Routing
(Part 2)

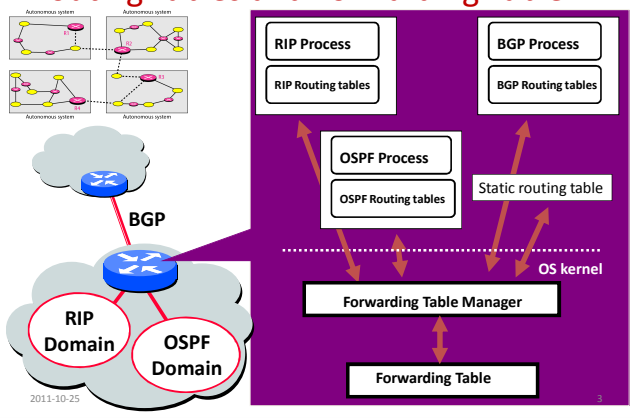
ETSF10 – Internet Protocols – 2011

Kaan Bür & Jens Andersson
Department of Electrical and Information Technology

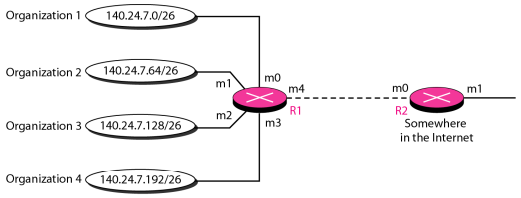
Routing

- Introduction
- Inside the Router §8.4
- Unicast Routing §22.3
  - Intradomain Routing
  - ***Detour → Forwarding Process*** §22.2
  - Interdomain Routing
- Multicast Routing §22.4
  - IGMP §21.3

Routing Tables and Forwarding Table



Forwarding: Address aggregation



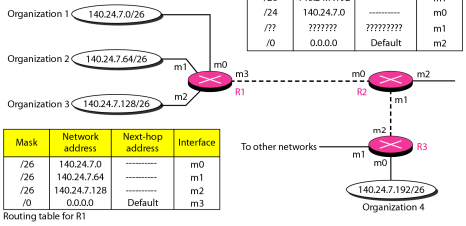
Mask	Network address	Next-hop address	Interface
/26	140.24.7.0	----------	m0
/26	140.24.7.64	----------	m1
/26	140.24.7.128	----------	m2
/26	140.24.7.192	----------	m3
/0	0.0.0.0	Default	m4

Routing table for R1

Mask	Network address	Next-hop address	Interface
/24	140.24.7.0	----------	m0
/0	0.0.0.0	Default	m1

Routing table for R2

Forwarding: Longest mask matching



Routing table for R2

Mask	Network address	Next-hop address	Interface
/26	140.24.7.192	----------	m1
/24	140.24.7.0	----------	m0
/??	???????	?????????	m1
/0	0.0.0.0	Default	m2

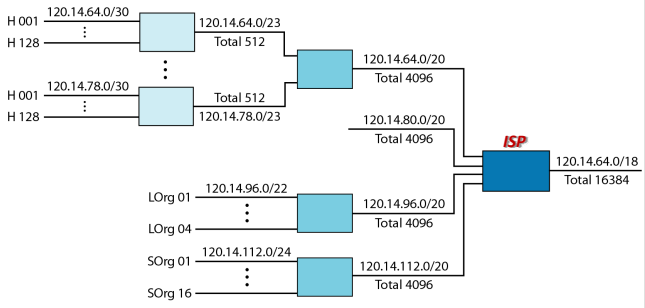
Mask	Network address	Next-hop address	Interface
/26	140.24.7.0	----------	m0
/26	140.24.7.64	----------	m1
/26	140.24.7.128	----------	m2
/0	0.0.0.0	Default	m3

Routing table for R1

Mask	Network address	Next-hop address	Interface
/26	140.24.7.192	----------	m0
/??	???????	?????????	m1
/0	0.0.0.0	Default	m2

Routing table for R3

Forwarding: Hierarchical routing

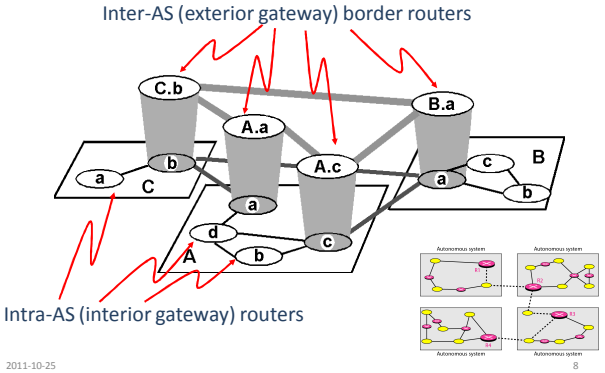


Routing

- Introduction
- Inside the Router §8.4
- **Unicast Routing §22.3**
  - Intradomain Routing
  - **Interdomain Routing**
- Multicast Routing §22.4
  - IGMP §21.3

2011-10-25 7

Internet AS Hierarchy



2011-10-25 8

Hierarchical Routing

- aggregate routers into regions, "autonomous systems" (AS)
- routers in same AS run same routing protocol
  - "intra-AS" routing protocol
  - routers in different AS can run different intra-AS routing protocol

Border Gateway Routers

- special routers in AS
- run intra-AS routing protocol with all other routers in AS
- also responsible for routing to destinations outside AS
  - run inter-AS routing protocol with other gateway routers

2011-10-25 9

Why different Intra- & Inter-AS routing?

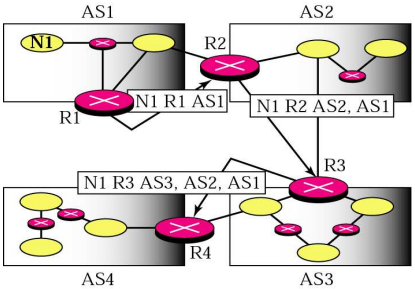
- Policy
  - Inter-AS: admin wants control over how its traffic routed, who routes through its net.
  - Intra-AS: single admin, so no policy decisions needed
- Scale
  - Hierarchical: saves table size, reduced update traffic
- Performance
  - Intra-AS: can focus on performance
  - Inter-AS: policy may dominate over performance

Internet Inter-AS routing: BGP

- Border Gateway Protocol: *de facto* standard
- Path Vector protocol:
  - Similar to *Distance Vector*
  - Border gateways broadcast to neighbours (peers) entire path (sequence of AS) to destination
  - BGP routes to networks (AS), not individual hosts

Path Vector Messages

- Same principle as distance vector routing



Path Vector Routing Table

Network	Next Router	Path
N01	R01	AS62, AS23, AS67
N02	R05	AS67, AS22, AS05, AS89
N03	R06	AS67, AS89, AS09, AS34
N03	R12	AS62, AS02, AS34

Network id

"Output port"

"Metric"
One of many
ATTRIBUTES

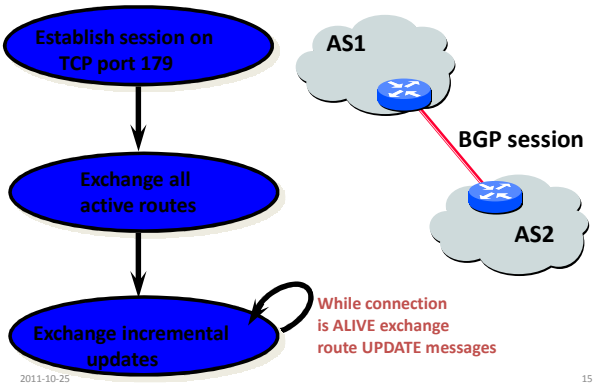
2011-10-25 AS = Autonomous System = Organisation 13

BGP Router Operations

- Receiving and filtering route advertisements from directly attached neighbor(s)
- Route selection
  - To route to destination X, which path (of several advertised) will be taken?
- Sending route advertisements to neighbours

2011-10-25 14

BGP Router Operations



2011-10-25 15

BGP Messages

- **OPEN**: opens TCP connection to peer and authenticates sender
- **UPDATE**: advertises new path (or withdraws old)
- **KEEPALIVE** keeps connection alive in absence of UPDATES; also ACKs OPEN request
- **NOTIFICATION**: reports errors in previous msg; also used to close connection

2011-10-25 16

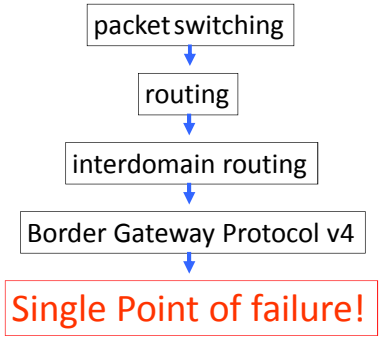
BGP Attributes

Value	Code	Reference
1	ORIGIN	[RFC1771]
2	AS_PATH	[RFC1771]
3	NEXT_HOP	[RFC1771]
4	MULTI_EXIT_DISC	[RFC1771]
5	LOCAL_PREF	[RFC1771]
6	ATOMIC_AGGREGATE	[RFC1771]
7	AGGREGATOR	[RFC1771]
8	COMMUNITY	[RFC1997]
9	ORIGINATOR_ID	[RFC2796]
10	CLUSTER_LIST	[RFC2796]
11	DPA	[Chen]
12	ADVERTISER	[RFC1863]
13	RCID_PATH / CLUSTER_ID	[RFC1863]
14	MP_REACH_NLRI	[RFC2283]
15	MP_UNREACH_NLRI	[RFC2283]
16	EXTENDED COMMUNITIES	[Rosen]
...		
255	reserved for development	

From IANA: http://www.iana.org/assignments/bgp-parameters

2011-10-25 17

Is There A Problem?



2011-10-25 18

Scary?

- **BGP is not guaranteed to convergeon a stable routing. Policy interactions could lead to "livelock" protocol oscillations.**
  See "Persistent Route Oscillations in Inter-domain Routing" by K. Varadhan, R. Govindan, and D. Estrin. ISI report, 1996
- **Corollary: BGP is not guaranteed to recover from network failures.**

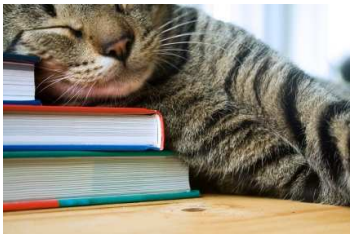
2011-10-25 19

To Do Now: "One Minute Paper"

- Routers / switches
  - Functions and architecture
- Routing / forwarding
  - Intra- vs. inter-domain
  - Distance vector vs. link state
- **What was the most important thing you've learnt so far? Why?**

2011-10-25 20

See you in 15' :)



- After the break
  - Multicast routing
  - IGMP

2011-10-25 21

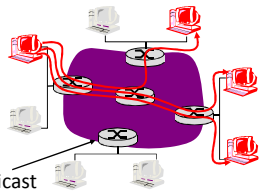
Routing

- Introduction
- Inside the Router §8.4
- Unicast Routing §22.3
  - Intradomain Routing
  - Interdomain Routing
- **Multicast Routing §22.4**
  - IGMP §21.3

2011-10-25 22

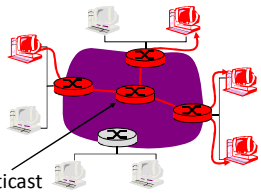
Multicast: One-to-many Routing

Unicast



Unicast routers forward multiple unicast datagrams

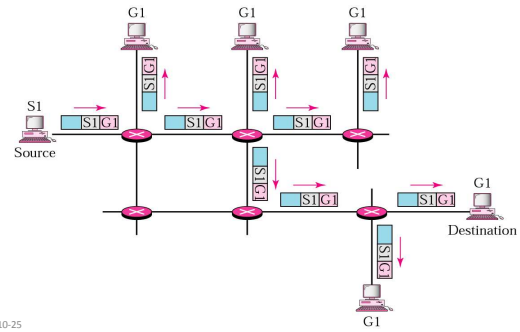
Multicast



Multicast routers (red) duplicate and forward multicast datagrams

2011-10-25 23

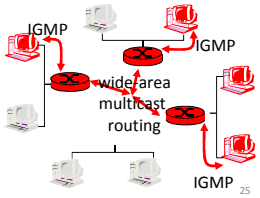
Source and Group Addresses



2011-10-25 24

Joining a Multicast Group

- **Local**: host informs local multicast router
  - IGMP (Internet Group Management Protocol)
- **Wide area**: local router interacts with other routers to build forwarding tree and receive multicast data flow
  - MOSPF, DVMRP, PIM-DM
  - CBT, PIM-SM



2011-10-25 25

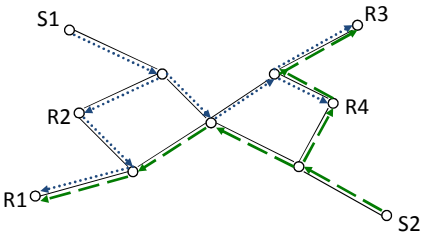
Multicast Routing Protocols

- Shortest path trees, again!
- In unicast routing
  - One path (on tree) used at a time
- In multicast routing
  - Whole tree used each time
  - Each source needs a tree

2011-10-25 26

Source-Based Tree

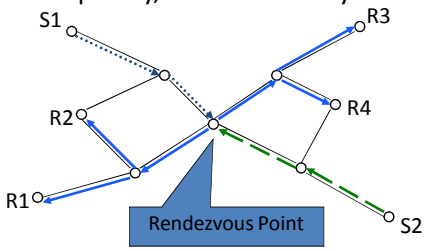
- One tree per source (at each router)
- One source per group
- High complexity, high efficiency



2011-10-25 27

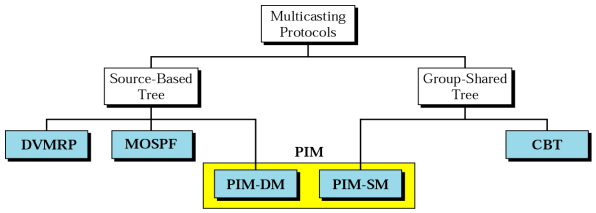
Group-Shared Tree

- One tree per group (at one router)
- Shared by multiple sources in group
- Lower complexity, lower efficiency



2011-10-25 28

Classification of Algorithms



2011-10-25 29

PIM

- Independent from unicast protocol
- Uses available routing info for path lookups
- Two modes:
  - Sparse Mode
  - Dense Mode

2011-10-25 30

PIM-SM

- Relatively few members assumed
- Trees are built on demand (when needed)
  - Group-shared trees with rendezvous points
- Methods for tree construction
  - Grafting
  - Pruning
- Can switch from group-shared to source-based if more efficient

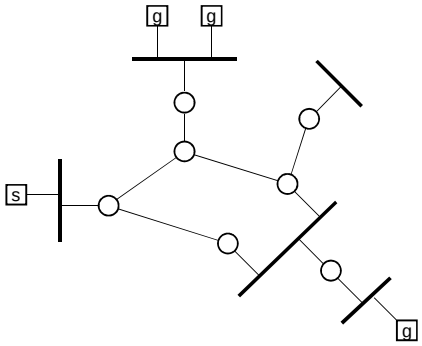
2011-10-25 31

PIM-DM

- All hosts assumed to be members
- Build source-based tree from source
- Routers without members prune tree
- Grafting used to add new members

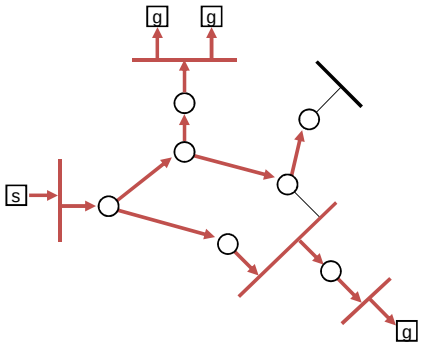
2011-10-25 32

Example Topology



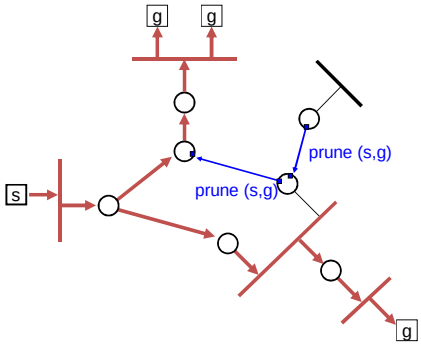
2011-10-25 33

Truncated Broadcast



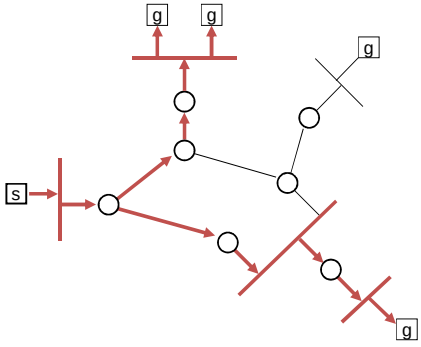
2011-10-25 34

Pruning



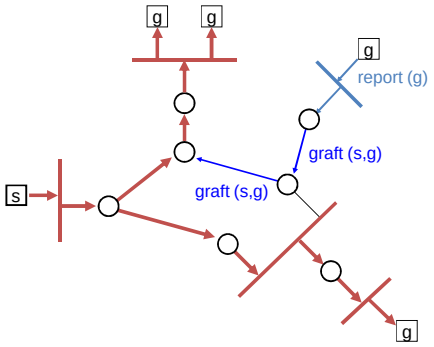
2011-10-25 35

Steady State after Pruning

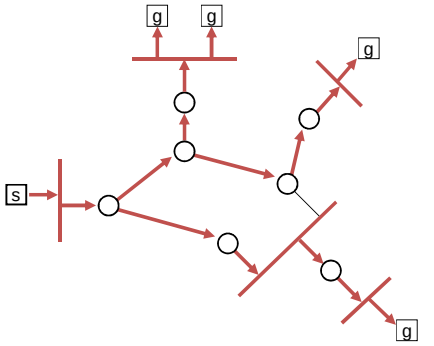


2011-10-25 36

Grafting on New Receivers

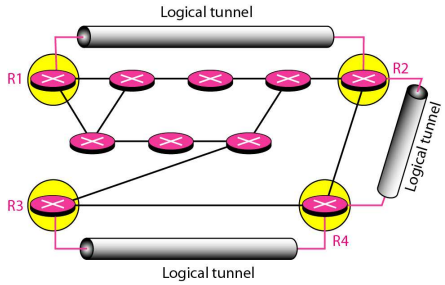


Steady State after Grafting



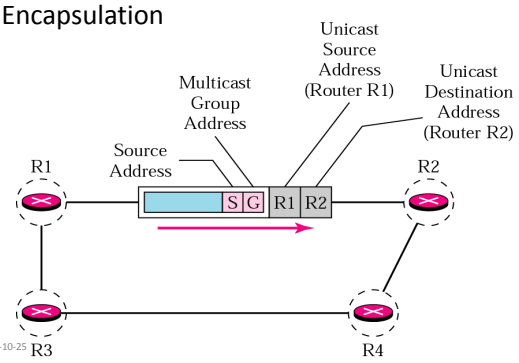
Logical Tunnelling

- Very few Internet routers can multicast
  - How to connect them?



Multicast Backbone (MBONE)

- Encapsulation



2011-10-25 40

Routing

- Introduction
- Inside the Router §8.4
- Unicast Routing §22.3
  - Intradomain Routing
  - Interdomain Routing
- **Multicast Routing §22.4**
  - **IGMP §21.3**

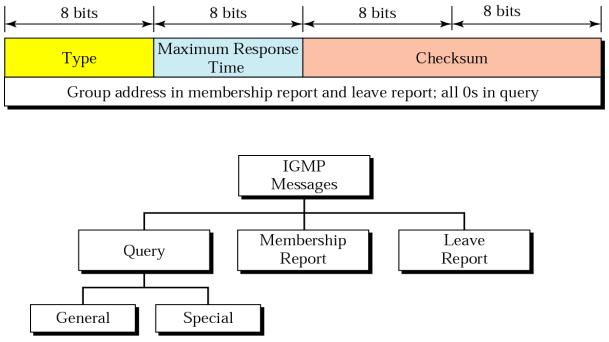
2011-10-25 41

Internet Group Management Protocol

- IGMP, runs on top of IP
- Not a multicast protocol
  - Complementary
  - Runs in the leaves of the network
- Manages group membership
  - Provides multicast router with info

2011-10-25 42

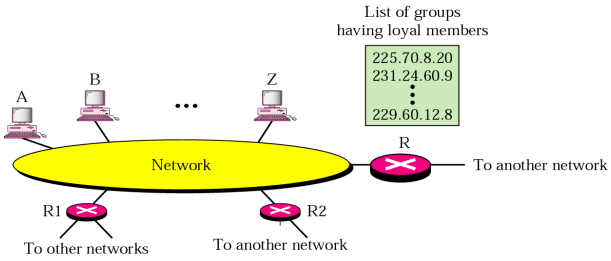
IGMP Message Format



2011-10-25 43

IGMP Operation

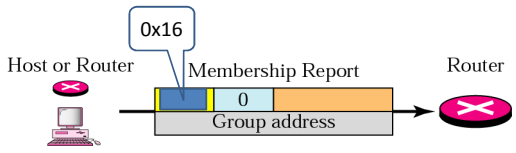
- Only one router distributes packets in a group
  - Other routers may be serving their networks



2011-10-25 44

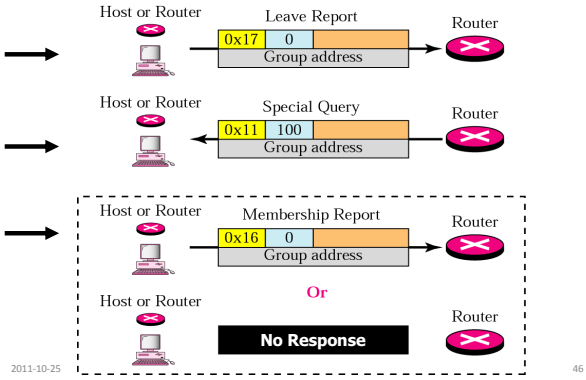
Joining a Group

- Request to router
  - Forwarded if first for a group



2011-10-25 45

Leaving a Group



IGMP General Query