

Muss gewisse Regeln beachten (politische, sicherheitstechnische, wirtschaftliche):

Kein Transitverkehr durch bestimmte autonome Systeme. Vom Pentagon ausgehender Datenverkehr darf nie über den Irak übertragen werden, Datenverkehr, der bei IBM beginnt und endet darf nicht über Microsoft führen...

Berücksichtigt nur andere BGP Router. 3 Kategorien von Netzen: **Stub-Netze** (haben nur eine Verbindung zum BGP-Graphen, können nicht für den Transitverkehr genutzt werden), **Mehrfachanschlussnetze** (Sollten für den Transitverkehr benutzt werden, soweit sie diesen nicht ablehnen) und **Transitnetze** (Sind bereit, Pakete von Dritten zu befördern, mit Einschränkungen). BGP=Abstandsvektorprotokoll: Buchführung über Pfade, kein Count-to-Infinity Problem.

Internet Multicasting

Beispiele permanenter Adressen: 224.0.0.1 alle Systeme in einem LAN, 224.0.0.2 alle Router in einem LAN, 224.0.0.5 alle OSPF Router in einem LAN

Mobiles IP (Ziele des Drafts für Ipv6)

Ziele: Jeder mobile Host muss seine Heimatadresse überall benutzen können. Keine Softwareänderungen in festen Hosts. Keine Änderungen der Router-Software und -Tabellen. Möglichst wenig Umwege bei der Übertragung. Kein Overhead, wenn der mobile Host an seinem Heimatort steht.

CIDR (Classless InterDomain Routing)

Problem: zu wenig Adressen

Lösungsansatz: Unterteilung der Klasse B Netze in Blöcke und Zonenunterteilung (der Welt) in der Klasse C, sowie 32-Bit Adressmasken, um eine Explosion der Routingtabellen zu verhindern.

DNS (Domain Name Service)

Auflösen von Namen in IP-Adressen (Domain-Namen sind benutzerfreundlich & Domain-Namen können bei einer Migration beibehalten werden). Als C/S Dienst über TCP/IP, UDP/IP realisiert, ergänzt durch selbstfalsifizierendes Cache-Protokoll.

Ipv6