

---

# Planung von Netzen mit MPLS VPN

---

Herbert Almus  
TU Berlin / EANTC AG

Benutzergruppe Netzwerke  
Herbsttagung, 27.-28.11.2003

# Unternehmensprofil EANTC AG

EANTC ist ein hochspezialisierter Dienstleister im Bereich Netzwerktechnologien. EANTC bietet unabhängige Tests zur Qualitätssicherung in Computernetzen.



EANTC, Berlin-Charlottenburg

## EANTC Geschäftsbereiche

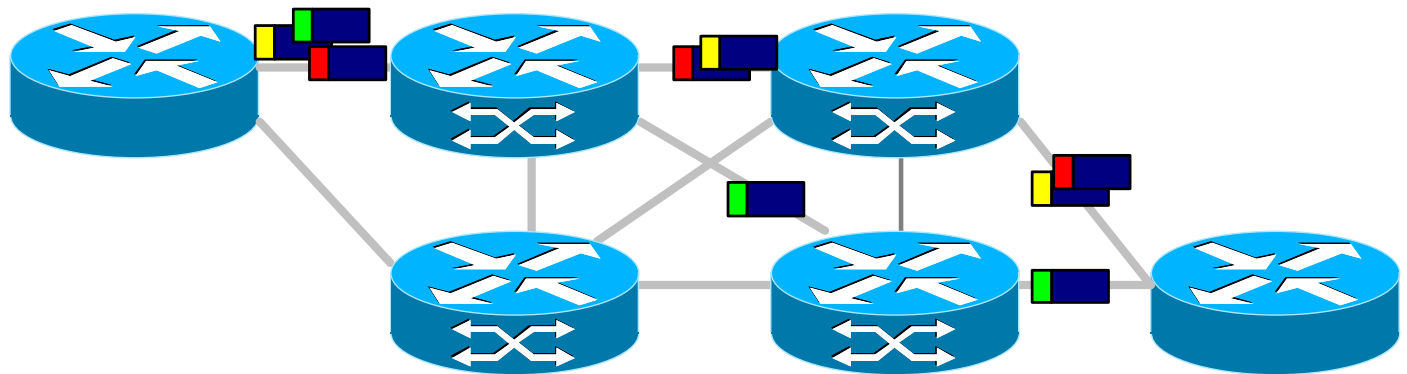
- Tests und Zertifizierung von Netzkomponenten für Hersteller
- Tests von High-Speed Enterprise / Service Provider Netzen und Netzwerk-Design Beratung
- Forschung und Entwicklung von Testmethoden und Analyse-Tools
- Herstellerunabhängige Technologieseminare (MPLS, Voice over Broadband, ATM)

Spin-off aus einem Forschungsschwerpunkt der TU Berlin  
Aktiengesellschaft seit 1999

Kooperationsvertrag regelt die Zusammenarbeit mit der TU Berlin

# Was ist MPLS?

- MPLS realisiert verbindungsorientiertes Label-Switching auf Basis von IP Routing und (zusätzlichen) Kontroll-Protokollen
  - Fokussiert auf IPv4 (derzeit)
  - Aber offen für andere Protokolle, z.B. IPv6, IPX, Appletalk, ....
- Nicht begrenzt auf eine bestimmte Layer 2 Technologie
  - Kann verstanden werden als Zwischenschicht (shim layer) zwischen L2 und L3 Protokollen,
  - Funktioniert über ATM, Ethernet, Frame Relay



# MPLS – Konzept

Route at the edge, switch in the core

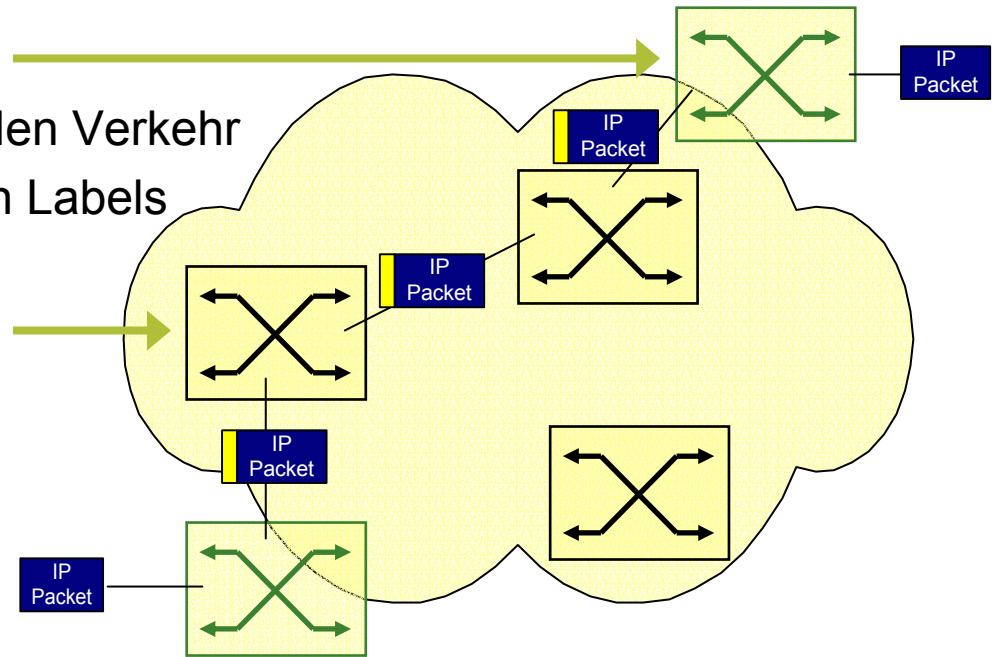
Stelle Paketen am Netzwerkübergang einen Prefix (Label) voran.

Label Edge Router (LER)

- ❑ Analysiert und klassifiziert den Verkehr
- ❑ Anforderung/Zuweisung von Labels

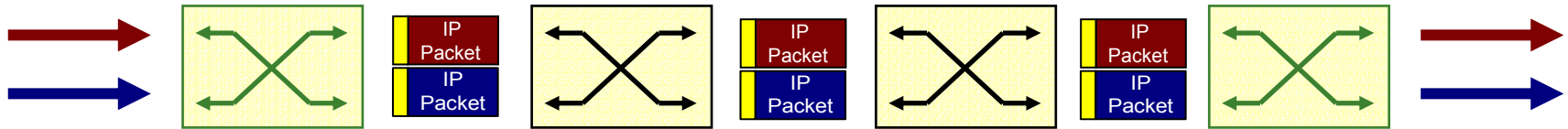
Label Switch Router (LSR)

- ❑ Label-Swapping und Forwarding im MPLS Netzwerk



# Forward Equivalence Class (FEC)

- Eine FEC spezifiziert eine Untermenge von Paketen, die durch einen Router in identisch behandelt werden.



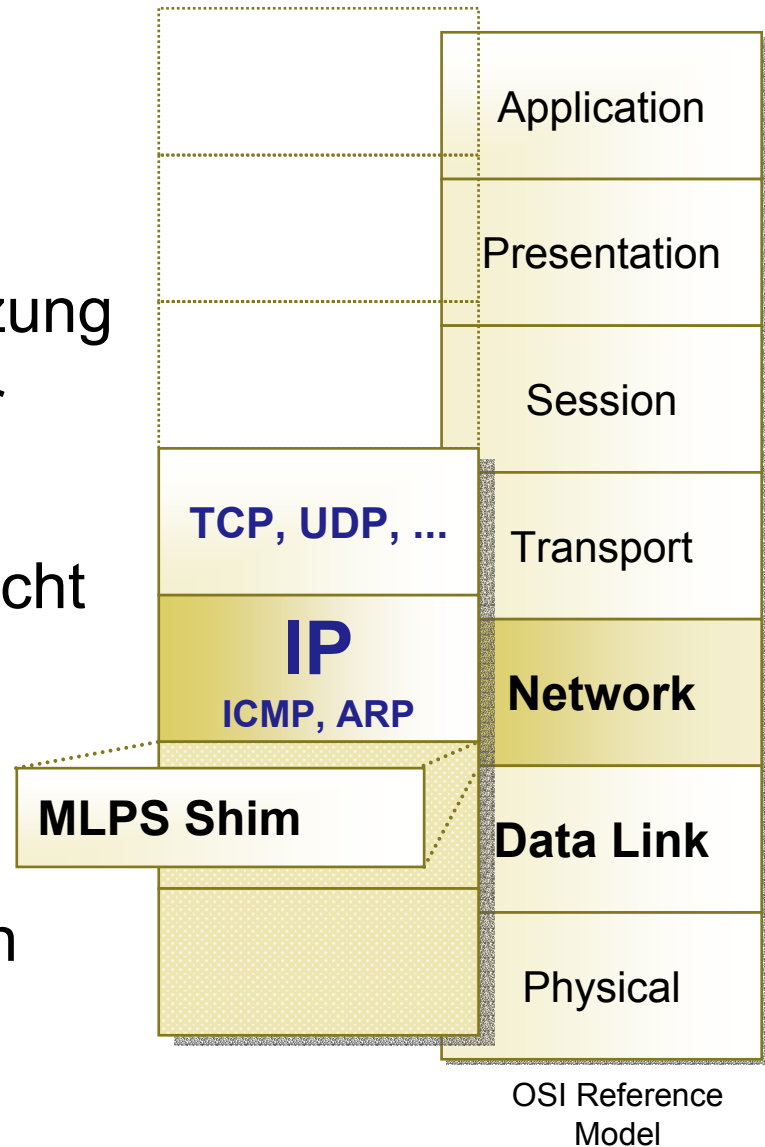
- Beispiele:
  - Bündelung von Paketen verschiedener Quellen und Zielen in eine FEC
  - Aufteilung des Verkehrs zwischen einer Quelle und einem Ziel anhand von QoS-Anforderungen

# MPLS Label

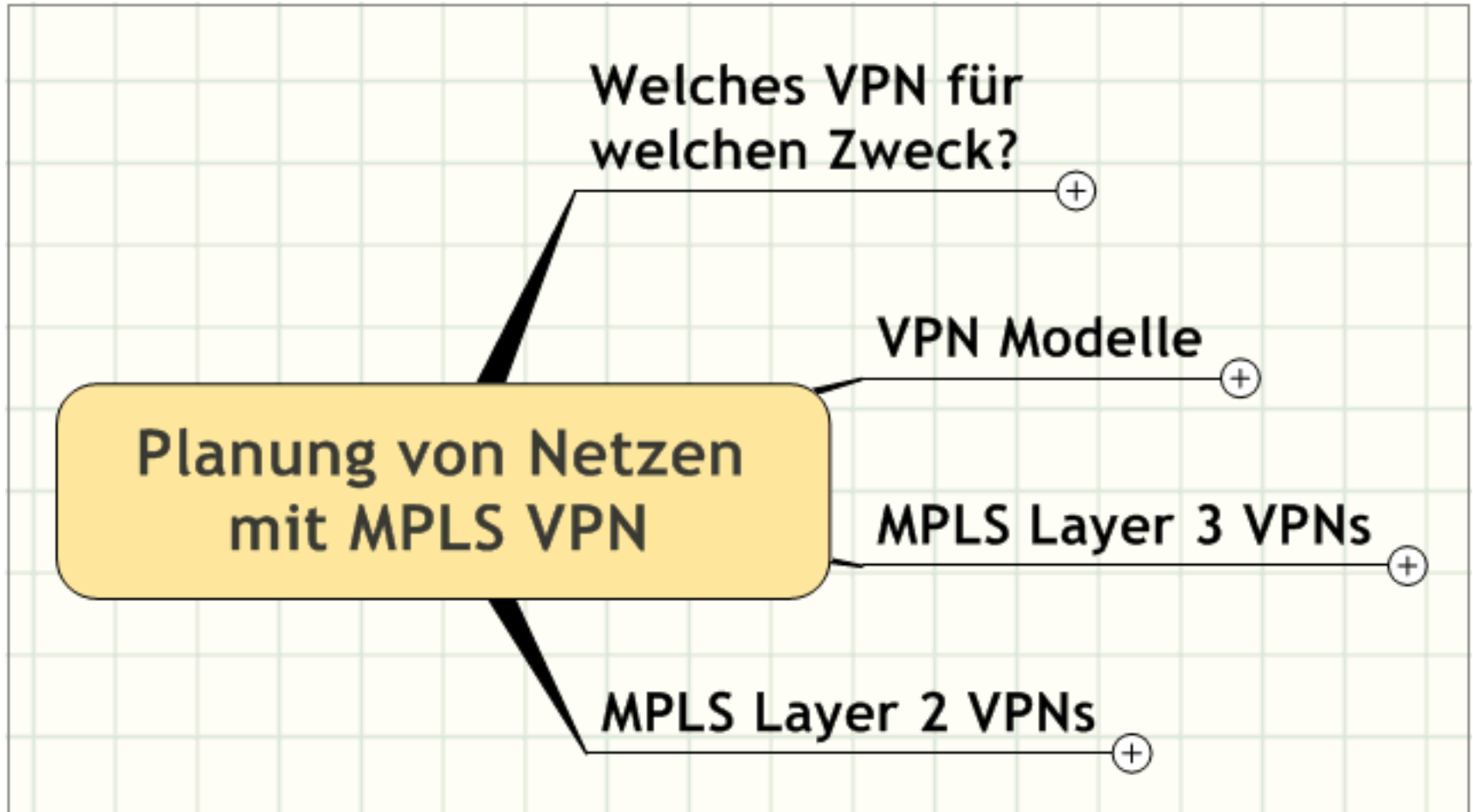
MPLS ist konzipiert für die Nutzung über verschiedenste Link Layer

Label können als Zwischenschicht (shim layer) zwischen L2 und L3 betrachtet werden

Label sollen integriert werden in die spezif. Eigenarten des verwendeten L2



# Planung von Netzen mit MPLS VPN



# Welches VPN für welchen Zweck?

**Welches VPN für welchen Zweck?**

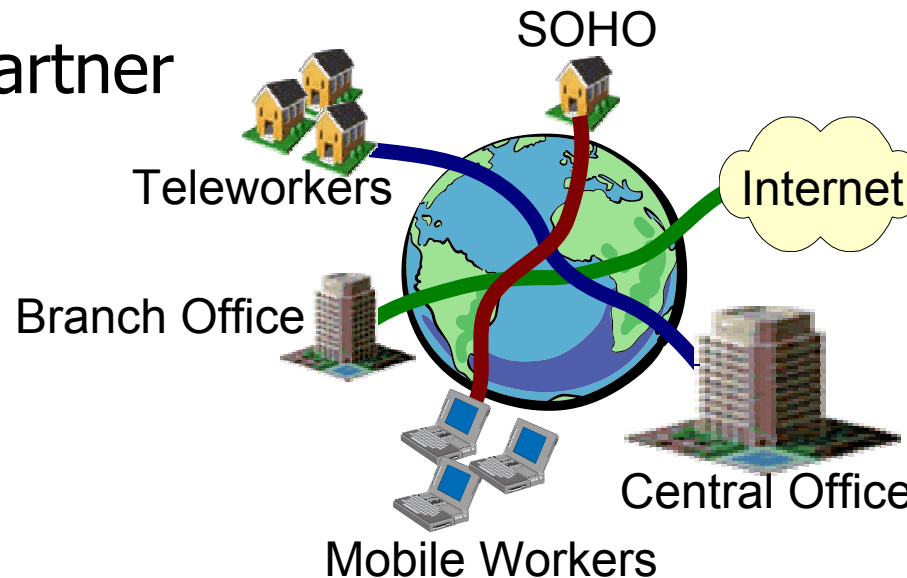
VPN Anwendungsbereich / was ist die primäre Nutzung? ⊕

Was sind die angestrebten Geschäftsziele? ⊕

Anforderungen / SLA ⊕

# VPN Anwendungsbereich / was ist die primäre Nutzung?

- MAN / WAN Intranet
  - Verbindung von Firmenstandorten
- Extranet
  - Zugang für Geschäftspartner
  - SOHO
- Access-Network
  - SOHO
  - Telearbeit

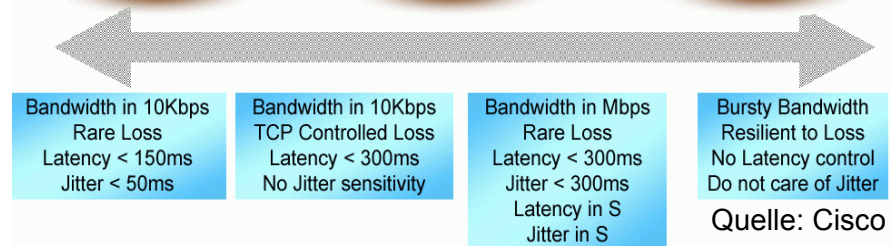


# Was sind die angestrebten Geschäftsziele?



- Kostenreduzierung bei Netzwerkverbindungen?
- Outsourcing von IT Infrastruktur?
- Welche Anwendungen mit welchen Anforderungen sollen im VPN genutzt werden?

- "traditionelle" IP-Nutzung
- Streaming Services
- Voice / Video Interaktiv
- Transaktionssysteme



# Anforderungen / SLA

- Sicherheit
- Skalierbarkeit
- Verfügbarkeit und Leistung
- Management / Wer betreibt das VPN?
- QoS / CoS

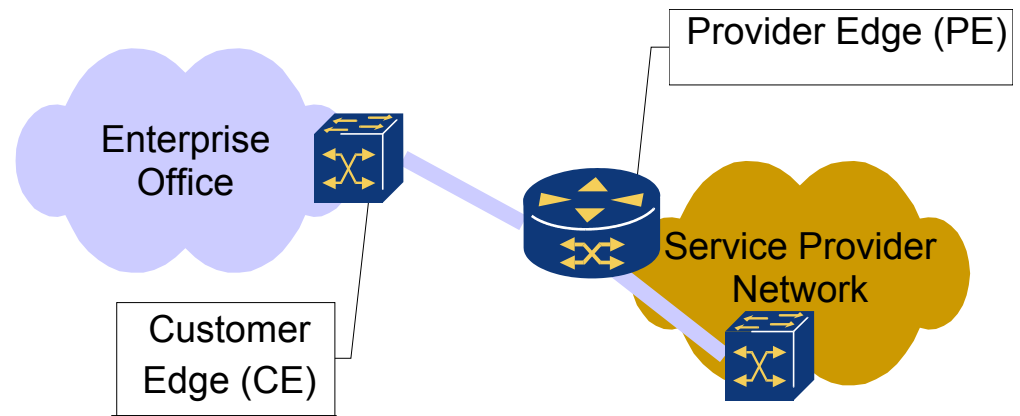
# Sicherheit

- Auswahl verschiedener Sicherheitsstufen
- Zuverlässige Trennung des Datenverkehrs
  - vergleichbar zu Frame Relay / ATM
  - Sicherstellung der korrekten Konfiguration
    - Sicheres Management
    - Authentifizierung der kommunizierenden Router
- Verschlüsselung
  - IPSec
  - Im Provider-Netz
  - End-to-End (Kundennetz)

# Skalierbarkeit / Leistung, Verfügbarkeit

- Keine Begrenzung
  - in der Anzahl der Standorte
  - Anzahl der VPNs, Überlappung von VPNs
- Einfache Anbindung neuer Standorte
- Nutzung privater Adressbereiche
- Leistungs- und Verfügbarkeits-Zusicherung in verschiedenen Stufen
  - Messzeitraum: Sekunden statt traditionell Tage, Wochen, ...

# Management / Wer betreibt das VPN?

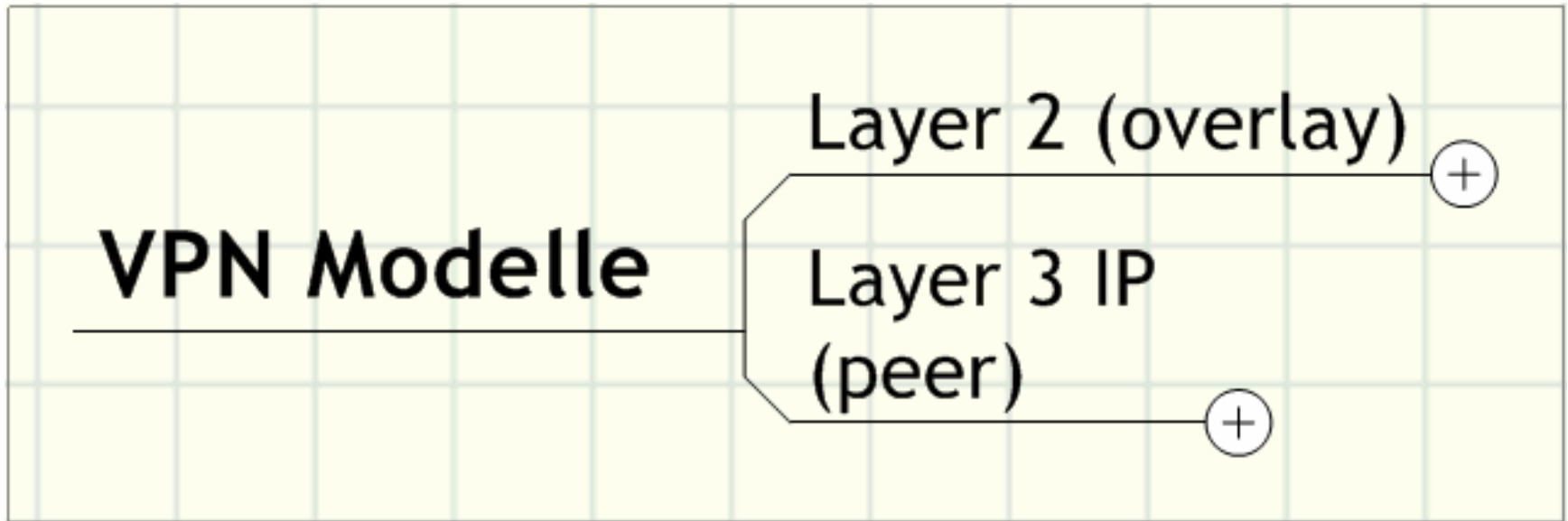


- Wem „gehört“ welches Equipment / Edge Devices?
  - Management durch eigene IT-Abteilung (in Kooperation)
  - Management durch Provider
- Technologieoptionen (MPLS, IPSec, SSL, ...)

# QoS / CoS

- Bandbreite
- Loss Rate
- Delay
- Jitter
- Aggregation von Datenströmen
- DiffServ-Support (expedited forwarding /assured forwarding)

# VPN Modelle



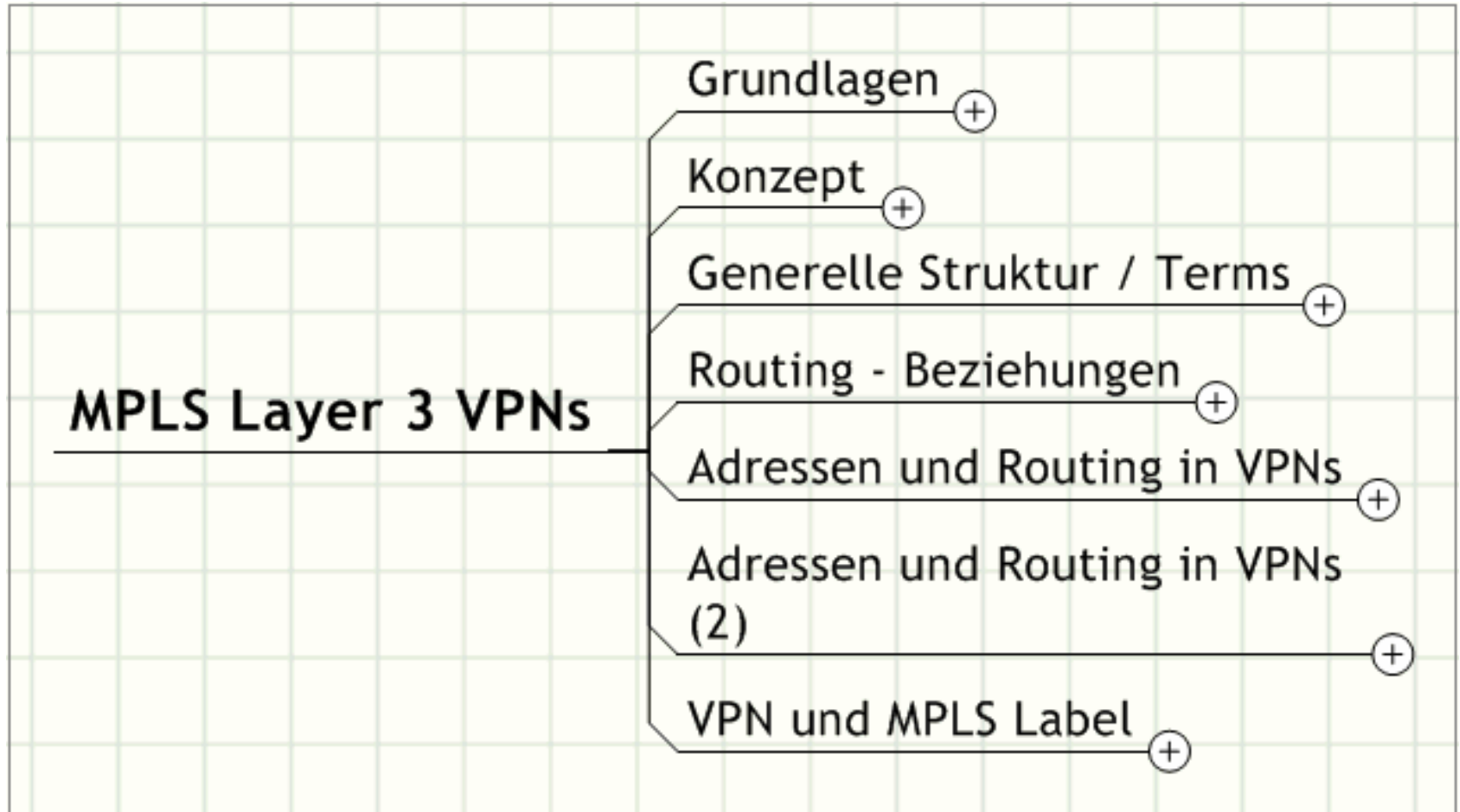
# Layer 2 VPNs (overlay)

- Point-to-Point Tunneling Protokoll (PPTP, RFC2637)
- Layer 2 Tunneling Protokoll (L2TP, RFC2661)
- Layer 2 MPLS VPNs (Martini / Kompella)
- MPLS Layer 2 VPN Vorteile
  - Kundeninterface auf Layer 2
    - eingeführt in Carrier-Netzen durch Frame Relay, ATM
  - Layer 2 PDUs transportiert über Layer 3 Multiprotokoll-Umgebung
    - LAN / MAN – Technologie über WAN
  - Transparente Dienste für die höheren Schichten

# Layer 3 IP VPNs (peer)

- IPSec (nur IP, verschlüsselt)
- Layer 3 MPLS (MP-BGP, RFC2547bis)
- Layer 3 MPLS VPN Vorteile
  - Kundeninterface auf Layer 3
  - liefert voll geroutete IP Netzwerklösung
  - Tunneling, Backbone hat keine Kenntnis über das IP-Netz des Kunden

# MPLS Layer 3 VPNs



# Grundlagen

- auch BGP/MPLS IP VPNs genannt
- basieren auf RFC2547bis
- aktuelle Weiterentwicklung:
  - [draft-ietf-l3vpn-rfc2547bis-01.txt](#)
- VPN Modell für SPs (Service Provider)

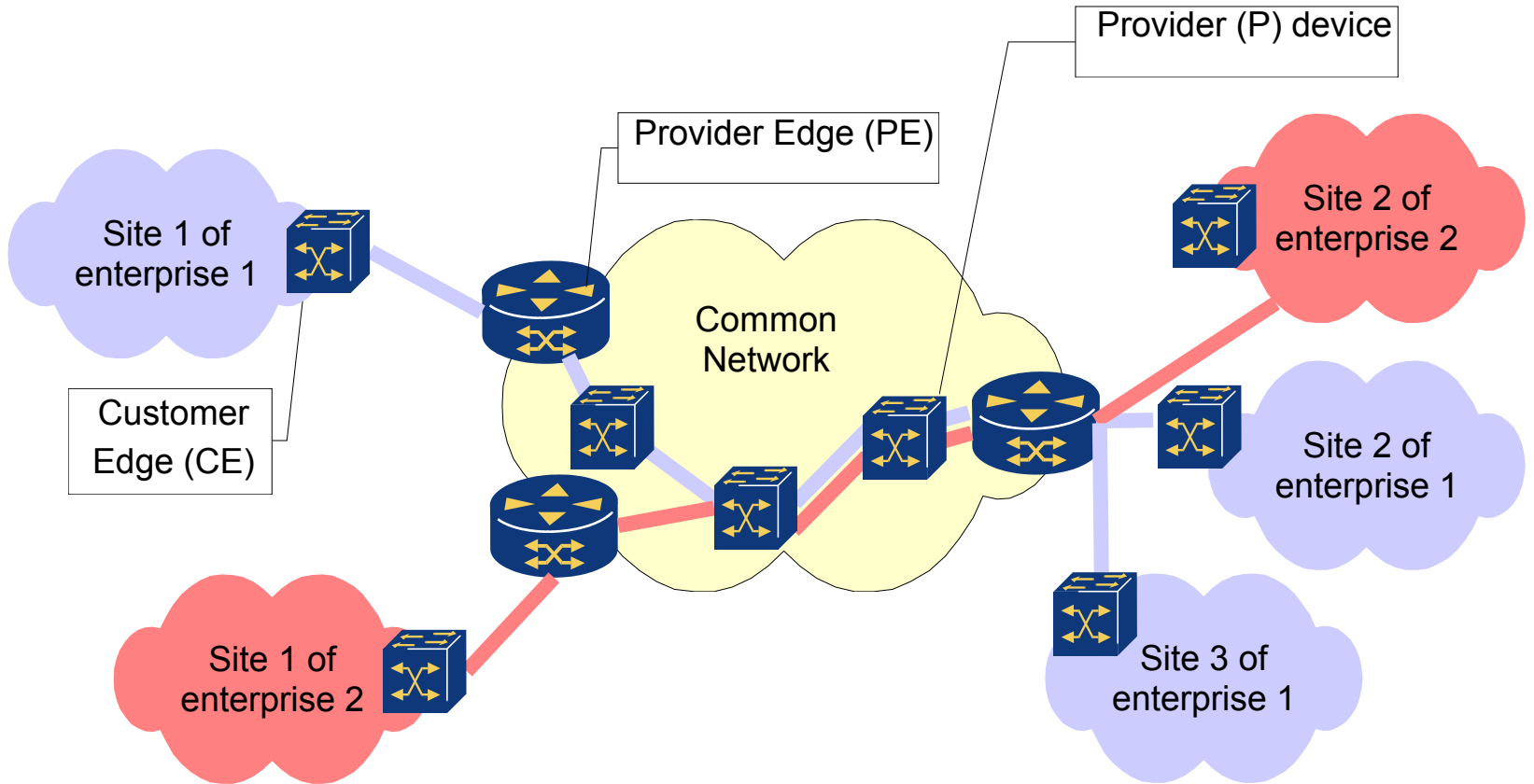
# Konzept

- IP Verbindung über einen gemeinsamen Backbone
- Standorte (Sites) haben nur Verbindung, wenn sie einem gemeinsamen VPN angehören
- Policies zur Bildung der VPNs definiert der Kunde
  - Umsetzung kann alleine durch den SP erfolgen oder in Kooperation mit ihm
  - Breite Palette an Policies möglich
    - volle Vermaschung
    - Routen bestimmter Sites über eine Dritte

# Generelle Struktur / Terms

- Jeder VPN-Standort muss mindestens ein Customer Edge Device (CE) stellen
  - CEs können Router oder Hosts sein
- Jedes CE Device ist mit einem oder mehreren Provider Edge Routern (PE) verbunden
- CE und PE Devices sind durch Attachment Circuits verbunden
  - Ethernet interface, PPP connection, ATM VCs, Frame Relay DLCIs
  - VLANs
  - IPSec Tunnel
- CE Devices gehören logisch zum Kundennetz, PE Devices zum Netz des SP
- Router im SP-Netz, die nicht mit CE Devices verbunden sind, nennt man "P Router"

# VPN nach RFC2547



# Routing - Beziehungen

- Ist das CE Device ein Router, dann ist es Routing Peer zum PE
- CE Router sind nicht Routing Peers von CE Routern anderer Standorte
  - Daher tauschen sie mit denen keine Routing Informationen aus
    - CE Router verschiedener Standorte brauchen sich nicht mal zu kennen
  - Der Kunde hat keinen Backbone (weder real noch virtuell) zu managen
    - Der Kunde hat generell nichts mit dem Inter-Site Routing zu tun
    - Anders formuliert, das VPN ist kein Overlay-Netz vom SP-Backbone

# Adressen und Routing in VPNs

- In getrennten VPNs können überlappende/identische Adressen verwendet werden
  - BGP-MP führt VPN-IPV4-Adressen ein, die eine IPv4-Adresse um einen Route Distinguisher (RD) erweitern
- Verkehr aus verschiedenen VPNs zu einem System, das all diesen VPNs angehört, kann über verschiedene Routen geführt werden!
  - Verkehr der Standorte B und C geht direkt zu A, Verkehr vom Standort D geht über den Firewall von B zu A
- Routing-Informationen über ein bestimmtes VPN werden nur in den PE-Routern benötigt (und vorgehalten), die mit Standorten dieses VPNs verbunden sind (Skalierbarkeit!)
  - "P Routers" brauchen keinerlei "ANY per VPN routing information" oder ähnliches

# Adressen und Routing in VPNs (2)

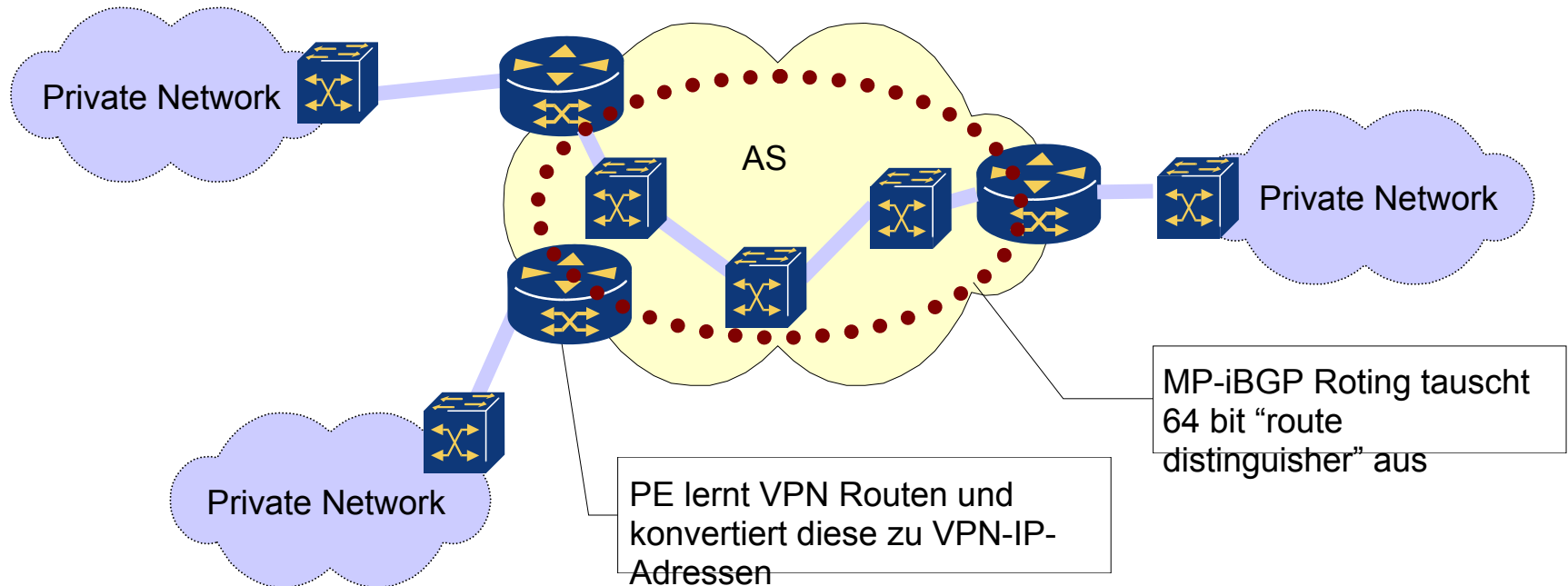
- PE Router bauen VRFs (VPN Routing and Forwarding tables) für jeden PE-CE attachment circuit auf
  - Im typischen Fall genau einen VRF pro CE Device
  - Pakete, die nicht über einen "attachment circuit" eingehen, werden anhand einer "default forwarding table" gerouted
- PE erkennt, über welchen "attachment circuit" ein Paket / Frame eintrifft
  - anhand des physikalischen Interfaces
  - durch Analyse des Layer 2 Header (z.B. DLCI bei Frame Relay)
  - VLAN tag values
  - ...

# VPN und MPLS Label

- Jede Route in einem VPN erhält ein MPLS Label
  - BGP-MP distributiert zu einer VPN-Route immer auch das zugehörige MPLS Label
- Bevor ein Kunden-Datenpaket an den SP Backbone übergeben wird, erhält es das Label, das im Kunden-VPN der besten Route zum Ziel entspricht
  - Vergabe durch "Ingress PE Router"
- Zusätzlich erhält das Datenpaket ein MPLS Label, über welches das Paket durch den Backbone geleitet wird
  - Ziel des Tunnelings durch den Backbone ist der "Egress PE Router"

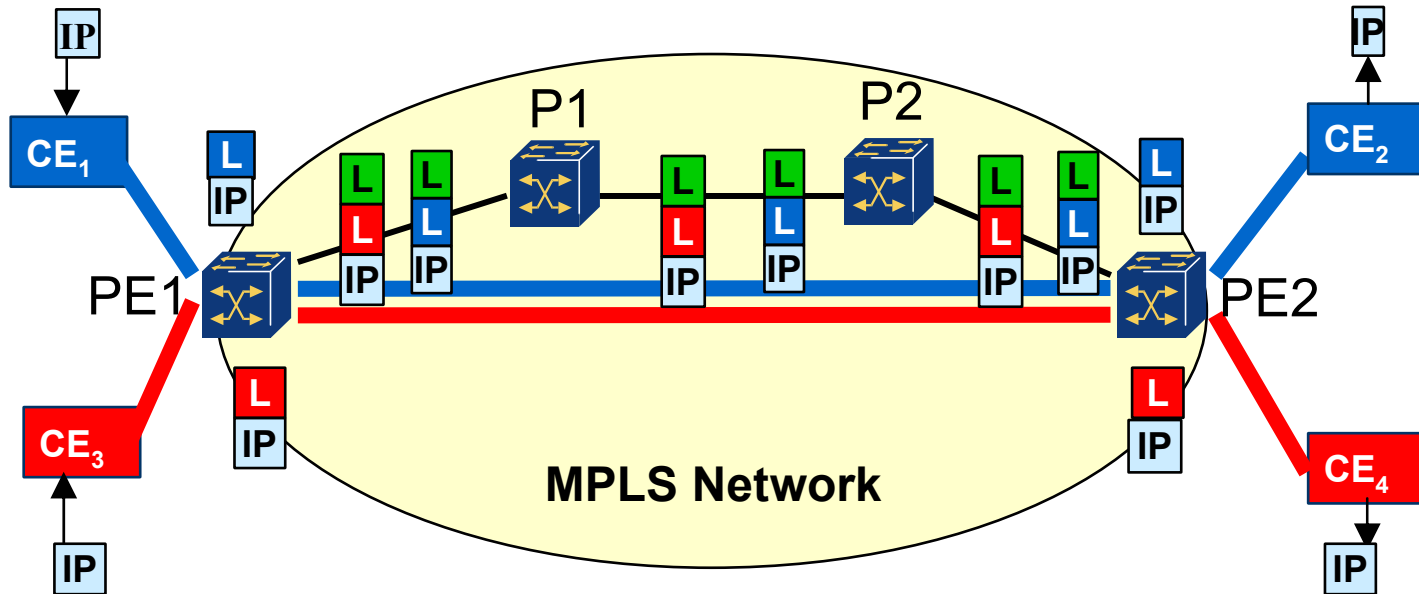
# VPN Route Distribution mit BGP

Provider Edge Router sind Teil eines gemeinsamen AS (Autonomous System), auf dem iBGP-MP läuft



iBGP-MP = interior Border Gateway Protocol / Multi-Protocol Extensionsh

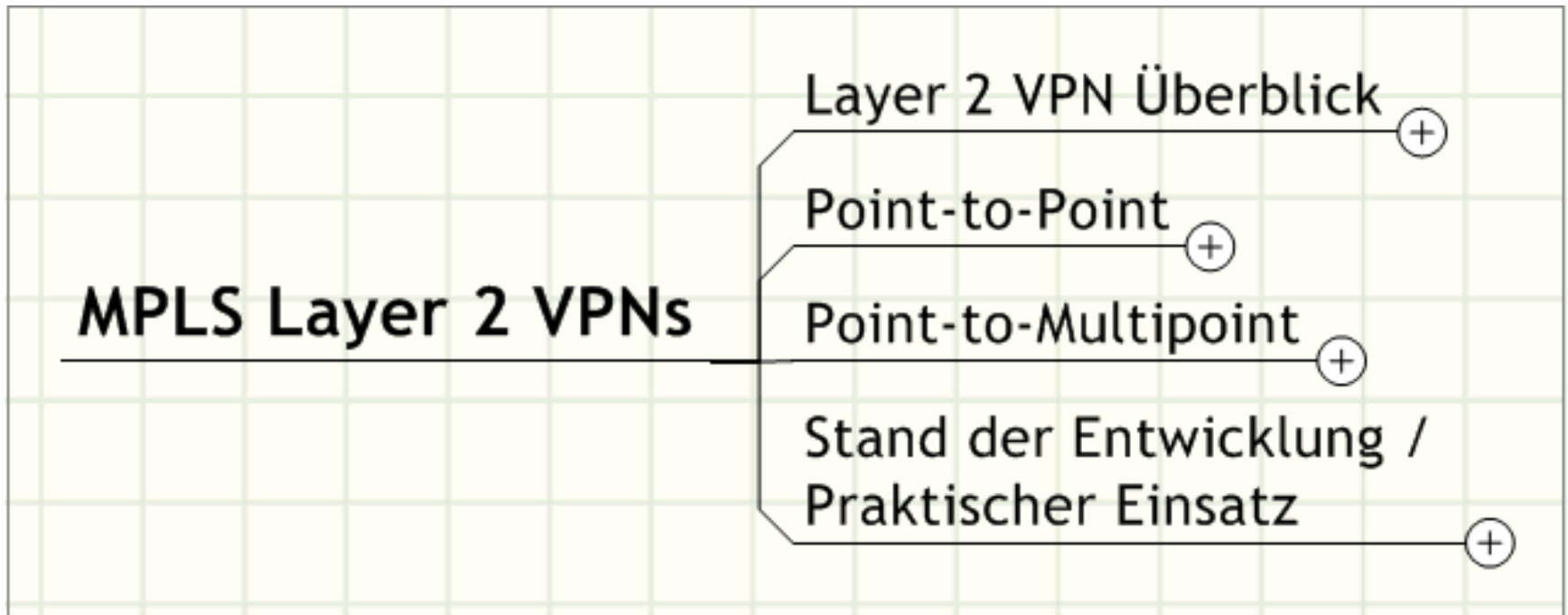
# Beispiel: VPN – Labelling



■ Label VPN A  
■ Label VPN B

■ Label zwischen PE1 und PE2

# MPLS Layer 2 VPNs



# Layer 2 VPN Überblick

## ■ Definition

- Eine geschlossene Gruppe mit Zugang basierend auf einer Adresse
  - Ethernet-Adresse, Frame Relay DLCI, ATM VC

## ■ Ziel

- Transport von Layer 2 PDUs über ein MPLS-Netzwerk

## ■ IETF: 2 konkurrierende" Work Groups

- Pseudo Wire Emulation Edge to Edge (PWE3) Group
  - ursprüngliche Basis: "Martini-Drafts"
- Provider-Provisioned VPN (PPVPN) group
  - ursprüngliche Basis: "Kompella-Drafts"

# Point-to-Point

- IETF: Pseudo Wire Emulation Edge to Edge (pwe3)
  - WG behandelt verschiedene Technologien: TDM, Frame Relay, ATM, SONET, Ethernet, ...
  - Overlay Modell
    - PE Router ist nicht Peer des CE Routers
    - PE Router mapped einfach nur Layer 2 Verkehr in den geeigneten Point-to-Point-Tunnel
- Pseudowire Setup and Maintenance using LDP
  - Encapsulation von Layer 2 Paketen mit einem "pseudowire header"
    - Ein "Demultiplexer field" (hier MPLS Label) wird vor dem Versand über den "Pseudowire" hinzugefügt
    - Am Endpunkt des "Pseudowires" ermöglicht ein Demultiplexer die Identifizierung, auf welchem "Pseudowire" das Paket eingetroffen ist
  - MPLS Label Switched Pathes (LSPs) werden als Tunnel für den Transport genutzt
    - Tunnel Label wird nur für den Transport durch das SP-Netzwerk verwendet (edge-to-edge)
- Wichtigste Drafts:
  - draft-martini-l2circuit-trans-mpls-07.txt
  - draft-martini-l2circuit-encapmpls-03.txt

# Point-to-Multipoint

- IETF: Layer 2 Virtual Private Networks (ppvpn)
- Fokus Ethernet: Virtual Private LAN Services (VPLS) over MPLS
- Draft Kompella
- Draft Laresse-VKompella
- (Derzeitige) Vereinfachungen / Einschränkungen

# IETF: Layer 2 Virtual Private Networks (ppvpn)

- 2 Drafts derzeit diskutiert
  - draft-kompella-ppvpn-vpls-01.txt
    - Basis bekannt als "Kompella Draft"
    - Funktionsstruktur im Kern ähnlich zum L3 VPN (RFC2547bis)
  - draft-lasserre-vkompella-ppvpn-vpls-02.txt
    - auch "Lasserre-VKompella Draft" genannt
- Konzept erlaubt, Frame Relay, ATM und Ethernet VLAN-basierte VPNs auch gemeinsam über eine IP/MPLS-Infrastruktur zu tunneln
- Klare Trennung der Funktionen
  - SP ist nur zuständig für Layer 2 Konnektivität
    - reduziert Komplexität und Aufwand für SP
    - keine Routing Tables (skaliert gut)
    - keine Berücksichtigung überlappender Adressbereiche erforderlich
  - Kunde verantwortlich für Layer 3 und Bridging

# Fokus Ethernet: Virtual Private LAN Services (VPLS) over MPLS

- auch als Transparent Lane Services (TLS) bekannt
- Realisierung aktuell über MPLS-Tunnel diskutierte, aber offen (z.B. GRE-Tunnel)
- Verbindet entfernt Standorte funktional betrachtet zu einem lokalen LAN
  - IP/MPLS Routing Protokolle ersetzen Spanning Tree
  - MPLS Labels ersetzen VLAN IDs
  - Realisiert Layer 2 Broadcast Domain
  - volle Unterstützung des Learning & Forwarding von MAC-Adressen
  - Länger nicht genutzte MAC-Adressen werden gelöscht (out aging)
- Kunden werden einer oder auch mehreren VPLS-Domains zugeordnet
  - Eine VPLS-Domain umfaßt alle zugehörigen PEs
  - PE Router werden hierzu erweitert um spezielle VPLS-Funktionen

# Draft Kompella

- Signallisierung über BGP
- Autodiscovery unterstützt
  - über BGP Session zwischen dem neuen PE und einem BGP Route Reflector
  - LSPs mit dem neuen PE können dann automatisch aufgebaut werden

# Draft Laresse-VKompella

- Signallisierung über LDP
  - Vorteil: weniger komplex als BGP
  - Nachteil für "BGP-Nutzer": weiters Protokoll!
- (derzeit) kein Autodiscovery
  - Neuer PE muss konfiguriert werden mit den Adressen aller anderen PEs einer VPLS-Instanz
  - Denkbar über verschiedenste Wege, z.B. LDAP Database etc.
  - Nachteil: arbeitsintensiv + fehleranfällig (Skalierbarkeit)
  - Vorteil: BGP nicht erforderlich, z.B. für SPs, die es bisher nicht nutzen

# (Derzeitige) Vereinfachungen / Einschränkungen

- Voll vermaschte Topologie erforderlich
  - erfordert für jede VPLS-Instanz  $n*(n-1)/2$  Point-to-Point-Pseudowires
  - Neuester Draft (Nov. 2003) beinhaltet "Hierarchical VPLS Model", verbessert Skalierbarkeit
- Multicast nicht direkt unterstützt (Verteilung per Broadcast)

# Stand der Entwicklung / Praktischer Einsatz

- "Öffentliche" Tests der Multi-Vendor-Interoperabilität
  - Initiator MPLS Forum
- Aktuelle Tests:
  - Test während der MPLS Conf Paris 2/2003
  - Test während der Supercomm, Atlanta (USA), 6/2003

# Test auf der MPLS Conf Paris 2/2003

- Testlabor: EANTC in Kooperation mit ETSI
  - Vortest im EANTC Lab
- Teilnehmer
  - Alcatel, Avici, Cisco, Data Connection, Ixia, NetPlane, Nortel, Quallabay, RAD Data Communications, Redback, Riverstone, Spirent
- Testbereiche
  - BGP/MPLS VPN + Skalierbarkeit
  - Ethernet/VLAN over MPLS + Skalierbarkeit
  - Fast ReRoute (FRR)
    - Detour Fast ReRoute Test
    - Facility Fast ReRoute Test
- Testergebnisse

# Testergebnisse MPLS Conf Paris 2/2003

Key Features Tested		Results
L2 VPNs	Interoperability Ethernet VLANs	OK
	Scalability 200 Ethernet VLANs	OK
	Data Transfer	OK
L3 VPNs	Interoperability LDP	Most combinations interoperable
	Interoperability MP-BGP	OK
	Scalability 255 VPNs	All implementations reached 255 VPNs
	Scalability 10-1000 routes per provider edge	All implementations reached 10 VPNs x 1000 routes (BGP/OSPF routing) plus 245 VPNs x 10 routes (static routing)
	Data Transfer through VPNs	A few combinations tested; no issues found
Fast Reroute Interoperability Detour and Facility Backup	Backup tunnels established in most cases, switch-over verified (< 50 ms, SDH grade resiliency)	
LDP over RSVP-TE Tunnel Interoperability	Unresolved configuration issues in a couple of cases	

# Tests Supercomm Atlanta (USA), 6/2003

- Testlabor: IOL der Uni New Hampshire
  - Vortest im IOL Lab
- Teilnehmer
  - Alcatel, Agilent, Cisco, Ixia, Juniper, Laurel, Marconi, Nortel, RADUSA, Riverstone, TiMetra, Vivace, Masergy, Finisar

# Testbereiche

- ATM, Frame Relay und Ethernet VLANs über MPLS nach PWE3 drafts
  - MPLS Signallisierung zum Tunnelaufbau inkl. Parameterdefinition
  - MPLS Data Encapsulation
    - getestete Typen: Ethernet (port und VLAN basierend), Frame Relay und ATM (transparent mode, cell mode und AAL5)
  - Bereitstellung der service-spezifischen Interfaces zum Kunden (CE)
  - Skalierbarkeit
- VPLS nach draft-laresse-vkompella-ppvvpn-vpls-04.txt
  - Aufbau voll-vermaschter VPLS-Tunnel
  - Transparenter Transport von Layer 2 PDUs
- BGP/MPLS IP VPNs entsprechend RFC 2547bis
  - Korrekter Aufbau der VPNs
  - Skalierbarkeit
- Fast ReRoute (FRR) Unterstützung im Core-Netzwerk
  - Detour Fast ReRoute Test
  - Facility Fast ReRoute Test

# Testergebnisse

- Grundlegende Funktionen interoperabel, deutliche Verbesserung im Vergleich zur Supercomm 2002
  - Aber einzelne Probleme traten auf, teilweise direkt vor Ort behoben, andere fließen in die Weiterentwicklung der Drafts ein
  - Zusammenfassung verfügbar über MPLS Forum Web-Server
- Testergebnisse Skalierbarkeit

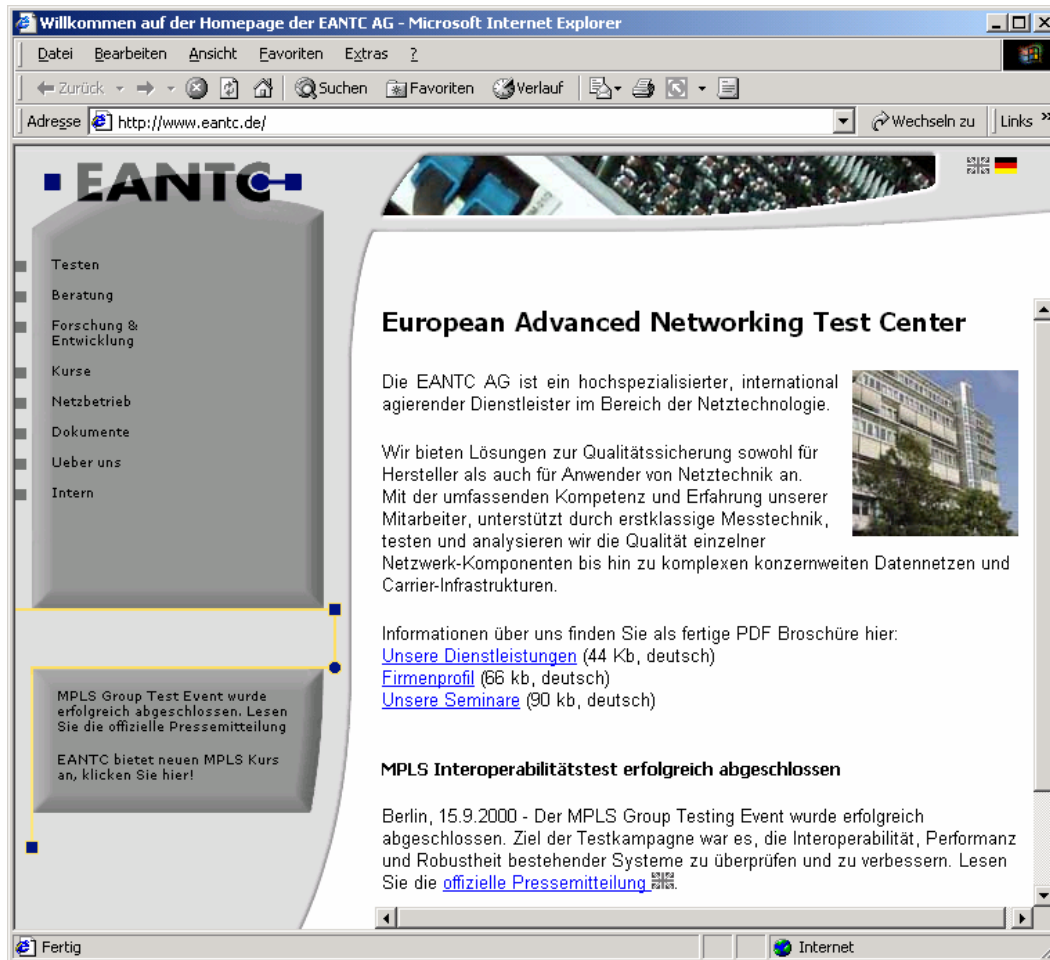
Quelle MPLS Forum

Service Type	Scalibility Number achieved per PE	Participating Companies with PE devices
BGP/MPLS VPNs	250	11
FR over MPLS	100	8
ATM over MPLS	100	6
Ethernet VLANs over MPLS	100	6
VPLS	1 at UNH, 100 at Supercomm	4

# Zusammenfassung: VPN Checkliste

	Layer 2 (ATM / FR)	IPSec	MPLS Layer 2 VPNs	MPLS Layer 3 VPNs
Provides security (VPN isolation)	✓ ✓	✓ ✓ ✓	✓	✓
Encryption	✗	✓ ✓ ✓	✗	✗
Scale for many end points (meshed)	✗	✓	✗	✓ ✓ ✓
Forwarding performance	✓ ✓	✗	✓ ✓	✓ ✓ ✓
Available from many carriers	✓ ✓ ✓	✓	✓ ✓ / ✗	✓ ✓
Provides quality of service	✓ ✓ ✓	✓ / ✗	✓ ✓	✓ ✓
Large-scale manageability	✓ ✓	✗	✗	✓
Interoperable with 3 <sup>rd</sup> party products	✓ ✓ ✓	✓ ✓	✓ ✓ / ✗	✓ ✓
Best suited for IP traffic	✗	✓ ✓ ✓	✗	✓ ✓ ✓
Suited for non-IP traffic	✓ ✓ ✓	✗	✓ ✓	✗

# Herzlichen Dank!



Für mehr Informationen  
steht unser Webserver  
zur Verfügung:

<http://www.eantc.de/>