

TCP/IP Architektur und TCP/IP Services for OpenVMS

Karl Bruns
Trainer/Consultant

OpenVMS and Networking
OSI, DECnet, X.25 and TCP/IP

Lessingstr. 1
D-86438 Kissing
Phone +49/8233/2938
Mobile +49/1717168148

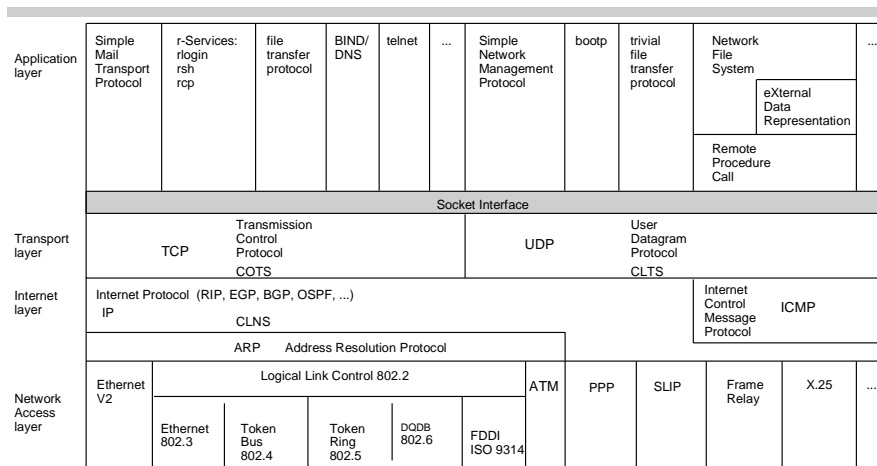
www.kbruns-training.de
karl.brun@web.de

DECUS IT - Symposium

4. April - 8. April 2005

Düsseldorf

TCP/IP-Architektur



Karl Bruns

TCP/IP Schichtenmodell

- 4 Schichtenmodell
- OSI : 7 Schichtenmodell
- Network Access Layer beinhaltet Physical Layer und Datalink Layer
- Application Layer beinhaltet OSI Session, OSI Presentation und OSI Application

HP TCP/IP Services for OpenVMS

- OpenVMS Implementierung der Industriestandard-architektur TCP/IP, Internet Services for OpenVMS Industry Standard 64 (I64) und OpenVMS Alpha
- Basierend auf Tru64 UNIX 5.1B, Kernel portiert von Berkeley Software Distribution BSD 4.3 und 4.4 mit Erweiterungen von HP
- Früherer Name : UCX Ultrix Connection
- Ab OpenVMS 7.0 : TCP/IP Services for OpenVMS
- OpenVMS 7.3-2 : Version 5.4
- OpenVMS 8.2 : Version 5.5
- Unterstützt IPv4 sowie IPv6

TCP/IP Services: Network Access Layer

- Ethernet für IPv4 und IPv6 (OpenVMS I64 und OpenVMS Alpha)
- FDDI für IPv4 und IPv6 (OpenVMS Alpha)
- Token Ring für IPv4 (OpenVMS Alpha)
- ATM für IPv4 (OpenVMS Alpha)
- PPP für WAN Verbindungen über serielle Schnittstellen

Network Access Layer: LAN

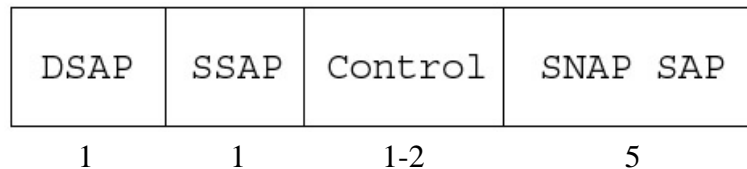
- IP über Ethernet V2

RFC894: A Standard for the Transmission of IP Datagrams
over Ethernet Networks
Ch. Hornig, April 1984

Type field %x0800

Network Access Layer: LAN

- Logical Link Control 802.2
 - Typenfeld im Ethernet V2 wird ersetzt durch ein Längenfeld
 - Einheitliche Struktur sämtlicher LAN entities
 - Service Access Point (SAP) ist Bestandteil von LLC
 - Aufbau des Headers :



Network Access Layer: LAN

- LLC1: Connection Less Service

Control: 03

DSAP = SSAP = %xFE ISO Internet 8473

DSAP = SSAP = %xE0 Novell IPX

DSAP = SSAP = %xF0 Netbeui

DSAP = SSAP = %x42 Spanning Tree

SNAP SAP adressiert herstellerpropertäre

Protokolle: 3 byte Herstellerkennung

2 byte Typenfeld aus Ethernet V2

z.B. DEC LAT 08-00-2B-60-04

Network Access Layer: LAN

- LLC2: Connection Oriented Service

einzigste Anwendung: X.25 Packet Layer Protocol

Bedeutung: IBM Token Ring
X.25 over LAN (CONS)

Keine Verwendung in TCP/IP

Network Access Layer: LAN

- IP über Ethernet 802.3

RFC948: Two Methods for the transmission of IP Datagrams over IEEE 802.3 Networks
Ira Winston, June 1985

1. Ethernet Compatible Method
„length field in a private manner“

2. IEEE 802.2 / 802.3 Compatible Method
DSAP=SSAP=96
Control=03

Network Access Layer: LAN

- IP über Ethernet 802.3

RFC1042: A Standard for the Transmission of IP Datagrams over
IEEE 802 Networks
J. Postel, J. Reynolds, Februar 1988

DSAP=SSAP=%xAA
Control 03
SNAP SAP=%x0000000800

Network Access Layer : WAN

- Point to Point Protocol PPP

- RFC 1171,1172,1331,1548,1661,1662
W. Simpson, Daydreamer, July 1994

- PPP has three components:

1. A method for encapsulating datagrams over serial links. PPP uses HDLC as a basis for encapsulating datagrams over point-to-point links. At this time, PPP specifies the use of asynchronous or synchronous duplex circuits, either dedicated or circuit switched

- .

2. An extensible Link Control Protocol (LCP) to establish, configure, and test the data-link connection.

- 3. A family of Network Control Protocols (NCP) for establishing and configuring different network-layer protocols. PPP is designed to allow the simultaneous use of multiple network-layer protocols.

PPP Protokollfeld

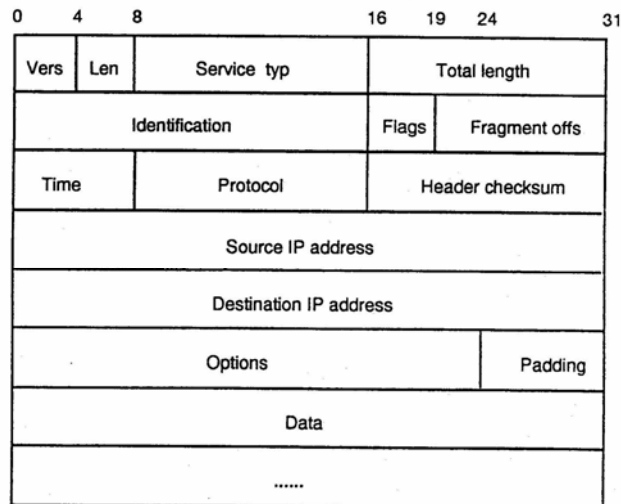
- PPP führt ein 2 byte Protokollfeld:

0021 IP
0023 OSI
0027 DECnet
002b IPX
0207 CDP
0031 Bridging PDU
003d Multilink PPP
8021 IPNCP
8023 OSINCP
802b IPXNCP
c021 LCP
c023 PAP
c223 CHAP

Internet Protocol IP

- RFC 791
- IPv4
- Protokoll zur Adressierung der Knoten
- IP adressiert Interfaces und nicht Knoten !
- Länge der Adresse 4 Byte: z. B. 151.15.60.231
- Adressraum: 4 Billionen
- Connection Less Network Service (CLNS)
- Kann fragmentieren, falls notwendig
- Wege finden über die Routingtabelle

IPv4 Datagram



Routing in TCP/IP

- Historisch: Gateway = Router
heute noch: Default Gateway
- Statisch/Dynamisch
- Dynamisches Routing innerhalb einer Domäne: Internet Gateway Protocol (IGP):
RIP RFC 1058 RIPv1, RIPv2
Integrated IS-IS RFC 1195
OSPF RFC 1131
- Es werden keine ES-IS Hello Messages ausgetauscht !
- Dynamischen Routing zwischen Domänen: Exterior Gateway Protocol (EGP)
EGP
BGP-4

TCP/IP Services: Internet Layer

- IP und ICMP für IPv4 und IPv6
- ARP nur für IPv4
- Neighbor Discovery Protocol findet Knoten im gleichen Segment, entdeckt Router, bildet IP-Adressen auf MAC Adressen ab für IPv6 und liefert Informationen über Wege zu aktiven Nachbarn
- Path MTU Discovery unterstützt TCP und UDP um IP Fragmentierung zu verhindern
- CIDR (Classless Inter Domain Routing) reduziert die Routingtabellen und erlaubt mehr Subnetze

TCP/IP Services: Routing

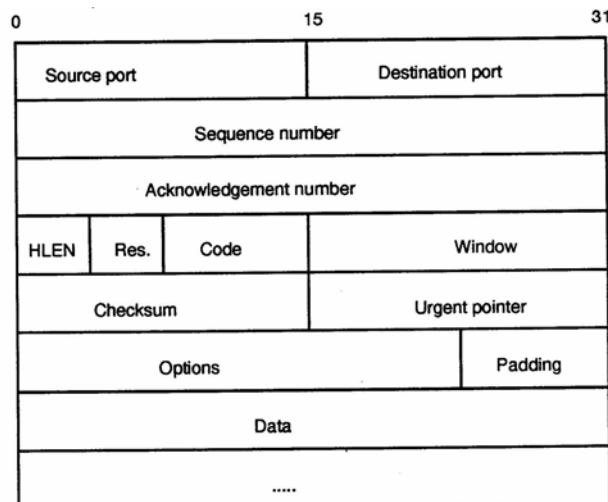
- RIP V1 und V2 für IPv4, RIPng für IPv6
- OSPF für IPv4
- EGP und BGP für IPv4
- Router Discovery für IPv4
- Routing Daemons werden dazu benötigt:
ROUTED und GATED für IPv4
RIPv2 benötigt GATED
TCPIP\$CONFIGURE erlaubt die Auswahl des Daemon
Achtung: GATED erzeugt ein Template, das editiert werden muss !

ip6rtrd für IPv6

TCP Transmission Control Protocol

- RFC 793
- TCP ist ein Transportprotokoll mit der Garantie einer fehlerfreien Übertragung
- Connection Oriented Transport Service (COTS)
- Segmentierung, Wiederausammensetzung, Paketnummerierung, Bestätigung,
- Absprache der Paketgrösse über option field
- Verbindungsaufbau über SYNC, SYNC-ACK und ACK

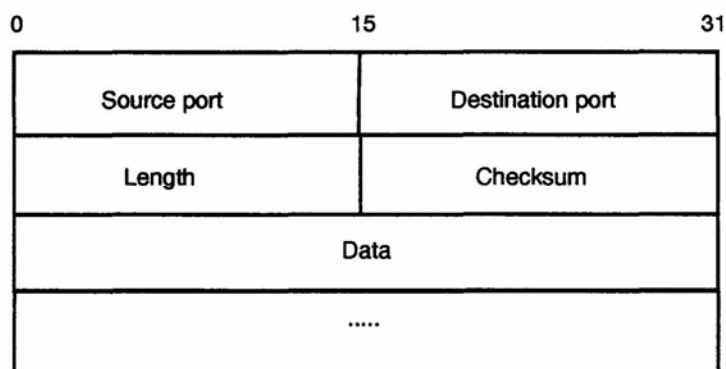
TCP Segment



User Datagram Protocol (UDP)

- Ungesichertes Transportprotokoll (CLTS)
- RFC 768
- Wird verwendet von Applikationen, die geringe Mengen Daten übertragen
- Garantie einer fehlerfreien Übertragung wird von der Applikation geliefert, z. B. bootp, tftp, DHCP
- Einzige Aufgabe: Portnummer um Applikation zu adressieren

UDP Packet



Socket Interface

- Programmierschnittstelle (API)
- Socket:

Source IP Address

Destination IP Address

Protocol Identifier

Source Port Number

Destination Port Number (well known port)

TCP/IP Services : Application layer

- Remote Computing: TELNET, IBM TN3270, rlogin
- Kerberos Authentication, Secure Shell (SSH)
- File Transfer: FTP, TFTP, rcp, SCP, SFTP
- FTP unterstützt OpenVMS Extended File Specifications für ODS-5 Platten
- Remote Printing mit LPR/LPD und TELNET print symbiont
- Network File System: NFS Client hat transparenten Zugriff zu Remote File Services, d.h. Dateien und Verzeichnisse sind wie lokale unabh. vom Betriebssystem, Hardware, etc. sichtbar.
- TCP/IP Services unterstützt NFS V2 und V3 für NFS Server, NFS Client V2, Server und Client nur für OpenVMS Alpha und IPv4

TCP/IP Services : Application layer

- NFS V3 unterstützt bis zu 1 Terabyte für eine OpenVMS Datei und 4 Gigabyte für ein Container File System.
- NFS Server kann über TCP oder UDP konfiguriert werden, für WAN Verbindungen wird TCP empfohlen.
NFS Client unterstützt nur UDP.
- NFS Server und Client unterstützen OpenVMS Extended File Specifications auch für ODS-5 Dateistrukturen.
- Mail Services: SMTP verwendet OpenVMS Mail Utility, folgende Mail Systeme arbeiten mit SMTP als mail repositories: Post Office Protocol (POP), IMAP Mail Server (nur für Alpha)
neu: Secure IMAP, Secure POP (beide benötigen SSL)

TCP/IP Services : Application layer

- Network Time Protocol: Zeitsynchronisation mit NTP V4.2 auch für IPv6
Verbesserung: NTP braucht nun nur noch 20 sec um 1 Sekunde zu korrigieren im Gegensatz zu 30 Minuten vorher !
- Management mit SNMP V2 beinhaltet einen Master Agent und Subagents für eine Management Station im Netzwerk
- Berkeley Internet Name Domain ist eine verteilte Datenbasis für die Übersetzung der IP Knotenamen auf IP Adressen und ersetzt die lokale Hostdatei. Der BIND 9 Server (nur Alpha) basiert auf Internet Software Consortium's (ISC) BIND 9.2.1 .
Dynamische Updates, BIND Server Cluster failover (mehrere BIND Master Server greifen auf eine gemeinsame DB zu)

TCP/IP Services : Application layer

- Load Broker ermöglicht Cluster Load Balancing, round robin scheduling und automatic failover (wie LAT schon vor 20 Jahren)
- Metric Server berechnet den DNS Load des Knoten, der mit logischen Namen beeinflusst werden kann.
- Portmapper kann Ports für Client Applikationen liefern
- Dynamic Host Configuration Protocol (DHCP) als Erweiterung von bootp liefert IP-Adressen, Masken, Default Gateway etc. für bootp clients. DHCP Server sowie Client werden unterstützt.
- DHCP kann dynamisch Einträge in BIND durchführen.
- Management Control Program: leider immer noch das alte UCX unter neuem Namen TCPIP ! Wünschenswert : NCL
- Tipp: verwende UNIX Kommandos, dazu
`$ @SYSS$MANAGER:TCPIP$DEFINE_COMMANDS`

TCP/IP Services : UNIX Kommandos

- Arp
- Finger
- Ifconfig
- Netstat
- Ping
- Ripquery
- Route
- Sysconfig
- Sysconfigdb
- Tcptrace
- Traceroute
- tcpdump

TCP/IP Services 5.4 : New Features

- **failSAFE IP:** redundante IP Adressen
- Verschiedene Interfaces erhalten die gleiche IP-Adresse
- OpenVMS Distributed Lock Manager stellt sicher, dass nur eine zu einer Zeit aktiv ist im Cluster
- Standby IP Adressen können auf den Interfaces eines Knotens oder im Cluster konfiguriert werden
- Ziel: Ausfall des Controllers, Verbindungsunterbrechung, fehlerhaftes Kabel oder toter Port am Switch soll erkannt werden
- failSAFE Service überwacht den Zustand jedes Interface über den Counter „Bytes received“
- Im Fehler werden IP Adressen und statische Routen des Interfaces entfernt, Standby Adressen werden aktiviert und die Routen auf andere Interfaces migriert

TCP/IP Services 5.4 : New Features

- Secure shell (SSH) v2 Client und Server Version 3.2 für secure login, rsh und Dateitransfer
- Secure Socket Layer (SSL) für POP
- ISC BIND 9.2.1 Server
- Tcpcdump
- IPv6 Software Update

TCP/IP Services V5.4: Scalable Kernel

- Klassischer TCP/IP Kernel ist Single Threaded
- Eine Operation in einer Zeiteinheit kann ausgeführt werden
- Synchronisation verwendet IOLOCK8 global spinlock
- Contention mit anderen IOLOCK8 Benutzern (DECnet,SCS,...)
- Keine gute Performance in Multi-CPU System
- Scalable Kernel: Skalierung linear mit hoher Anzahl CPUs
Reduzierung von MPSYNCH nahe 0, wenig CPU Contention
keine Verwendung von IOLOCK8
Transmit und Receive in Parallelverarbeitung
- Einschalten vor dem Start von TCP/IP:
DEFINE/SYSTEM/EXEC TCPIP\$STARTUP_CPU_IMAGES
„PERF=ALL“

TCP/IP Services V5.4: Performance

- BG Device handling: über 10000 BG devices
sinnvoll für Web Server
sysconfig: ovms_unit_maximum > 9999 (<32K)

dazu: faster UCB creation and deletion
sysconfig: ovms_unit_fast_credel einschalten
- Telnet/rlogin Driver Performance
Multi-CPU systeme verwenden keine IOLOCK8
Support for multiple concurrent I/O thru TN Devices
Overhead für Verwaltung der TN Devices wurde reduziert

TCP/IP Services V5.4: Performance

- NFS Verbesserungen:
- NFS Server kann Directory und Name Caches effizienter einsetzen
SYSCONFIGTAB.DAT ovms_xqp_plus_enabled
- Internal Performance Improvements
neue Hashalgorithmen
verbesserter Code
Erhöhung der NFS threads
- NFS kann eigene CPU in Multi-CPU Systemen verwenden

TCP/IP Services V5.5 : New Features

- IPv6 Updates: Konfiguration von 6to4 tunnels, Routen für 6to4 relay router, automatic tunnels, IPv6 over IPv6 manual tunnels und manuelle Routen
- failSAFE Support für IPv6
- PWIP Driver Support für IPv6
- NFS Server unterstützt case-sensitive file lookups:
für OpenVMS-to-OpenVMS mode: \$ SET PROC/CASE_LOOKUP
- NFS symbolic links support für Client und Server
- NTP V4.2: unterstützt auch IPv6 Adressen
- TCPDUMP V3.7.2: Upgrade von V2.2, verwendet libpcap V0.8.3 API
siehe www.tcpdump.org
TCPIP HELP TCPDUMP
- TCP/IP Kernel: Scalable Kernel ist nun der Standard Kernel

TCP/IP Services Zukunft

- IPSEC EAK H2 CY04
 - Updated SSH, SSL, Kerberos
 - BIND 9 resolver
 - FTP large file support > 2 GB files
 - Weitere IPv6 Funktionalitäten
-
- HP OpenVMS Network Transport Home Page:
<http://www.hp.com/products/OpenVMS>
 - Product Management: Lawrence.Woodcome@hp.com