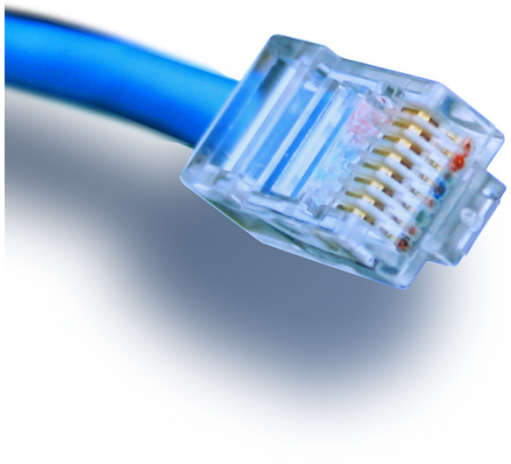
HOUSE OF TECHNOLOGY

– en del af mercantec+



Spanning Tree

- mulighed for redundans på Ethernet!

Netteknik 1

Hvorfor Spanning Tree?

HOUSE OF TECHNOLOGY

– en del af mercantec

- Driftsikkerhed & redundans!
  - Moderne firmanetværk kræver en større og større grad af **stabilitet og driftsikkerhed**, fordi flere og flere arbejdsrutiner og -processer bliver baseret på brugen af IT og netværk
  - Ved at indføre **redundans** i firmaets netværk på alle de kritiske enheder, typisk enhederne som er placeret på distributionslaget, kan man eliminere alle single-point-of-failure situationer fuldstændig
  - Ved samtidig at opretholde en effektiv **netværksovervågning** samt et godt **IT beredskab** tilknyttet firmaet kan man holde 'nedetiden' meget, meget lav på firmanetværket
  - I dette kapitel kigger vi nærmere på **Spanning Tree protokollerne** samt de teknikker som benyttes til at skabe **redundans** på Ethernet enheder og netværk

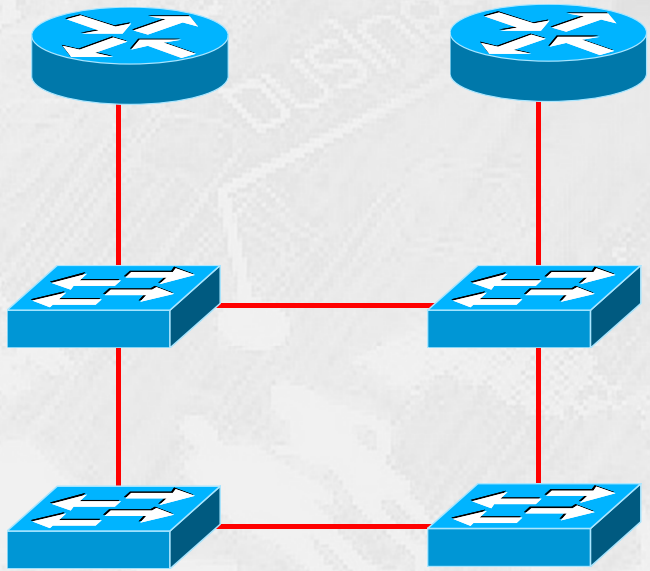
Spanning Tree Protocol

HOUSE OF TECHNOLOGY

– en del af mercantec

- **Spanning Tree Protocol,** STP, er en **Ethernet switch-teknologi** der er indført for at kunne håndtere de forskellige problemer der opstår på et **switched LAN netværk**, når man kobler det op med flere veje mellem de enkelte switche (redundans).

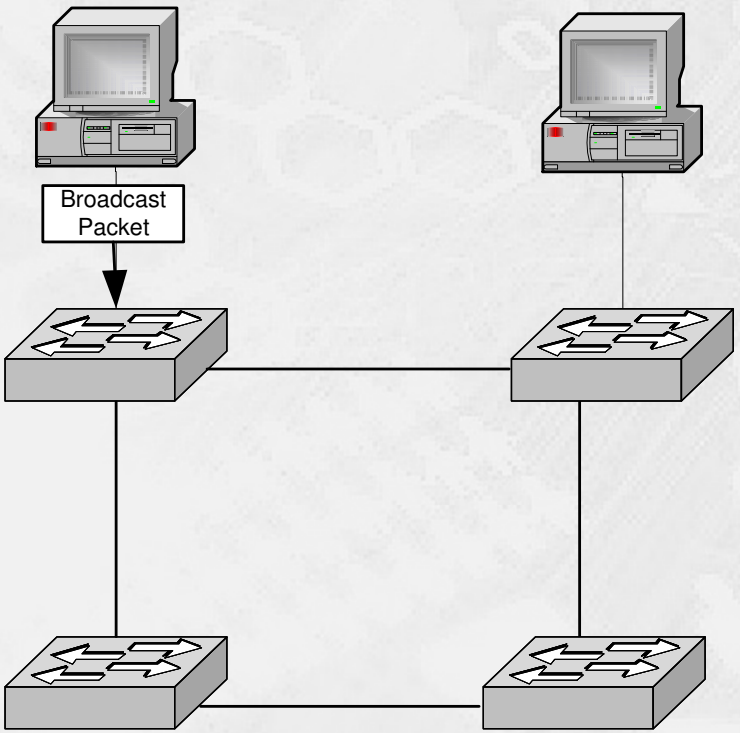
- Problemerne kan være:
  - Broadcast storms
  - Dublerede pakker
  - Ustabile MAC-adresse tabeller …



Problemet **broadcast storm**

HOUSE OF TECHNOLOGY

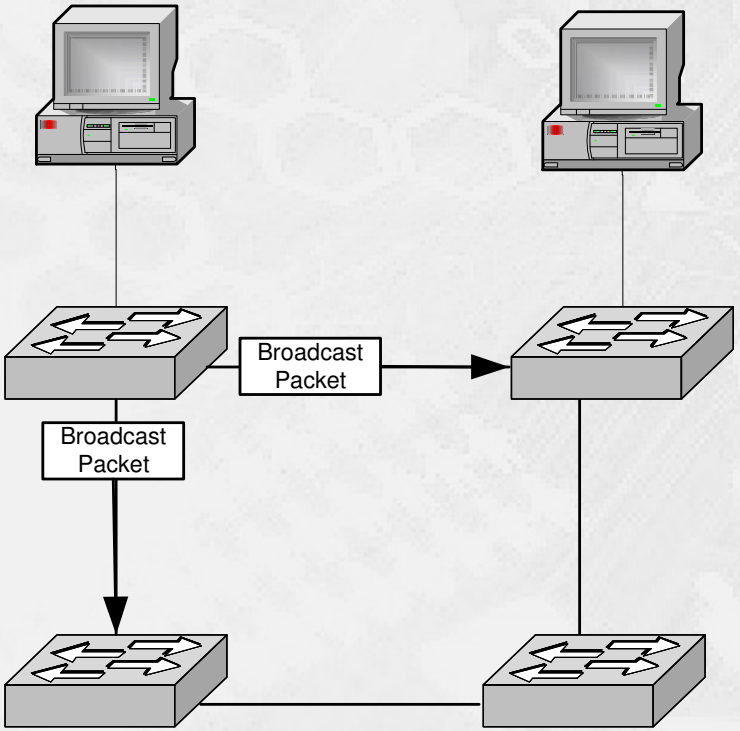
– en del af mercantec



Problemet **broadcast storm**

HOUSE OF TECHNOLOGY

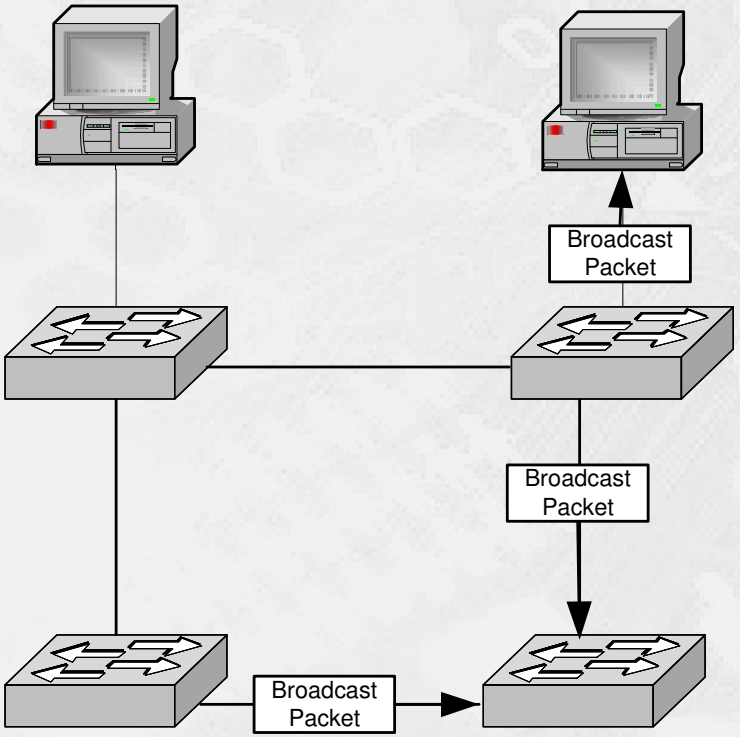
– en del af mercantec



Problemet **broadcast storm**

HOUSE OF TECHNOLOGY

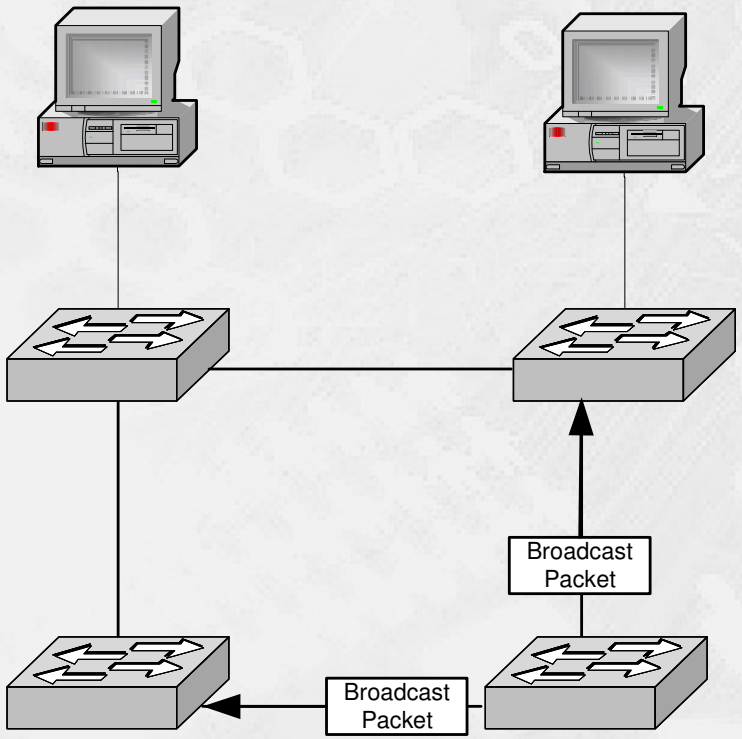
– en del af mercantec



Problemet **broadcast storm**

HOUSE OF TECHNOLOGY

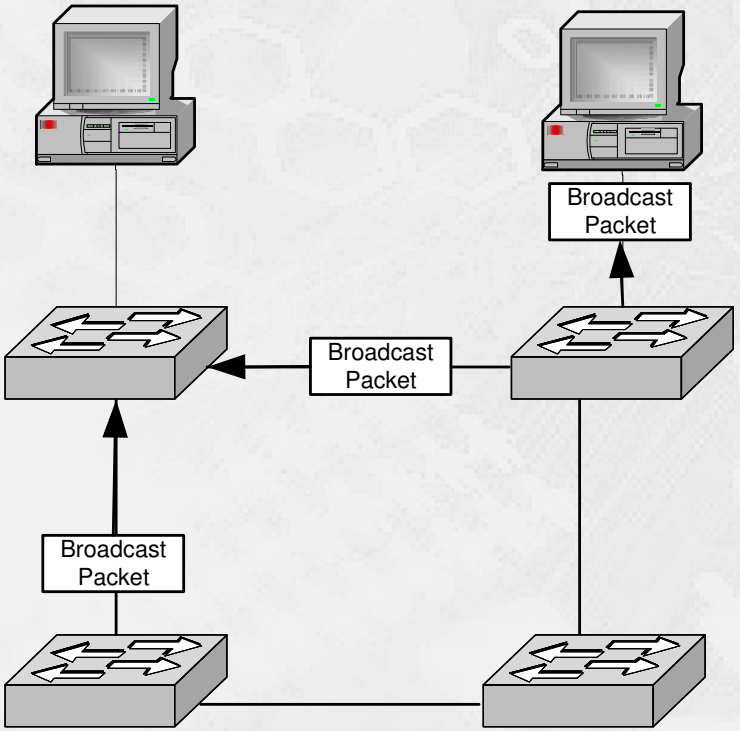
– en del af mercantec



Problemet **broadcast storm**

HOUSE OF TECHNOLOGY

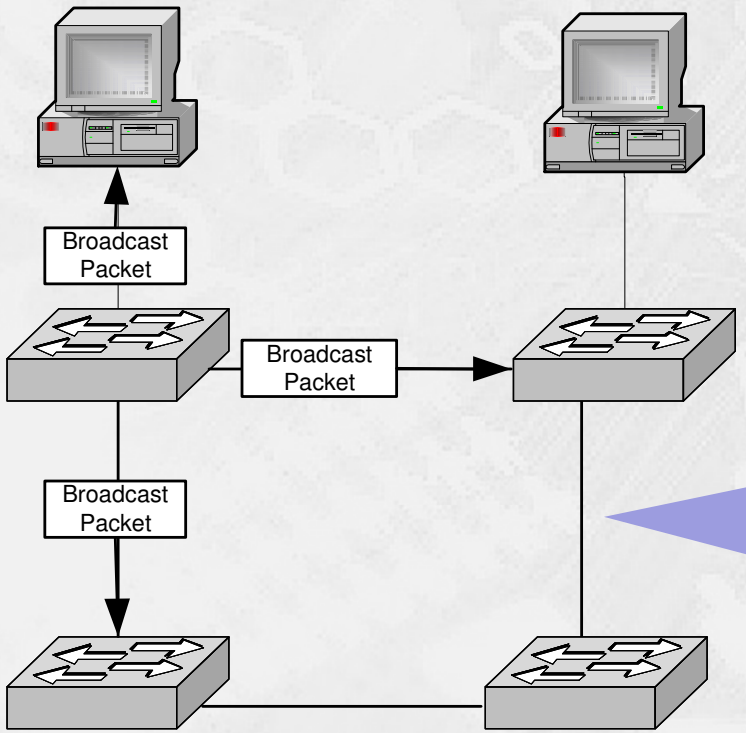
– en del af mercantec



Problemet **broadcast storm**

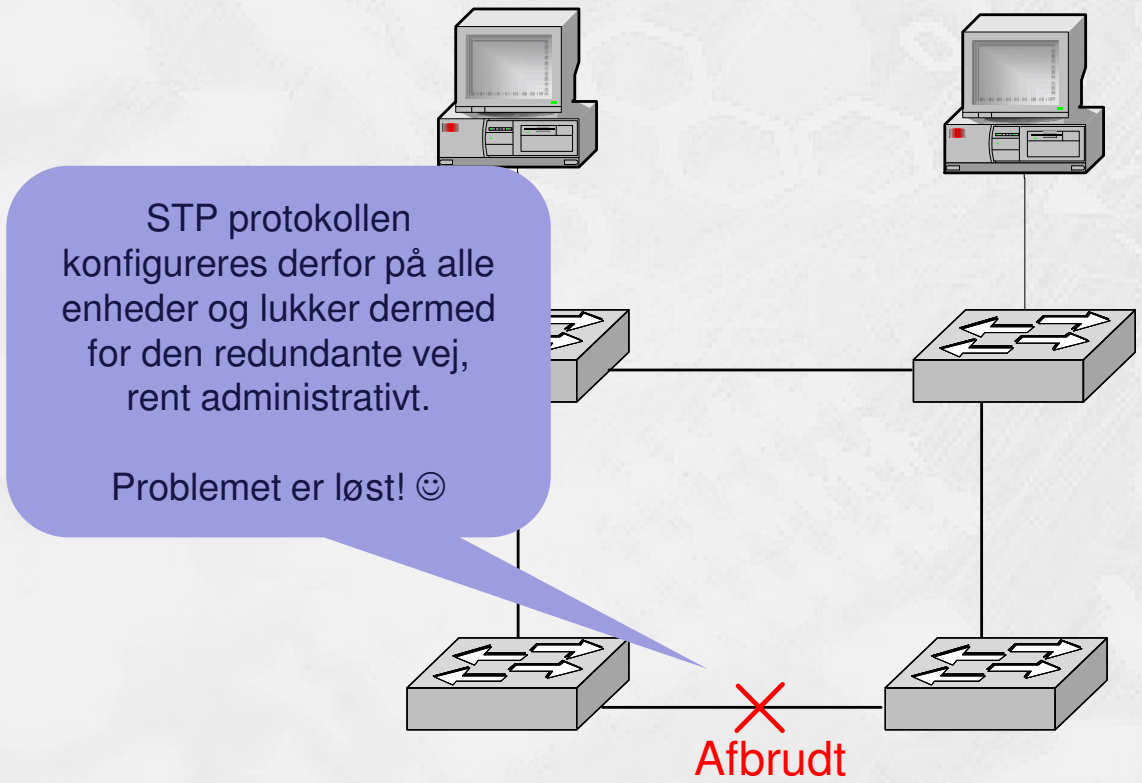
HOUSE OF TECHNOLOGY

– en del af mercantec



Hmmm, Ethernet standarderne på lag 2 mangler vist et '**maksimalt antal hop**' felt i protokollen, f.eks. som TTL feltet der er i en ip datapakke … ☺

Problemet **broadcast storm**

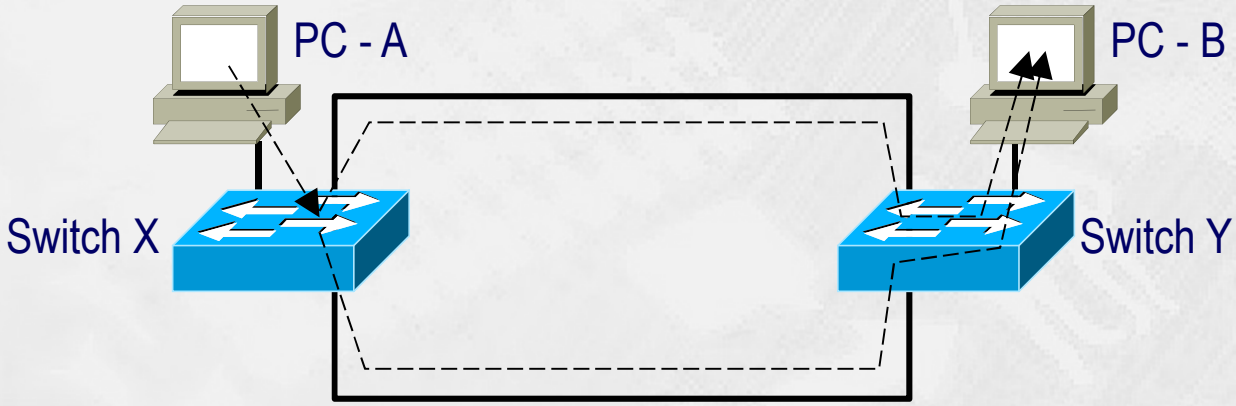


Problemet **dublerede pakker**

HOUSE OF TECHNOLOGY

– en del af mercantec

- Et andet problem med redundante forbindelser mellem switche er **dublerede pakker**. Se følgende eksempel:
  - MAC adresse på PC-B er slettet (time out) på begge switche
  - PC-A har stadig PC B's MAC adresse i sin ARP cache
  - PC-A sender en pakke til PC-B, så vil Switch X sende pakken til alle sine porte fordi den ikke kender MAC adr. på PC-B
  - På samme måde vil Switch Y sende pakken til alle sine porte, hvilket betyder at **PC-B modtager samme pakke flere gange**. Ikke godt … ;-(

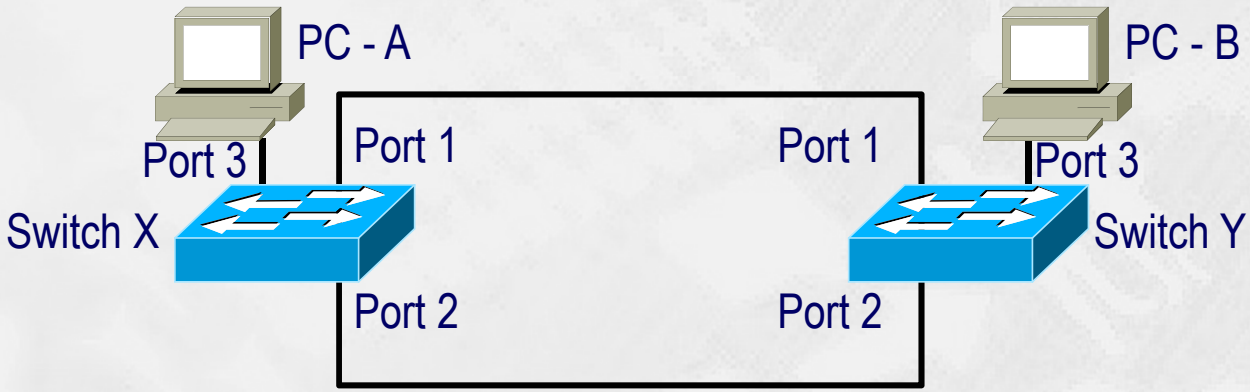


Problemet **ustabile MAC tabeller**

HOUSE OF TECHNOLOGY

– en del af mercantec

- Ved at indføre redundante forbindelser mellem switche kan de 'se' den samme MAC adresse på flere porte og bygger derfor **ustabile MAC tabeller**:
  - MAC adressen på PC-B er slettet (time out) på begge switche
  - PC-A sender en pakke til PC-B
  - Switch X lærer at MAC adressen for PC-A sidder på Port 3 og derefter sendes pakken ud på alle andre porte (flooding) for at finde PC-B
  - Switch Y lærer nu at MAC adressen for PC-A sidder på både Port 1 og 2, og derefter floodes pakken efter PC-B. Hvilket betyder at Switch X nu lærer at MAC adr. for PC-A også sidder på dens Port 1 og 2! **Der bliver uoverensstemmelse i switch MAC tabellen, og trafikken fra PC-A til PC-B kører i ring** … ikke godt! ;-(

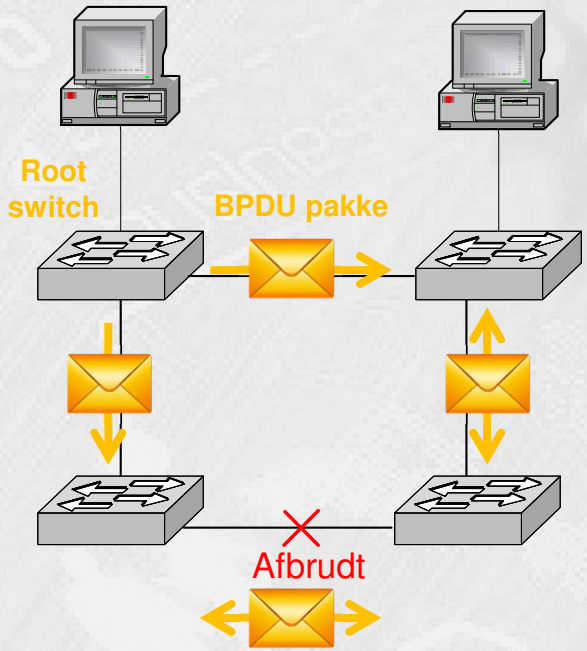


Spanning Tree begreber – **BPDU**

HOUSE OF TECHNOLOGY

– en del af mercantec

- STP protokollen tilbyder redundans, eller flere fysiske veje mellem enhederne:
  - Sender og modtager opdateringer (✉ = **BPDU**'er) for at opbygge og overvåge bedste vej gennem nettet - kontinuerligt
  - Finder den mest effektive vej mellem to enheder
  - Kun én af vejene mellem enhederne holdes åben
  - Alle alternative veje holdes administrativt lukkede …

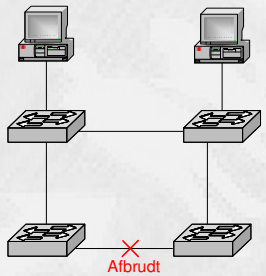


STP parametre

HOUSE OF TECHNOLOGY

– en del af mercantec

- Spanning Tree Protocol arbejder ud fra følgende **parametre**:
  - Afstanden til én udvalgt **'Root Switch'** måles konstant fra de andre switche og bedste vej gennem nettet vælges ud fra bl.a.:
    - Liniehastigheder (10Mbps, 100Mbps …)
    - Liniestatus (ændringer giver nye veje)
  - Enheder i et netværk med STP aktiveret sender hvert andet sekund små datapakker, Bridge Protocol Data Units, til hinanden for at etablere et hierarki mellem sig, med det formål at **opbygge et loop-frit netværk**.
  - Højeste prioritet er, at **loops skal altid afbrydes!**



STP – **protokol varianter**

HOUSE OF TECHNOLOGY

– en del af mercantec

- Der er mange varianter af spanning tree protokollerne:
  - **STP** (Spanning Tree Protocol)
    - Dens første standard, IEEE802.1D - senere kaldet CST, er fra år 1990
    - Understøtter slet ikke VLANs og den er lang tid om at konvergere
  - **RSTP** (Rapid Spanning Tree Protocol)
    - Dens første standard, IEEE802.1w, er fra år 2001
    - Understøtter heller ikke VLANs, men den konvergerer meget hurtigere end STP
  - **802.1D-2004**
    - Standarden, IEEE802.1D-2004, er fra år … 2004! ☺
    - Understøtter heller ikke VLANs, men det er en opdateret RSTP som helt erstatter STP

STP – **protokol varianter** (fortsat)

HOUSE OF TECHNOLOGY

– en del af mercantec

- **PVST** (Per VLAN Spanning Tree)
  - Ciscos første version af deres proprietære spanning tree variant
  - Har samme funktion som STP, men håndterer desuden VLANs via ISL trunking protokollen samt en del andre Cisco switch features
- **PVST+** (Per VLAN Spanning Tree Plus)
  - Ciscos forbedrede PVST protokol
  - Har samme funktion som PVST, men håndterer desuden VLANs via 802.1Q trunking samt en del nyere Cisco switch features, bl.a. load balancing
  - Ulempen er klart en større CPU belastning på maskinen ved mange VLANs
- **Rapid PVST+** (Rapid Per VLAN Spanning Tree Plus)
  - Ciscos forbedrede PVST+ protokol
  - Har samme funktion som PVST+, men tilføjer bl.a. hurtigere konvergens

STP – **protokol varianter** (fortsat)

HOUSE OF TECHNOLOGY

– en del af mercantec

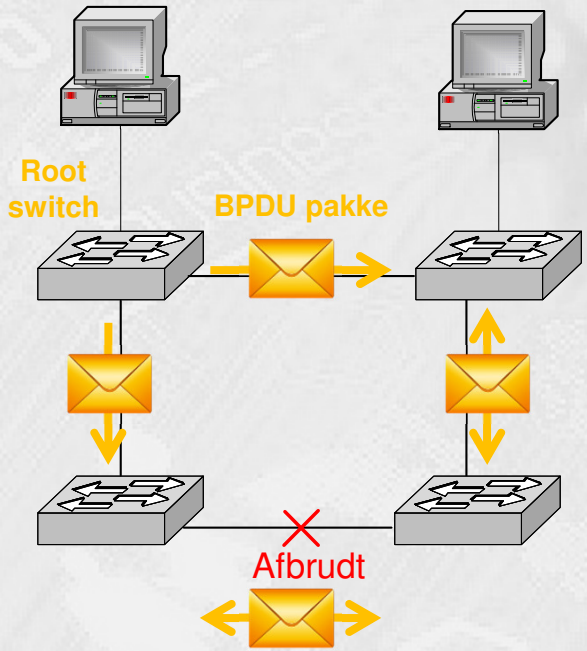
- **MSTP** (Multiple Spanning Tree Protocol)
  - Dens IEEE navn er IEEE802.1s
  - Er oprindelig introduceret som en tilføjelse til IEEE802.1Q i 1998
  - Er baseret på RSTP og konvergerer dermed hurtigt
  - Er opbygget efter Ciscos MISTP og giver mulighed for at anvende VLANs
  - Dens første reelle IEEE standard er IEEE802.1s-2002
  - MSTP er i dag direkte implementeret i standarden IEEE802.1Q-2005
  - Systemadministratoren der konfigurer netværket skal opdele MSTP området i 'regions' og 'instances', som hver især kan indeholde et antal VLANs
  - MSTP benytter kun ét BPDU format og dette gør protokollen kompatibel med bl.a. RSTP, som blot opfatter en MSTP region som én enkelt RSTP enhed
- **MST** (Multiple Spanning Tree)
  - Ciscos egen version af multiple spanning tree protokollen
  - Har stort set samme funktioner som MSTP, men mange flere Cisco features

Sådan virker Spanning Tree

HOUSE OF TECHNOLOGY

– en del af mercantec

- Som vi har set tidligere sender switche i et redundant LAN netværk hvert 2. sekund BPDU pakker fra root switchen og ned gennem nettet. Formålet er at bygge et loop-frit net samt at overvåge linje status.
- BPDU pakker sendes til andre switche med IEEE802.1D multicast adresserne:
  - 01-80-C2-00-00-00
  - 01-80-C2-00-00-10
- Switche som ikke deltager i STP sender blot BPDU'er videre …

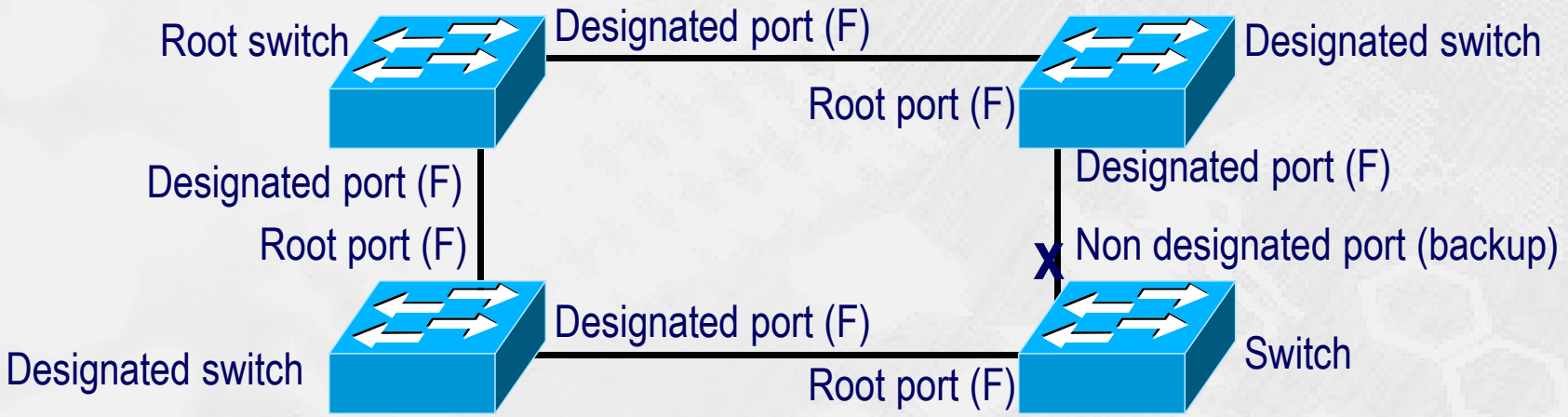


STP – og switchens opgaver

HOUSE OF TECHNOLOGY

– en del af mercantec

- BPDU'erne giver switchene en del information og de udfører så følgende opgaver:
  - Vælger en Root switch for netværket (kun én pr. netværk).
  - Beregner den korteste / bedste vej fra den selv til Root switchen.
  - Vælger hvilken switch der er tættest på Root, på hvert LAN segment. En switch som håndtere al kommunikation mellem LAN segmentet og Root, den kaldes "designated switch".
- Switche der ikke er Root vælger den port der er nærmest Root, den defineres som Root-port og overfører data (F).
- Porte der er med i Spanning-Tree kaldes designated ports og de overfører data (F), ikke designated porte blokeres.



STP – **valg af root switch**

HOUSE OF TECHNOLOGY

– en del af mercantec

- I ethvert Spanning Tree vælges der kun én Root Switch
- Switchene udsender BPDU'er indtil Root Switchen er valgt
- Root Switchen er den Switch med den laveste Priority værdi i Bridge Identifier feltet
  - Er alle Priority-værdier lige vælges switch med laveste MAC-adresse værdi automatisk til Root.
- Kampen for at blive Root Switch er kontinuerlig og BPDU'er udsendes kontinuerligt
- Ændringer i topologien genererer straks en ny STP udregning
- Hvilken switch bliver Root Switch på tegningen?

STP – **protokollens felter**

HOUSE OF TECHNOLOGY

– en del af mercantec

- **Bridge Identifier, BID**, er et 8 Byte felt, som er delt i to og som bruges til at identificere afsenderen samt Root switchen med:

Prioritet 2 Byte	MAC Adresse 6 Byte

- Prioritets feltet indeholder:
  - En administrativ værdi mellem 0 og 65535 (Default = 32768).
  - Den mindste værdi i forhold til andre switche = højeste prioritet
- MAC Adresse feltet indeholder:
  - Switchens / VLAN'ets MAC adresse.

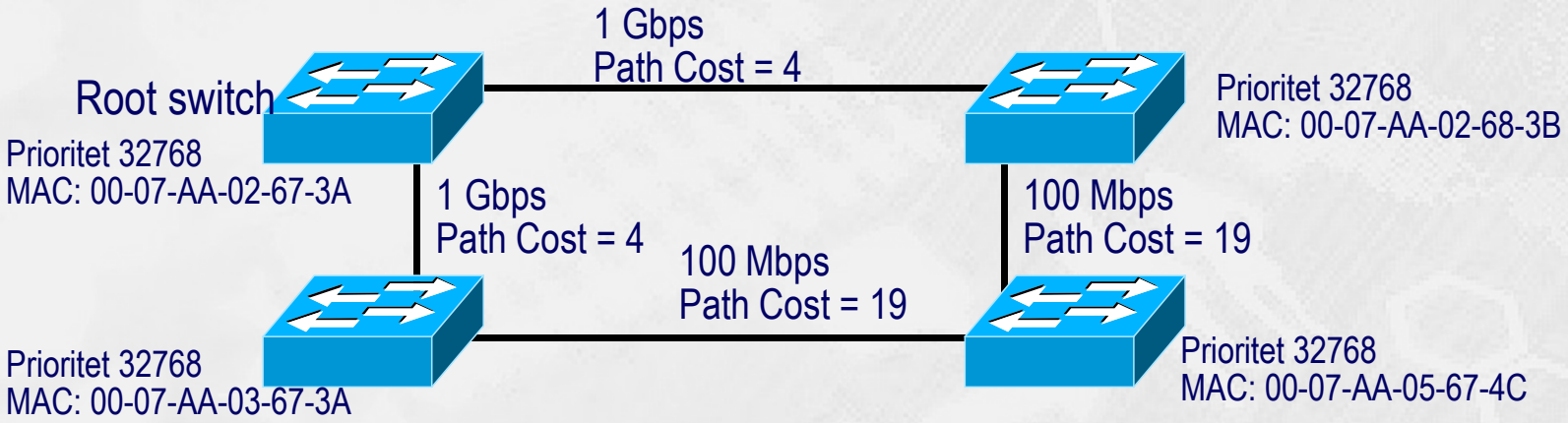
STP – **path cost**

HOUSE OF TECHNOLOGY

– en del af mercantec

- En transmissionslinje mellem to switche har en '**Path Cost**' baseret på linjens hastighed og værdierne kan ændres rent administrativt

HASTIGHED	PATH COST IEEE spec.
10 Mbps	100
100 Mbps	19
1 Gbps	4
10 Gbps	2

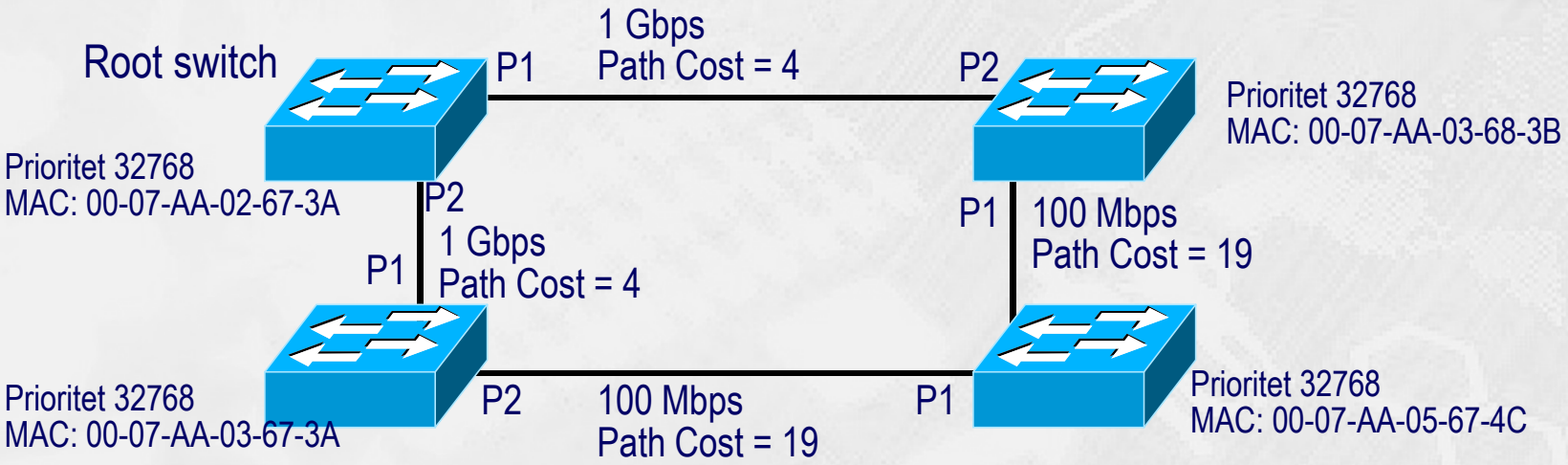


STP – **path cost** (fortsat)

HOUSE OF TECHNOLOGY

– en del af mercantec

- Hvis der findes flere alternative veje frem til Root Switchen, så vælges den korteste - baseret på Path Cost værdier
- Hvis der er flere veje med samme Path Cost, så vælges den vej hvor afsender har lavest BID
- Hvis flere veje har samme afsender BID, så vælges det laveste portnummer



STP – **protokollens felter**

HOUSE OF TECHNOLOGY

– en del af mercantec

Switch
MAC adr. 00-C0-7F-A4-86-7A

Eksempel på BPDU felter:

BPDU felter

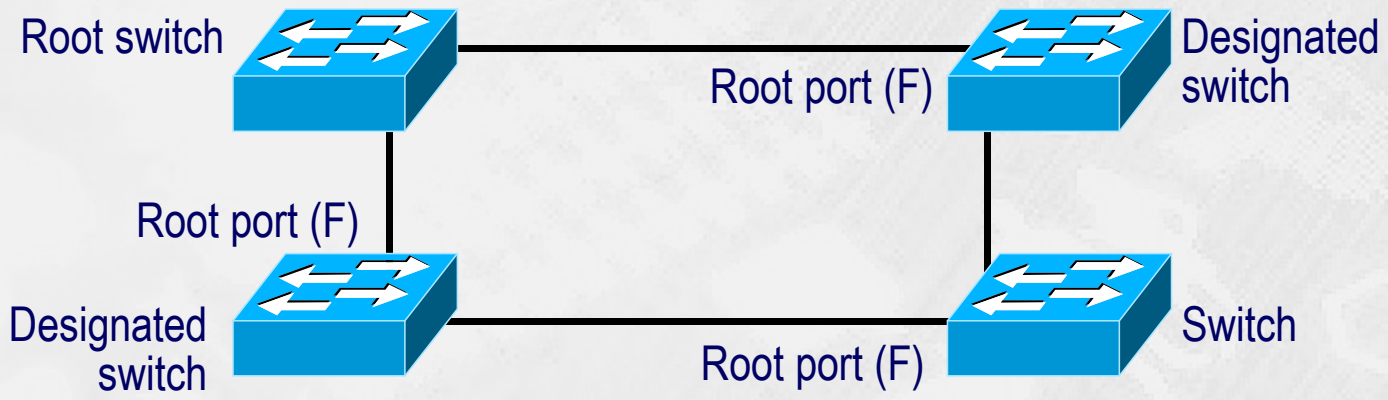
Root BID	32768 00-AB-CD-45-B5-D7	Root bridge identifikations nr.
Root Path Cost	4	Path cost til root bridge
Afsender BID	32768 00-C0-7F-A4-86-7A	Afsenderens identifikations nr.
Port ID	Afsender port nr. 13	Hvilken port fra afsender switchen kom denne BPDU fra

STP – **root port** på en switch

HOUSE OF TECHNOLOGY

– en del af mercantec

- Den port på en Designated switch som har den korteste vej set op mod root-switchen det er en **Root-port**, og den kan overføre data (F = forward)
  - Hvis der er flere porte mod root-switchen vil porten med den **laveste STP-port priority** blive valgt som Root-port
  - Hvis portene har samme STP-port priority vil porten med **det laveste portnummer** vælges

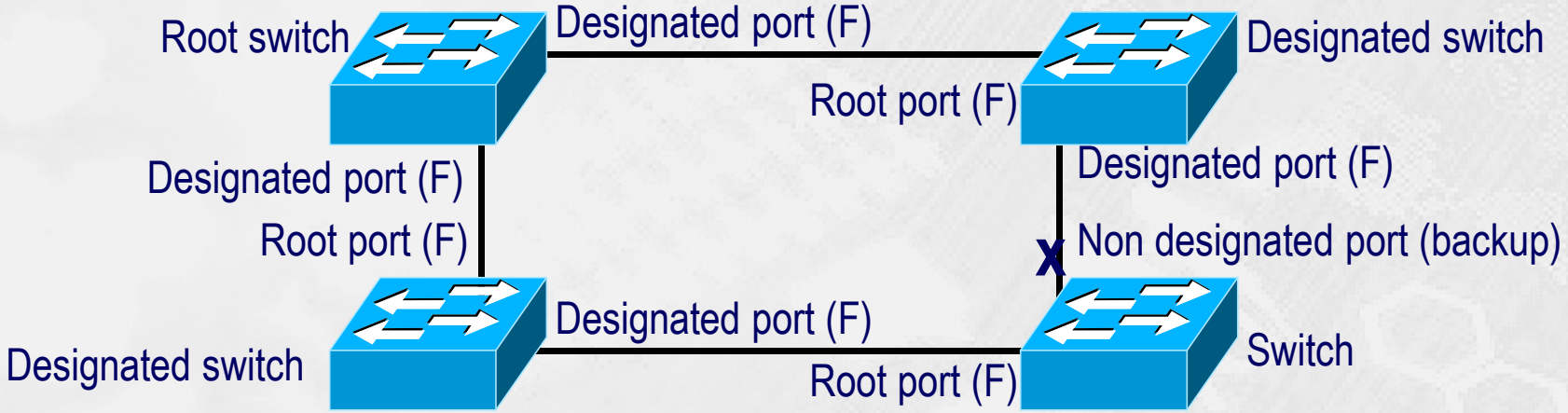


STP – **designated switchporte**

HOUSE OF TECHNOLOGY

– en del af mercantec

- Porte der er med i Spanning-Tree kaldes **Designated ports** og de kan overføre data (F = forward).
- Porte der ikke er med i Spanning Tree kaldes **Non-designated** porte.
  - De blokeres så data pakker ikke kan modtages eller sendes og de lærer heller ikke MAC adresser.
  - De bruges som backup og aktiveres kun hvis den aktive forbindelse svigter
- Når en port er blokeret fortsætter den med at modtage BPDU'er på porten (for at kunne se hvis nettet skal rekonfigureres), men den sender dem ikke videre.

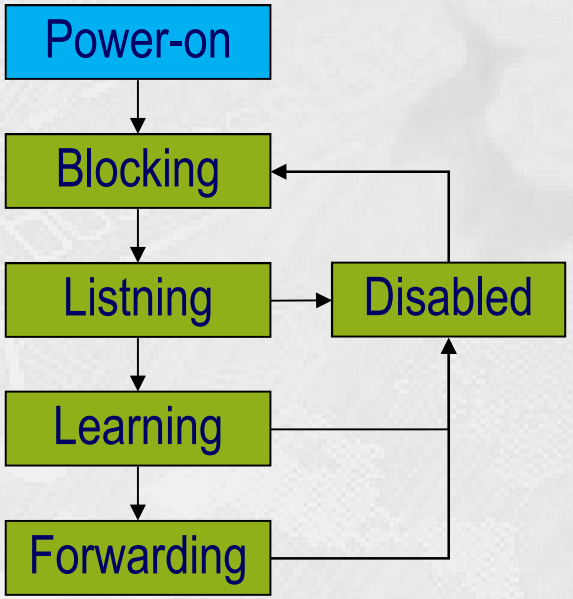


STP – **interface states**

HOUSE OF TECHNOLOGY

– en del af mercantec

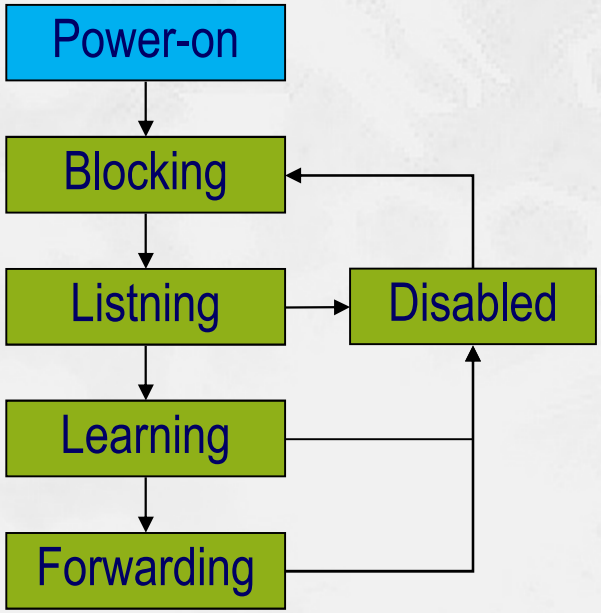
- STP leverer en loop-fri vej gennem hele lag 2 broadcast domænet og denne vej læres bl.a. via de informationer der er i BPDU'erne som udveksles mellem switchene
- Hver eneste switchport gennemløber ved power on fem mulige 'port states' samt tre BPDU tidsforløb
- Forwarding state er det eneste tidspunkt en port kan flytte data, men hvis porten gik direkte fra slukket og til forwarding kunne der opstå loops i nettet – ikke godt
- Spanning Tree udfører ved power on opstart altid følgende per VLAN:
  - Vælger root bridge
  - Vælger root port på non-root bridge
  - Vælger Designated port på hvert segment
  - Resterende porte gives status Alternate



STP – **interface states** (fortsat)

HOUSE OF TECHNOLOGY

– en del af mercantec



- **Blocking state** (max-age 20 sek.)
  - Lytter til BPDU'er
  - Data pakker sendes og modtages ikke
  - Porten lærer ikke MAC adr.
- **Listening state** (forward delay = 15 sek.)
  - Lytter til BPDU'er
  - Alle porte bliver i denne state indtil Root Swich er valgt.
  - Data pakker afvises, lærer ikke MAC adresser.
  - Non-designated porte blokeres
- **Learning state** (forward delay = 15 sek.)
  - Lytter til BPDU'er
  - Data pakker modtages for at lære MAC-adresser
  - Data pakker sendes ikke
- **Forwarding state**
  - Lytter til BPDU'er
  - Lærer MAC-adresser
  - Sender og modtager datapakker
- **Disabled state**
  - porte der er administrativt lukkede (Ikke STP)

STP – **konvergens i nettet**

HOUSE OF TECHNOLOGY

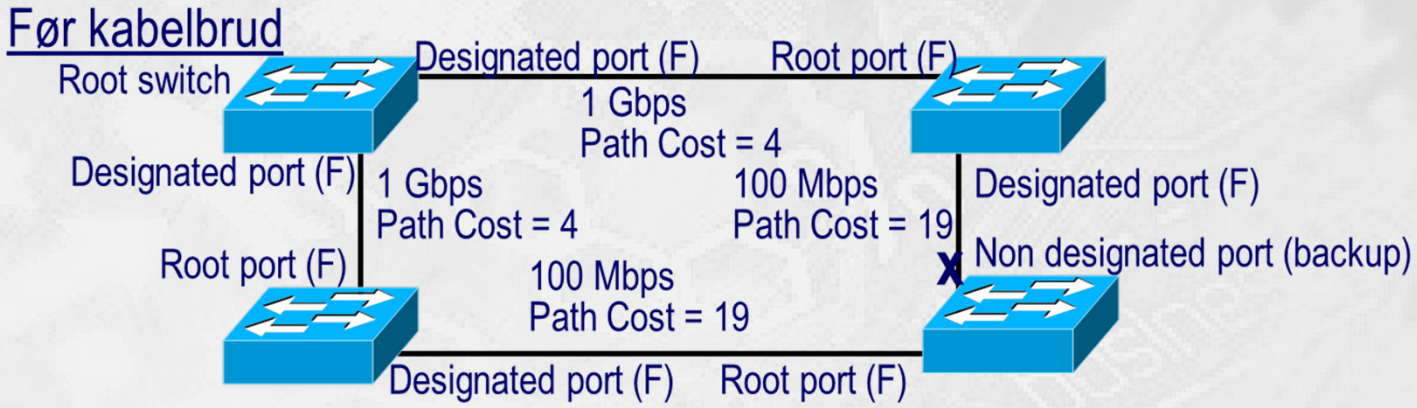
– en del af mercantec

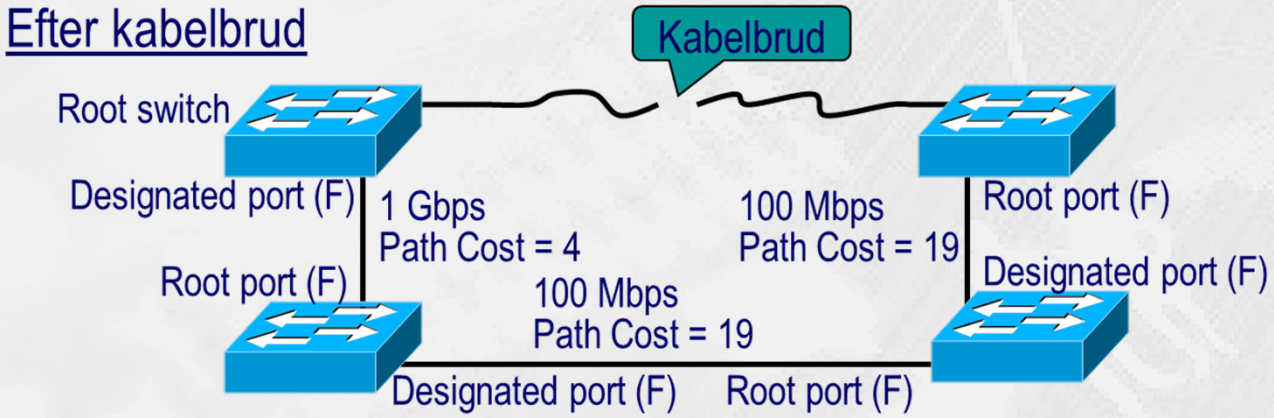
- Et switched netværk er konvergeret når alle portene enten er i forward eller blokeret state
- Hvis der sker en topologi ændring (fx aktiv forbindelse brydes) skal switchene igen beregne hvordan det nye Spanning Tree skal se ud
- Konvergens tiden til forward state kan blive op til 50 sekunder, hvor der ikke sendes data pakker
- Konvergens tiden består af:
  - Max-age tid på 20 sek.
  - Listen forward delay på 15 sek.
  - Learning forward delay på 15 sek.

STP – eksempel på **fejl i nettet**

HOUSE OF TECHNOLOGY

– en del af mercantec



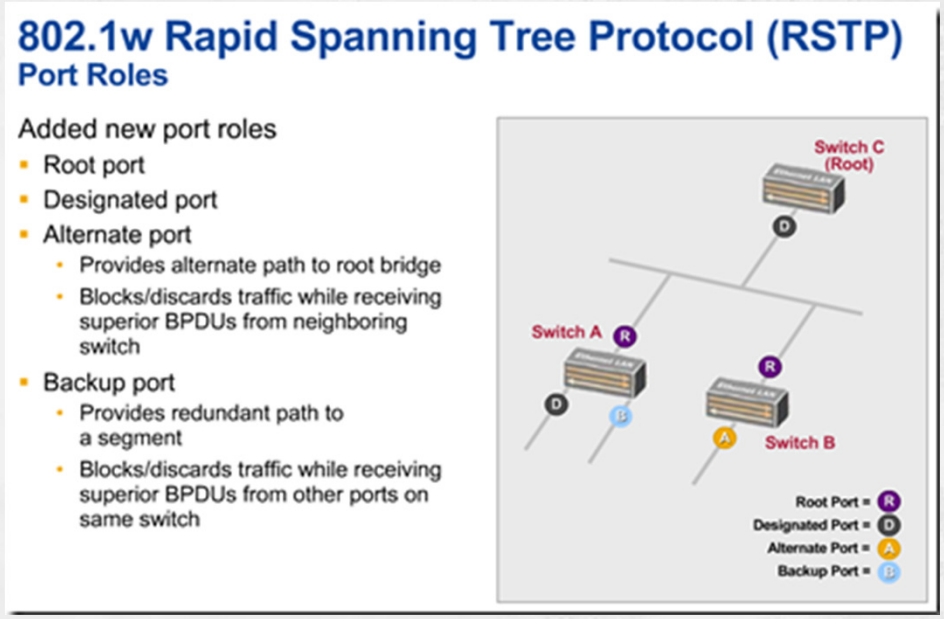


STP - **Rapid Spanning Tree**

HOUSE OF TECHNOLOGY

– en del af mercantec

- Rapid Spanning Tree Protocol er udviklet for at få **hurtigere konvergens**:
  - Netværkets konvergenstid må ikke vare længere en 15 sek. ifølge standarden IEEE 802.1w

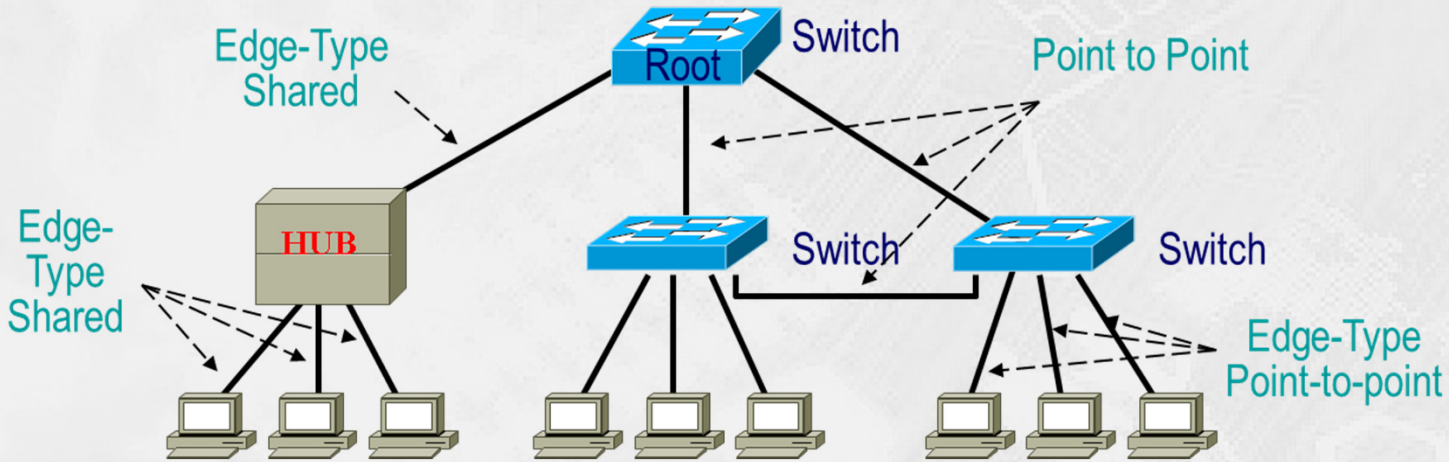


Rapid STP - **ændringer fra STP**

HOUSE OF TECHNOLOGY

– en del af mercantec

- Definition af link type som kan starte forward af data hurtigt.
  Link typerne er:
  - Point to Point, Edge-Type og Shared



Rapid STP - ændringer (fortsat)

HOUSE OF TECHNOLOGY

– en del af mercantec

- Switche i et RSTP netværk kan nu generere BPDU'er i stedet for at videresende Root switch BPDU'er
- Bloking state er omdøbt til Discarding state. Portens rolle er nu en Alternate (skiftende) port, Discarding porten kan blive den designated port hvis den designated port på LAN segmentet fejler.
- Mistes der 3 BPDU betragtes linien som fejlende og her ved man hvilken linie det er, nemlig den mellem de to switche. Før var det sådan at det kunne være en af linierne op til Root switchen der var afbrudt.

STP – Ciscos **PVST+**

HOUSE OF TECHNOLOGY

– en del af mercantec+

- Cisco udviklede deres egen version af RSTP protokollen, fordi de havde behov for en Spanning Tree protokol **som kunne håndtere VLANs**:
  - Den nye protokol kom til at hedde **Per VLAN Spanning Tree Plus, PVST+**
  - Den er baseret på Ciscos egen PVST protokol, som er udviklet ud fra standarden STP
  - Den er hurtig til at konvergere – som RSTP
  - Den kan implementere lag 2 load balancing
    - Dette er muligt fordi der kører en separat, uafhængig instans af STP per VLAN, som kan blokeres eller tillades individuelt på hver port
    - F.eks. kan halvdelen af et firmas VLANs køre på den ene fysiske linje, mens den anden halvdel flyttes over på den redundante.
  - Omkostningerne ved PVST+ er bl.a. en større CPU load

STP – Ciscos PVST+ (fortsat)

HOUSE OF TECHNOLOGY

– en del af mercantec

- Ofte har en administrator brug for at udvælge en bestemt switch til root, måske fordi den er bedst placeret eller fordi der skal konfigureres load balancing. Se Ciscos NetAcad – fine eksempler!

- Cisco CLI kommandoer:

  - En bestemt switch og VLAN ønskes sat som root:
    - **spanning-tree vlan** *vlan-id* **root primary**
  - En bestemt switch og VLAN ønskes sat som secondary:
    - **spanning-tree vlan** *vlan-id* **root secondary**
  - En bestemt switch og VLAN ønskes givet en priority værdi manuelt:
    - **spanning-tree vlan** *vlan-id* **priority** *value*

  - Kontrollér indstillingerne bagefter med kommandoen:
    - **show spanning-tree**

STP – Ciscos PVST+ (fortsat)

HOUSE OF TECHNOLOGY

– en del af mercantec

- Ciscos switch features **Portfast** samt **BPDU Guard** er udviklet specielt med henblik på et problemfrit redundant lag 2 miljø med spanning tree aktiveret
- En pc workstation eller en server på et fast kablet firmanet er normalt koblet direkte på én bestemt switchport, og de flyttes sjældent til en ny port
  - Ciscos PortFast konfigureres på alle disse workstation- og serverporte, hvilket betyder at porten altid springer direkte fra blocking state til forwarding state
  - Der kobles normalt aldrig en anden switch på disse porte, men skulle det ske ville det være katastrofalt, så derfor aktiveres Ciscos BPDU Guard på portene. Dette sikrer at porten øjeblikkeligt lukkes ned hvis der modtages blot en enkelt BPDU, og porten skal manuelt lukkes op igen af en administrator