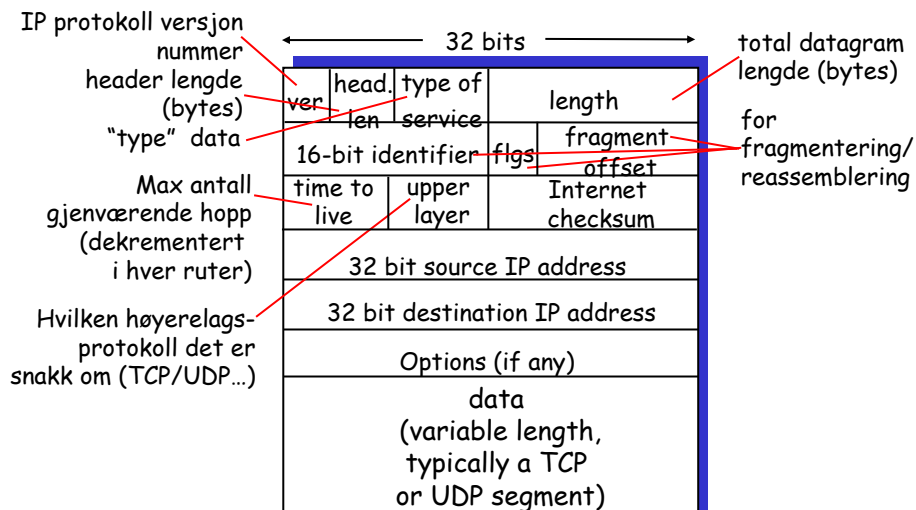


## Ruting i Internett

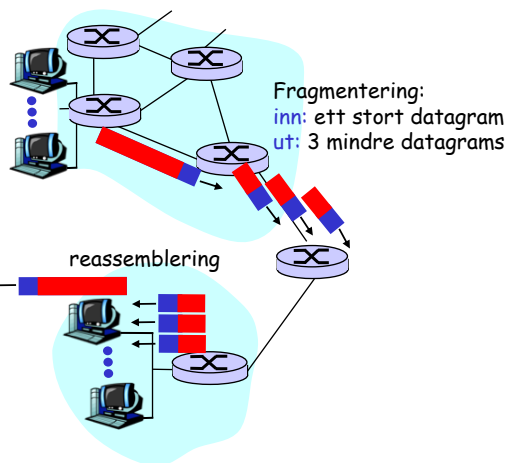
- Det globale internettet består av sammenkoblede **Autonomous Systems (AS)** :
  - **Stub AS**: liten organisasjon
  - **Multihomed AS**: stor organisasjon (ingen gjennomgangstrafikk)
  - **Transit AS**: provider
  
- Ruting i to nivåer:
  - **Intra-AS**: AS administrator er fri til å velge
  - **Inter-AS**: unik standard

## IP datagram format



## IP Fragmentering & Reassemblering

- Nettverkslinker har MTU (max.transfer unit) - største mulige link-nivå ramme.
  - Forskjellige linker - forskjellige MTUer
- store IP datagram deles ("fragmenteres") i nettet
  - ett datagram blir flere datagram
  - "reassembleres" bare ved destinasjonen
  - IP header bits blir brukt til å identifisere, og ordne relaterte fragmenterte pakker.



## IP fragmentering og reassemblering

length	ID	fragflag	offset
=4000	=x	=0	=0

Ett stort datagram blir til flere mindre datagram

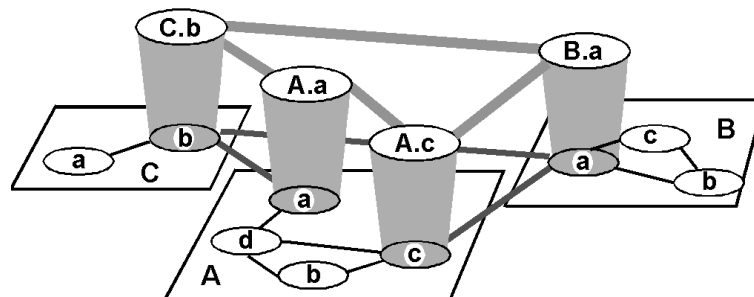
length	ID	fragflag	offset
=1500	=x	=1	=0
=1500	=x	=1	=1480
=1040	=x	=0	=2960

## ICMP: Internet Control Message Protocol

- Brukes av vertsmaskiner, rutere, gateways til å kommunisere nettnivå informasjon
- feilrapportering: unåbar maskin, nettverk, port, protokoll
  - echo request/reply (brukt av ping)
- nettverkslag "over" IP:
  - ICMP meldinger pakkes som IP datagram
- ICMP melding: type, code plus første 8 bytes av IP datagrammet som førte til feil

Type	Code	description
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control - not used)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

## Internettets AS hierarki



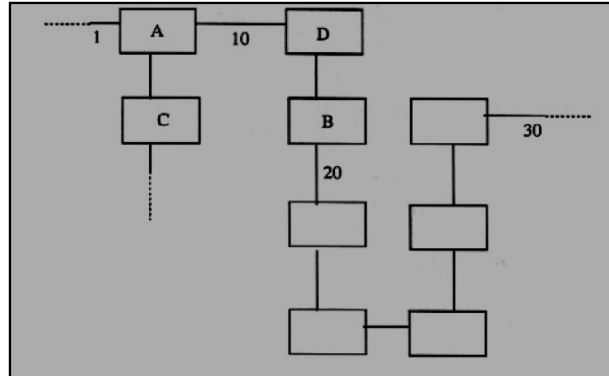
## Intra-AS Routing

- ❑ Også kjent som **Interior Gateway Protocols (IGP)**
- ❑ Mest vanlige IGP'er:
  - RIP: Routing Information Protocol
  - OSPF: Open Shortest Path First
  - IGRP: Interior Gateway Routing Protocol (Cisco propr.)

## RIP ( Routing Information Protocol)

- ❑ Distansevektor
- ❑ Inkludert i BSD-UNIX Distribusjon i 1982
- ❑ Distanse mål: antall hopp (max = 15 hopp)
  
- ❑ Distansevektor: utveksles hvert 30 sek via en *Response Message* (også kalt **Advertisement**)
- ❑ Hvert Advertisement inneholder opp til 25 destinasjonsnett

## RIP (Routing Information Protocol)



Destinasjonsnett	Neste Ruter	Ant hopp til dest.
1	A	2
20	B	2
30	B	7
10	--	1

[ simula . research laboratory ] ... Routing i Internett

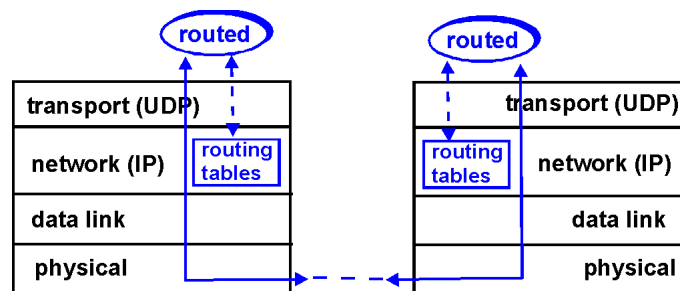
## RIP: Håndtering av link-feil

- ❑ Dersom intet advertisement hørt etter 180 sek, anta at nabo eller link er død
- ❑ Alle stier via naboen blir invalidert; nye advertisements sendt til de andre naboene.
- ❑ Naboene sender i sin tur ut advertisements dersom deres tabeller endret seg.
- ❑ Informasjon om feilende linker propagerer hurtig.
- ❑ "Poison reverse " brukes til å hindre ping-pong loops (uendelig distanse = 16 hops)

## RIP Tabellprosessering

- ❑ RIP ruting tabeller blir vedlikeholdt av en **applikasjonsprosess** kalt route-d (daemon)
- ❑ Advertisements innkapsles i UDP pakker (pålitelig levering er ikke påkrevet; advertisements repeteres jevnlig)

## RIP Tabellprosessering



## RIP Tabelleksempel

RIP tabelleksempel

(ved ruteren *giroflee.eurocom.fr*):

- ❑ Tre tilkoblede klasse C nettverk (LANs)
- ❑ Ruterer kjener bare stier knyttet til LANs
- ❑ Default router i bruk for å komme inn og ut
- ❑ Multicast adresse: 224.0.0.0
- ❑ Loopback interface (til debugging)

## RIP Tabelleksempel

Destination	Gateway	Flags	Ref	Use	Interface
127.0.0.1	127.0.0.1	UH	0	26492	lo0
192.168.2.	192.168.2.5	U	2	13	fa0
193.55.114.	193.55.114.6	U	3	58503	le0
192.168.3.	192.168.3.5	U	2	25	qaa0
224.0.0.0	193.55.114.6	U	3	0	le0
default	193.55.114.129	UG	0	143454	

## OSPF (Open Shortest Path First)

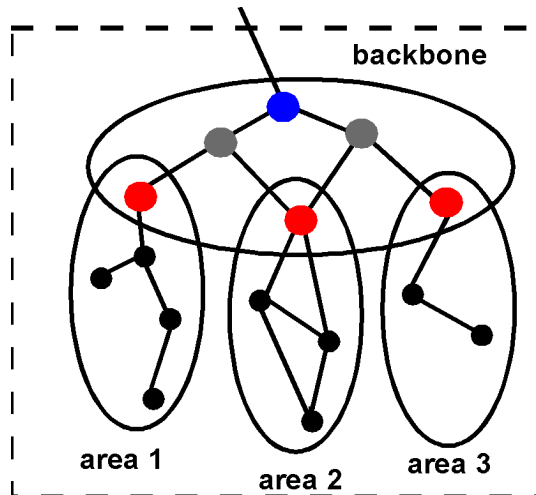
- ❑ "open": Offentlig tilgjengelig
- ❑ Bruker Link State algoritmen
  - LS pakkesdisseminering
  - Topology kart i hver node
  - Stitregning ved hjelp av Dijkstra's alg
- ❑ OSPF advertisement inneholder en entry per naboruter.
- ❑ Advertisements dissemineres til **hele** Autonomous System.

## OSPF "avanserte" egenskaper (ikke i RIP)

- ❑ **Sikkerhet**: alle OSPF meldinger blir autentisert (for å hindre "hacking"); TCP forbindelser blir brukt
- ❑ **Multiple stier** med samme kostnad tillatt (bare en sti i RIP)
- ❑ For hver link, multiple kost metrikker for forskjellige anvendelser (feks., satellite link kostnad er "lav" for best effort; høy for real time)
- ❑ Integrert uni- og **multicast** støtte:
  - Multicast OSPF (MOSPF) bruker samme topologidatabase som OSPF
- ❑ **Hierarkisk** OSPF i store domener.



## Hierarkisk OSPF



[ **simula** . research laboratory ]

Ruting i Internett 17

## Hierarkisk OSPF

- To-nivå hierarki: local area og backbone.
- Link-state advertisements forlater ikke sitt domene.
- Noder i hvert domene har detaljert domenetopologi; de kjenner bare retningen (shortest path) til noder i andre domener.
- **Area Border rutere** "sammenfatter" avstander til nettverk i domenet, og informerer andre Area Border routers.
- **Backbone rutere** kjører en OSPF ruting alg begrenset til backbone
- **Boundary rutere** kobler seg til andre AS'er.

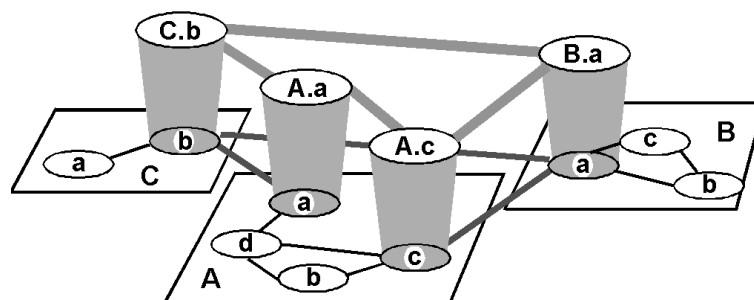
[ **simula** . research laboratory ]

Ruting i Internett 18

## EIGRP (Enhanced Interior Gateway Routing Protocol)

- ❑ CISCO proprietær; etterfølger til RIP (midten av 80-tallet)
- ❑ Distansevektor, som RIP
- ❑ flere kost-metrikker (forsinkelse, båndbredde, pålitelighet, last osv.)
- ❑ bruker TCP til utveksling av informasjon
- ❑ routingtabeller utveksles bare når kostnad endrer seg.
- ❑ Uendelig løkke unngås ved Distributed Updating Alg. (DUAL).
- ❑ I DUAL, etter at en distanse har øket, fryses routingtabellene inntil alle noder som påvirkes av endringen har hørt om den.

## Inter-AS ruting



## Inter-AS routing (forts)

- BGP (Border Gateway Protocol): de facto standard
- Path Vector protokoll: utvidelse av Distance Vector
- Hver Border Gateway kringkaster til naboene (peers) hele stien (dvs, sekvens av AS'er) til destinasjonen
- F. eks, Gateway X kan lagre følgende sti til destinasjon Z:

$$\text{Path (X,Z)} = X, Y_1, Y_2, Y_3, \dots, Z$$

## Inter-AS ruting (forts)

- Anta at Gwy X sender sin sti til Gwy W
- Gwy W velger om den vil benytte stien tilbudt av Gwy X, på bakgrunn av kostnad, politikk (\$\$\$\$) eller for å unngå rutingløkker.
- Dersom Gwy W velger stien tilbudt av Gwy X:

$$\text{Path (W,Z)} = w, \text{Path (X,Z)}$$

Merk: valg av stier skjer ikke så mye basert på kostnad, (f.eks, # AS hopp), men mest basert på administrative og politiske avgjørelser (f.eks., ikke send pakker gjennom konkurrentens AS)

## Inter-AS ruting (forts)

- Utrveksling av BGP meldinger vha. TCP.
- **OPEN** melding åpner TCP forbindelsen og autentiserer senderen.
- **UPDATE** melding avetterer ny sti (eller trekker tilbake gammel)
- **KEEPALIVE** melding holder forbindelsen i live i mangel av UPDATES; tjener også som ACK til en OPEN forespørsel
- **NOTIFICATION** melding rapporterer om feil i forrige melding. Benyttes også til å terminere en forbindelse.

## Hvorfor forskjellig Intra- og Inter-AS ruting ?

- **Politikk:** Inter må kunne forholde seg til "politiske" avgjørelser (hvilken provider man skal benytte seg av, unngå osv.). Intra er inneholdt og kontrollert av *en* organisasjon, så ingen politiske avgjørelser er nødvendige
- **Skala:** Inter gir et ekstra nivå i reduksjon av tabellstørrelse og oppdateringstraffikk.
- **Ytelse:** Intra er rettet mot ytelsesmetriker; trenger å holde "kostnader" lave. I Inter er det vanskelig å propagere ytelsesmetriker effektivt. Dessuten er politiske betraktninger mer meningsfulle her.

**Man trenger begge!**

## Noen kommentarer til slutt

- ❑ Det vi har forelest til nå er ikke hele historien.
- ❑ Det er mange svakheter ved internettet som man forsøker å bøte på ved tilleggs-standarder.
- ❑ Noen av disse er i vanlig bruk.
- ❑ "MPLS" gir bedre kontroll med stier og mer effektiv svitsjing. Dette er i svært vanlig bruk i kjærnenett.
- ❑ Tjenestekvalitet er et svakt punkt, men søkt løst ved standarder som "Intserv" og "Diffserv". Diffserv er forholdsvis vanlig i bruk.