Routing

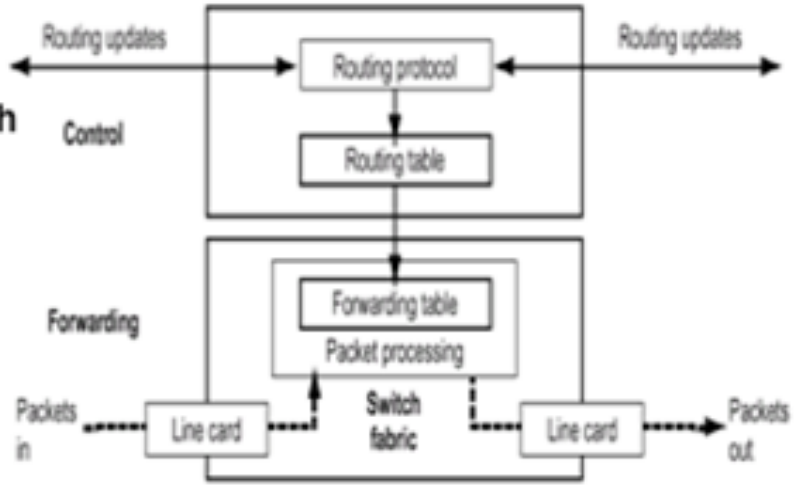
Intradomain (RIP, OSPF, IS-IS)

Outline

- **Introduction: terminology, concepts**
- **Hierarchical routing,**
  - **Intra-domain, Inter-domain**
  - **Choice of routing protocol, characteristics**
  - **RIP, OSPF**
- **OSPF vs IS-IS**
  - **Similarities, terminology, transport**
  - **Packet encoding, Area Architecture, Database granularity, Neighbor Adjacency Establishment, AAA, MPLS-TE support**
  - **For Service Providers**

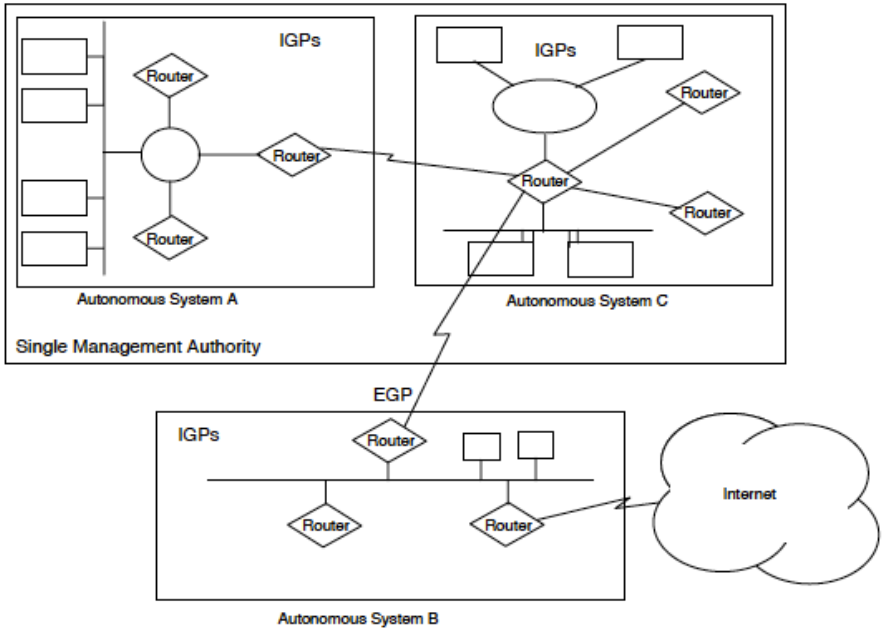
First: *Routing* vs Forwarding

- **Routing**: building and maintaining the routing table
  - Connectionless services - per packet
  - Routing protocols build the routing tables: Router control plane
    - <dst, hop count, next hop>
- **Forwarding** is in the data path
  - Packet processing
  - In-out on physical line cards



Organization of the Internet – autonomous systems and hierarchical routing

- An **autonomous system (AS)**:
  a domain (collection of IP networks and routers) under the control of one entity (or sometimes more) that presents a common routing policy to the Internet



- **Hierarchical routing**
  - **Intra-domain** (interior) routing:
    within an AS, Interior Gateway Protocol
  - **Inter-domain** (exterior) routing:
    between AS, Exterior Gateway Protocol
  - Reduces the routing information to be kept within an AS
    - Uses ”standard router” (default router) towards other AS’s
    - Simple for all to know which route to be applied towards the external

The interconnection of autonomous routing domains allow for packets to be sent anywhere

Routers implement routing algorithms through routing protocols

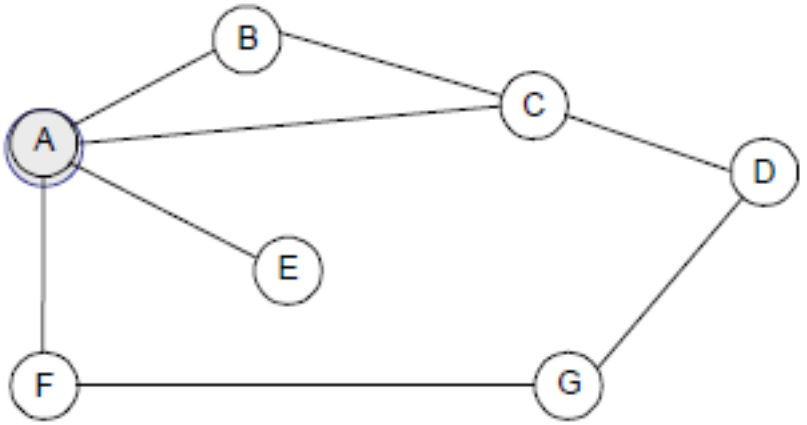
- Static vs dynamic
- Source routing vs hop-by-hop
- One vs multiple paths
- Flat vs hierarchical
- “Distance vector” vs “Link state”

- Hierarchical routing
  - Intra-domain (interior) routing vs. Inter-domain (exterior) routing

Routing protocols build the routing tables …

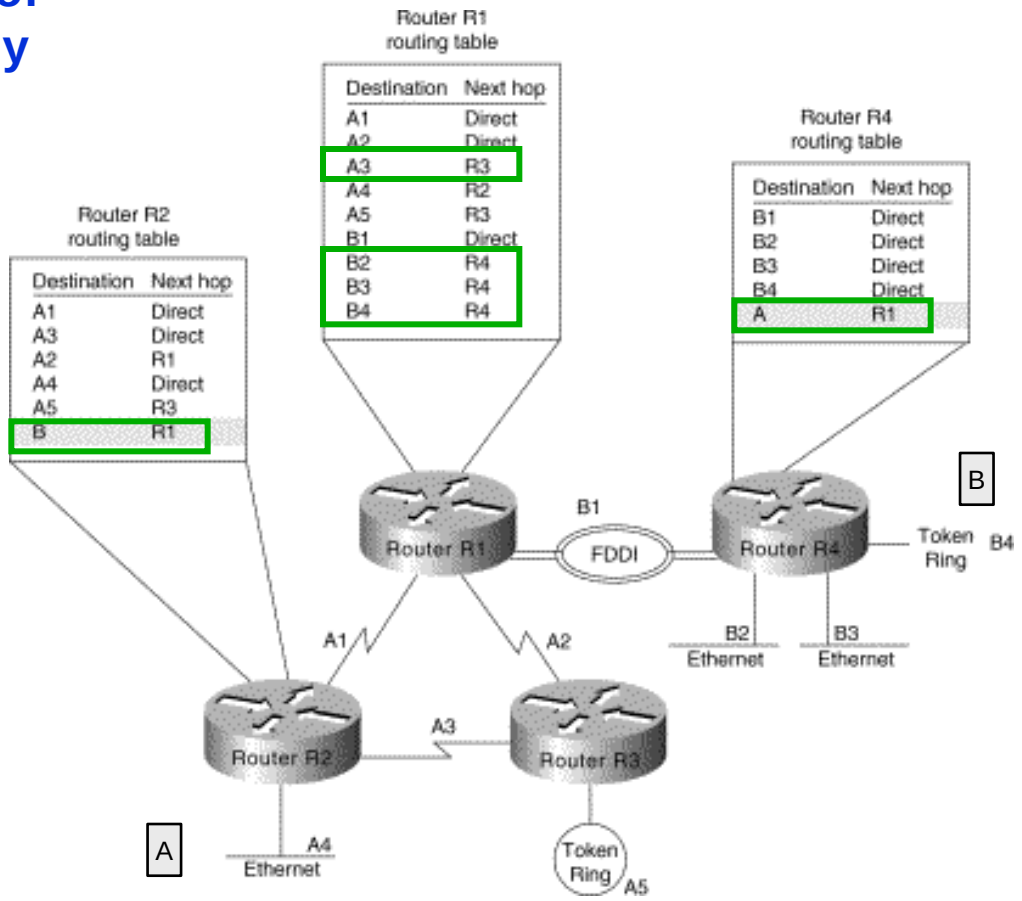
- The routing table at A, lists – *at a minimum* – the next hops for the different destinations

Dest	Next Hop
B	B
C	C
D	C
E	E
F	F
G	F



Routing protocols build the routing tables …

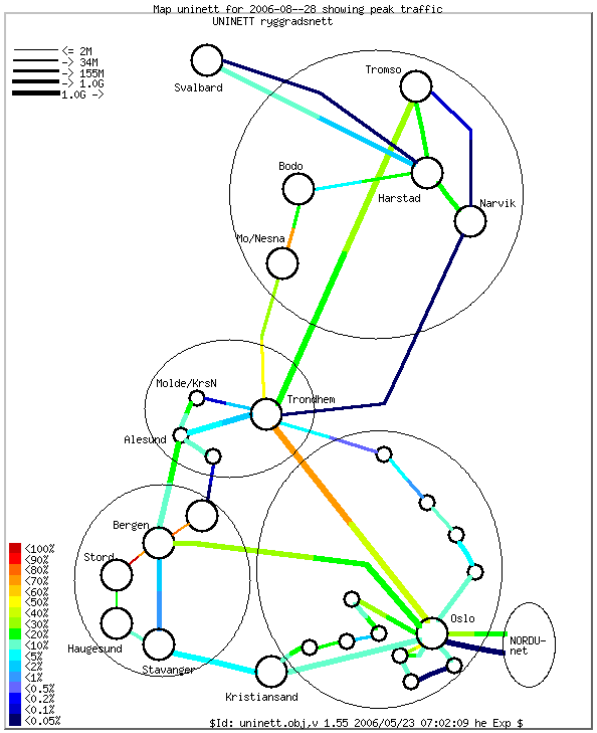
- **Routing protocols are run between routers to maintain a correct picture of the dynamic network topology**
- **Shortest path is found through distributed and dynamic routing algorithms**



… based on the cost of the links

- **Path length, number of hops**
  - $\Sigma cost_i$ i= links on relevant path
- **Bandwidth**
- **Line cost**
- **Reliability**
  - **Bit errors**
- **Delay**
- **Load**
  - **Queue length**
  - **CPU load**
  - **Packets per sec**

⇨ **and … combinations**



Outline

- Introduction: terminology, concepts
- Hierarchical routing,
  - Intra-domain, Inter-domain
  - Choice of routing protocol, characteristics
  - RIP, OSPF
- OSPF vs IS-IS
  - Similarities, terminology, transport
  - Packet encoding, Area Architecture, Database granularity, Neighbor Adjacency Establishment, AAA, MPLS-TE support
  - For Service Providers

Hierarchical routing reduces the routing information to be kept within domains

- An autonomous system (AS) is a domain (collection of IP networks and routers) under the control of one entity (or sometimes more) that presents a common routing policy to the Internet

- Intra-domain routing is within an AS
  - Ignores the Internet outside the autonomous system
  - Interior Gateway Protocols (IGP's) such as RIP, OSPF and IS-IS
  - Uses default routes towards other AS's

- Inter-domain routing is between AS's
  - Assumes that the Internet consists of a collection of interconnected AS's
  - Exterior Gateway Protocols (EGP's) such as BGP (Border Gateway Protocol)

The choice of a routing protocol depends on network complexity, size, and administrative policies

- **Scalability to large environments**
- **Stability during outages**
- **Speed of convergence**
- **Metrics**
- **(Support for VLSM – variable length subnet mask)**
- **Vendor interoperability**

Dinstance Vector Algorithm

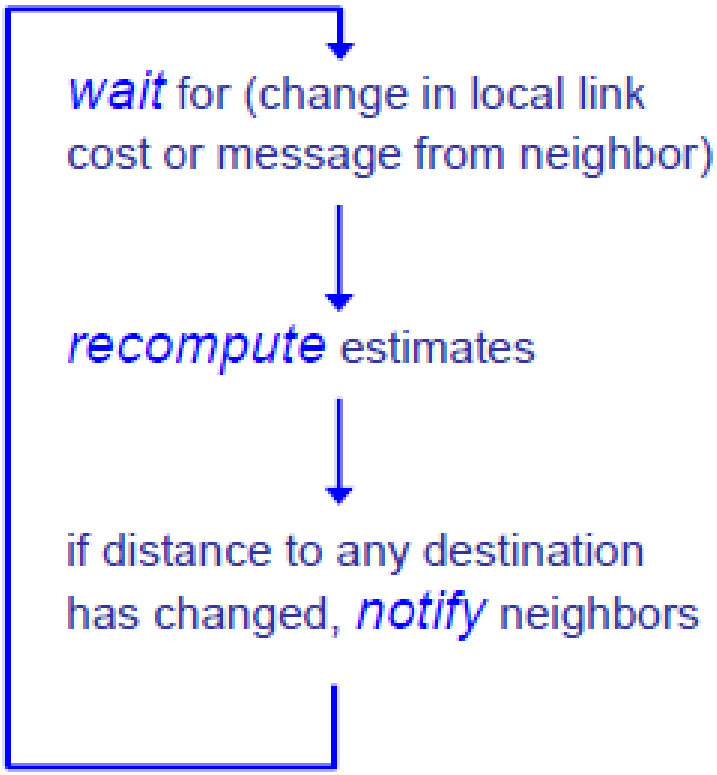
Iterative, asynchronous: each local iteration caused by:

- Local link cost change
- Distance vector update message from neighbor

Distributed:

- Each node notifies neighbors *only* when its DV changes
- Neighbors then notify their neighbors if necessary

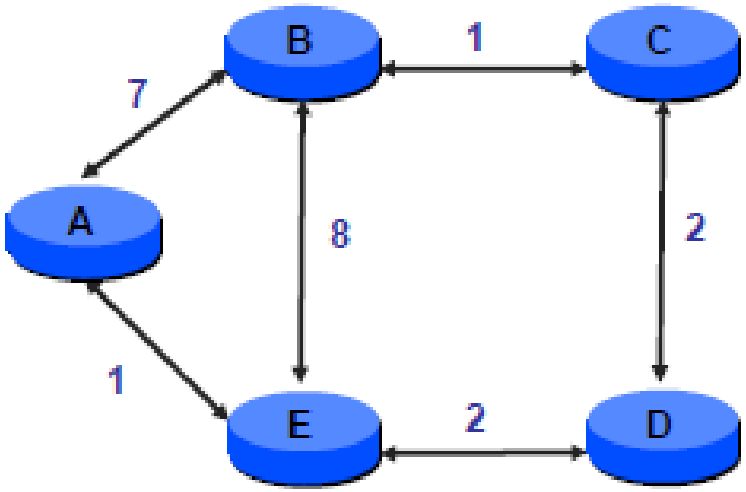
Each node:



Distance Vector – Aspects maintained by node

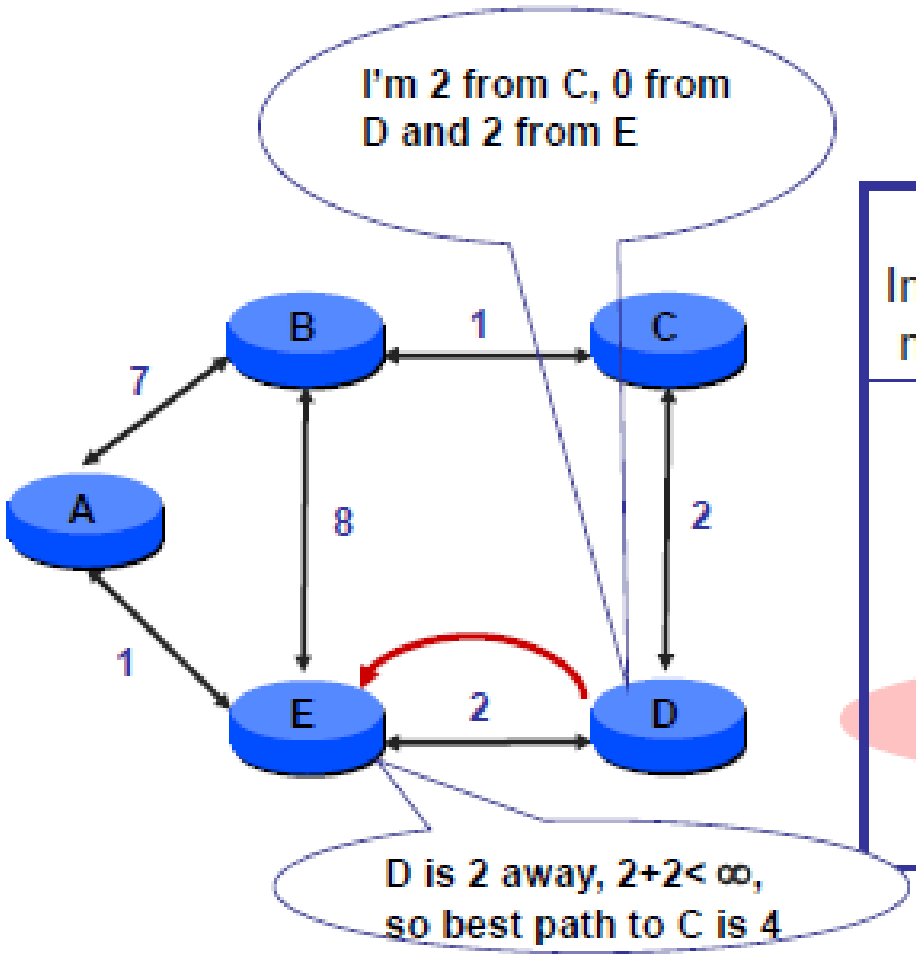
- $c(x,v)$ = cost for direct link from x to v
  - Node x maintains costs of direct links $c(x,v)$
- $D_x(y)$ = estimate of least cost from x to y
  - Node x maintains distance vector $\mathbf{D}_x = [D_x(y): y \in N]$
- Node x maintains its neighbors' distance vectors
  - For each neighbor v, x maintains $\mathbf{D}_v = [D_v(y): y \in N]$
- Each node v periodically sends D_v to its neighbors
  - And neighbors update their own distance vectors
  - $D_x(y) \leftarrow \min_v\{c(x,v) + D_v(y)\}$ for each node $y \in N$

Example – Intial State



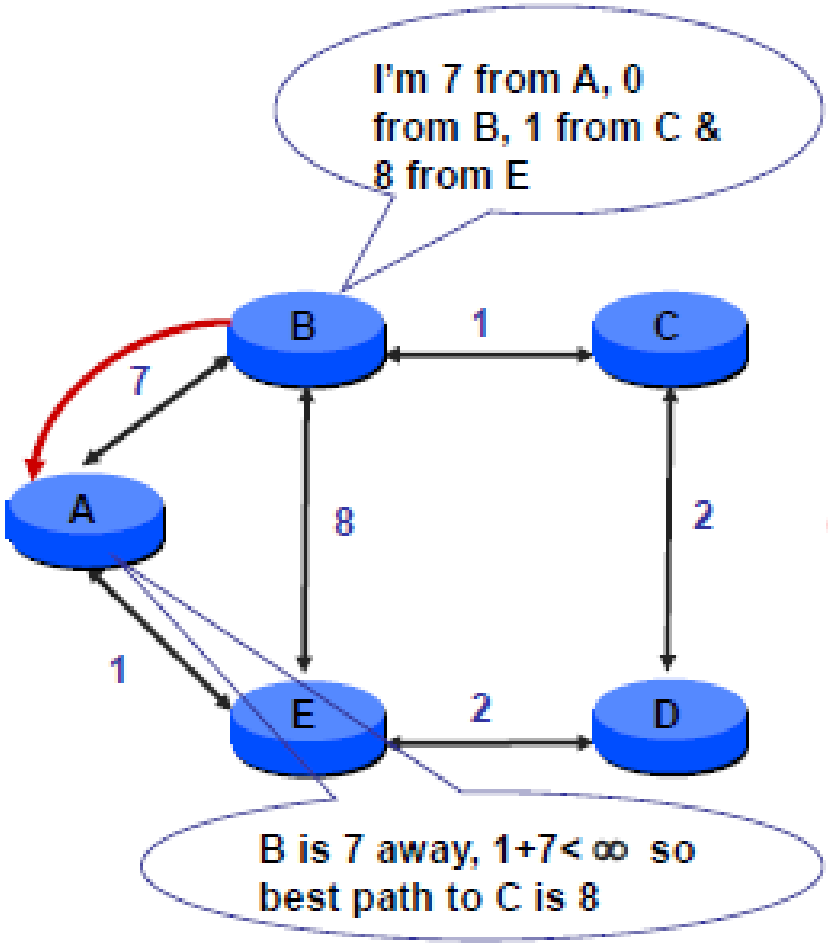
Info at node	Distance to Node				
	A	**B**	**C**	**D**	**E**
A	0	7	∞	∞	1
B	7	0	1	∞	8
C	∞	1	0	2	∞
D	∞	∞	2	0	2
E	1	8	∞	2	0

D sends vector to E



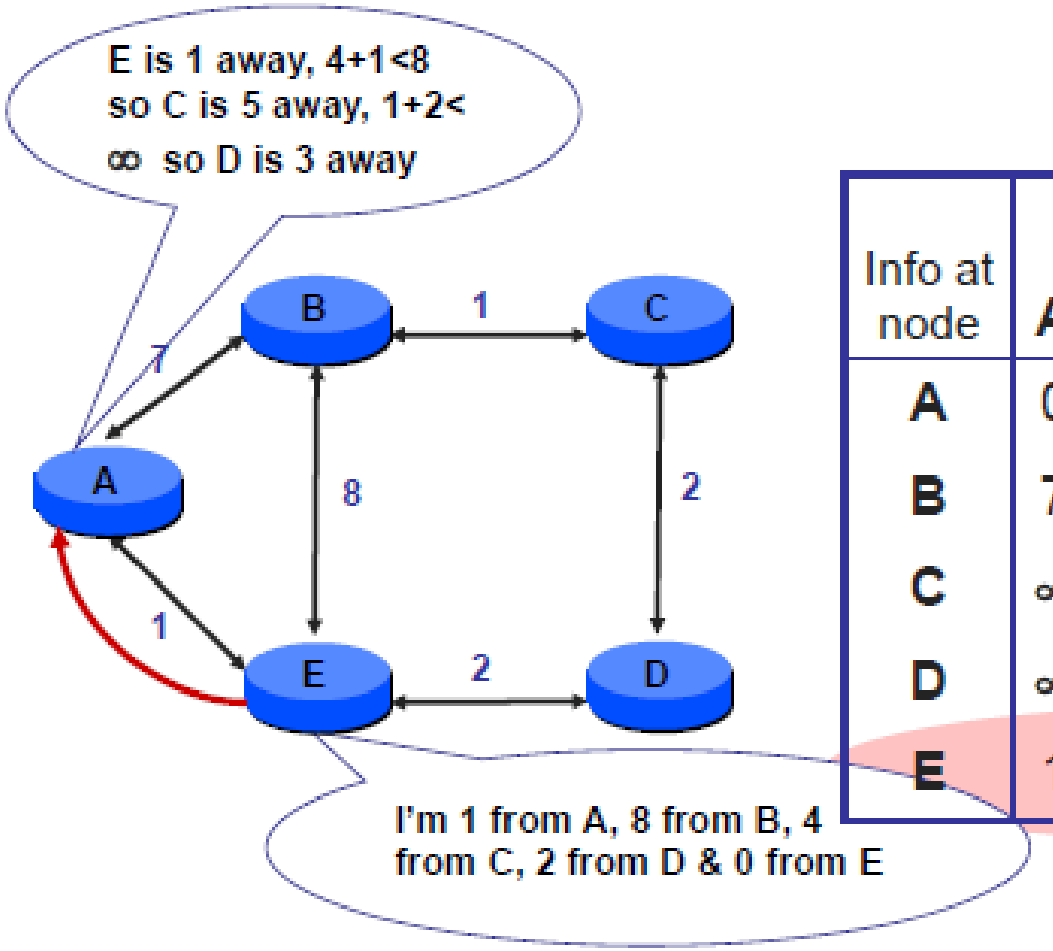
Info at node	Distance to Node				
	A	B	C	D	E
A	0	7	∞	∞	1
B	7	0	1	∞	8
C	∞	1	0	2	∞
D	∞	∞	2	0	2
E	1	8	**4**	2	0

B sends vector to A



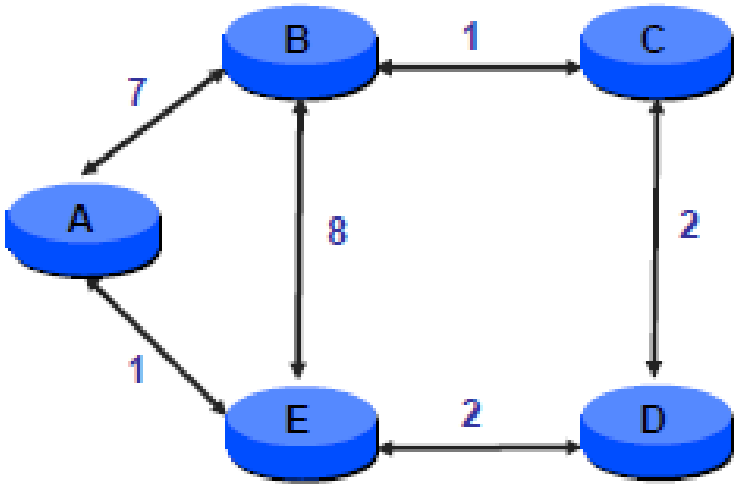
Info at node	Distance to Node A	B	C	D	E
A	0	7	8	∞	1
B	7	0	1	∞	8
C	∞	1	0	2	∞
D	∞	∞	2	0	2
E	1	8	4	2	0

E sends vector to A



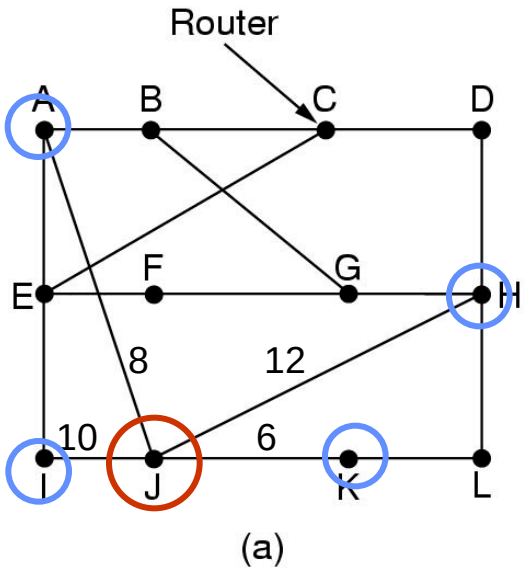
Info at node	Distance to Node A	B	C	D	E
A	0	7	5	3	1
B	7	0	1	∞	8
C	∞	1	0	2	∞
D	∞	∞	2	0	2
E	1	8	4	2	0

And so on...until convergence

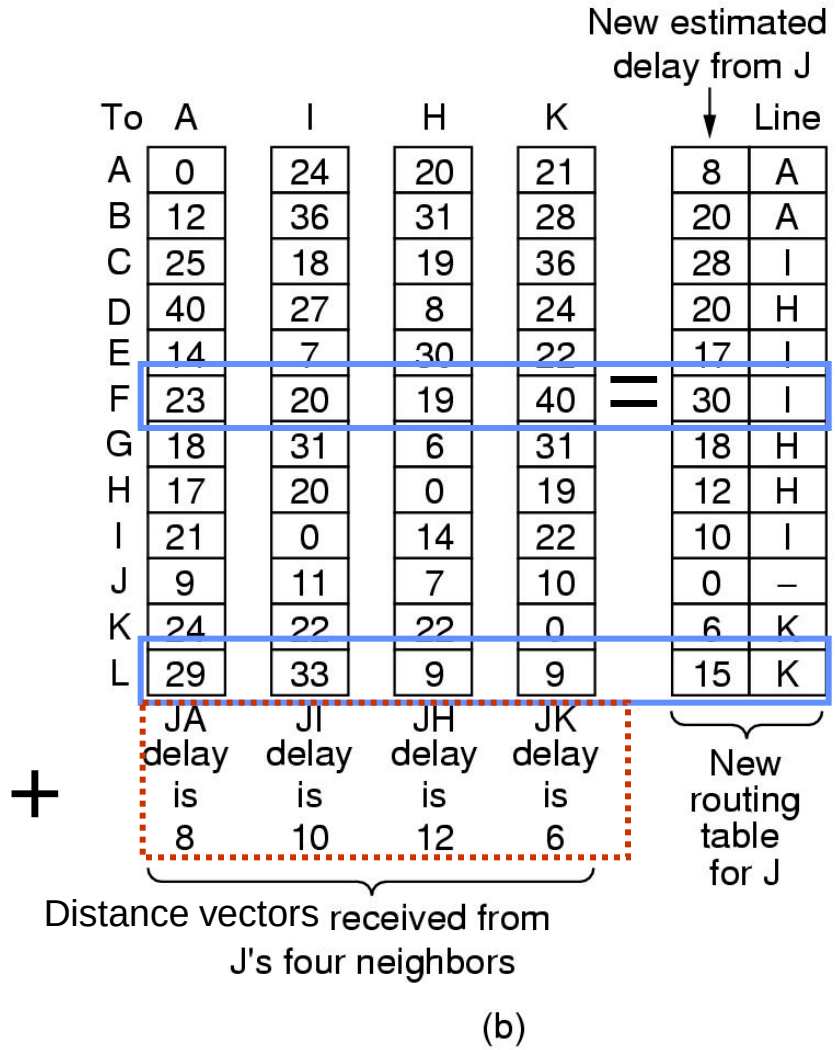


Info at node	Distance to Node				
	A	B	C	D	E
A	0	6	5	3	1
B	6	0	1	3	5
C	5	1	0	2	4
D	3	3	2	0	2
E	1	5	4	2	0

In ”distance vector” each node sends its distance vector to all its neighbors



(a)



(b)

Outline

- Introduction: terminology, concepts
- Hierarchical routing,
  - Intra-domain, Inter-domain
  - Choice of routing protocol, characteristics
  - RIP, OSPF
- OSPF vs IS-IS
  - Similarities, terminology, transport
  - Packet encoding, Area Architecture, Database granularity, Neighbor Adjacency Establishment, AAA, MPLS-TE support
  - For Service Providers

Distance vector vs. link state properties

	RIPv2	OSPF
Type	distance-vector	link-state
Convergence time	slow	fast
VLSM	yes	yes
Bandwidth consumption	high	low
Resource consumption	low	high
Multi-path support	no	yes
Scales well	no	yes
Proprietary	no	no

VLSM – variable length subnet mask

RIPv2 (Route Information Protocol)

- Distributed with BSD (Berkley Software Distribution) UNIX
- Use distance-vector algorithm with updates
  - exchanged every 30. seconds
  - when routing table is changed
  - max 25 routing entries
- For small and medium size networks
  - based on hop-count, 16= unlimited
  - link cost = 1
  - Long convergence time
- Supports authentication
- Supports CIDR, VLSM
- Support more IP address families
- RFC2453

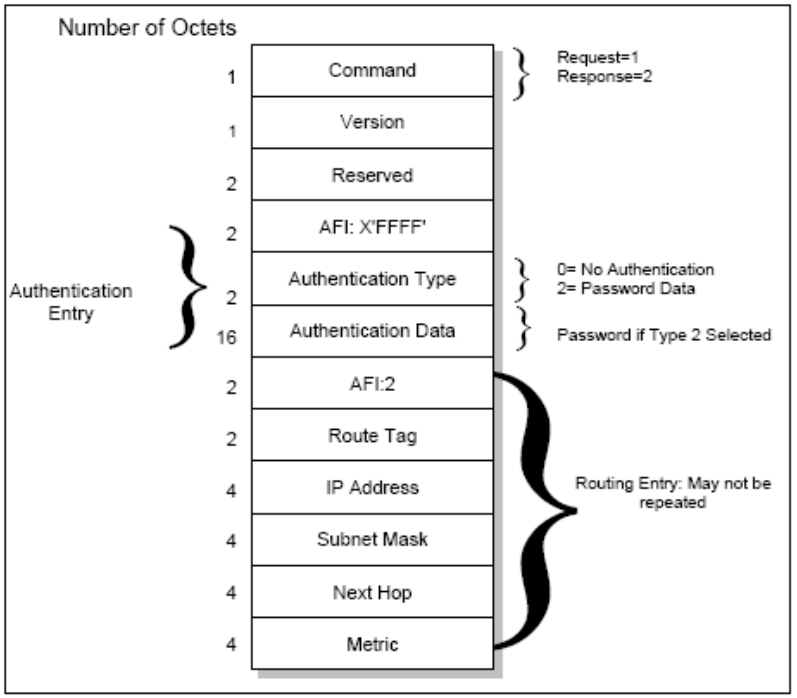
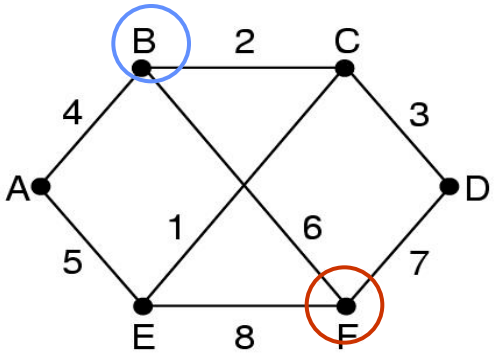


Figure 5-9 RIP-2 packet format

In “link state” each node advertises the state of directly connected links to all nodes

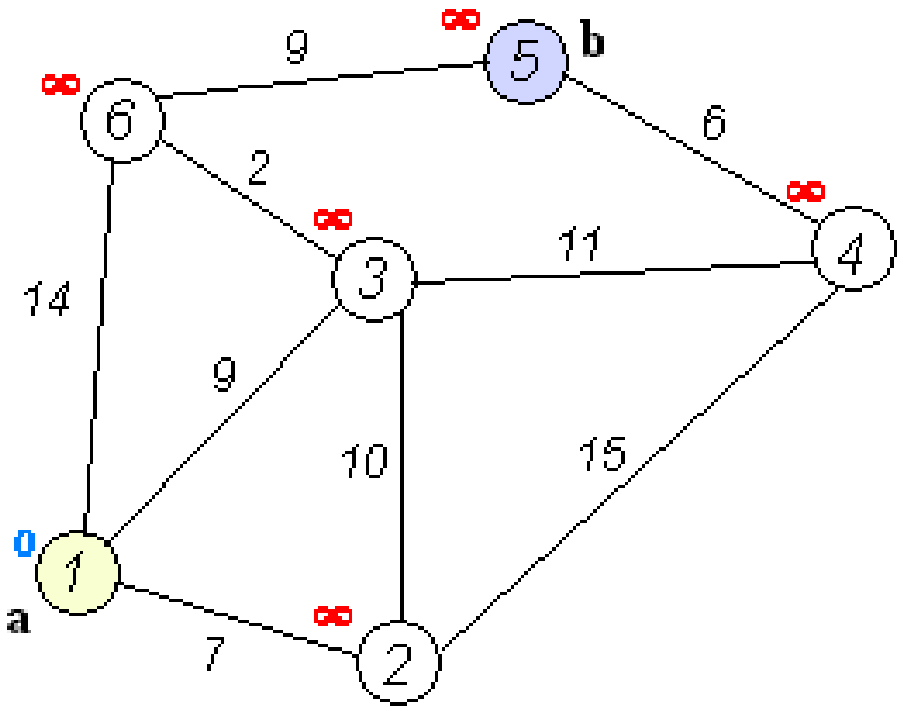
- **Each node has a complete map of the topology**
- **Link state announcement**
  - **id of source node**
  - **cost of each directly connected link**
- **Reliable “flooding”**
  - **Sequence number (SEQNO)**
  - **“time-to-live” (TTL) for this packet**
- **Calculate routing table based on link state database (Dijkstras Shortest Path First)**



A		Link: B		State: C		D		Packets: E		F	
Seq.		Seq.		Seq.		Seq.		Seq.		Seq.	
Age		Age		Age		Age		Age		Age	
B	4	A	4	B	2	C	3	A	5	B	6
E	5	C	2	D	3	F	7	C	1	D	7
		F	6	E	1			F	8	E	8

Djikstra

- **Build a «tree» from the starting point**



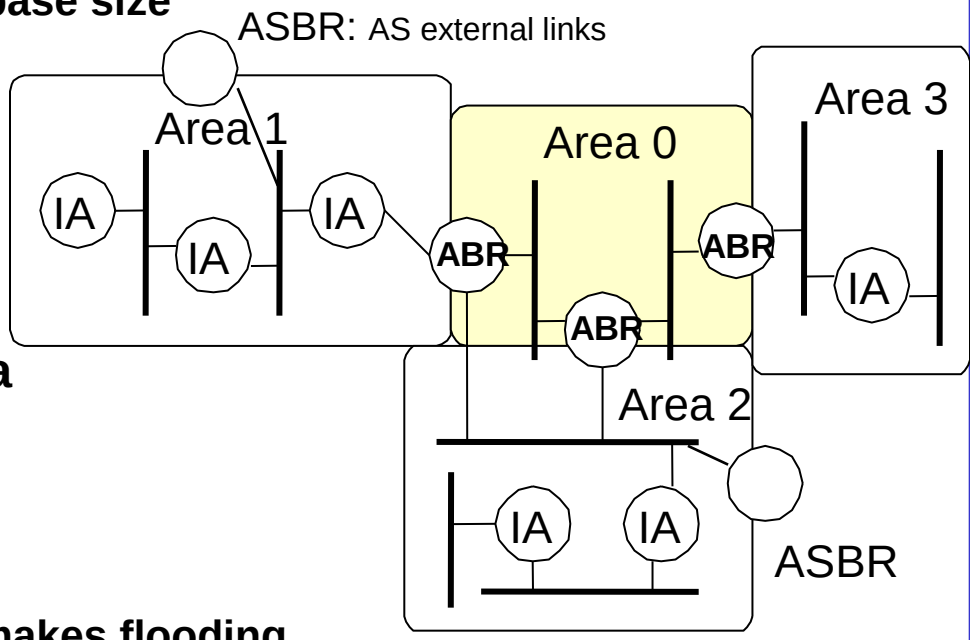
OSPF (Open Shortest Path First) is the most important link state protocol

- Based on ”link-state” algorithm
- Mechanisms for load sharing
  - Equal cost load balancing
- Network hierarchy
  - Areas - router need to know how to get to correct area
  - Stub areas – to reduce the size of the link state database within the area
- Mechanisms for authentication
  - 8-byte password
- Faster convergence time – instantaneous propagation of topology changes
- Support CIDR and VLSM for efficient IP address allocation
- Extendable: multicast, QoS, address resolution etc
- RFC2328

http://www.hurgh.org/articles.php?article=ospf_tutorial

Hierarchy: OSPF nodes and routing areas

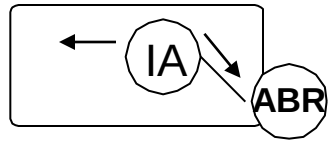
- **Area: additional hierarchy = set of routers configured to exchange routing info**
  - **reduce link state (topology) database size**
  - **reduce CPU processing**
  - **limit number of link state updates**
  - **backbone area – area 0**
- **Intra area routers (IA) exchange routing info within an OSPF area**
- **Area border routers (ABR) connect to two or more areas**
  - **topology DB per area**
  - **summarizes route information - makes flooding and route calculation more scalable**
- **AS boundary routers (ASBR) exchange reachability with other routing domains**



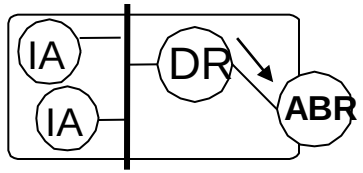
The link state database in each node is built from flooding the set of link state advertisements (LSAs)

- An LSA describes an individual network component (router, segment or external destination)

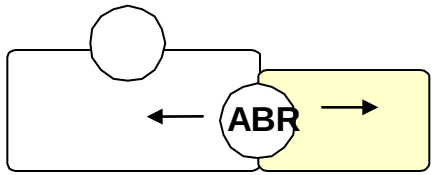
- **Router LSAs** generated by every OSPF router describes state of the router's interfaces (links) within the area. Flooded throughout the area.



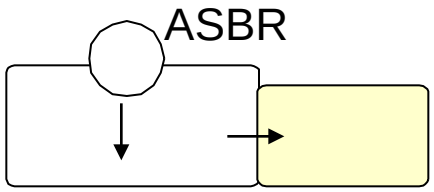
- **Network LSAs** generated by the DR (designated router) list the routers connected to a multi-access network. Flooded throughout the area.



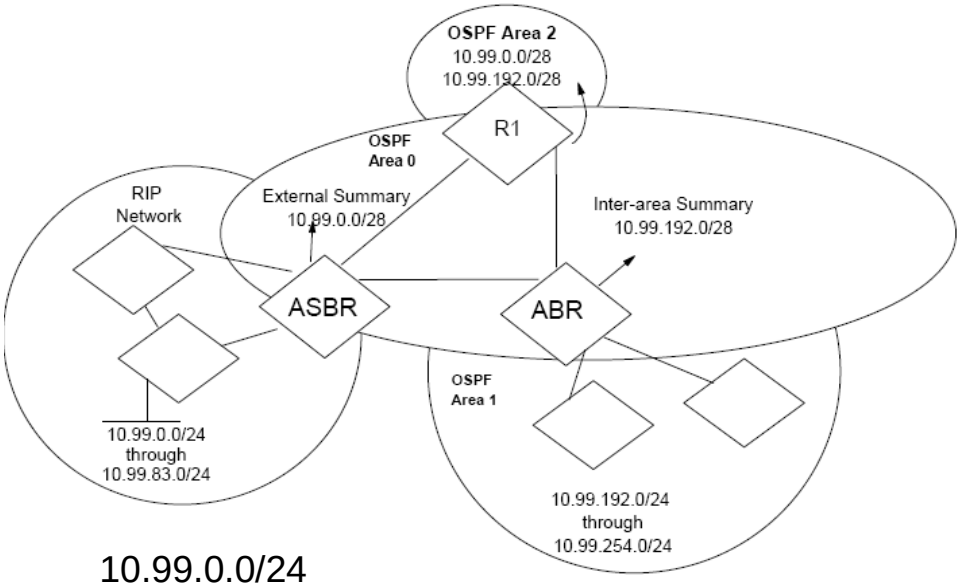
- **Summary LSAs** (Type-3 and Type-4) generated by ABR (area border router) are used to exchange reachability information between areas
  - Type-3 describes routes to destinations in other areas within the OSPF network (inter-area destinations)
  - Type-4 describes routes to ASBRs



- **AS external LSAs** generated by ASBR (AS boundary router) describe routes to destinations external to the OSPF network. Flooded throughout all areas in the OSPF network.



OSPF route summarization consolidates routing entries into a single advertisement



10.99.0.0/24
through
10.99.**6**3.0/24

10.99. 0000 0000
10.99. 0011 1111

64 subnets

/18

- **Performed by a border router**
- **Inter-area route summarization**
  - **ABR perform route summarization on route advertisements originating within the area**
  - **Summarized route announced into the backbone area**
  - **The aggregated route is further announced to other areas**
- **External route summarization is performed by ASBR and applies to external routes injected into the OSPF network**

Outline

- **Introduction: terminology, concepts**
- **Hierarchical routing,**
  - **Intra-domain, Inter-domain**
  - **Choice of routing protocol, characteristics**
  - **RIP, OSPF**
- **OSPF vs IS-IS**
  - **Similarities, terminology, transport**
  - **Packet encoding, Area Architecture, Database granularity, Neighbor Adjacency Establishment, AAA, MPLS-TE support**
  - **For Service Providers**

Comparing ISIS and OSPF

- **Both are Link State Routing Protocols using the Dijkstra SPF Algorithm**
- **So what's the difference then?**
- **And why do ISP engineers end up arguing so much about which is superior?**

IS-IS & OSPF: Similarities

- **Both are Interior Gateway Protocols (IGP)**
  - **IGP = Intra-domain**
  - **They distribute routing information between routers belonging to a single Autonomous System (AS)**
- **With support for:**
  - **Classless Inter-Domain Routing (CIDR)**
  - **Variable Subnet Length Masking (VLSM)**
  - **Authentication**
  - **Multi-path**
  - **IP unnumbered links (useful for dynamically created interface, such as Virtual-Access based on Virtual-Templates, routing across PPPoE & PPPoA interfaces)**

IS-IS and OSPF Terminology

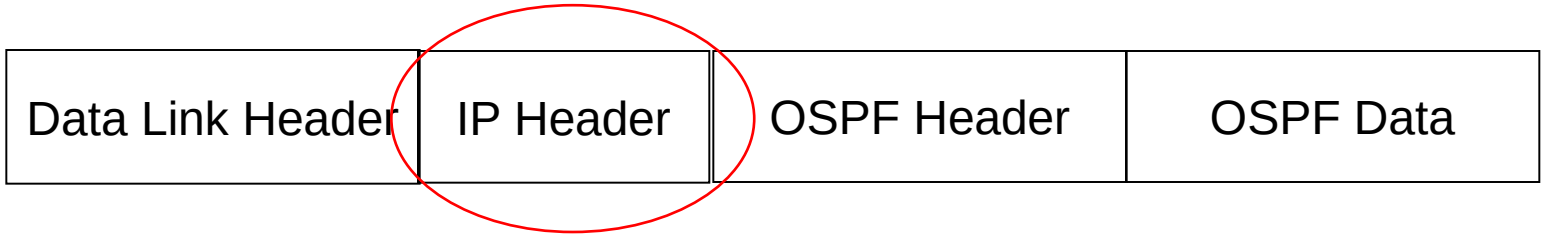
OSPF	ISIS
Host	End System (ES)
Router	Intermediate System (IS)
Link	Circuit
Packet	Protocol Data Unit (PDU)
Designated router (DR)	Designated IS (DIS)
Backup DR (BDR)	N/A (no BDIS is used)
Link-State Advertisement (LSA)	Link-State PDU (LSP)
Hello packet	IIH PDU
Database Description (DBD)	Complete sequence number PDU (CSNP)

IS-IS and OSPF Terminology (Cont.)

OSPF	ISIS
Area	Sub domain (area)
Non-backbone area	Level-1 area
Backbone area	Level-2 Sub domain (backbone)
Area Border Router (ABR)	L1L2 router
Autonomous System Boundary Router (ASBR)	Any IS

Transport

- **OSPF uses IP Protocol 89 as transport**



- **IS-IS is directly encapsulated in Layer 2**

Outline

- **Introduction: terminology, concepts**
- **Hierarchical routing,**
  - **Intra-domain, Inter-domain**
  - **Choice of routing protocol, characteristics**
  - **RIP, OSPF**
- **OSPF vs IS-IS**
  - **Similarities, terminology, transport**
  - **Packet encoding, Area Architecture, Database granularity, Neighbor Adjacency Establishment, AAA, MPLS-TE support**
  - **For Service Providers**

OSPF vs IS-IS

- **Packet Encoding**
  - **OSPF is "efficiently" encoded**
    - **Positional fields**
    - **Holy 32-bit alignment provides tidy packet pictures, but not much else**
    - **Only LSAs are extensible (not Hellos, etc.)**
    - **Unrecognized LSA types not flooded (though opaque LSAs can suffice, if implemented universally, and IS-IS-like encoding can provide good granularity)**
  - **IS-IS is mostly Type-Length-Value encoded**
    - **No particular alignment**
    - **Extensible from the start (unknown types ignored but still flooded)**
    - **All packet types are extensible**
    - **Nested TLVs provide structure for more granular extension (though base spec does not use them; OSPF is starting to do so)**

OSPF vs IS-IS

- **Area Architecture**
  - **Both protocols support two-level hierarchy of areas (to reduce SPF graph complexity, and potentially to allow route aggregation)**
  - **OSPF area boundaries fall within a router**
    - **Interfaces bound to areas**
    - **Router may be in many areas**
    - **Router must calculate SPF per area**
  - **IS-IS area boundaries fall on links**
    - **Router is in only one area, plus perhaps the L2 backbone (area)**
    - **Biased toward large areas, area migration**
    - **Requires router per area (unless multiple virtual routers are implemented)**
    - **Historically proven somewhat difficult for users to grasp**
    - **Little or no multilevel deployment (large flat areas work so far)**

OSPF vs IS-IS

- **Database Granularity**

  - **OSPF database node is an LSAdvertisement**
    - **LSAs are mostly numerous and small (one external per LSA, one summary per LSA)**
    - **Network and Router LSAs can become large**
    - **LSAs grouped into LSUpdates during flooding**
    - **LSUpdates are built individually at each hop**
    - **Small changes can yield small packets (but Router, Network LSAs can be large)**

  - **IS-IS database node is an LSPacket**
    - **LSPs are clumps of topology information organized by the originating router**
    - **Always flooded intact, unchanged across all flooding hops (so LSP MTU is an architectural constant--it must fit across all links)**
    - **Small topology changes always yield entire LSPs (though packet size turns out to be much less of an issue than packet count)**
    - **Implementations can attempt clever packing**

OSPF vs IS-IS

- **Neighbor Establishment**
  - **Both protocols use periodic multicast Hello packets, "I heard you" mechanism to establish 2-way communication**
  - **Both protocols have settable hello/holding timers to allow tradeoff between stability, overhead, and responsiveness**
  - **OSPF requires hello and holding timers to match on all routers on the same subnet (side effect of DR election algorithm) making it difficult to change timers without disruption|**
  - **IS-IS requires padding of Hello packets to full MTU size under some conditions (to detect media with MTUs smaller than 1497 bytes). This has been deprecated in more recent implementations**
  - **OSPF requires routers to have matching MTUs in order to become adjacent (or LSA flooding may fail, since LSUpdates are built at each hop and may be MTU-sized)**

OSPF vs IS-IS

- **Neighbor Adjacency Establishment**
  - **The goal is to synchronize databases**
  - **The method is tell your neighbor everything you've got**
  - **You (or your neighbor) will figure out what you're missing and make sure that you get it**
  - **Each protocol's approach is driven by database granularity**
  - **OSPF uses complex, multistate process to synchronize databases between neighbors**
    - **Intended to minimize transient routing problems by ensuring that a newborn router has nearly complete routing information before it begins carrying traffic**
    - **Accounts for a significant portion of OSPF's implementation complexity**
    - **Partially a side effect of granular database (requires many DBD packets)**

OSPF vs IS-IS

- **Authentication and Security**
  - Both support cryptographic authentication
  - OSPF really needs this (packet bombs)
  - Successful IGP attacks will be catastrophic (or worse, subtle)
  - Use packet filtering, particularly with OSPF

- **MPLS Traffic Engineering extensions**
  - Protocols carry around TE link information (available bandwidth, link color, etc.) on behalf of MPLS but don't use the data themselves
  - TE functionality is identical for the two protocols (by design)
  - TE functions are IGP-independent, so mechanisms ought to be identical

Outline

- Introduction: terminology, concepts
- Hierarchical routing,
  - Intra-domain, Inter-domain
  - Choice of routing protocol, characteristics
  - RIP, OSPF
- OSPF vs IS-IS
  - Similarities, terminology, transport
  - Packet encoding, Area Architecture, Database granularity, Neighbor Adjacency Establishment, AAA, MPLS-TE support
  - For Service Providers

For Service Providers

- **Which IGP should an ISP choose?**
  - **Both OSPF and ISIS use Dijkstra SPF algorithm**
  - **Exhibit same convergence properties**
  - **ISIS less widely implemented on router platforms (NOT TRUE ANYMORE)**
  - **ISIS runs on data link layer, OSPF runs on IP layer**

- **Biggest ISPs tend to use ISIS – why?**
  - **In early 90s, Cisco implementation of ISIS was much more solid than OSPF implementation – ISPs naturally preferred ISIS**
  - **Main ISIS implementations more tuneable than equivalent OSPF implementations – because biggest ISPs using ISIS put more pressure on Cisco to implement "knobs"**

For Service Providers

- Moving forward a decade
  - Early OSPF implementation substantially rewritten
    - Became competitive with ISIS in features and performance
  - Router vendors wishing a slice of the core market need an ISIS implementation as solid and as flexible as that from Cisco
    - Those with ISIS & OSPF support tend to ensure they exhibit performance and feature parity

- OSPF
  - Rigid area design – all networks must have area 0 core, with sub-areas distributed around
  - Suits ISPs with central high speed core network linking regional PoPs
  - Teaches good routing protocol design practices

- ISIS
  - Relaxed two level design – L2 routers must be linked through the backbone
  - Suits ISPs with “stringy” networks, diverse infrastructure, etc, not fitting central core model of OSPF
  - More flexible than OSPF, but easier to make mistakes too