

MPLS

Grunnlag for fremtidig IP-nett

Ingvild Sorteberg
Baseline Communications as

ingvild.sorteberg@baseline.no

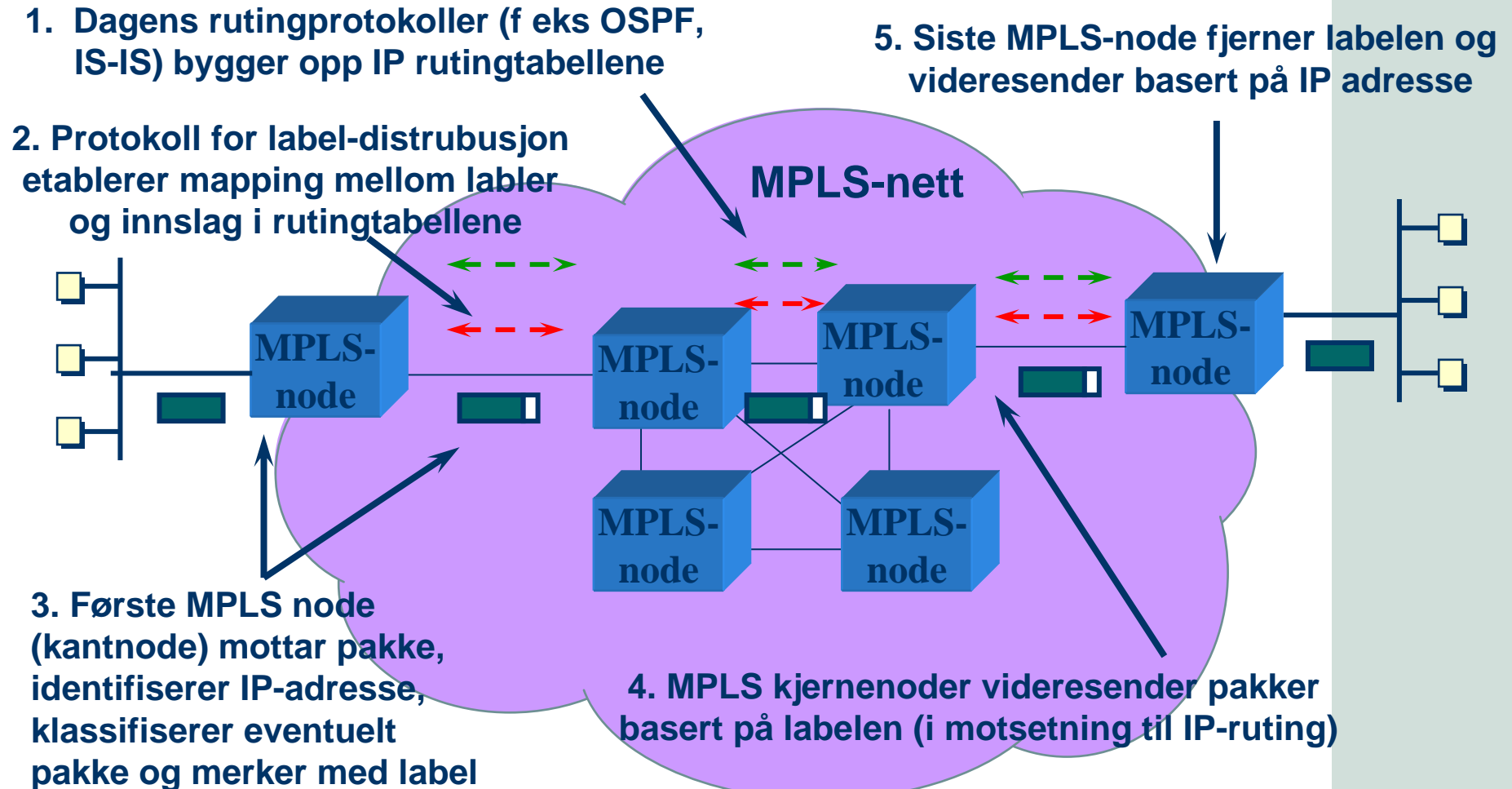
Innhold

- MPLS – en kort innføring
- Hva kan MPLS brukes til?
 - ◆ VPN
 - ◆ IPv4 – IPv6 overgang
 - ◆ Traffic engineering
 - ◆ Beskyttelse og rask feilhåndtering
- GMPLS

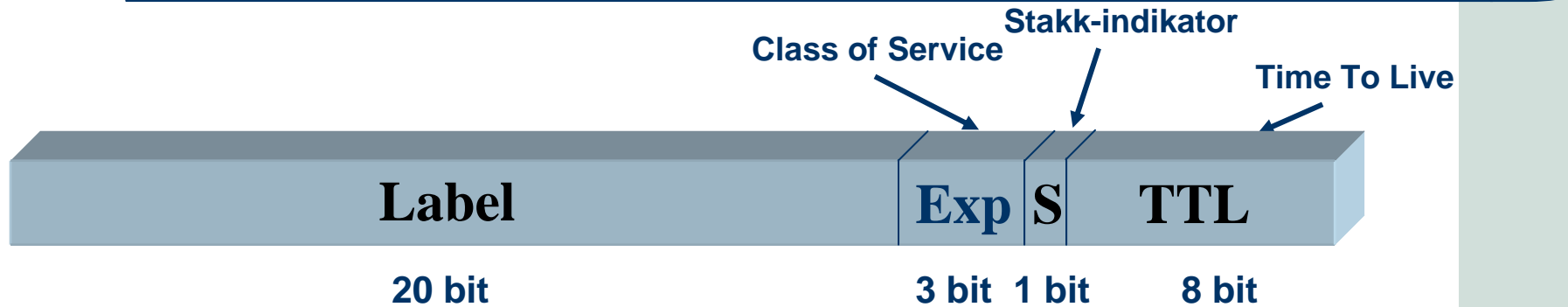
Hvorfor trenger vi MPLS?

- Kombinerer IP-ruting og adressering med linklags-svitsjing ("Best of both worlds")
- Mindre overhead
 - ◆ Ikke behov for IP-adresse oppslag i transittnodene
 - ◆ Bedre ytelse (?)
- Bedre kontroll
 - ◆ Tillater eksplisitt engineering av stier/kanaler slik at trafikken kan balanseres bedre i nettet
 - ◆ Raskere reruting
- Nye tjenester
 - ◆ VPN med tjenestekvalitet
- Enklere integrasjon mellom IP og ATM / levere lags teknologier
 - ◆ Slipper VCeer alle-til-alle
 - ◆ Slipper to nivåer av ruting og adressering

En liten forsmak...



MPLS label



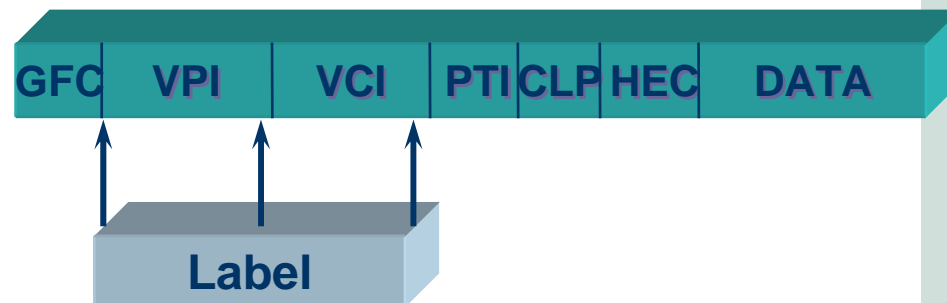
Mellom PPP og IP pakkehodet
f eks POS (Packet over SONET/SDH)



Label mellom lag 2 og 3
f eks mellom LAN MAC og IP pakkehodet

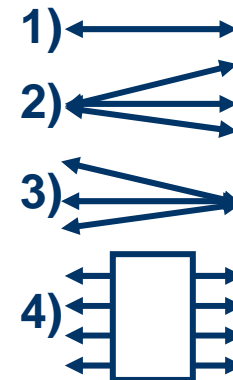
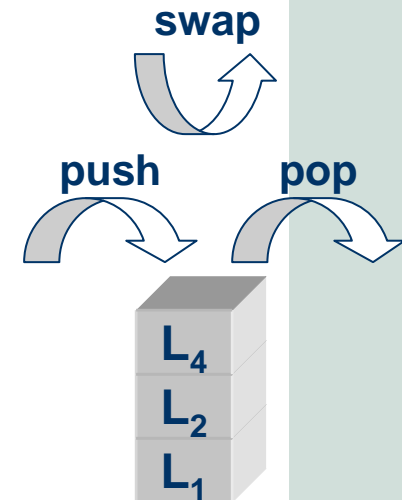


Label i ATM cellehodet

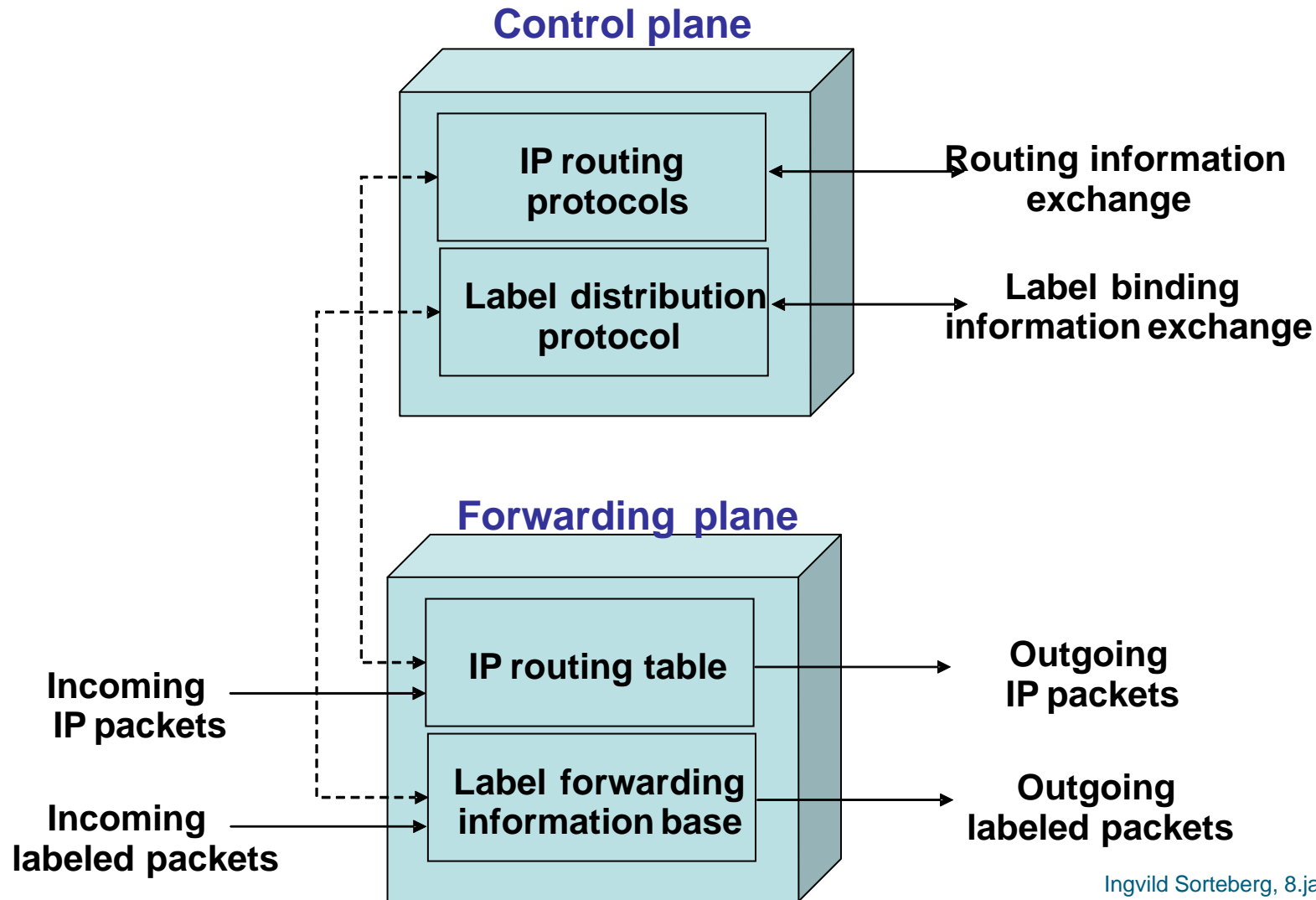


Label-stakk

- Hver pakke kan inneholde en stakk (LIFO) av labler
- Operasjoner: push, pop, swap, noop
- Videresending basert på øverste label i stakken
- Tillater mange trafikk mønstre
 - ◆ Kan brukes til sammenslåing og splitting av strømmer
- Gir skalerbarhet og effektivitet, men kompleksiteten øker



MPLS kontroll og videresendingsplan



Tilordning av labler

- Mapping mellom label og trafikkstrøm
- Datatrafikk-drevet
- Kontroll-drevet
 - ◆ topologi-drevet
 - trigges av innslag i rutetabellen
 - styres derved av standard ruteprotokoller
 - ◆ reservasjoner
 - trigges f eks av mottak av RSVP RESV melding
 - ◆ statisk konfigurering
 - for trafikkstyring

Label distribusjon

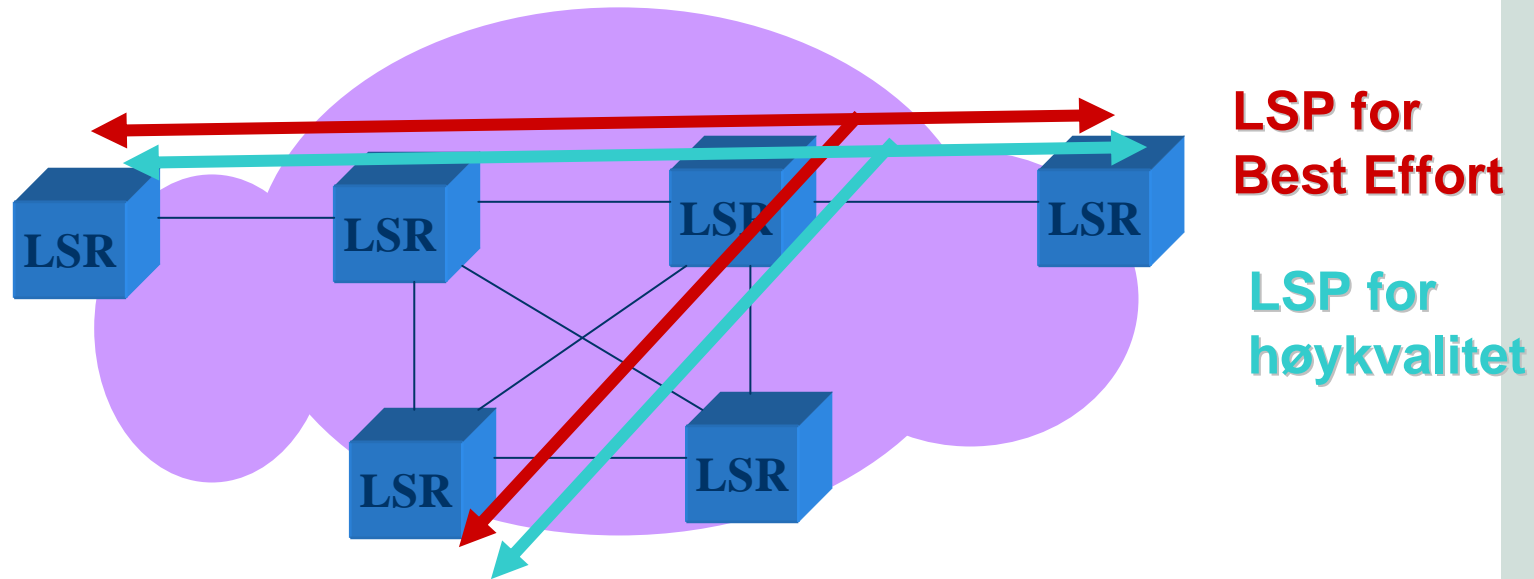
- LSR (Label Switched Routers) forhandler og informerer hverandre om hvilke labler som skal benyttes
- Flere ulike protokoller blir standardisert for label distribusjon
 - ♦ LDP (Label Distribution Protocol) benyttes for implisitt label distribusjon
 - ♦ MP-BGP (MultiProtocol BGP) benyttes for label distribusjon og VPN etablering
 - ♦ RSVP-TE og CR-LDP benyttes for eksplisitt label distribusjon i forbindelse med traffic engineering
- Bakgrunnen for de mange protokollene er for å dekke ulike behov (+ ulike leverandørers favoritter)

MPLS og DiffServ

Følgene muligheter kan støttes

- Alle DiffServ klassene rutes på ulike LSPer
 - ◆ Exp feltet i MPLS hodet benyttes for å angi drop precedence
- Alle DiffServ klasser rutes på samme LSP
 - ◆ Exp feltet i MPLS hodet benyttes for å angi hvilken CoS pakken tilhører
- Kombinasjoner av de to metodene
 - ◆ Noen DiffServ klasser rutes separat og andre kombineres, f eks tale trafikk rutes på en egen LSP og all annen trafikk deler en LSP
 - ◆ Ulike deler av nettet benytter ulike metoder

Tjenestekvalitet - eksempel



- En LSP pr IP tjenesteklasse (CoS)
- Båndbreddeallokering pr. klasse
- Aggregering pr klasse
- IP CoS-basert køing/betjening og intelligent kasting

MPLS versus DiffServ

- DiffServ er en mekanisme på lag 3 (IP-laget)
- MPLS er en protokoll på lavere lag, mellom lag 2 og 3
- DiffServ tilbyr en mulighet til å kontrollere videresendingen av pakker, ut over "best-effort"
- MPLS tilbyr bedre kontroll over topologien og hvordan pakkene videresendes enn ved bruk av kun en "shortest path" videresending.
- De har fordeler hver for seg eller sammen

MPLS and multicast

- Påmelding til multicast gruppe, etablering av multicast tre
 - ◆ Ingen kopling mellom multicast og MPLS signaleringen
 - ◆ Multicast overføringen initierer ikke etablering av MPLS stier, men benytter eventuelt eksisterende MPLS stiene.
 - ◆ Hvis MPLS stier ikke eksisterer kan disse etableres implisitt dersom MPLS implementasjonen støtter trafikk drevet LSP etablering
- Skiller seg fra vanlig bruk av MPLS der LSPer er forhåndsetablert
- Krever svært rask LSP etablering for å fungere tilfredsstillende
- Veldig dynamiske multicast trær vanskeliggjør ressursreservering siden det er vanskelig å forutse behovet for ressurser
- Utmelding av multicast grupper eller nedkopling av multicast trær signaleres ved bruk av IGMP til den nærmeste multicast ruterer
 - ◆ Siden det ikke er noen kopling mellom multicast og MPLS signalering vil dette ikke føre til en nedkopling av LSPen
 - ◆ Implisitte LSPer koples ikke ned, dette kan føre til at label tabellen går full
 - ◆ En timer kan assosieres med implisitte LSPer, men det er vanskelig å sette verdien på en slik timer

MPLS anvendelsesområder

- Hva kan MPLS brukes til?
 - ◆ VPN
 - ◆ IPv4 – IPv6 overgang
 - ◆ Traffic engineering
 - ◆ Beskyttelse og rask feilhåndtering
 - ◆ GMPLS – sammenkoping mellom IP og optikk

MPLS VPN-løsninger

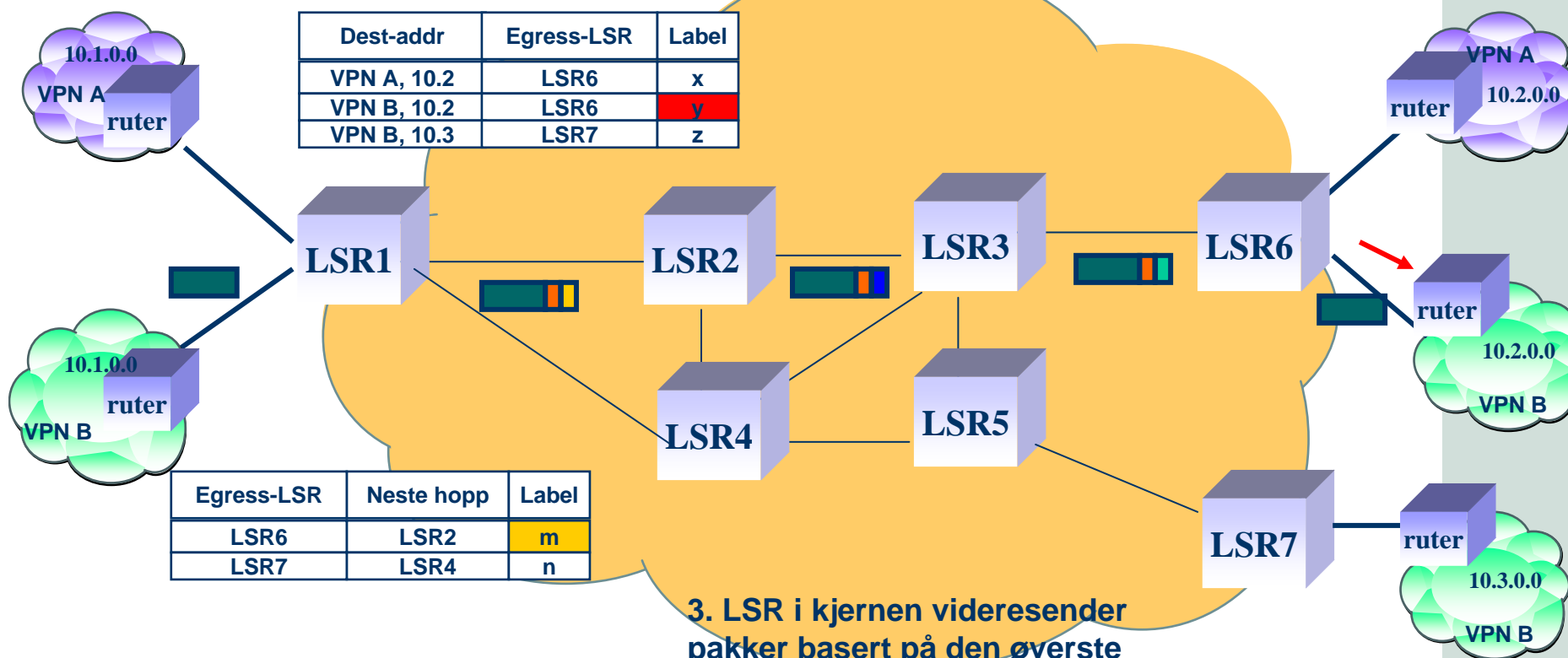
- Forhåndsoppsatte stier for hvert VPN
- MP-BGP4 (MultiProtocol BGP) benyttes for å distribuere ruteinformasjon innen et VPN
- Hvert VPN har separate rutetabeller (kjenner kun til ruter innenfor eget VPN)
- Pakkene rutes ikke på IP underveis, men på MPLS-labeler
- Kan benytte private IP-adresser internt i et VPN
- Skaleringsproblemer i transportnettet
 - ◆ Stort antall MPLS-stier
 - ◆ Rutetabellene blir veldig store (en rutetabell per VPN)
- Standardiseres av IETF Provider Provisioned VPN (PPVPN) WG

MPLS VPN-løsning

1. Ved mottak av pakke identifiserer ingress (kantnoden) LSR1 avsender og mottakeren, og finner egress (kantruter) LSR6 samt label for det aktuelle VPN-et

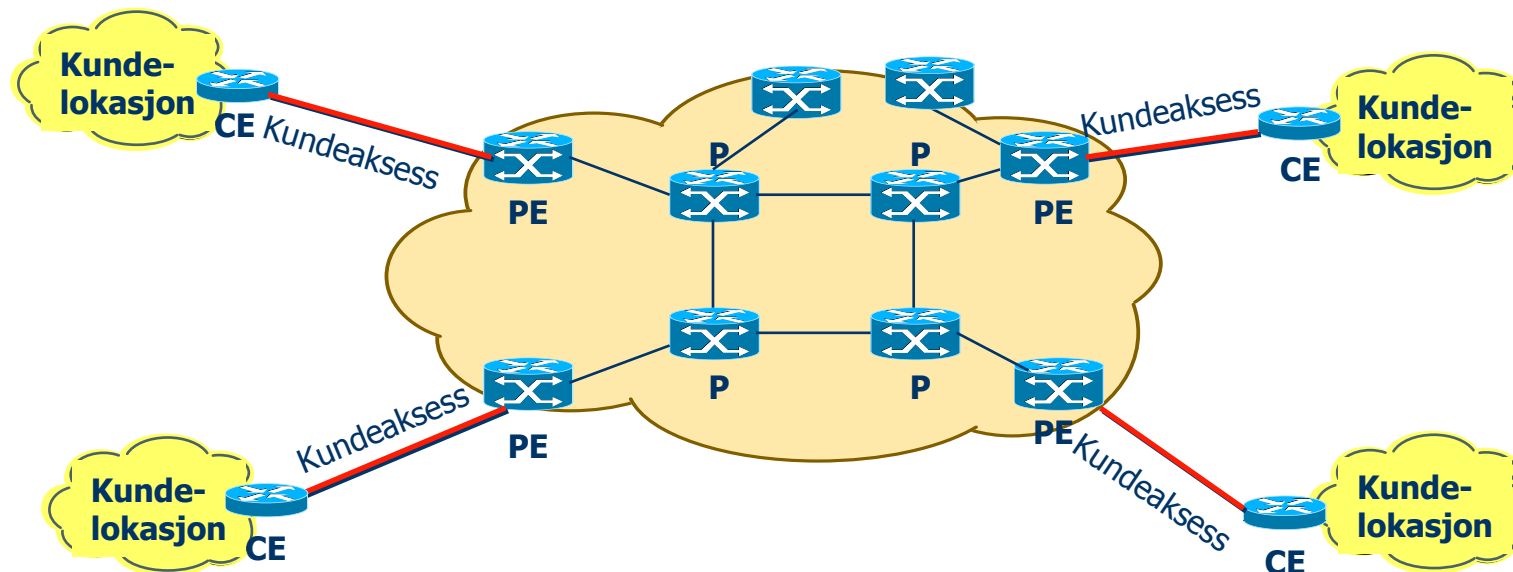
2. Så finnes neste hopp (LSR2) på veien mot kantruter (LSR6), samt tilhørende label; push label

4. Egress (kantruter) LSR6 fjerner den interne labelen; pop label; og videresender basert på neste label i stakken; pop label



3. LSR i kjernen videresender pakker basert på den øverste labelen i stakken

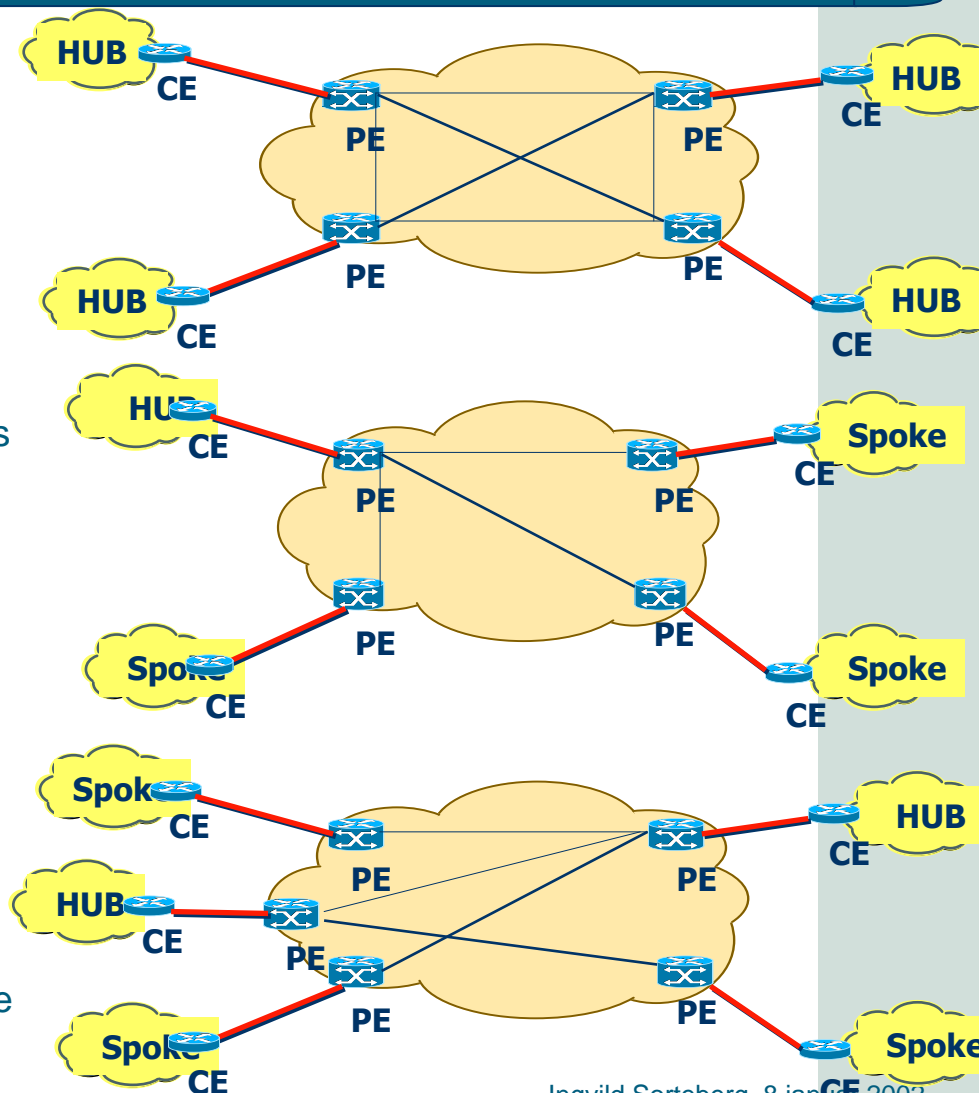
Arkitektur for et MPLS VPN



- PE-ruteren mottar ruteinformasjon fra CE og legger det i den korrekte VRF-tabellen (VPN Routing and Forwarding)
- Ruteinformasjonen PE-ruteren får fra kjernenettet legges i den globale rutetabellen
- MP-BGP benyttes for å distribuere ruteinformasjon om kundenettene mellom VPN-lokasjoner (PE rutere)
- Ved å benytte ulike rutetabeller for hvert VPN slipper VPN-kundene å benytte offisielle IPv4 adresser
- En VPN-lokasjon kan være medlem i flere VPNe

MPLS VPN topologier

- Alle-til-alle kommunikasjon (fully meshed)
 - ♦ Alle lokasjonene kan kommunisere direkte med hverandre
- Alle-til-en kommunikasjon (hub and spoke)
 - ♦ En lokasjon defineres som hub og alle de andre (en eller flere) defineres som spoke-lokasjoner
 - ♦ Spoke-lokasjoner kan ikke kommunisere direkte mellom seg, men må gå via hub-lokasjonen
- Kombinasjon av alle-til-alle og alle-til-en
 - ♦ To eller flere lokasjoner defineres som hub - disse vil da ha alle-til-alle kommunikasjon
 - ♦ en eller flere lokasjoner defineres som spoke og kan kun kommunisere med en eller flere av hub-lokasjonene (ingen andre spoke-lokasjoner)



IPv4 – IPv6 overgang

- MP-BGP støtter også tunnelling av IPv6 over IPv4 basert MPLS nett
- MP-BGP benyttes mellom IPv4-IPv6 kanrutere (PE rutere) og utveksler IPv6 ruteinformasjon
- Kan sammenlignes med det å etablere et IPv6 VPN, men det er ikke behov for separate VPN forwarding og rutetabeller siden ruteinformasjonen lagres i den globale rutetabellen

Utfordringer for IP nettet

1. Trafikkvekst
2. Effektiv utnyttelse av ressurser
3. Differensiert tjenestekvalitet

To religioner!!

Det vil være nok båndbredde (IP)

“Best Effort” er hurtigst!

“TE-tools may have a negative impact on performance. Hence sometime it is preferable to throw bandwidth at congestion problems”

Ferguson & Huston, “Quality of Service”.

Ressursknapphet medfører krav om avanserte QoS-mekanismer

Applikasjoner stiller krav til

forsinkelse

bandbredde

jitter (variasjon i forsinkelse)

som må garanteres.

Behov for differensiering av kunder

massemarked versus

bedriftsmarked

MPLS trafikkstyring (Traffic Engineering (TE))

Målsetning

- Redusere kostnader gjennom
 - ◆ Mer effektiv utnyttelse av båndbredde ressursene
 - ◆ Unngå situasjoner der deler av nettet er overbelastet mens andre deler er underbelastet
- Definisjon av Traffic Engineering (ref. TE framework, IETF draft standard)
 - ◆ *“that aspect of Internet network engineering dealing with the issue of performance evaluation and performance optimization of operational IP networks”.*
- For å oppnå dette er det behov for å måle, karakterisere, modellere og kontrollere trafikken nettet

Hovedmålsetningen er kostnadsbesparelser!!

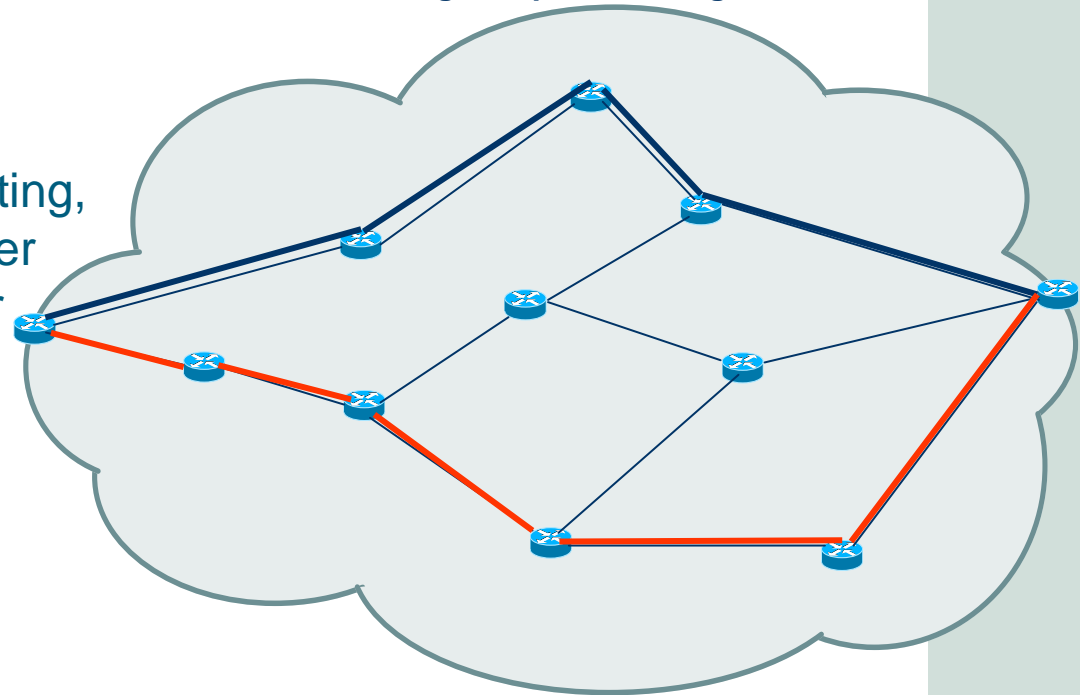
MPLS TE

- IP ruting åpner kun for bruk av “korteste vei” og lastdeling mellom veier med samme kost
 - ◆ Dvs all trafikk rutes den samme veien selv om denne er overbelastet og andre veier gjennom nettet er underbelastet.
- MPLS TE tilbyr eksplisitt ruting
 - Bedre utnyttelse av nettressursene
 - Bidrar til å sikre tjenestekvalitet, kan rute ulike tjenesteklasser ulikt
 - Kan differensiere mellom kunder

MPLS TE

- Dagens IP-nett bruker ATM for TE.
- Eksplisitt ruting kan benyttes til trafikkstyring, f eks QoS/policy og lastdeling
- IP source routing kan benyttes for eksplisitt ruting, men lite utbredt og krever tung prosessering i hver ruter
- MPLS er mer effektiv

- Rute valgt av IP rutingprotokollen
- Alternativ vei for best effort trafikken dersom belastningen på hovedlinken overstiger et predefinert grense

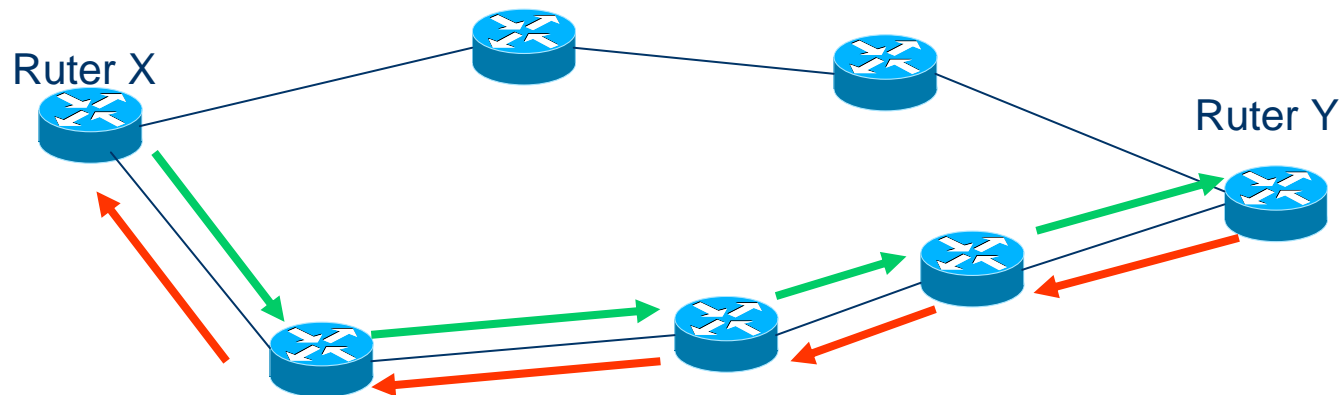


MPLS TE byggeklosser

- Utvidelser av IS-IS og OSPF ruteprotokollene
 - ◆ Distribusjon av informasjon om MPLS stier, ressurser, policier
- Signaleringsprotokoller
 - ◆ RSVP-TE
 - basert på RSVP
 - utvidelser for etablering og nedkopling av MPLS stier
 - ◆ CR-LDP
 - basert på LDP
 - utvidelser for eksplisitt ruting og ressurservering
- Ressursreservasjon og tilgangskontroll
 - ◆ Ressurser reserveres for de enkelte stiene (ikke per sesjon, men per klasse eller per samling av klasser)
 - ◆ Nye MPLS stier etableres kun dersom det er tilstrekkelig ressurser tilgjengelig

MPLS TE eksempel ved bruk av RSVP-TE

- PATH melding med forespørsel om etablering av en 40 Mbit/s MPLS sti
- ← RESV melding som bekrefter etableringen av en 40 Mbit/s MPLS sti



- Hver ruter foretar en sjekk om det er tilstrekkelig ressurser tilgjengelig
- Hvis ikke sendes en PathError melding tilbake til ruter X
- Path og Resv melding sendes med jevne mellomrom for å vedlikeholde LSPen (soft state)

MPLS TE og DiffServ

- MPLS TE kan benyttes for å bedre båndbredde utnyttelsen uavhengig av tjenestekvalitet (non DiffServ aware MPLS TE)
- DS-TE (DiffServ aware MPLS TE) tilbyr
 - ◆ Båndbredde begrensninger per trafikkklasse
 - ◆ Constraint based routing per klasse (max 4 klasser)
 - ◆ Admission kontroll per klasse gjennom å reservere båndbredde per klasse
- DS-TE benyttes der det er behov for bedre optimalisering av båndbredde ressursene
 - ◆ Medfører økt routing og signaleringstrafikk
 - ◆ Usikkert hvordan det skalerer

MPLS beskyttelse og feilhåndtering (Protection switching and restoration)

- Målsetning
 - ◆ Redusere nedetid og dermed hindre pakketap
 - ◆ Typiske feilrettingstider
 - SDH nettet – 50 ms
 - IP nettet – 15-30 sekunder (ofte lengre, avhenger av topologi og størrelse)
 - Telefoni – nedetid >150 ms er merkbart etter 2 sekunder koples forbindelsen ned
 - Gbit/s linker – nedetid på bare titalls ms forårsaker tap av store datamengder (mange forbindelser blir berørt)

Restoration (1)

- IGP restoration
 - ◆ Feildeteksjon typisk 10-15 sekunder, kan reduseres gjennom å redusere intervallet mellom Hallo-meldingene
 - ◆ Konvergering av rutetabellene avhenger av topologi og størrelse på nettet
 - Sende oppdatert linkstatus
 - Kalkulere ny rutetabell (ofte gjør ruterne dette før de videresender oppdateringsmeldinger)
 - ◆ Simuleringer har vist at det er mulig å redusere IGP konvergenstiden til 1-3 sekunder
- LSP restoration
 - ◆ Etablere en alternativ LSP
 - ◆ Feildeteksjon avhenger av lag 1 og 2 teknologi – benytter signalering/feilmeldinger fra lag 1 og 2
 - ◆ Informere MPLS ingress noden som kalkulerer og etablere en ny vei
 - ◆ Typisk tid for feilretting er 1-3 sekunder

Restoration (2)

- Link og node restoration
 - ◆ Midlertidig etablere en vei rundt problemet mens MPLS ingress noden finner den mest optimale veien
 - ◆ Benytte en predefinert backup link(er)
 - ◆ Feilhåndteringstid 50 ms
 - ◆ Bruk av predefinerte backup linker forutsetter at det er mulig å stakke labeler (labelene for reservelinken følger med)
- Path protection
 - ◆ Preetablere reservestier (LSPer) ende-til-ende som ikke benyttes med mindre en feil detekteres (hot standby)
 - ◆ Liten gevinst sammenlignet med LSP restoration
 - Feilmeldingene må signaleres tilbake til MPLS ingress noden
 - Sparer kun kalkulasjon av en ny MPLS sti

MPLS TE – implementasjonssteg

1. Måle og karakterisere trafikken dagens trafikk
2. Definere kostnadsdriverne
 - Båndbreddekost (aksess, kjernenett, utenlandsforbindelser)
 - Utstyr
 - Drift
 - Annet?
3. Målsetning – hva skal vi optimalisere for
 - Bedre utnyttelse av eksisterende ressurser
 - Bedre kvalitet for enkelte eller alle brukergrupper/applikasjoner
 - Lavere forsinkelse
 - Høyere tilgjengelighet
 - Lavere pakketap
4. Ta i bruk enkle TE mekanismer i deler av nettet for å få en forståelse av konsekvensene
5. Trafikkmålinger for å verifisere at de ønskede virkningene oppnås og for å tilpasse løsningen til endringer i trafikkbildet
6. Vurdere behovet for en bredere implementering (større deler av nettet, mer avanserte funksjoner)

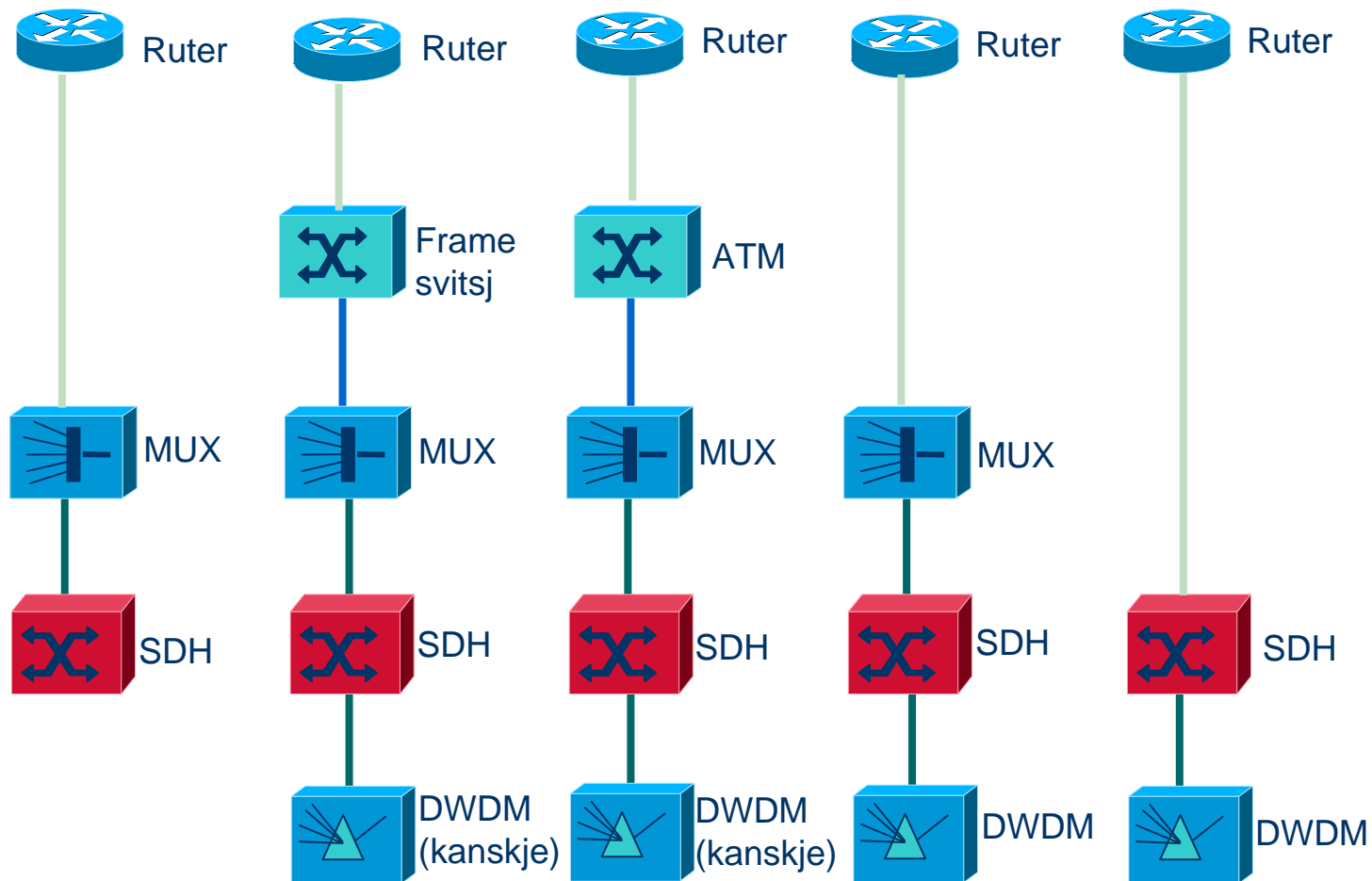
MPLS TE - oppsummering

- Vil bidra til større kontroll
- Viktig dersom nettet skal levere garantert tjenestekvalitet
- Usikkert hva og hvor mye som trengs
- Mange opsjoner – usikkert hvor godt utstyr fra ulike leverandører vil samspille

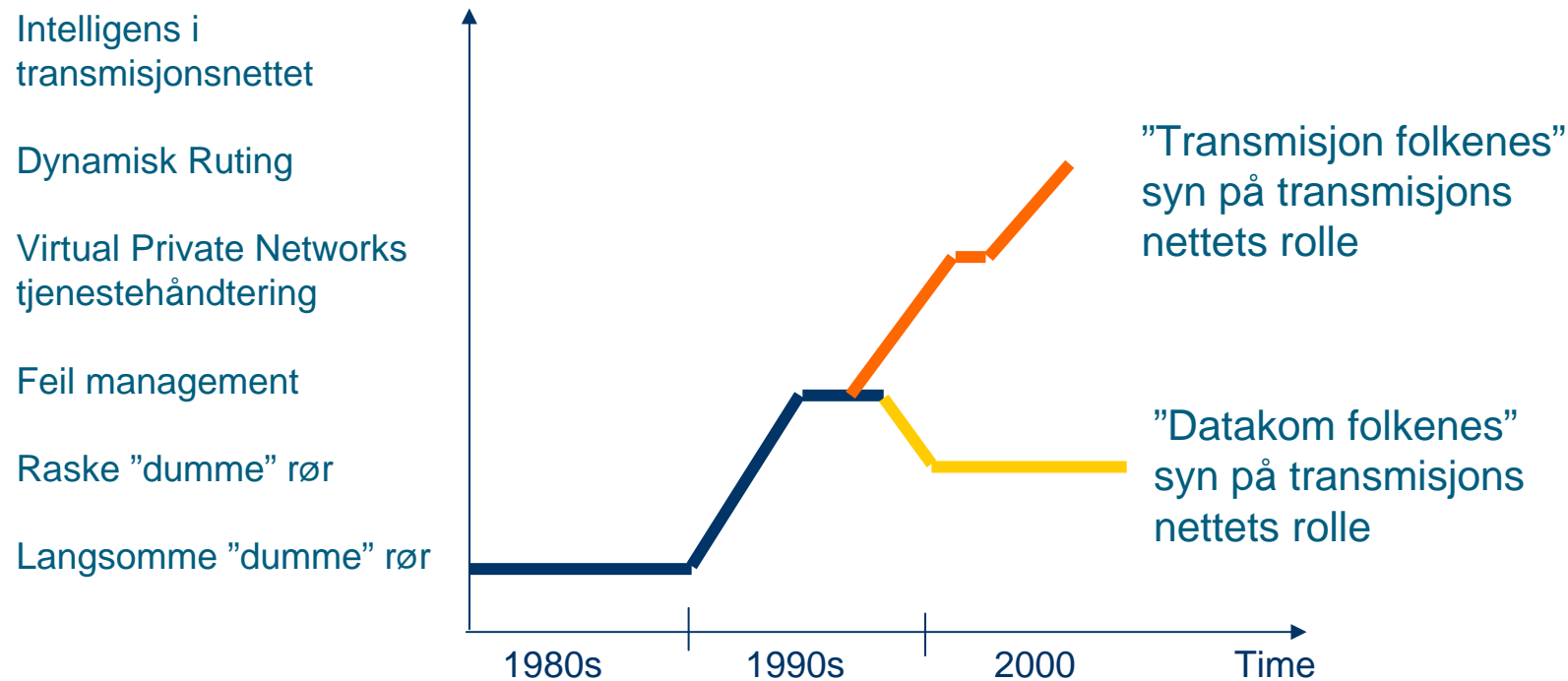
Uløste problemer

- Mangler interdomene løsninger
 - ◆ Usikkert om det blir standardisert
 - ◆ Krever avtaler mellom operatørene
- Multicast støttes ikke på en effektiv måte

Utviklingen av "utstyrslagene"



Hva skal de ulike lagene tilby?



Overlay versus peer modell

- OIF (Optical Interworking Forum) støtter en overlay modell
 - ♦ To ulike kontroll plan, et for det optiske kjernenettet og et mellom kantnodene og kjernenodene
 - UNI signaleringen standardiseres av OIF
 - ♦ Skjuler topologien i det optiske nettet
- IETF standardiserer GMPLS (Generalized MPLS) som støtter en peer modell
 - ♦ Ett kontrollplan på tvers av lagene og det administrative domenet
 - ♦ Viser topologien i det optiske nettet
- Peer modellen er et supersett av overlay modellen
 - ♦ Kan benytte en kombinasjon av de to ved å skru av topologiinformasjons utveksling på enkelte kantnoder

GMPLS (Generalized MPLS)

- Utvide MPLS TE til det optiske laget
- Utvider MPLS til å støtte mer enn kun pakkesvitsjing
 - ◆ SDH svitsjing
 - ◆ Bølgelengde svitsjing
 - ◆ Fysisk portsvitsjing
- Peer modell
- Integrerer IP og optikk lagene
- Åpner for nye forretningsmodeller og effektivisering
 - ◆ Raskere og enklere levering av tjenester
 - ◆ Mer effektiv management

GMPLS mekanismene

- IGP utvidelser
- Forwarding adjacency
 - ◆ Annonserer LSPer inn i IGP
- LSP hierarki
 - ◆ Aggregering av LSPer
- Constraint-based routing
- Signaleringsutvidelser
- Link bundling
 - ◆ Økt skalerbarhet gjennom å slå sammen linker (annonserer de som en enkelt link)
- OIF optisk UNI signalerings utvidelser
- LMP (Link Management Protocol)

LMP

- Benyttes mellom naboroder
- Etablerer og vedlikeholder en signalleringskanal
 - ◆ Utveksler kontrollinformasjon, label distribusjon, topologi informasjon
 - ◆ Status meldinger
 - ◆ Feilrapporter
 - ◆ Verifiserer link konnektivitet

GMPLS - oppsummering

- Åpen standard
- Reduserer kompleksiteten i nettet
 - ◆ Færre bokser
- Prosesseringskapasiteten i ruterne vokser langsommere enn båndbredden
 - ◆ Krever optiske krysskopplere/svitjer i kjernenettet
- Ruterne kan "se" underliggende infrastruktur
 - ◆ Unngår n^2 problemet
- Utnytter MPLS TE erfaringer
- Kan velge modell (peer eller overlay)
- Åpner for raskere og mer dynamisk levering av båndbredde
- Nettet optimaliseres for Internettrafikk

Sammendrag

- Levering av garantert og differensiert tjenestekvalitet krever bedre kontroll av trafikken i nettet
 - ◆ MPLS TE nødvendig for å støtte ulike markedssegmenter og tjenester i ett nett
 - ◆ MPLS TE vil bidra til bedre kostnadskontroll
- Enklere integrasjon mellom IP og underliggende lag
 - ◆ Slipper flere nivåer av ruting, adressering og feilhåndtering

MPLS danner grunnlaget for neste generasjons IP nett

MPLS standardising....

- IETF - Internet Engineering Task Force
 - ◊ <http://www.ietf.org/>
- MPLS WG
 - ◊ <http://www.ietf.org/html.charters/mpls-charter.html>
- Traffic Engineering WG
 - ◊ <http://www.ietf.org/html.charters/tewg-charter.html>
- Provider Provisioned VPN (PPVPN) WG
 - ◊ <http://www.ietf.org/html.charters/ppvpn-charter.html>