18C3-Workshop

Switched Ethernet

Probleme mit Ethernet...

- Hubs/Repeater sind i.A. reine Signalverstärker auf Layer I.

- Hubs besitzen i.A. keinerlei Intelligenz um auf Netzwerkprobleme angemessen reagieren zu können.

- CSMA/CD ist mehr als unbrauchbar bei größeren Netzwerken (Datendurchsatz, Leitungslänge, Paketgröße, …)

- Hubs lassen sich nur sehr eingeschränkt managen (Prioritäten, Adressfilter, …)

- Keine Sicherheit gegen das Mitlesen (”Sniffen”) fremder Daten (aber: NeedToKnow [3Com], “Smart Hubs” CD-Signal)

Was wünscht man sich?

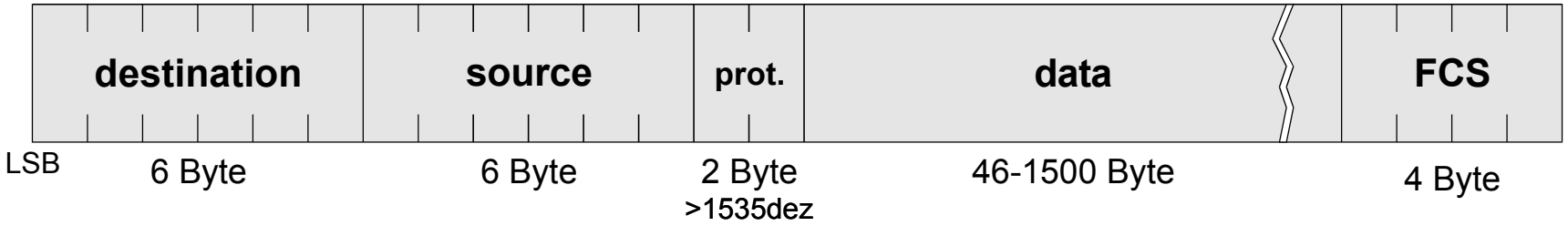
- Man will so selten wie möglich CSMA/CD verwenden müssen
- Bessere Auslastung des Netzwerkes, weniger “unnützer” Datenverkehr
- Netzwerk-Standardprobleme sollen automatisiert gelöst werden (Schleifenauflösung, Backup-Leitungen, etc.)
- Netzwerk-/Datensicherheit soll gewährleistet werden
- Das Netzwerk soll in seinen Knotenpunkten managebar werden
- Das Ganze soll transparent für die bisherigen Netzteilnehmer bleiben.

Switched Ethernet

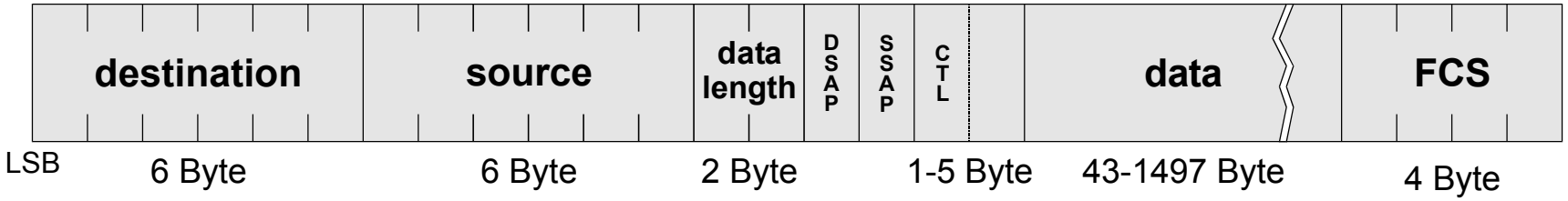
(via transparentem Bridging, IEEE 802.1d)

Paketformate

Ethernet Frame (nach Xerox (Intel, Digital))



IEEE 802.3 Frame



DSAP = Destination Service Access Point
SSAP = Source Service Access Point
CTL = CTL

Adressformat

LSB b0

Global(0) / Local(1) - Bit

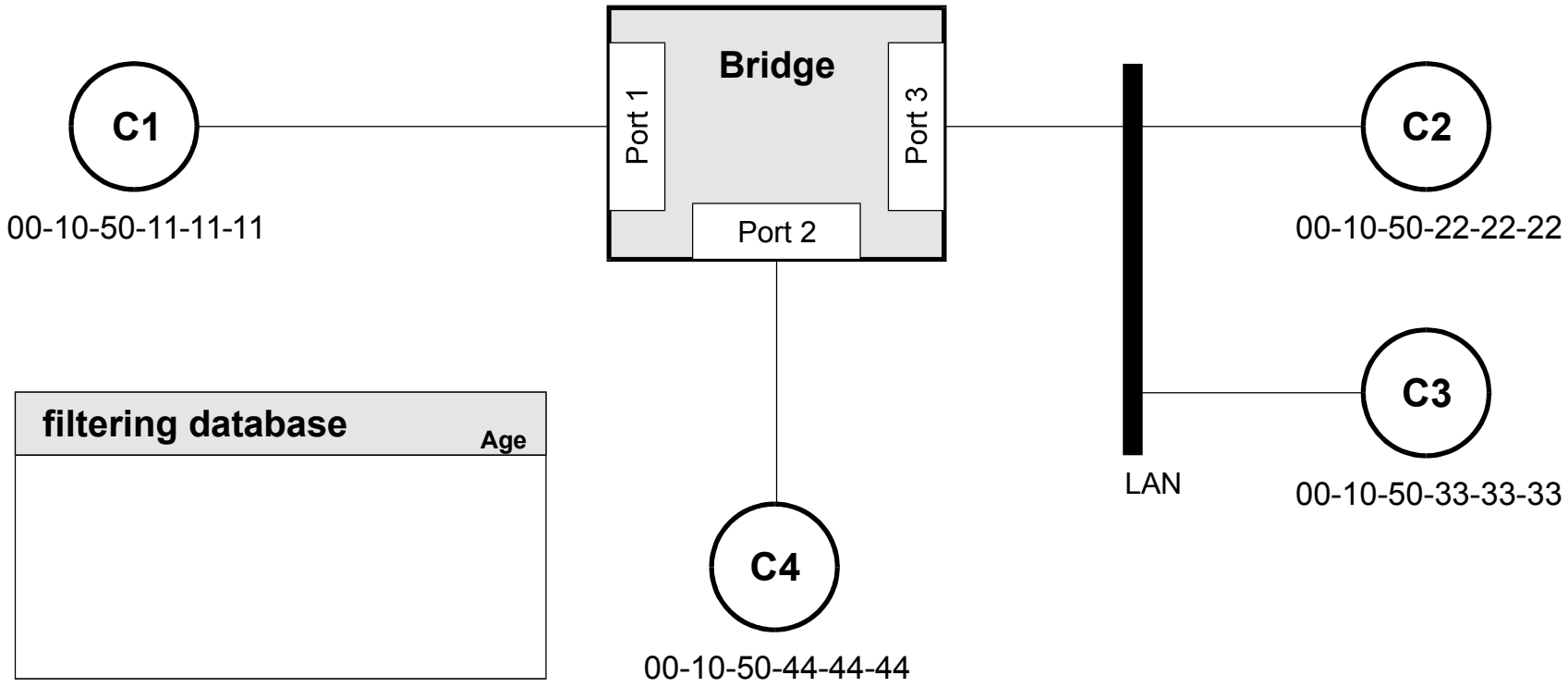
DMac: **Group(1) / Individual(0) - Bit**

SMac: **Transparent Bridging(0) / Source Route Bridging(1) - Bit**

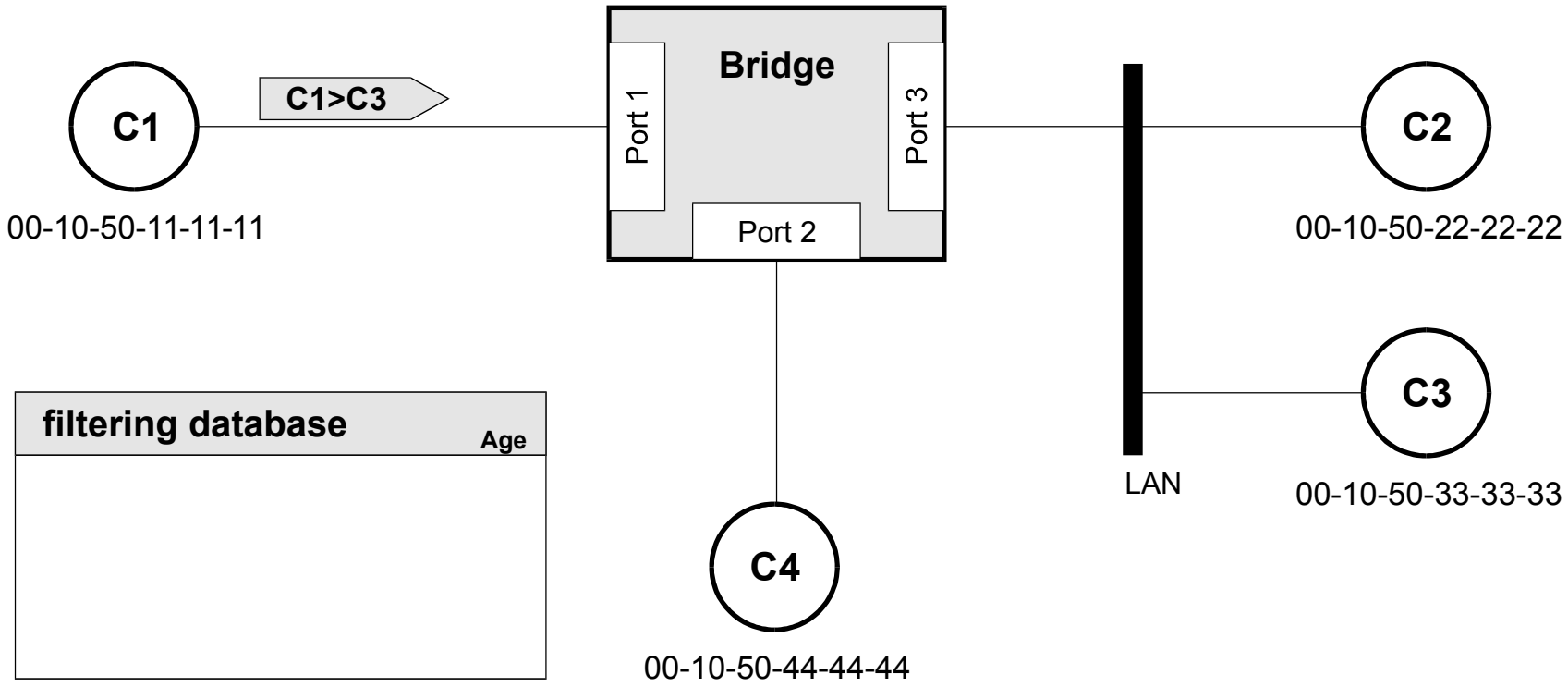
z.B. 01-80-c2-00-00-00
Bridge Group Address (STP)

OUI = Organizationally Unique Identifier

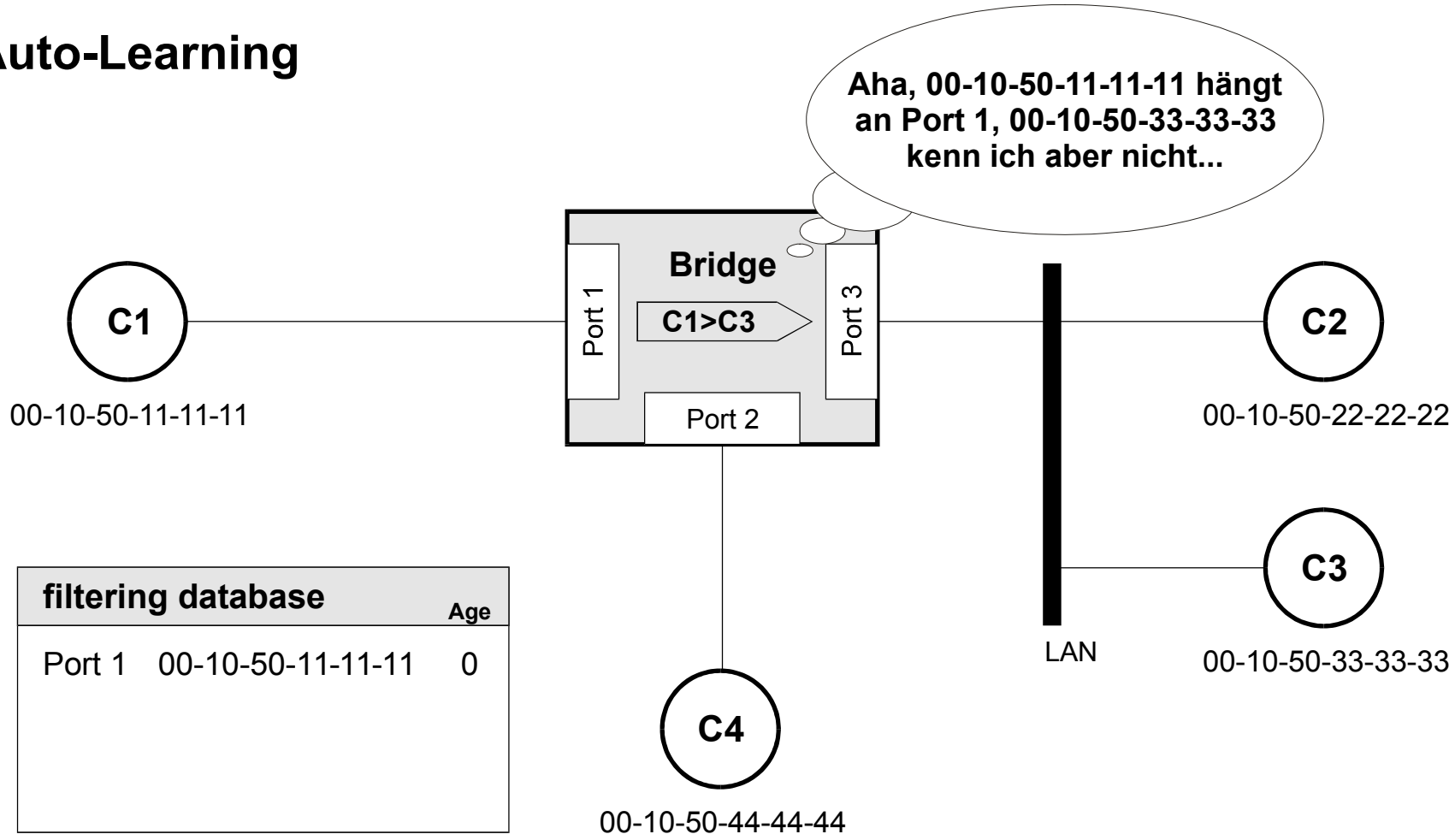
Auto-Learning



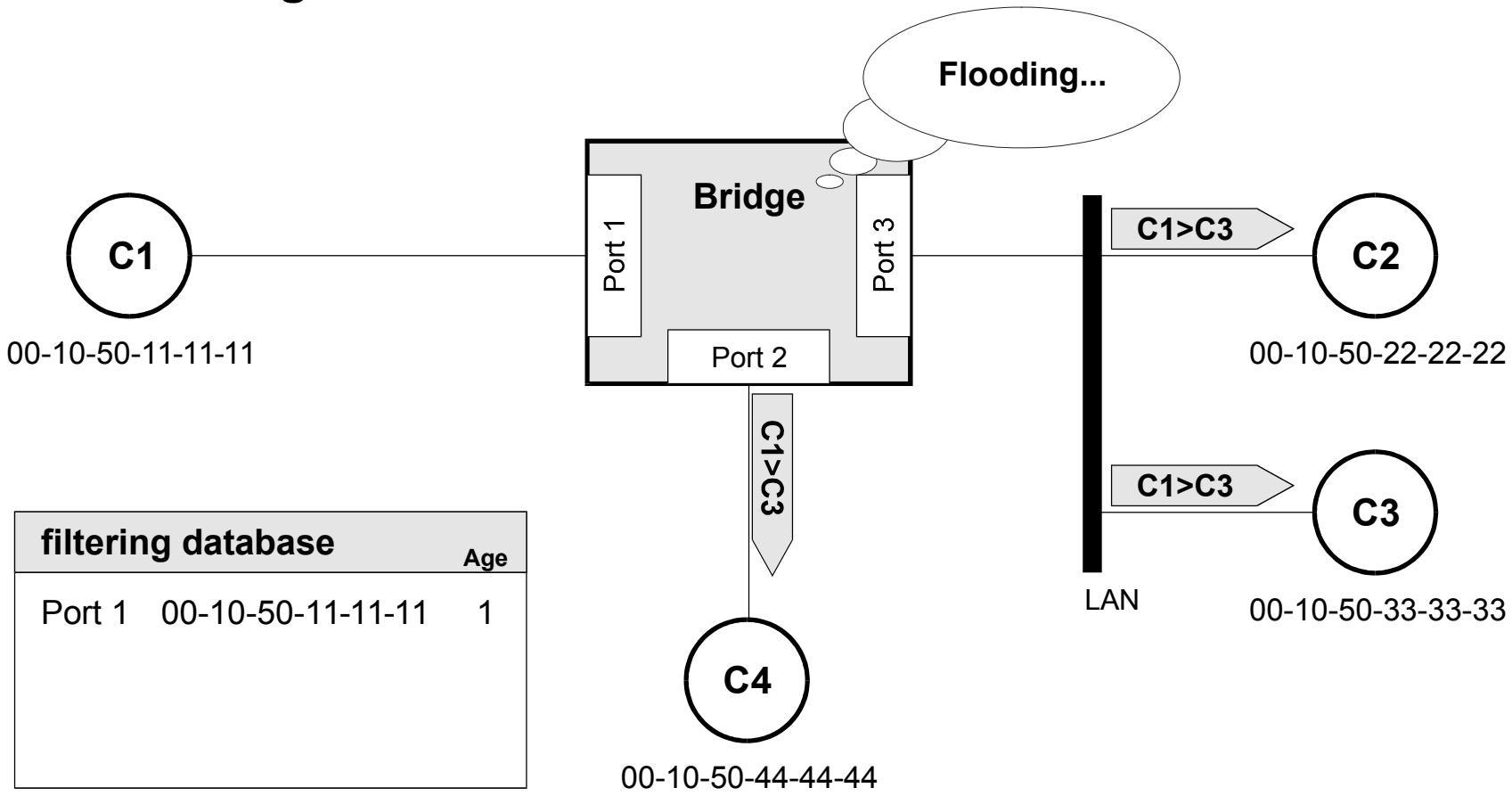
Auto-Learning



Auto-Learning

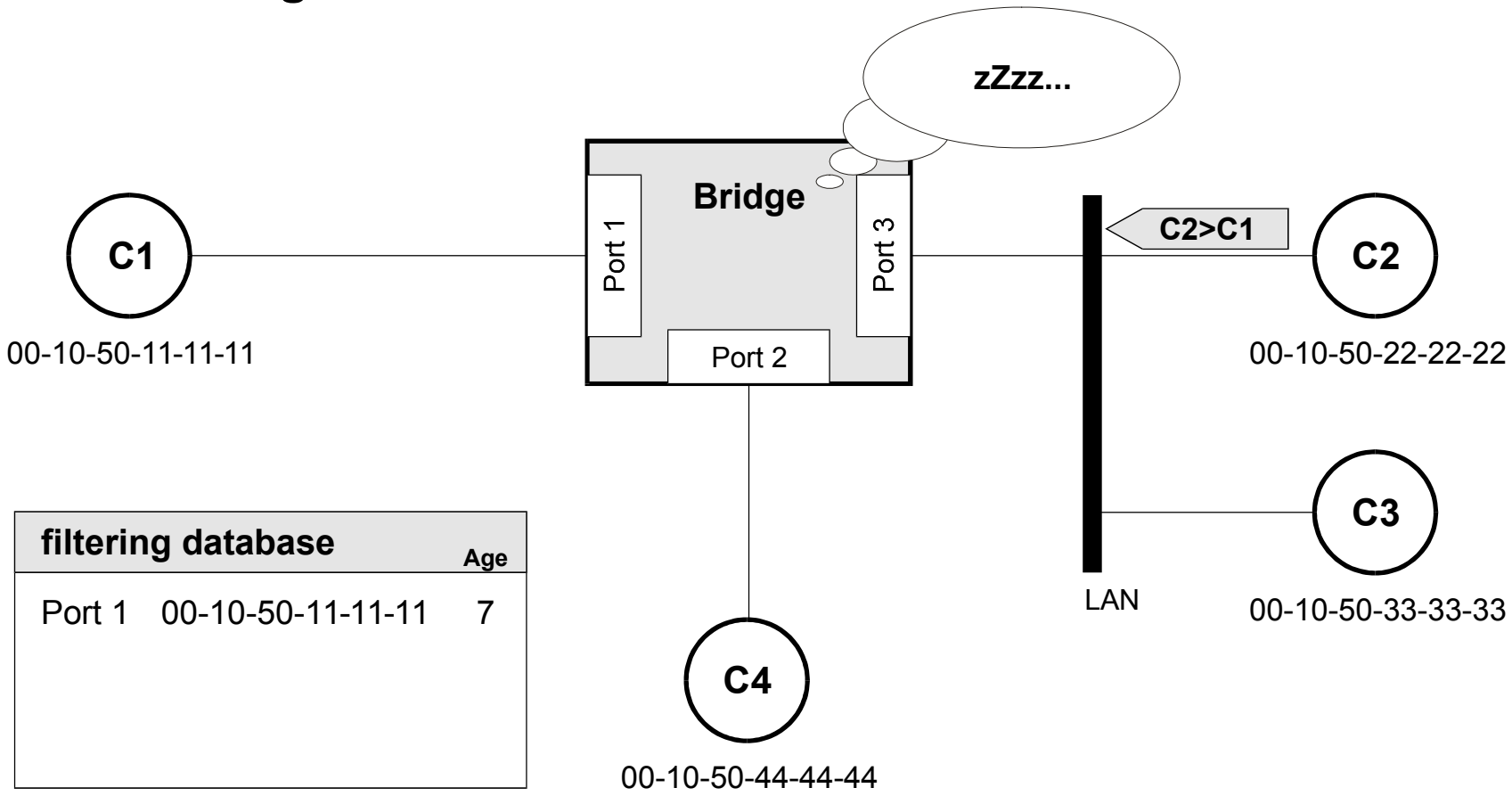


Auto-Learning

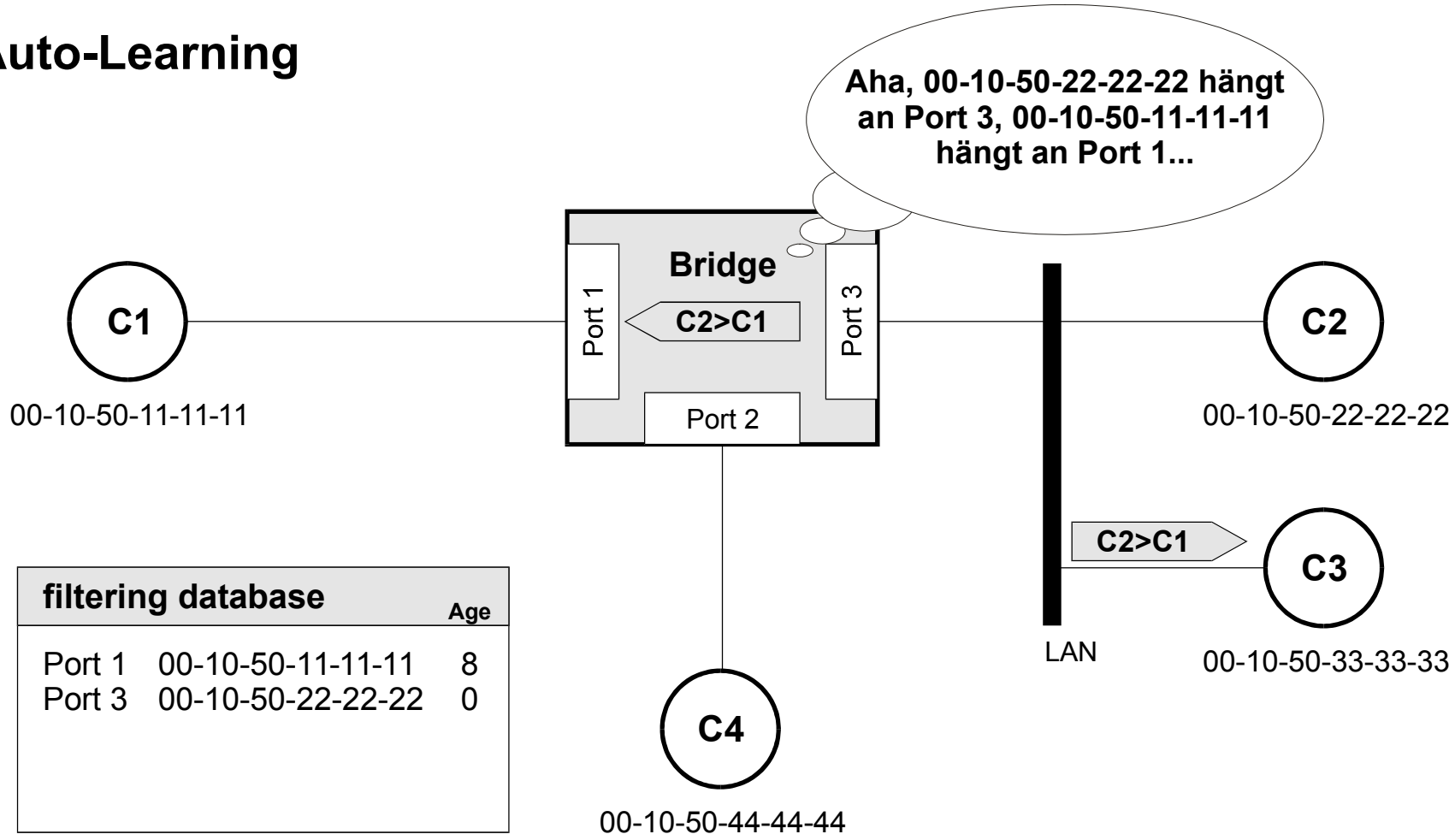


Achim Friedland <mail@ahzf.de>

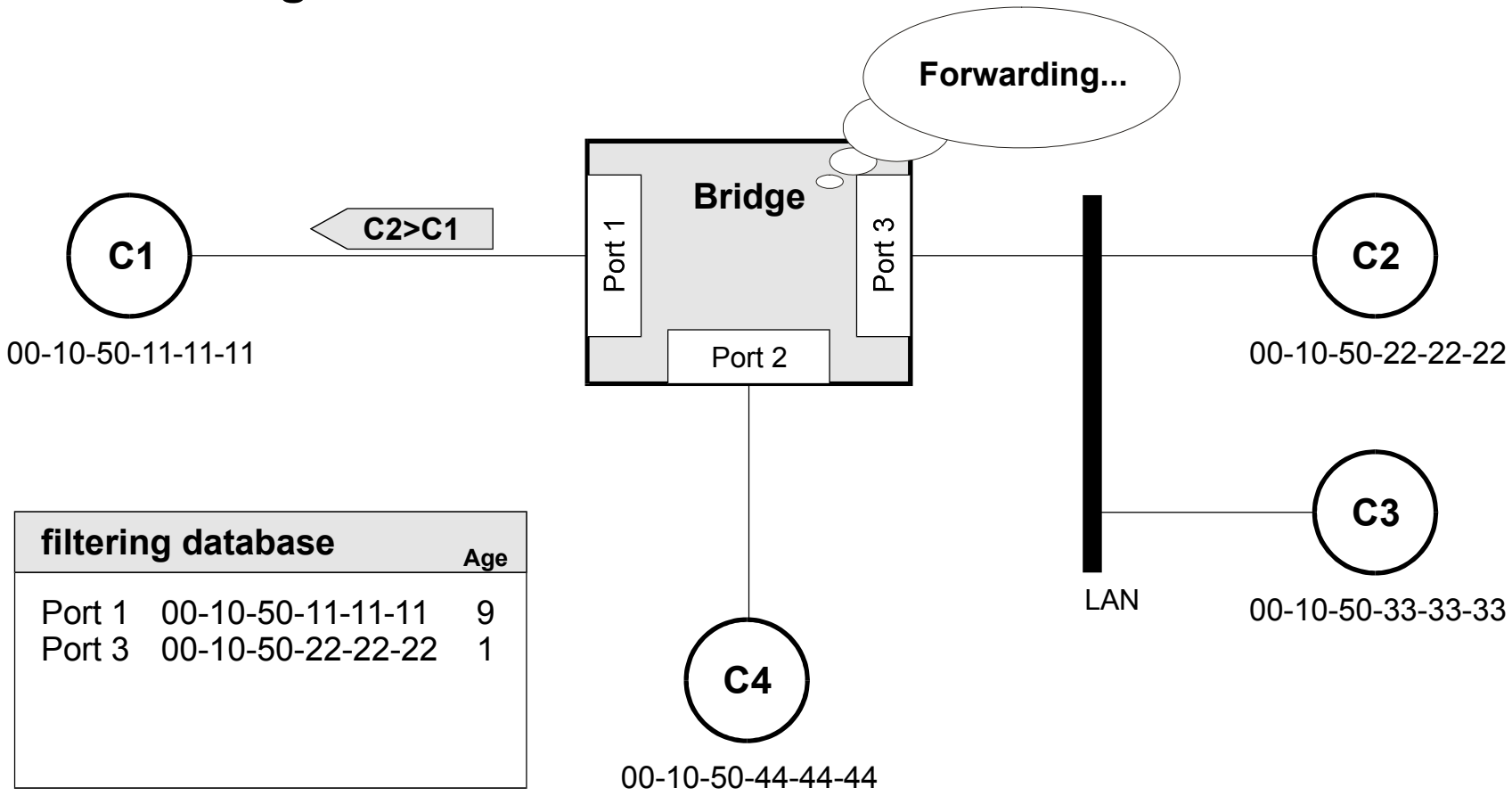
Auto-Learning



Auto-Learning

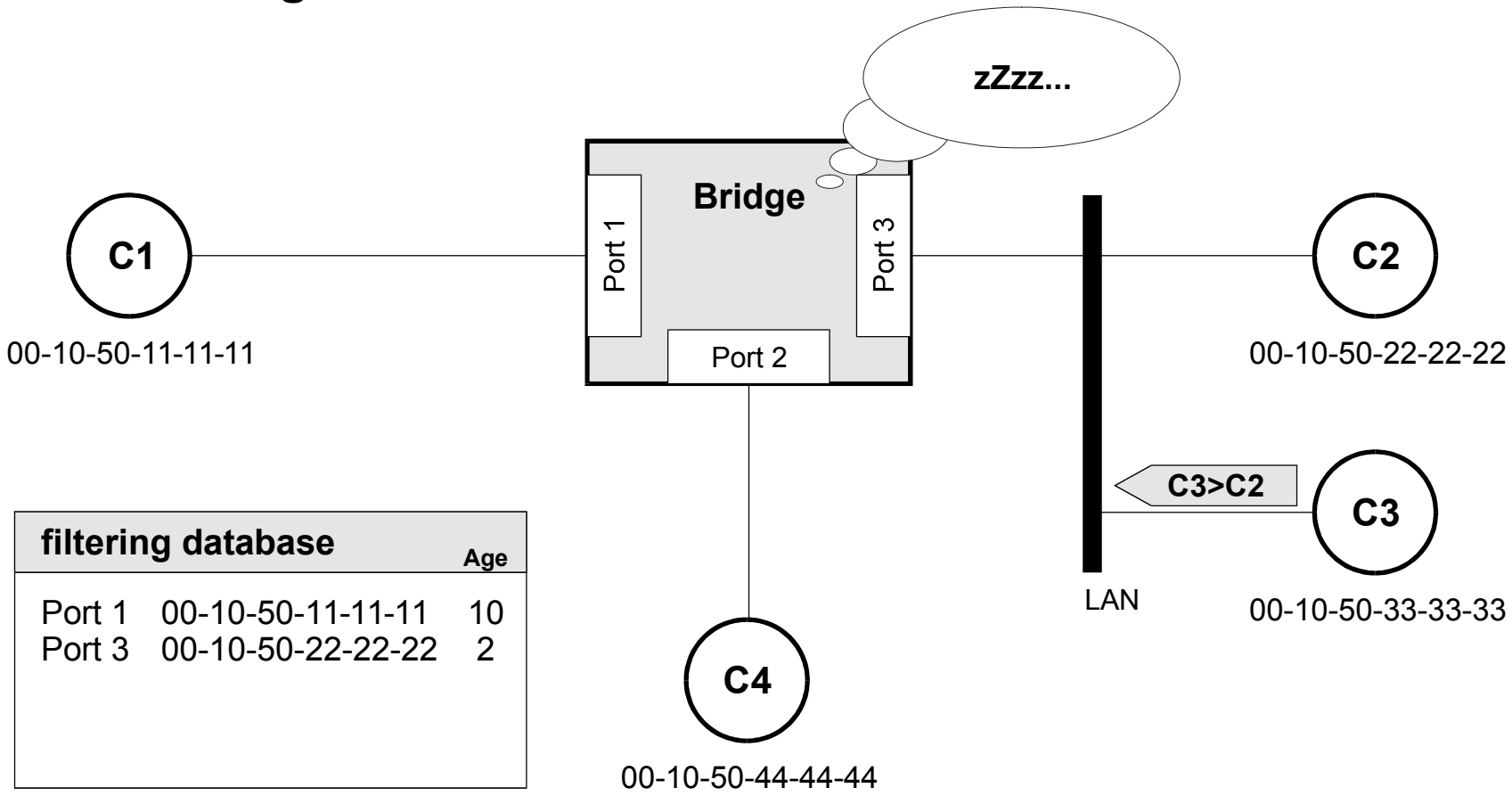


Auto-Learning



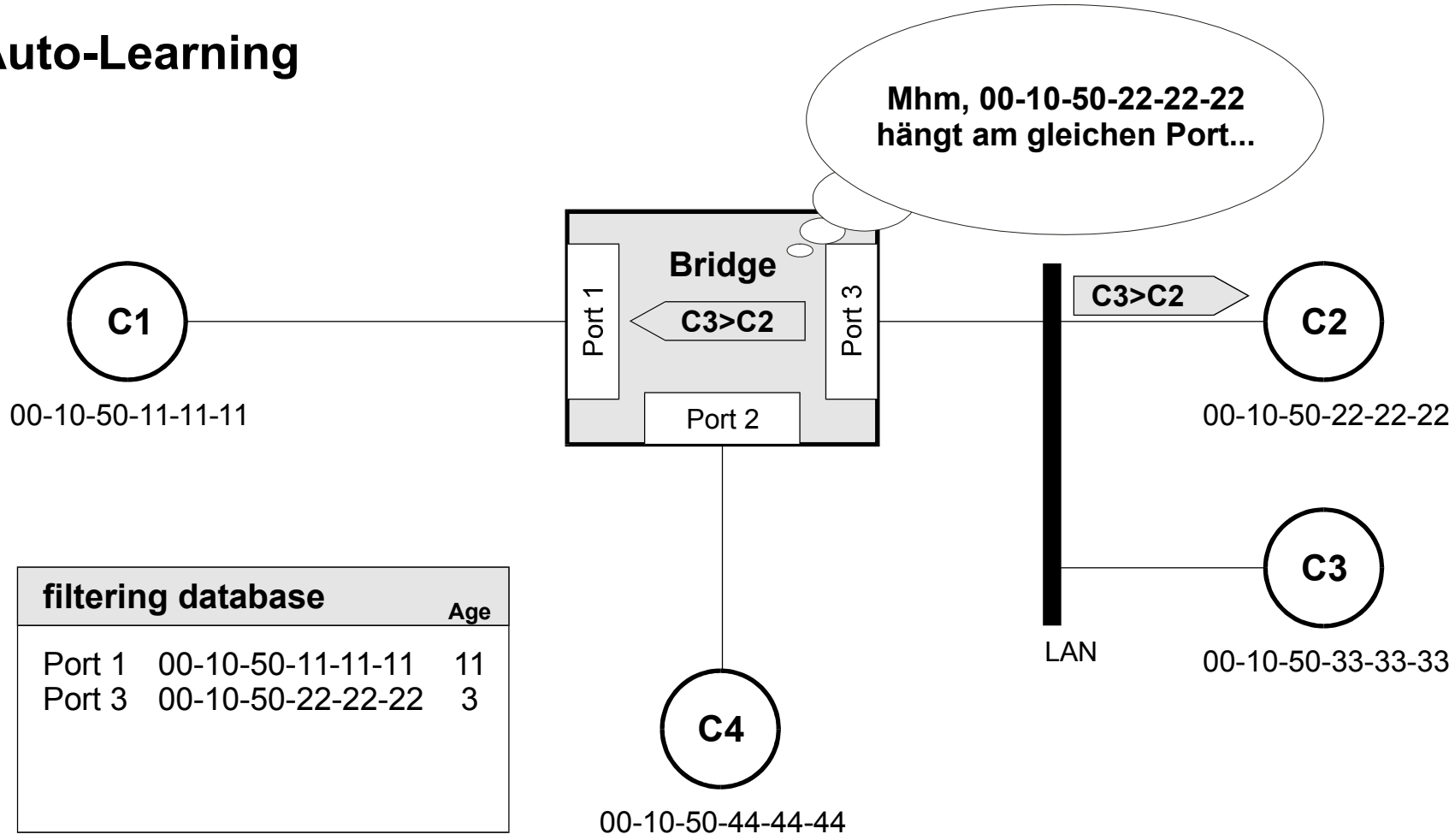
filtering database		Age
Port 1	00-10-50-11-11-11	9
Port 3	00-10-50-22-22-22	1

Auto-Learning



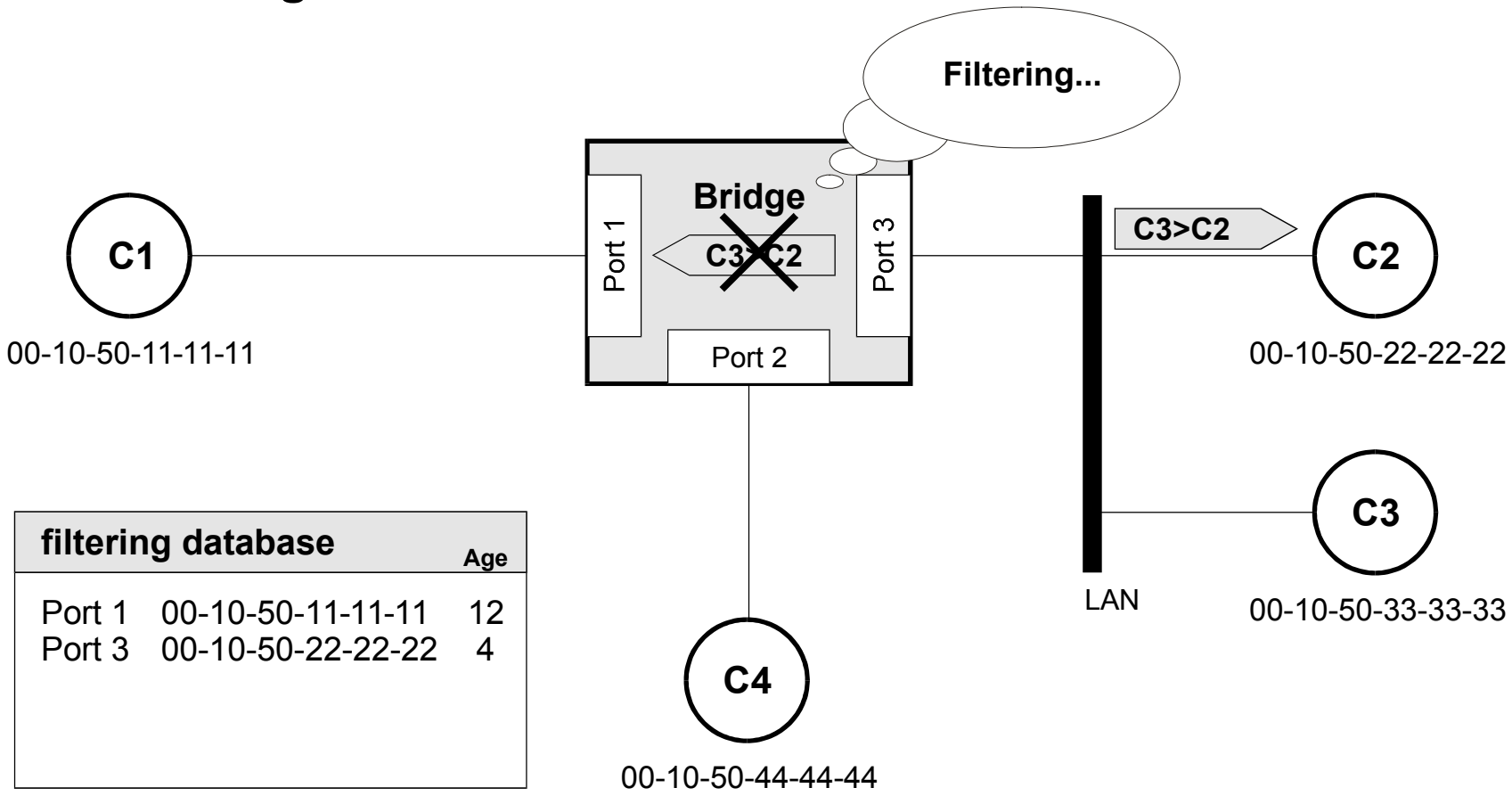
filtering database		Age
Port 1	00-10-50-11-11-11	10
Port 3	00-10-50-22-22-22	2

Auto-Learning



filtering database		Age
Port 1	00-10-50-11-11-11	11
Port 3	00-10-50-22-22-22	3

Auto-Learning



filtering database		Age
Port 1	00-10-50-11-11-11	12
Port 3	00-10-50-22-22-22	4

Auto-Learning - Vorteile

- Sofern Rechner direkt an Switches angeschlossen werden und eine Full-Duplex Verbindung möglich ist, entfällt CSMA/CD ;)

- Sofern Rechner direkt an Switches angeschlossen werden bekommen sie nur noch Pakete die für sie bestimmt sind...
  Es ist kein Mitlesen/"Sniffen" fremder Daten möglich (erstmal!)

- In voll geswitchen Netzwerken kann theo. jeder Rechner mit seiner vollen "Bandbreite" kommunizieren ohne andere Rechner zu beeinflussen. (Abhängig von der Backplane der Bridge)

- Die gelernten MAC-Adressen werden mit einem Timer versehen und nach Ablauf dessen gelöscht.

- Auto-Learning mit Switches (transparenten Bridges) funktioniert auch mit mehreren kaskadierten Switches.

Auto-Learning - Nachteile

- Im geswitchten Ethernet können Pakete verlorengehen, ohne dass der Versender darüber informiert wird... Egal, da Ethernet sowieso für gar nichts garantiert (anders: Token Bus/Ring)...

- Broadcasts/Multicasts werden in handelsüblichen Bridges immer “geflutet”... (besser: Extended Filtering Rules/IEEE 802.1d)

- MAC-Adressen können statisch an mehrere Ports eingetragen werden. Dynamisch sollten sie jedoch nur an einem Port eingetragen werden.

-
  Handelübliche Implementierungen lassen es zu, dass eine MAC-Adresse an mehreren Ports gelernt wird. Das kann man dazu ausnutzen über Bridges hinweg zu sniffen. Als Gegenmassnahme könnten statische MAC-Filter an den Ports definiert werden.

Switching Strategien

Store-and-Forward

Das Ethernet-Frame wird komplett in den Cache der Bridge eingelesen, die CRC wird berechnet, verglichen und danach wird entschieden wohin das Frame weitergeleitet werden soll...
Sicheres Verfahren, aber lange Verzögerung

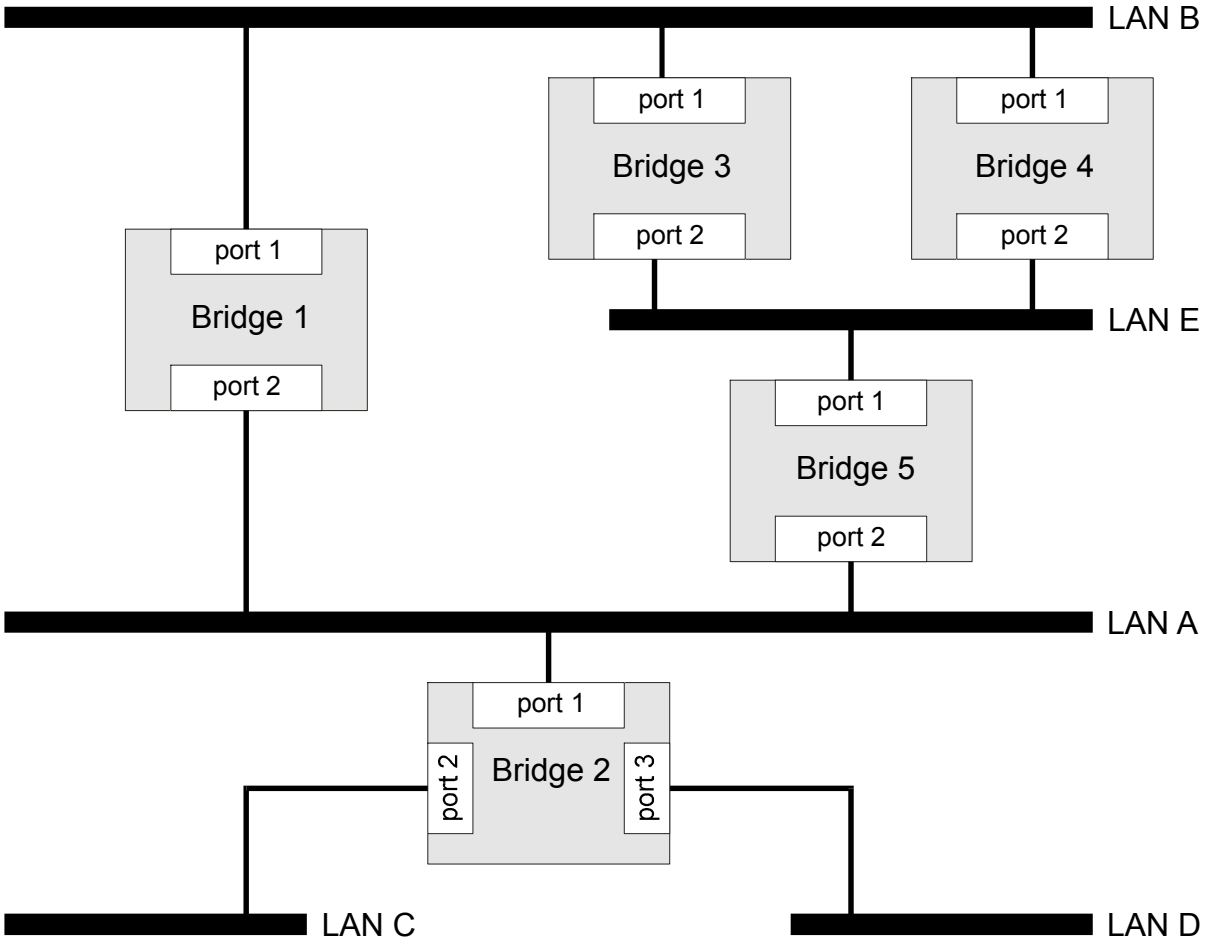
Cut-Through

Nur die Destination MAC-Adresse wird eingelesen, die Zielports werden ausgewählt und der Rest des Frames wird direkt weitergeleitet...
Deutlich kürzere Verzögerungen aber möglicherweise Weiterleitung defekter Pakete. Funktioniert nicht bei Verbindungen zwischen Ports mit unterschiedlichen Geschwindigkeiten.

Fragment-Free

Die ersten 64 Byte werden gelesen, wenn okay, dann wie Cut-Through

Redundante Netzwerktopologien



Redundante Netzwerktopologien

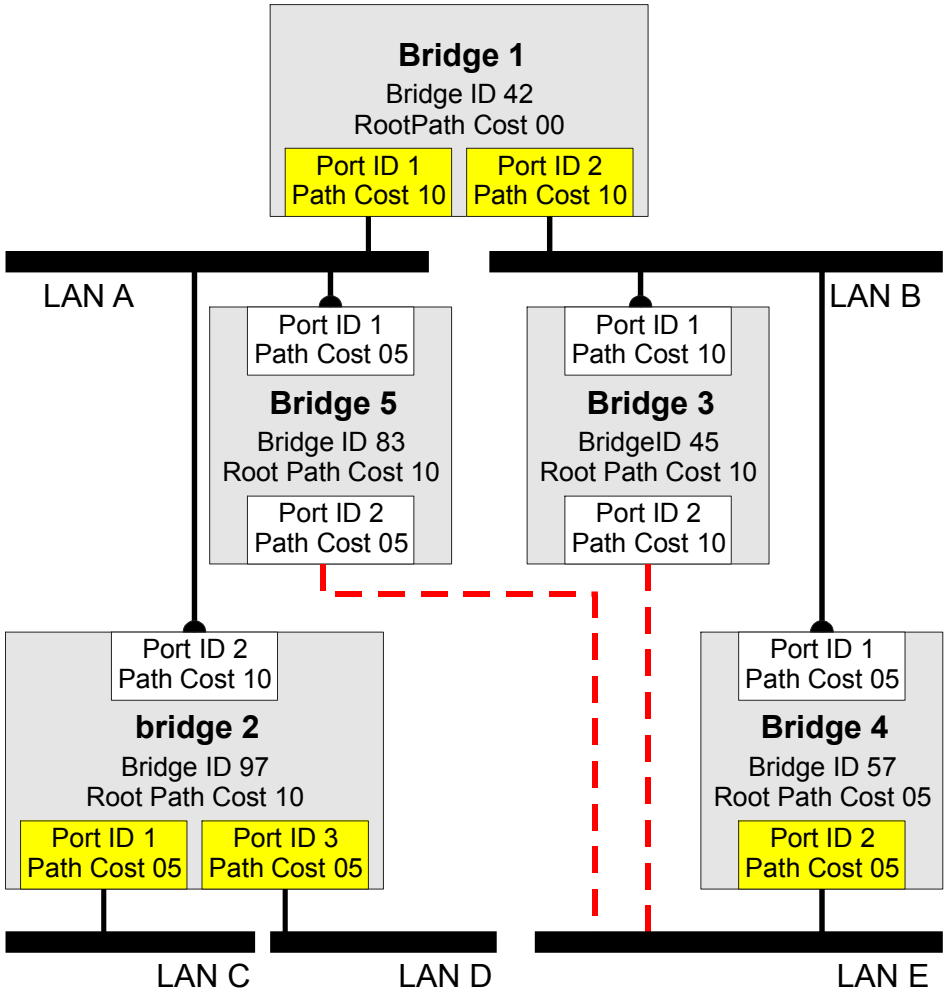
- Man will bewusst redundante Netzwerkpfade verwenden, die sobald der Hauptlink ausfällt automatisch aktiv werden können.
- “Ungeplante” Redundanzen (sogenannte Schleifen) sollen automatisch erkannt und zu “guten” Redundanzen gemacht werden.
- Das Ganze soll deterministisch ablaufen und reproduzierbare Ergebnisse liefern.
- Für die Netzwerkendgeräte soll das Verwenden eines redundanten Pfades transparent bleiben (anders: Source Route Bridging)

Spanning-Tree-Algorithmus

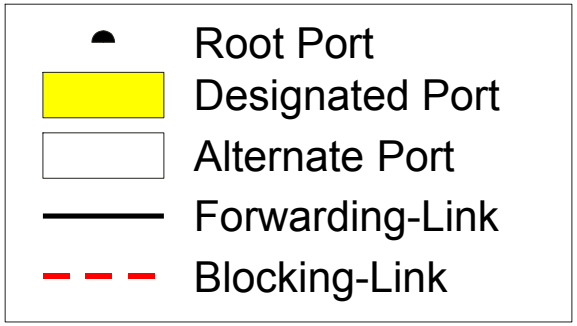
Algorhyme
I think that I shall never see
A graph more lovely than a tree.
A tree whose crucial property
Is loop-free connectivity.
A tree that must be sure to span
So packets can reach every LAN.
First, the root must be selected.
By ID, it is elected.
Least-cost paths from root are traced.
In the tree, these paths are placed.
A mesh is made by folks like me,
Then bridges find a spanning tree.

- Radia Perlman

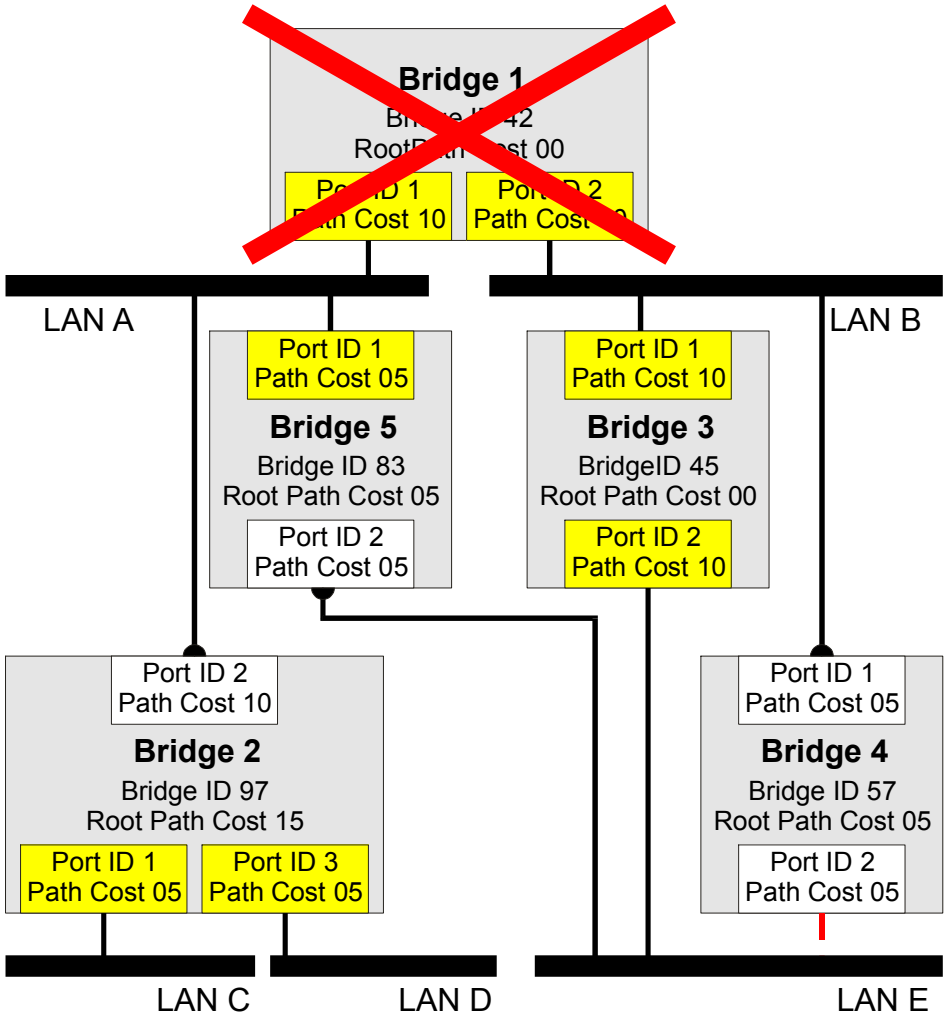
Spanning-Tree-Algorithmus



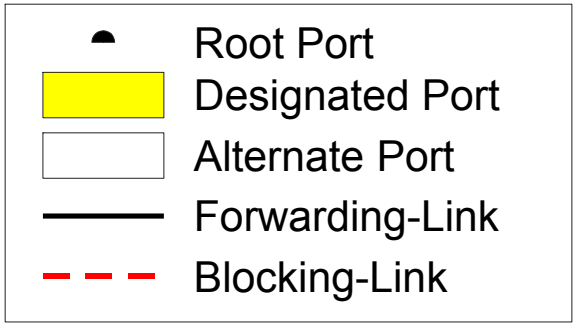
1. Bridge 1 wird Root-Bridge
2. Bridge 1 wird Designated Bridge für LAN A and LAN B
3. Bridge 2 wird Designated Bridge für LAN C and LAN D
4. Bridge 4 wird Designated Bridge für LAN E



Spanning-Tree-Algorithmus



1. Bridge 3 wird Root-Bridge
2. Bridge 3 wird Designated Bridge für LAN B und LAN E
3. Bridge 5 wird Designated Bridge für LAN A
4. Bridge 2 wird Designated Bridge für LAN C and D



Spanning-Tree-Algorithmus

- SPA läuft an... Alle Bridges schalten ihre Ports in den Blocking Modus Es werden nur Configuration BPDUs ausgetauscht.

- Alle Bridges senden C-BPDUs in die angeschlossenen LANS aus, um sich als mögliche ROOT Bridge bekannt zu machen.

- Enthält eine Bridge eine C-BPDU mit niedrigerer Bridge ID announced sie diese in Zukunft als ihre ROOT Bridge.

- Die Bridge mit der niedrigsten Bridge ID wird schließlich zur ROOT Bridge des gesamten Netzwerkes.

- Alle Bridges berechnen den kürzesten Pfad von sich zur Root Bridge und wählt einen Port aus über den sie die Root Bridge mit minimalen Pfadkosten erreichen kann (ROOT Port).

Spanning-Tree-Algorithmus

- Sind mehrere Bridges an einem LAN angeschlossen, so vergleichen sie ihre Pfadkosten zur ROOT Bridge. Die mit den niedrigeren Werten wird Designated Bridge auf dem entsprechenden LAN. Im Zweifelsfall entscheidet die Bridge ID, dann die Port ID.

- Alle ROOT Ports und die Designierten Ports werden in einen Pre-Forwarding Modus geschaltet (Listening, Learning). Nachdem einigermaßen sichergestellt ist, dass allen Bridges die Topologie bewußt ist wird in den Forwarding Modus geschaltet.

- Ca. Alle 2 Sek. macht sich die ROOT Bridge via C-BPDUs auf all ihren direkt angeschlossenen LANs bekannt. Alle Bridges die diese erhalten senden ebenso ihre C-BPDUs über ihre designierten Ports aus.

- Empfangene und nicht erneut empfangene C-BPDUs werden nach einem Timeout (ca. 20 Sek) gelöscht. Es wird angenommen, dass ein Fehler vorliegt und ein neuer Weg zur ROOT Bridge berechnet.

Spanning-Tree-Algorithmus - Topologieänderungen

- Schaltet eine Bridge den Blockierungs Modus eines ihrer Ports an oder aus. So teilt sie diese Topologieänderung der via ROOT Port erreichbaren Bridge mit. Diese teilt es ihrer mit... usw.

- Gleichzeitig setzt diese Bridge eine Bestätigung über die empfangene TÄ-Mitteilung in ihren weiteren C-BPDUs.

- Empfängt die ROOT Bridge eine Topologieänderungsnachricht, so sendet sie ebenfalls eine Bestätigung derer.

- Typischerweise braucht ein STA ca. 30-35 Sekunden

IEEE 802.1d Configuration BPDU

Okt.	7	6	5	4	3	2	1	0	
2	protocol id (0000)								
1	version id (00)								
1	bpdu type (00)								
1	TCA	reserviert						TC	
2	root priority								
6	root id								
4	pathcost to root								
2	bridge priority								
6	bridge id								
1	port priority								
1	port id								
2	message age								in 1/256 secs
2	max age (20 sec.)								in 1/256 secs
2	hello time (2 sec.)								in 1/256 secs
2	forward delay (15 sec.)								in 1/256 secs

IEEE 802.1d Topoloy Change BPDU

Okt.	7	6	5	4	3	2	1	0
2	protocol id (0000)							
1	version id (00)							
1	bpdu type (80hex)							

Bridge PDUs werden immer an die Ethernet Multicast Adresse 01-80-C2-00-00-00 gesand. Als Source dient jeweils die MAC-Adresse des Bridge Ports über den die BPDU versand wurde.

Angriffe auf Spanning-Tree?

- Man kann sich recht einfach zur ROOT Bridge des Netzwerkes ernennen und durch erfundene Topologie Änderungen das Netzwerk mehr mit Spanning-Tree als mit Forwarding beschäftigen.

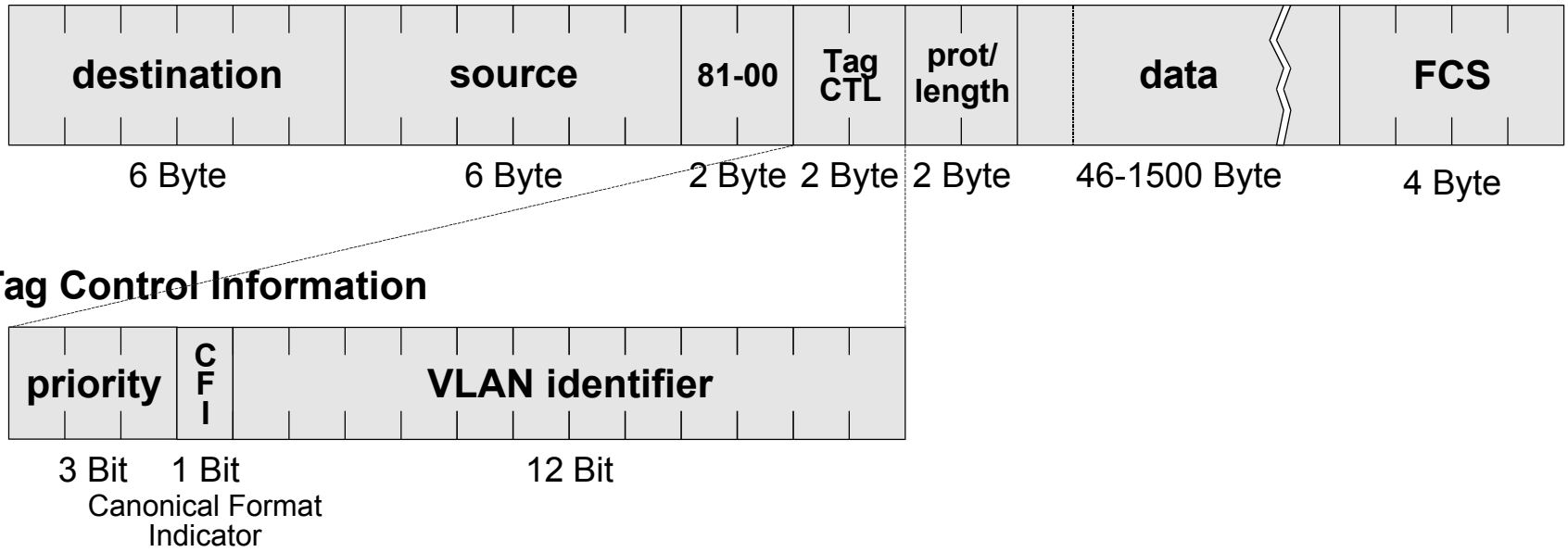
- In einem gewissen Umfang ist es möglich gefälschte C-BPDUs zu versenden um alle anderen Bridges zum Blocken zu verleiten.

- Beides sind recht unbekannte Angriffe die netterweise mit recht wenig Zutun große Verwirrung stiften können und dabei kaum auffallen.

- Als Gegenmassnahme kann Spanning Tree für alle außer Interbridgeports deaktiviert werden.

Virtuelle LANs

Ethernet Frame mit VLAN Markierung nach IEEE 802.1q (One-Level Tagging)



- VLAN-ID fließt in die CRC Berechnung ein (hohe Verzögerung möglich)
- Normalerweise ein Spanning-Tree für jedes VLAN...
- VLAN ist auch auf Basis IEEE 802.10 (ISL) möglich

IEEE 802.1 Standards

802.1s	Multiple Spanning Trees
802.1w	Rapid Reconfiguration of Spanning Tree
802.1x	Port Based Network Access Control
802.1D	MAC bridges
802.1G	Remote MAC bridging
802.1Q	Virtual LANs
802.1v	VLAN Classification by Protocol and Port

Literatur

RadiaPerlman
Bridges, Router, Switches und Internetworking Protocols
Addision-Wesley

Matthew Naugle
Network Protocol Handbook
McGraw-Hill Series on Computer Communication

Dr. Franz-Joachim Kauffels
Lokale Netze
Datacom/MITP