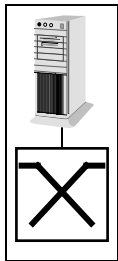


Netze und Protokolle für das Internet



6. Multicast-Kommunikation im Internet

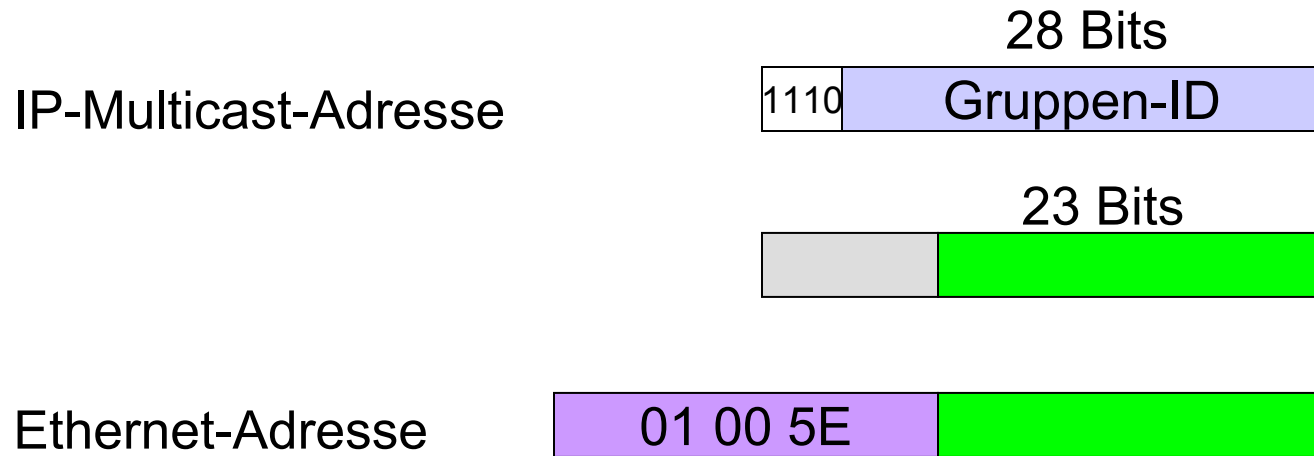
Inhalt

- Multicast-Anwendungen
- IPv4-Multicast-Adressierung
- Internet Group Management Protocol
- Intra-Domain Multicast Routing
 - MOSPF, PIM-DM
 - Core-Based Trees: PIM-SM
- Inter-Domain Multicast-Routing
 - Multicast Source Discovery Protocol
 - Border Gateway Multicast Protocol
- Multicast Backbone (MBone)
- Scoping
- Multicast Address Allocation (Malloc)
- neue Multicast-Ansätze
 - EXPRESS, REUNITE, Small Group Multicast, Endsystem-Multicast

Multicast-Anwendungen

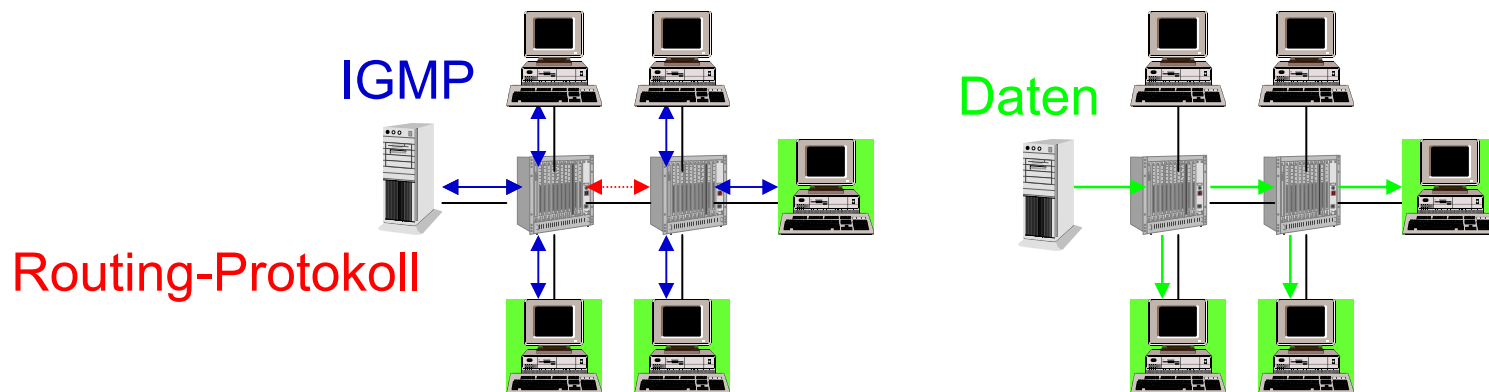
- **Gruppenkommunikation**
 - Audio/Video-Konferenzen
 - Computer-Supported Cooperative Work
 - Push-Technologien
(Software- und Informationsverteilung, Cache-Updates)
 - Paralleles Rechnen
- **Kontrollprotokolle**
 - Routing-Protokolle
 - Verteilen der Routing-Informationen an Router-Multicast-Adresse
 - IP Version 6
 - Adressauflösung
 - automatische Adresskonfiguration

IPv4-Multicast-Adressierung



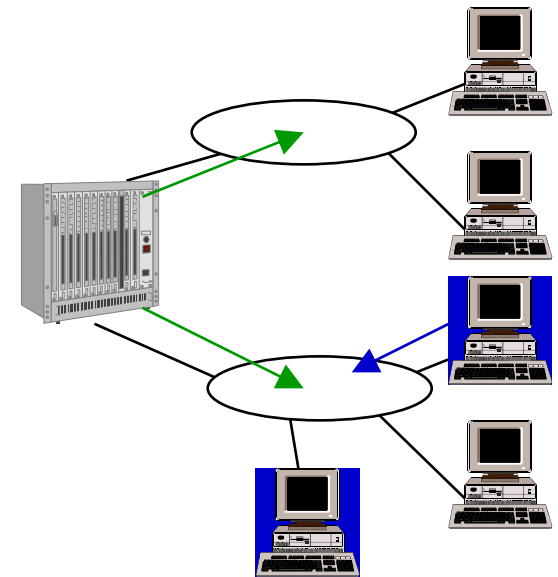
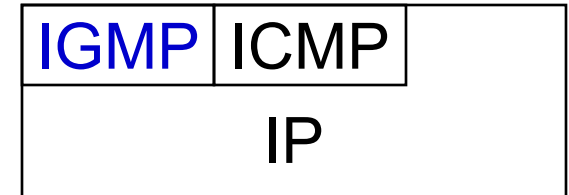
Multicast-Modell

- Sender sendet Multicast-Paket an durch eine IP-Multicast-Adresse gekennzeichnete Empfängergruppe
- Sender kennt die Empfänger nicht.
- Jedes Endsystem kann einer Gruppe beitreten.
- UDP unterstützt Multicast, TCP jedoch nicht
⇒ Multicast-Kommunikation ist Datagramm-orientiert
- Aufbau eines Multicast-Baums



Internet Group Management Protocol Version 1

- Router müssen wissen, ob lokale Gruppenmitglieder existieren
- Nachrichtentypen
 - Membership Query
 - periodisch (z.B. 1/min) durch Router (Querier) an „all hosts“-Gruppe
 - ein Querier pro physikalischem Netz (Router mit niedrigster IP-Adresse)
 - Membership Report
 - Antwort eines Hosts zur Anzeige der Gruppenzugehörigkeit
 - sollte sofort nach Beitritt von einem Endsystem gesendet werden
- Dienstprimitive
 - `JoinHostGroup (address , interface)`
⇒ Senden von Reports
 - `LeaveHostGroup (address , interface)`
⇒ kein Senden von Reports
 - Implementierung über `setsocketopt`



IGMP Membership Query
IGMP Membership Response

IGMP Versionen 2 und 3

■ IGMPv2

- Endsystem setzt Flag, falls es auf das letzte Membership Query geantwortet hat.
- Bei gesetztem Flag sendet das Endsystem, welches die Gruppe verlässt, eine *Leave Group* Nachricht.
- Router sendet anschliessend ein *gruppenspezifisches* Membership Query.
- Router kann damit schneller feststellen, dass keine weiteren Gruppenmitglieder existieren.

■ IGMPv3

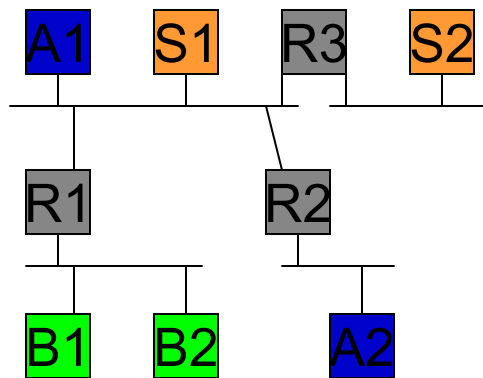
- Source-Filtering
 - Fähigkeit eines Systems, Multicast-Pakete eines oder mehrerer Sender zu verlangen
- neues Primitiv
 - **IPMulticastList** (**socket, interface, multicast-address, filter-mode, source-list**)
ersetzt JoinHostGroup und LeaveHostGroup
- *gruppen- und quellenspezifische* Query-Nachrichten
- Report-Nachrichten können ausgewählte Quellen enthalten.

Intra-Domain Multicast Routing

- Link-State-Routing
 - Jeder Router besitzt Informationen über jeden Link der Domäne und berechnet die komplette Netztopologie der Domäne.
 - Dijkstra's Shortest Path Algorithmus
 - Multicast Open Shortest Path First (MOSPF)
- Distanz/Vektor-Routing
 - Flood and Prune
 - Router wissen, ob bzw. auf welchem Interface Multicast-Daten für eine Kombination (Sender, Gruppe) weitergeleitet werden müssen.
 - Distance Vector Multicast Routing Protocol (DVMRP)
 - Protocol Independent Multicast (PIM) - Dense Mode (DM)
 - PIM basiert auf unterliegendem Unicast-Routing-Protokoll:
kein Verteilen von Multicast-Routing-Informationen erforderlich
- Core-Based Trees
 - keine Sender-spezifischen Bäume
 - PIM - Sparse Mode (SM)

MOSPF

- Multicast-Erweiterung von OSPF
- Group-membership link state advertisements: [group attached_network]
- Router gewinnen Informationen aus IGMP
- Verteilen der Multicast-Daten nach Dijkstra's Shortest-Path-Algorithmus



R: Router

S: Sender

A: Mitglied der Gruppe A

B: Mitglied der Gruppe B

S2→B: R3 leitet Paket weiter

S1→B: R3 leitet Paket nicht weiter

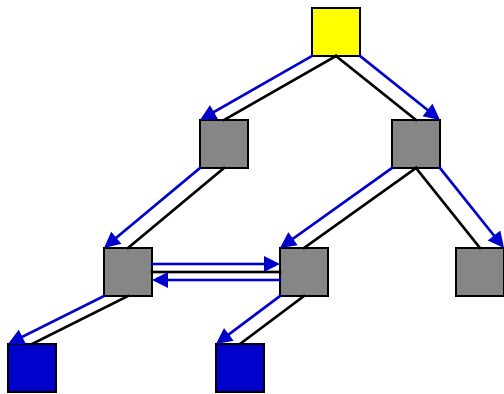
S1→B: R1 leitet Paket weiter

S1→A: R1 leitet Paket nicht weiter

PIM-DM

▪ Funktionsweise

- Router empfängt Multicast-Paket von S an G
- falls Eingangs-Interface = Ausgangs-Interface für Unicast-Pakete an S:
 - Weiterleiten des Pakets auf allen anderen Interfaces, sofern keine Prune-Nachricht empfangen wurde
 - sonst: Prune(S,G) an Eingangs-Interface



▪ Prune (S,G)

- wird auch gesendet, falls keine Gruppenmitglieder in einem Blatt-Link existieren

▪ Graft (S,G)

- Aufheben des Prunings (auch periodisch)
- Quittierung durch Graft Ack

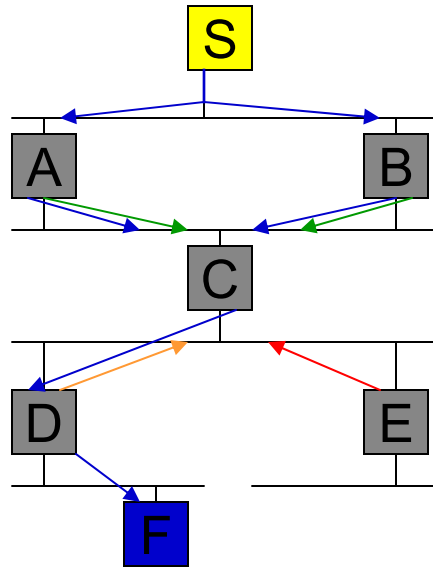
▪ Join (S,G)

- Aufheben von Prune(S,G) durch Nachbar-Router am selben Link

▪ Assert (S,G)

- Distanzanzeige eines Routers zum Sender zur Kollisionserkennung

PIM-DM: Protokollablauf



- Router A/B empfangen Multicast-Paket am Ausgangs-Interface und senden **Assert** an vorreservierte Multicast-Adresse mit Angabe der Distanz zum Sender: Die nachfolgenden Pakete werden nur über A oder B nach C weitergeleitet
- **Prune, Join**: C leitet **Datenpakete** nur an D weiter, nicht an E.

Core-Based Trees

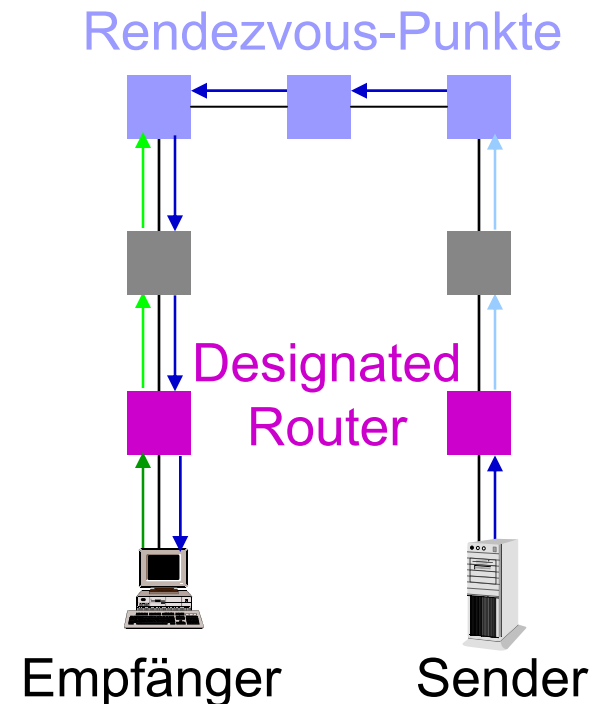
- PIM-DM, DVMRP und MOSPF
 - Berechnung von Routing-Bäumen für jede Kombination Sender/Multicast-Gruppe
 - Vorteil
 - optimierte Wege von einer Quelle zu den einzelnen Empfängern
 - Nachteile
 - erhöhter Verbrauch von Netzressourcen
 - hohe Anzahl von Routing-Einträgen
- Core-Based Trees (z.B. PIM-SM)
 - ein einziger Baum für eine Multicast-Gruppe, der von allen Sendern benutzt wird
 - Vorteile
 - weniger Routing-Tabellen-Einträge
 - geringere Netzbelastung
 - Nachteil
 - nicht optimale Wege zwischen Sender und einzelnen Empfängern

PIM-SM

- PIM-SM basiert auf Core-Based-Tree Konzept.
- Zu jeder Gruppe gehören ein oder mehrere Rendezvous-Punkte (RP), die den Kern des Multicast-Baums bilden.
- Designated Router (DR) stellen Verbindung der Gruppenmitglieder zu den RPs her.

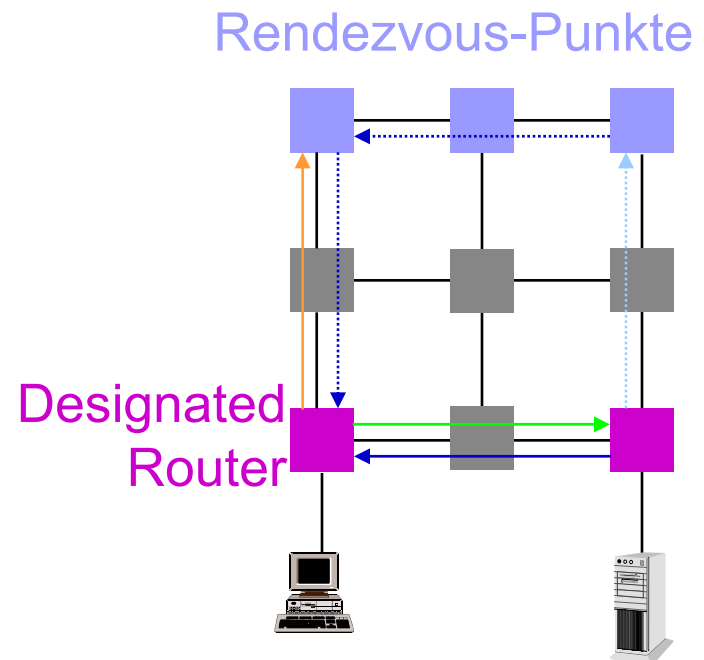
PIM-SM: Protokollablauf

- Beitreten eines Empfängers
 - IGMP Membership Report
 - PIM-Join/Prune an RP (periodisch von DR)
- Senden an Multicast-Gruppe
 - DR kapselt Daten in PIM-Register ein und leitet sie an RP weiter.
 - RP kapselt Daten aus und verteilt sie entlang des Multicast-Baums.
 - Router leiten Daten nur über die Interfaces weiter, über die Join/Prune empfangen wurde.



PIM-SM: Quellenspezifischer Baum

- Designated Router mit Gruppenmitgliedern oder RP kann (z.B. bei hoher Datenrate) das Etablieren eines quellenspezifischen Baums durch Senden eines **quellenspezifischen Joins** an Sender initiieren
- PIM-Router des Senders leitet dann die **Daten** direkt an den PIM-Router des Empfängers
- **Prune**-Nachricht zum RP



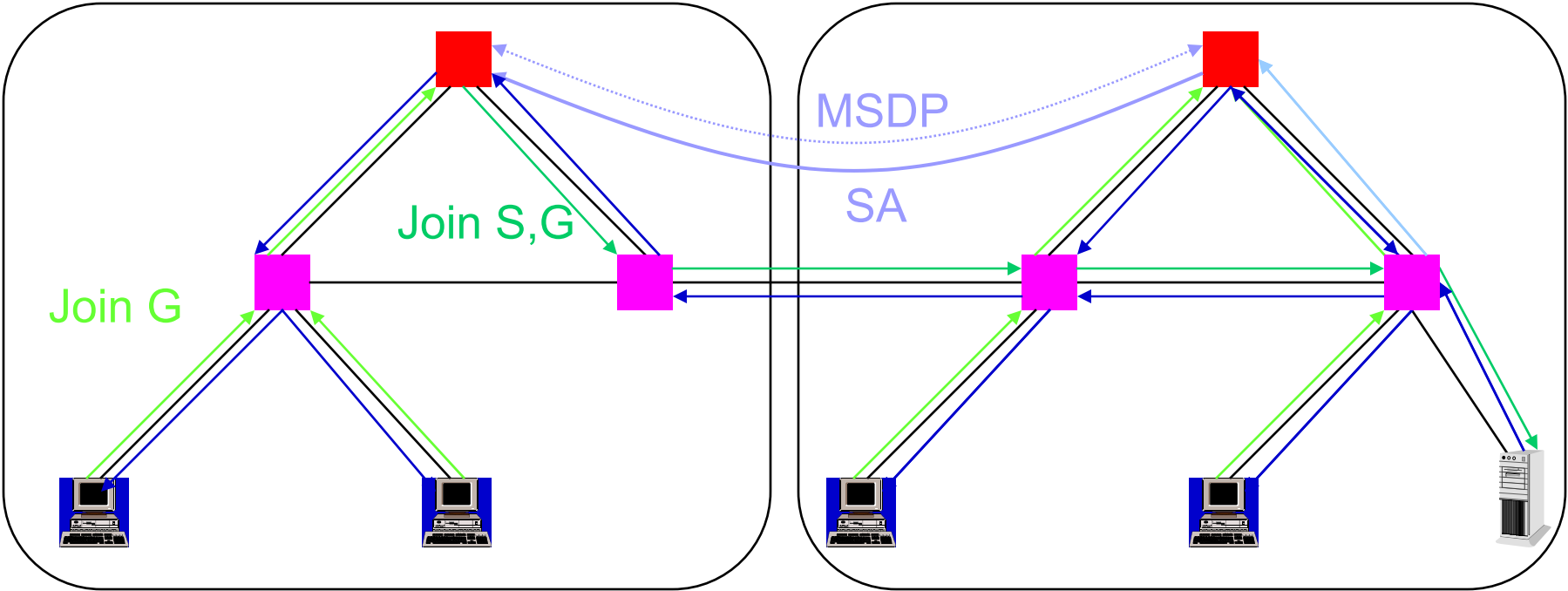
Inter-Domain Multicast-Routing

- Multiprotocol-BGP (MBGP)
 - Problem: PIM nimmt an, dass Unicast-Rückwärtspfad für Multicast geeignet ist: Join-Nachrichten können nicht-Multicast-fähige Router erreichen.
 - Multiprotokollerweiterung von BGP-4 erlaubt Aufsetzen von Unicast- und Multicast-fähigen Routen
- Multicast Source Discovery Protocol (MSDP)
 - Mechanismus, um PIM-SM-Domänen zu verbinden
- Border Gateway Multicast Protocol (BGMP)
 - Core-Based Tree Protokoll

Multicast Source Discovery Protocol (MSDP)

- Ziel: Unabhängigkeit der zu verbindenden Domänen ohne zentrale(n) RP(s) in einer einzigen Domäne (Robustheit !)
- Etablieren von TCP-Verbindungen zwischen MSDP-fähigen RPs verschiedener Domänen
- Austausch von **Source-Active**-Nachrichten (SA) bei aktiven Sendern an eine Multicast-Gruppe
 - Sender-Quelladresse
 - Gruppen-Zieladresse
 - RP-Adresse
- Beim Empfang einer SA-Nachricht und gleichzeitiger Existenz von Gruppenmitgliedern sendet RP ein quellenspezifisches Join in Richtung der Quelle.
- Zwischenlösung, da nicht skalierbar !
 - Jede Domäne muss über neue Quellen informiert werden.
 - Am Anfang müssen Daten in SA-Pakete eingekapselt werden.

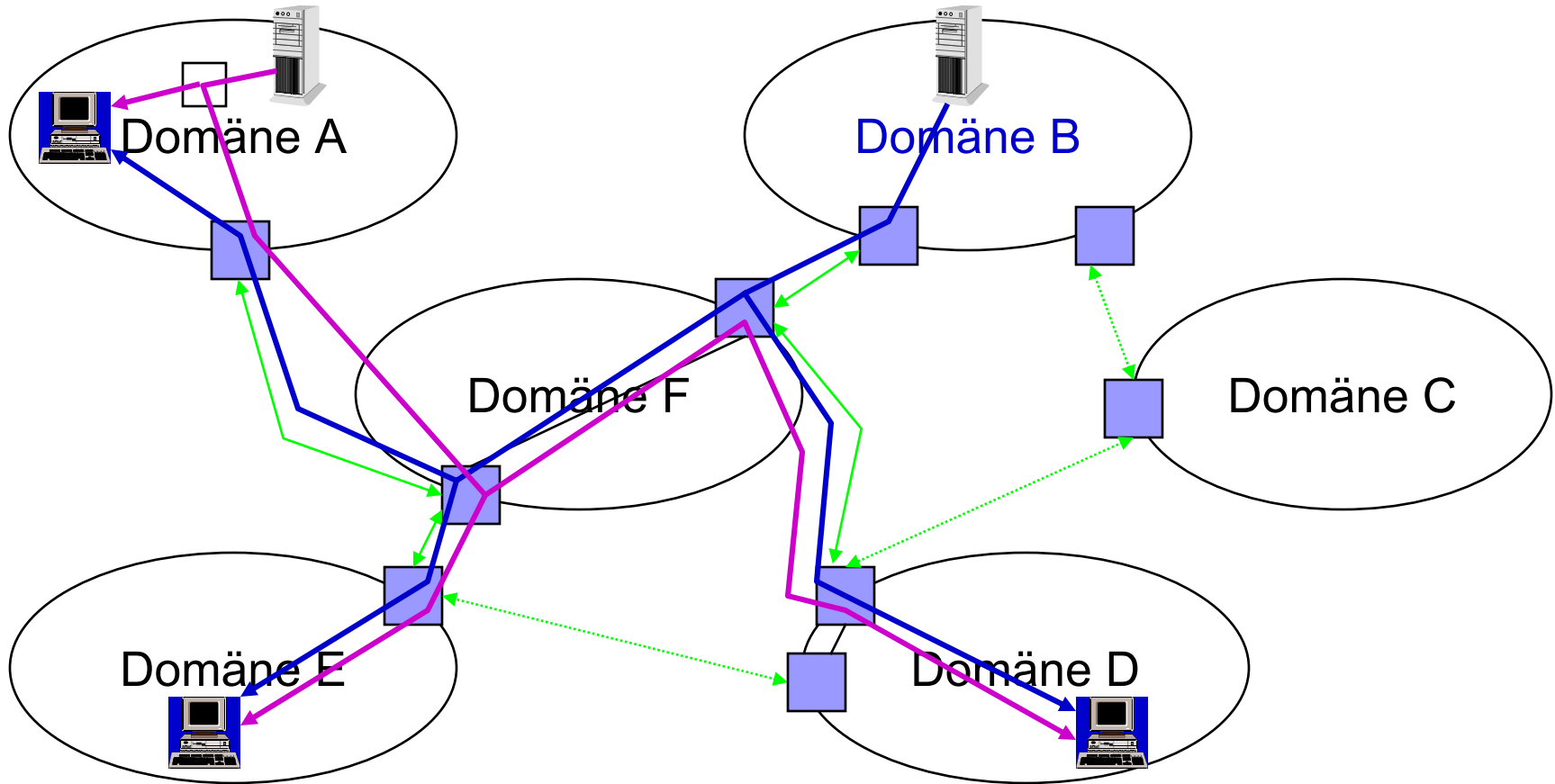
Beispiel: MSDP



Border Gateway Multicast Protocol (BGMP)

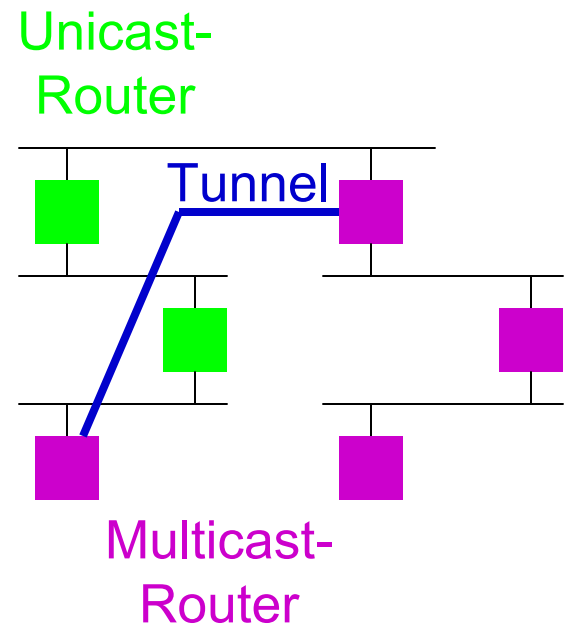
- Shared Trees für aktive Multicast-Gruppen
- Quellenspezifische Verteilung möglich
- basiert auf Konzepten von Core-Based Trees
- TCP-Verbindungen zwischen Border-Routern v. Domänen
- Zu jeder Multicast-Adresse gehört eine Root-Domäne, z.B. die Domäne, welche die Multicast-Adresse allokiert hat.
- Ablauf
 - Empfänger tritt einer Multicast-Gruppe bei.
 - Border-Router (Teilnehmer des Intra-Domain Multicast Routing-Protokolls) sendet gruppenspezifische Join-Nachricht in Richtung der Root-Domäne.
 - Erzeugen von gruppenspezifischen Forwarding-Einträgen in Border-Routern
 - Weiterleiten von Multicast-Daten entsprechend diesen Einträgen

Beispiel: BGMP



Multicast Backbone (MBone)

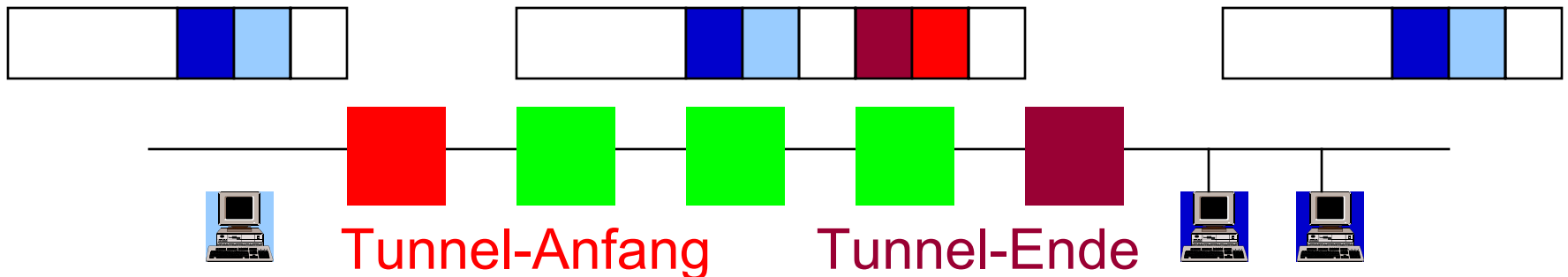
- Inbetriebnahme des ersten Multicast-Routers zur Übertragung einer IETF-Konferenz (1992)
- Technik: IGMP, oft noch DVMRP, zunehmend PIM und MSDP
- Aufbau eines Multicast-Overlay-Netzes
 - Automatische Erkennung benachbarter Multicast-Router
 - Verbindung von Multicast-Routern über Tunnels, falls diese nicht an ein gemeinsames Netz angeschlossen sind
 - IP-Multicast-Paket wird am Anfang des Tunnels in IP-Paket eingekapselt (IP-Quelladresse) und am Ende (IP-Zieladresse) wieder ausgekapselt.
 - manuelle Tunnel-Konfiguration
 - Subnetze ohne Multicast-Router sind nicht ans Mbone angeschlossen



Tunneln

```
tunnel 129.13.42.100 192.13.42.101 metric 1  
threshold 32 rate_limit 500
```

- Konfiguration von
 - Tunnelanfang und -ende
 - Kosten des Tunnels
 - Schwellwert für TTL-Wert
 - maximale Bandbreite für Multicast-Verkehr
- Tunnel erstreckt sich im allgemeinen über mehrere Router: TTL-Wert wird um mehr als 1 dekrementiert, falls das Paket über einen Tunnel geleitet wird.



Scoping

- Scoping = Begrenzen der Reichweite einer Multicast-Gruppe
 - Sicherheit
 - Mehrfachverwendung einer Multicast-Adresse
 - Leistungsaspekte bei Flooding-Protokollen
- Varianten
 - TTL-Scoping
 - Administratives Scoping

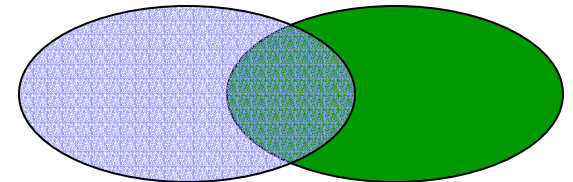
TTL-Scoping

- Dekrementieren von TTL-Werten alleine ist zu ungenau
- Konfigurieren von Schwellwerten in den Border-Routern von Regionen
- Beispiel:
 - Paket soll innerhalb von Europa gesendet werden
 - Schwellwert in Border-Routern von/zur Europa: 64
 - Schwellwerte zwischen nationalen Netzen: 48
 - TTL-Wert des Pakets < 64 , z.B. 63: erlaubt das Passieren von 15 Hops bevor es an einem internationalen Link gelöscht wird.
- Probleme
 - Konfiguration überlappender Regionen
 - Konsistente Wahl der Schwellwerte

Administratives Scoping

- erlaubt die Konfiguration von Multicast-Adressbereichen, die in Border-Routern auf beiden Seiten nicht weitergeleitet werden
- Die höchstwertigen 255 Adressen (/24) in jeder Region sind für Scope-relative Adresszuordnungen reserviert.
 - z.B. 239.255.255/24 bei IPv4 local scope
 - relative Zuordnung 0 für SAP (Session Announcement Protocol)
 - relative Zuordnung 1 = 239.255.255.254

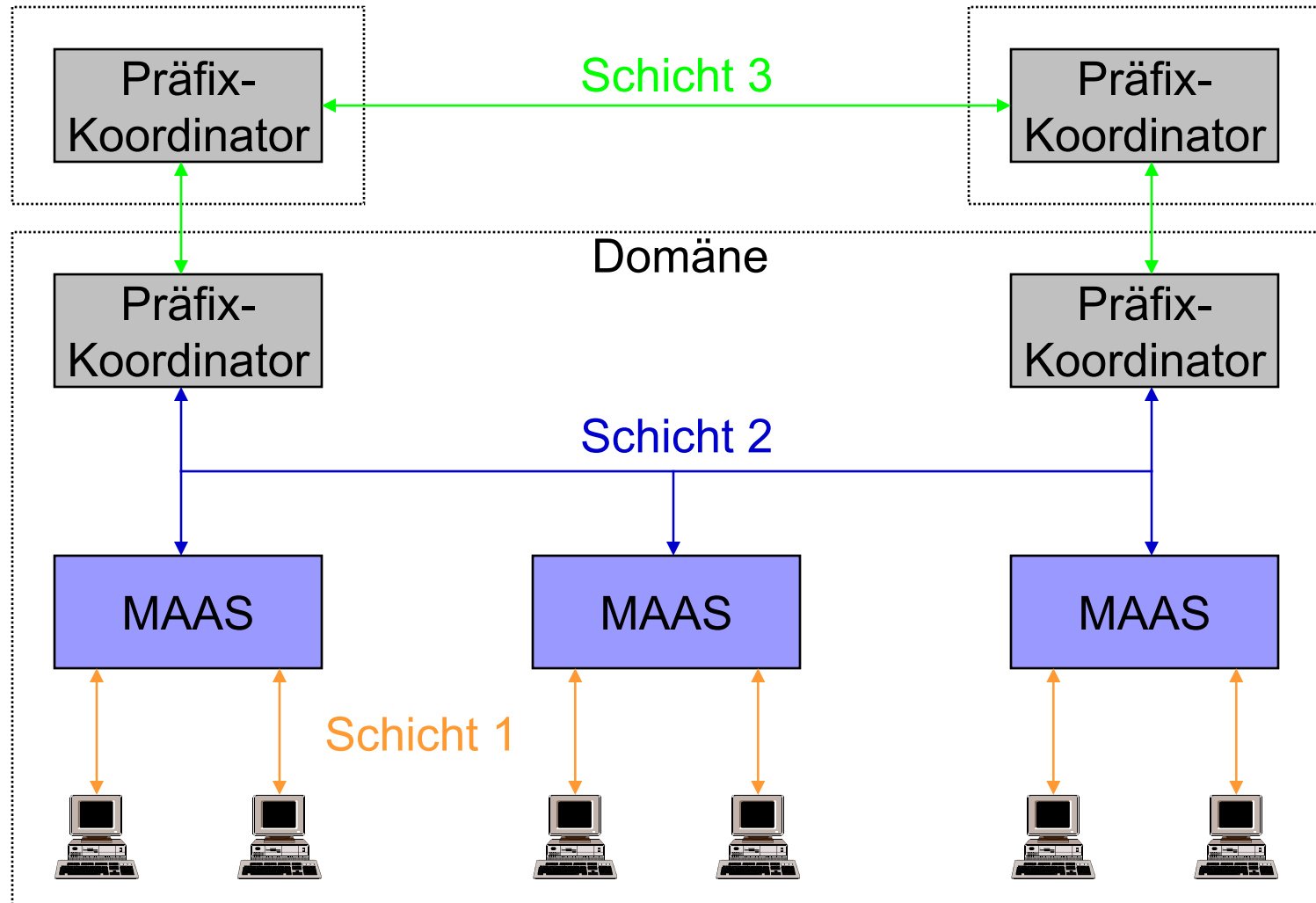
IPv4 local scope	239.255/16
IPv4 organization local scope	239.192/14
IPv4 link-local scope	224.0.0/24
IPv4 global scope	224/8



Multicast Address Allocation (Malloc)

- setzt administratives Scoping voraus
- Allokation der Multicast-Adressen
 - statisch
 - für spezifische Protokolle mit well-known Adressen, z.B. NTP
 - permanent
 - Scope-relativ
 - Die höchsten 256 Adressen eines Bereichs sind für relative Zuordnungen reserviert.
 - für Infrastruktur-Protokolle, die eine Adresse in jedem Scope benötigen
 - Zuordnen von Offsets mit permanenter Lebenszeit
 - dynamisch
 - für die meisten Anwendungen
 - Reservierung auf Verlangen
 - begrenzte Lebenszeit mit expliziten Verlängerungen

Malloc-Architektur



Malloc-Schichten

1. Protokoll zum Anfordern einer Multicast-Adresse durch Client von Multicast Address Allocation Server (MAAS)
 - Beispiel: Multicast Address Dynamic Client Allocation Protocol (MADCAP): Request/Response-Protokoll (vgl. DHCP)
2. Intra-Domain-Protokoll zur Koordination von Adressallokationen (Vermeiden von Adressduplikaten)
 - Beispiel: Multicast Address Allocation Protocol (AAP)
 - Anzeige freier Adressen durch Präfix-Koordinatoren
 - MAAS fordert Adressbereiche von Präfix-Koordinatoren an.
3. Inter-Domain-Protokoll zur Allokation von Adressbereichen und Zuordnung zu Präfix-Koordinatoren
 - Beispiel: Multicast Address-Set Claim Protocol (MASC)
 - Hierarchie von MASC-Domänen (AS)
 - Anfordern von Adressen bei übergeordneter Ebene
 - Alternative: statische Allokation abhängig von AS-Nr. (233.X.X)

Adressallokation

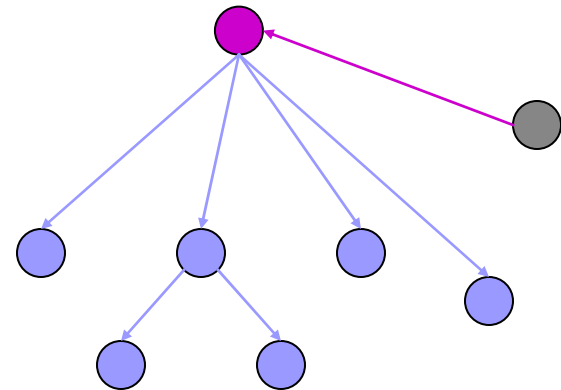
- Lokalisierung eines MAAS über manuelle Konfiguration, Service Location Protocol (SLP) oder Malloc-Schicht-1-Mechanismus
- Client fordert Multicast-Adresse bei MAAS an.
- MAAS wählt unbenutzte Adresse aus dem für den gewünschten Scope verfügbaren Bereich.
- MAAS stellt mit Schicht-2-Mechanismen sicher, dass die ausgewählte Adresse nicht kollidiert.
- Beim Auslaufen von Multicast-Adressen fordert der Präfix-Koordinator mit Schicht-3-Mechanismen neue Adressbereiche an.

Probleme von IP Multicast

- Sicherheit
 - Beliebige Empfänger können sich in Multicast-Strom einhängen (→ Verschlüsselung).
 - Beliebige Sender können an globale Multicast-Adresse senden:
Gefahr von Denial-of-Service-Attacken
- Skalierbarkeit
 - Multicast-Router haben Routing-Einträge der Form:
[Sender, Multicast-Adresse] → Ausgangs-Interfaces
- mangelnde Unterstützung durch ISPs
 - Multicast erfordert das Aufsetzen von zusätzlichen Routing-Protokollen.

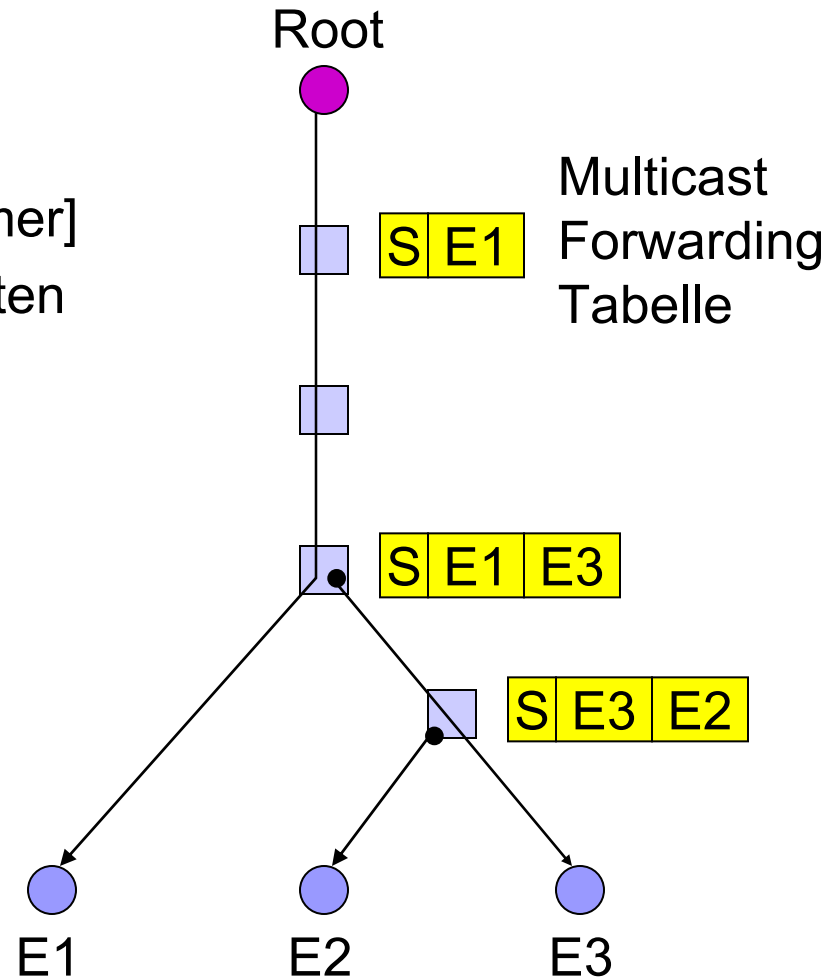
EXPRESS

- Explicitely Request Single Source Multicast
- Multicast-Kanal wird über [Quelle, IP-Multicast-Adresse] identifiziert
→ globale Eindeutigkeit der Kanal-ID
- Empfänger müssen sich explizit in Kanal anmelden.
- Nur Quelle darf an Gruppe senden.
- Subcasting



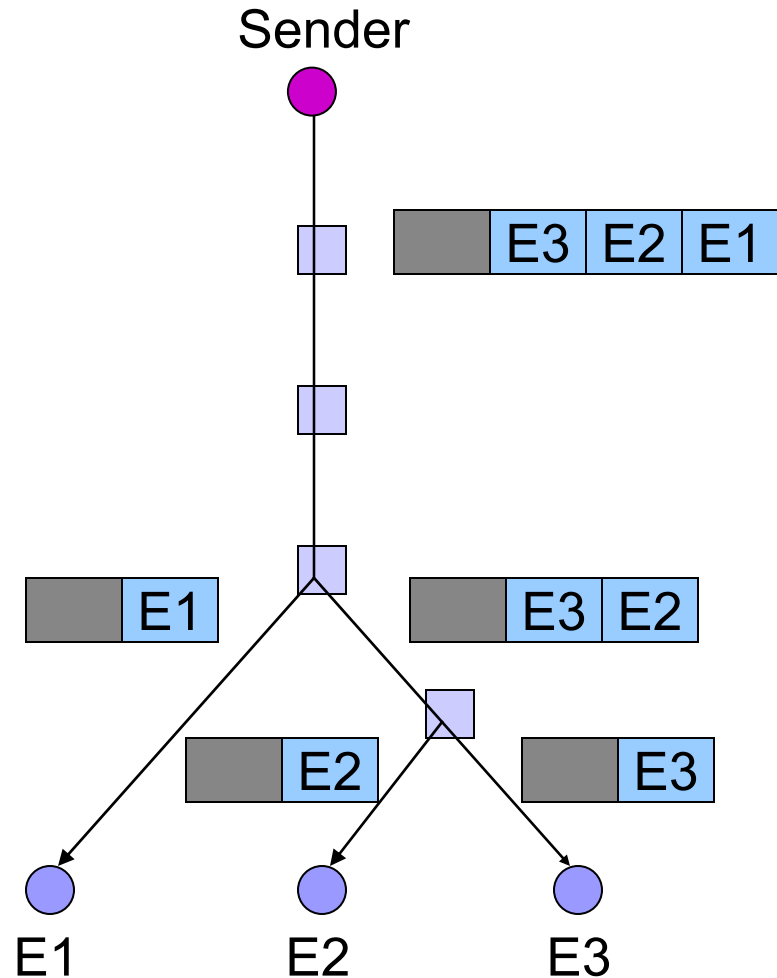
REUNITE

- Recursive Unicast Approach to Multicast
- Identifikation einer Gruppe durch [Root-IP-Adresse, Root-Port-Nummer]
- Root = Sender oder spezieller Knoten
- Aufbau eines Multicast-Baums ausgehend von Root
- Sender senden Pakete zu Root, Root kann Zugangskontrolle durchführen.
- Multicast-Forwarding-Einträge nur in Verzweigungspunkten
- Kontrollnachrichten: JOIN (\uparrow) / TREE (\downarrow)



Small Group Multicast

- Sender erzeugt Paket mit Adressliste (IP-Unicast-Adressen)
- Adressliste in IPv4-Option oder IPv6-Erweiterungs-Header
- SGM-Knoten extrahieren IP-Adressen, die über ausgehende Links erreichbar sind und behalten diese in der jeweiligen Adressliste
- nur für kleine Gruppen skalierbar
- Problem (vgl. REUNITE): Integration mit IP-Multicast-Anwendungen



Endsystem-Multicast

- Implementierung von
 - Gruppenverwaltung
 - Paket-Replikationin Endsystemen
- Selbstorganisierende Gruppe (vgl. P2P-Netze)
- Protokollkomponenten
 - Gruppenverwaltung
 - Optimierung der Overlay-Netz-Struktur
 - direkte Konstruktion eines Verteilbaums aus Kenntnis über Nachbarn
 - Bilden von vermaschten Strukturen und Berechnen von Shortest Path Spanning Trees ausgehend von Sendern.

