Routing
in
Switched Networks

Reference:
Chapter 12 -Stallings

Lecture 10 Revision

- Question - Give some reasons for using fragmentation and reassembly.

Lecture 10 Revision

- Question - Give some reasons for using fragmentation and reassembly.
- Answer - (1) The communications network may only accept blocks of data up to a certain size. (2) Error control may be more efficient with smaller PDU size. With smaller PDU's, fewer bits need to be retransmitted when a PDU suffers an error. (3) More equitable access to shared transmission facilities, with shorter delay, can be provided. (4) A smaller PDU size may mean that receiving entities can allocate smaller buffers. (5) An entity may require that data transfer comes to some sort of closure from time to time, for checkpoint and restart/recovery operations.

Lecture 10 Revision

- Question - List the requirements for an internetworking facility.

Lecture 10 Revision

- Question - List the requirements for an internetworking facility.
- Answer - (1) Provide a link between networks. At minimum, a physical and link control connection is needed. (2) Provide for the routing and delivery of data between processes on different networks. (3) Provide an accounting service that keeps track of the use of various networks and routers and maintains status information. (4) Provide the services just listed in such a way as not to require modifications to the networking architecture of any of the constituent networks. This means that the internetworking facility must accommodate a number of differences among networks.

Lecture 10 Revision

- Question - What are the pros and cons of limiting reassembly to the end point as compared to allowing en route reassembly.

Lecture 10 Revision

- Question - What are the pros and cons of limiting reassembly to the end point as compared to allowing en route reassembly.
- Answer - If intermediate reassembly is not allowed, the datagram must eventually be fragmented to the smallest allowable size along the route. Once the datagram has passed through the network that imposes the smallest-size restriction, the fragments may be unnecessarily small for later networks, degrading performance. On the other hand, intermediate reassembly requires compute and buffer resources at the intermediate routers. Furthermore, all fragments of a given original datagram would have to pass through the same intermediate node for reassembly, which would prohibit dynamic routing.

Lecture 10 Revision

- Question - Explain the function of the three flags in the IPv4 header.

Lecture 10 Revision

- Question - Explain the function of the three flags in the IPv4 header.
- Answer - The **More bit** is used for fragmentation and reassembly. If this bit is 0, then either there has been no fragmentation of this packet or this is the last fragment. If this bit is 1, then this packet has been fragmented and this is not the last fragment. The **Don't Fragment bit** prohibits fragmentation when set.

Introduction

- A key design issue in switched networks is that of routing
  - These networks include circuit switching, packet switching, frame relay, and ATM networks
- In general terms, the routing function seeks to design routes through the network for individual pairs of communicating end nodes such that network is used efficiently

Routing in Circuit Switching Networks

- In a large circuit-switching network, many of the circuit connections will require a path through more than one switch
- When a call is placed, the network must devise a route through the network from calling subscriber to called subscriber
  - This route passes through some number of switches and trunks
- There are 2 main requirements for the network's architecture that bear on the routing strategy:

Routing in Circuit Switching Networks Contd.

- Efficiency
  - It is desirable to minimise the amount of equipment (switches and trunks) in the network, subject to the ability to handle that amount of load
  - The load requirement is usually expressed in terms of a busy-hour traffic load
    - This is simply the average load expected over the course of the busiest hour of use during the course of a day
  - From a functional point of view, it is necessary to handle that amount of load
  - From a cost point of view, we would like to handle that load with minimum

Routing in Circuit Switching Networks Contd.

- Resilience
  - Although a network may be sized for the busy hour load, it is possible for the traffic to surge temporarily above that level (during a major storm)
  - It will also be the case that, from time to time, switches and trunks will fail and be temporarily unavailable
  - We would like the network to provide a reasonable level of service under such conditions

- The key design issue that determines the nature of the tradeoff between efficiency and resilience is the routing strategy

Routing in Circuit Switching Networks Contd.

- Traditionally, the routing function in public telecommunications networks has been quite simple
  - In essence, the switches of a network were organised into a tree structure, or hierarchy.
  - A path is constructed by starting at the calling subscriber, tracing up the tree to the first common node, and then tracing down the tree to the called subscriber
  - To add some resilience to the network, additional high-usage trunks were added that cut across the tree structure to connect exchanges with high volumes of traffic between them

Routing in Circuit Switching Networks Contd.

- In general, the above mentioned is a static approach
  - As the routing scheme is not able to adapt to changing conditions, the network must be designed to meet some typical heavy demands
- It is difficult to analyse varying demands, which leads to oversizing and therefore inefficiency
- In terms of resilience, the fixed hierarchical structure with supplemental trunks any respond poorly to failures

Routing in Circuit Switching Networks Contd.

- To cope with the growing demand on public telecommunications networks, virtually all providers have moved away from static hierarchical approach to a dynamic approach
- A dynamic routing approach is one in which routing decisions are influenced by current traffic conditions
- Typically, the circuit switching nodes have a peer relationship with each other rather than a hierarchical one i.e. All nodes are capable of performing the same function

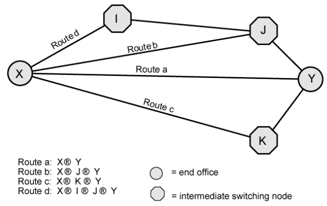
Routing in Circuit Switching Networks Contd.

- In such an architecture, routing is both more complex and more flexible
  - It is more complex because the architecture does not provide a ‘natural’ path or set of paths based on hierarchical structure
  - It is more flexible because more alternative routes are available
- A form of routing used in circuit-switching networks is known as alternate routing
  - In this routing scheme, the possible routes to be used between two end offices are predefined

Routing in Circuit Switching Networks Contd.

- It is the responsibility of the originating switch to select the appropriate route for each call
- Each switch is given a set of preplanned routes for each destination, in order of preference
  - If a direct trunk connection exists between two switches, this is usually the preferred choice
  - If this trunk is unavailable, then the 2^{nd} choice is to be tried, and so on
- The routing sequences reflect an analysis based on historical traffic patterns and designed to optimise the use of network resources

Routing in Circuit Switching Networks Contd.



(a) Topology

Time Period	First route	Second route	Third route	Fourth and final route
Morning	a	b	c	d
Afternoon	a	d	b	c
Evening	a	d	c	b
Weekend	a	c	b	d

(b) Routing table

Routing in Packet Switching Networks

- The primary function of a packet-switching network is to accept packets from a source station and deliver them to a destination station
- To accomplish this, a path or route through the network must be determined
  - Generally more than one route is possible
    - Thus, a routing function must be performed
    - The requirements for this function include
      - Correctness, simplicity, fairness, optimality, robustness, stability, and efficiency

Routing in Packet Switching Networks Contd.

» Robustness has to do with the ability of the network to deliver packets via some route in the face of localised failures and overloads

» The designer who seeks robustness must cope with competing requirements for stability

» A tradeoff also exists between fairness and optimality;

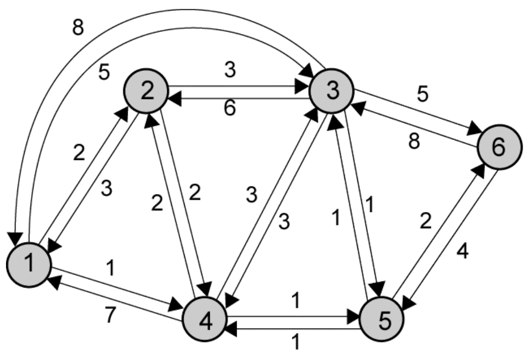
» Some performance criteria may give higher priority to the exchange of packets between nearby stations compared to an exchange between distant stations

» Routing techniques involves some processing overhead at each node and often a transmission overhead as well, both of which impair network efficiency

Routing in Packet Switching Networks Contd.

- Various design elements that contribute to a routing strategy are:
  - Performance Criteria
    - The selection of a route is generally based on some performance criterion
      - The simplest criterion is to choose the minimum-hop route through the network
      - A generalisation of the minimum-hop criterion is least-cost routing
    - In either the minimum-hop or least-cost approach, the algorithm for determining the optimum route for any pair of stations is relatively straightforward, and the processing time would be about the same for either computation

Routing in Packet Switching Networks Contd.



Routing in Packet Switching Networks Contd.

- Decision Time and Place
  - Two key characteristics of a routing decision are the time and place that the decision is made
  - Decision time is determined by whether the routing decision is made on a packet or virtual circuit basis
    - When the internal operation of the network is datagram, a routing decision is made individually for each packet
    - For internal virtual circuit operation, a routing decision is made at the time the virtual circuit is established
      - » In the simplest case, all subsequent packets using that virtual circuit follow the same route

Routing in Packet Switching Networks Contd.

– The term decision place refers to which node or nodes in the network are responsible for the routing decision

» Most common is distributed routing, in which each node has the responsibility of selecting an output link for routing packets as they arrive

» For centralised routing, the decision is made by some designated node, such as a network control centre

» A third alternative, used in some networks, is source routing

- The decision time and decison place are independent design variables

Routing in Packet Switching Networks Contd.

– Network Information Source and Update Timing

- Most routing strategies require that decisions be based on knowledge of the topology of the network, traffic load, and link cost
- However, some strategies use no such information and yet manage to get packets through
- With distributed routing, in which the routing decision is made by each node, the individual node may make use of only local information from adjacent nodes, such as the amount of congestion experienced at that node
- In the case of centralised routing, the central node typically makes use of information obtained from all nodes

Routing in Packet Switching Networks Contd.

- The concept of information update timing is a function of both the information source and the routing strategy
- If no information is used, there is no information to update
- If only local information is used, the update is essentially continuous
  - That is, an individual node always knows its local condition
- For all other information source categories, update timing depends on the routing strategy
  - For a fixed strategy, the information is never updated
  - For an adaptive strategy, information is updated from time to time to enable the routing decision to adapt to changing conditions

Routing in Packet Switching Networks Contd.

- A large number of routing strategies have evolved for dealing with the routing requirements of packet-switching networks
- Many of these strategies are also applied to internetwork routing
- Four key routing strategies are:

– Fixed Routing

- A single permanent route is configured for each source-destination pair of nodes in the network

Routing in Packet Switching Networks Contd.

- The routes are fixed, or at least only change when there is a change in the topology of the network
  - Thus, the link costs used in designing routes cannot be based on any dynamic variable such as traffic
  - They could, however, be based on expected traffic or capacity
- Fixed routing can be implemented using a central routing matrix, to be stored perhaps at a network control centre
  - In a routing matrix, it is not necessary to store the complete route for each possible pair of nodes

» Rather, it is sufficient to know, for each par of nodes, the identity of the first node on the route

Routing in Packet Switching Networks Contd.

CENTRAL ROUTING DIRECTORY

		From Node					
		1	2	3	4	5	6
To Node	1	—	1	5	2	4	5
	2	2	—	5	2	4	5
	3	4	3	—	5	3	5
	4	4	4	5	—	4	5
	5	4	4	5	5	—	5
	6	4	4	5	5	6	—

Node 1 Directory

Destination	Next Node
2	2
3	4
4	4
5	4
6	4

Node 2 Directory

Destination	Next Node
1	1
3	3
4	4
5	4
6	4

Node 3 Directory

Destination	Next Node
1	5
2	5
4	5
5	5
6	5

Node 4 Directory

Destination	Next Node
1	2
2	2
3	5
5	5
6	5

Node 5 Directory

Destination	Next Node
1	4
2	4
3	3
4	4
6	6

Node 6 Directory

Destination	Next Node
1	5
2	5
3	5
4	5
5	5

Routing in Packet Switching Networks Contd.

- From the above routing matrix, routing tables can be developed and stored at each node
  - Each node needs only to store a single column of the routing directory
    - » A node's directory shows the next node to take for each destination
- With fixed routing, there is no difference between routing for datagrams and virtual circuits
  - All packets from a given source to a given destination follow the same route
- The advantage of fixed routing is its simplicity, and it should work well in a reliable network with stable load
- Its disadvantage is its lack of flexibility – does not react to network congestion or failures

Routing in Packet Switching Networks Contd.

– Flooding

- A packet is sent by a source node to every one of its neighbours
- At each node, an incoming packet is retransmitted on all outgoing links except for the link on which it arrived
- Eventually, a number of copies of the packets will arrive at the destination
- The packet must have a some unique identifier so that the destination node knows to discard all but the first copy
- This technique requires no network information whatsoever
- Unless something is done to stop the incessant retransmission of packets, the number of packets in circulation grows without bound

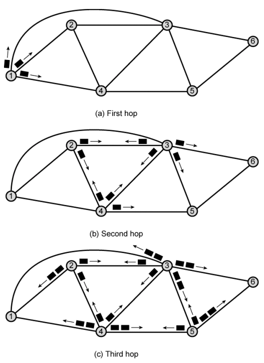
Routing in Packet Switching Networks Contd.

- One way to prevent the above situation is for each node to remember the identity of those packets it has already transmitted
  - When a duplicate copies arrive they are discarded
- A simpler technique is to include a hop count field with each packet
  - The count originally be set to some maximum value, such as the diameter of the network
  - Each time a node passes on a packet, it decrements the count by one
    - » When the count reaches zero, the packet is discarded
- The flooding technique has 3 remarkable properties:

Routing in Packet Switching Networks Contd.

  - All possible routes between source and destination are tried. If at least one path between source and destination exists, a packet will always get through
  - Because all routes are tried, at least one copy of the packet arrive at the destination will have used a minimum-hop route
  - All nodes are directly or indirectly connected to the source node are visited
- Because the flooding technique is highly robust, it could be used to send emergency messages
- Flooding can also be useful for the dissemination of important information to all nodes
- The principal disadvantage of flooding is the high traffic load that it generates, which is directly proportional to the connectivity of the network

Routing in Packet Switching Networks Contd.



(a) First hop

(b) Second hop

(c) Third hop

Routing in Packet Switching Networks Contd.

– Random Routing

- A node selects only one outgoing path for retransmission of an incoming packet
- The outgoing link is chosen at random, excluding the link on which the packet arrived
- If all links are likely to be chosen, then a node may simply utilise outgoing links in a round-robin fashion
- Random routing has the simplicity and robustness of flooding, with far less traffic load

Routing in Packet Switching Networks Contd.

- A refinement of the above technique is to assign a probability to each outgoing link and to select the link based on that probability
  - The probability could be based on data rate, in which case we have

$$P_i = \frac{R_i}{\sum R_i}$$

Where P_i = probability of selecting link I
R_i = data rate on link I

- Like flooding, random routing requires the use of no network information
- Because the route taken is random, the actual route will typically not be the least cost route nor the minimum-hop route

Routing in Packet Switching Networks Contd.

– Adaptive Routing

- The routing decisions that are made change as conditions on the network change
  - The conditions that influence routing decisions are:
    - » Failure
    - » Congestion
- In virtually all packet-switching networks, some sort of adaptive routing techniques is used
- For adaptive routing to be possible, information about the state of the network must be exchanged among the nodes

Routing in Packet Switching Networks Contd.

- The drawbacks associated with adaptive routing, compared to fixed routing are:
  - The routing decisions are more complex
    - » Therefore the processing burden on the network nodes increases
  - In most cases, adaptive strategies depend on status information that is collected at on place but used at another
    - » There is a tradeoff between the quality of the information exchanged and the amount of the overhead
    - » The more information that is exchanged, and the more frequently it is exchanged, the better will be the routing decisions
    - » On the other hand, this information is itself a load on the constituent networks, causing a performance degradation

Routing in Packet Switching Networks Contd.

- An adaptive strategy may react too quickly, causing congestion-producing oscillations, or too slowly, being irrelevant

- Despite the above drawbacks, adaptive routing strategies are by far the most prevalent, for 2 reasons:
  - An adaptive routing strategy can improve performance, as seen by the network user
  - An adaptive routing strategy can aid in congestion control
    - » As it tends to balance loads, adaptive routing can delay onset of severe congestion
- By and large, adaptive routing is an extraordinarily complex task to perform properly

Routing in Packet Switching Networks Contd.

- A convenient way to classify adaptive routing strategies is on the basis of information source:
  - Local
    - » A node routes each packet to the outgoing link with the shortest queue length
    - » This would have the effect of balancing the load on outgoing links; however, some outgoing links may not be headed in the correct general direction
  - Adjacent nodes
  - All nodes
    - » Both strategies, adjacent and all nodes, are commonly used
    - » They take the advantage of information that each node has about delays and outages
    - » Such adaptive strategies can be either distributed or centralised

Routing in Packet Switching Networks Contd.

- First Generation Routing
  - The original routing algorithm, designed in 1969, was a distributed adaptive algorithm using delay as the performance criterion
  - For this algorithm, each node maintains two vectors:

$$D_i = \begin{bmatrix} d_{i1} \\ d_{i2} \\ . \\ . \\ d_{iN} \end{bmatrix} \qquad S_i = \begin{bmatrix} s_{i1} \\ s_{i2} \\ . \\ . \\ s_{iN} \end{bmatrix}$$

  - Where Di = delay vector for node I
    d_{ij} = current estimate of minimum delay from node I to node j
    N = Number of nodes in the network
    S_i = successor node vector for node I
    s_{ij} = the next node in the current minimum-delay route from I to j

Routing in Packet Switching Networks Contd.

- Periodically (every 128ms), each node exchanges its delay vector with all of its neighbours
- On the basis of all incoming delay vectors , a node k updates both of its vectors as follows:

$$d_{kj} = \min_{i \in A} \left[d_{ij} + l_{ki} \right]$$

$s_{kj} = i$ using i that minimises the preceding expression

- Where A = set of neighbour nodes for k
  l_{ki} = current estimate of delay from k to i

Routing in Packet Switching Networks Contd.

- The estimated link delay is simply the queue length for that link
  - Thus, in building a new routing table, the node will tend to favour outgoing links with shorter queues
  - This tends to balance the load on outgoing links
    - However, as the queue lengths vary rapidly with time, the distributed perception of the shortest route could change while a packet is en route
    - This could lead to a thrashing situation in which a packet continues to seek out areas of low congestion rather than aiming at the destination
- The major shortcomings of the above algorithm were:
  - It did not consider line speed, merely queue lengths
    - Thus higher-capacity links were not given the favoured status they deserved

Routing in Packet Switching Networks Contd.

- Queue length is , in any case, an artificial measure of delay, because some variable amount of processing time elapses between the arrival of a packet at a node ad its placement in an outbound queue
- The algorithm was not very accurate
  - » In particular, it responded slowly to congestion and delay increases

- Second Generation Routing
  - The new algorithm was also a distributed one, using delay as the performance criterion, but the difference was significant
    - Rather than using queue length as a surrogate for delay, the delay was measured directly

Routing in Packet Switching Networks Contd.

- At a node, each incoming packet was timestamped with an arrival time; a departure time was recorded when the packet was transmitted
  - » If a positive acknowledgment is returned, the delay for the packet was recorded as the departure time minus the arrival time plus transmission time and propagation delay
  - » The node must therefore know the link data rate and propagation time
  - » If a negative acknowledgement comes back, the departure time is updated and the node tries again, until a measure of successful transmission delay is obtained
- Every 10 seconds, the node computes the average delay on each outgoing link
  - » If there are any significant changes in delay, the information is sent to all other nodes using flooding
  - » Each node maintains an estimate of delay on every network link; when information arrives, it recomputes its routing table

Routing in Packet Switching Networks Contd.

- Experience with this second strategy indicated that it was more responsive and stable than the previous one
- However, as the load on the network grew, a shortcoming in the new strategy began to appear, and it was revised in 1987
- The problem was the assumption that the measured packet delay on a link is a good predictor of the link delay encountered after all nodes reroute their traffic based on this reported delay
  - Thus, it is an effective routing mechanism only if there is some correlation between the reported values and those actually experienced after re-routing
  - This correlation tends to be rather high under light and moderate traffic loads, but there is little correlation under heavy loads

Routing in Packet Switching Networks Contd.

- Therefore, immediately after all nodes have made routing updates, the routing tables are obsolete
- The ARPANET designers concluded that the essence of the problem was that every node was trying to obtain the best route for all destinations, and these efforts conflicted
- It was concluded that under heavy loads, the goal of routing should be to give the average route a good path instead of attempting to give all routes the best path
- The designers decided that it was unnecessary to change the overall routing algorithm
  - Rather, it was sufficient to change the function that calculates link costs

Routing in Packet Switching Networks Contd.

- The calculation begins with measuring the average delay over the last 10 seconds
- The value is then transformed with the following steps:
  - – Using a simple single server queueing model, the measured delay is transformed into an estimate of link utilisation
    - » From queueing theory, utilisation can be expressed as a function of a delay as follows:

$$\rho = \frac{2(T_s - T)}{T_s - 2T}$$

Where ? = link utilisation
T = measured delay
T_s= service time

» The service time was set at network-wide average packet size (600bits) divided by the data rate of the link

Routing in Packet Switching Networks Contd.

– The results was then smoothed by averaging it with the previous estimate of utilisation

U(n+1) = 0.5 * p(n+1) + 0.5* U(n)

Where U(n) = average utilisation calculated at sampling time n
p(n) = link utilisation measured at sampling time

» Averaging increases the period of routing oscillations, thus reducing routing overhead

– The link cost is then set as a function of average utilisation that is designated to provide a reasonable estimate of cost while avoiding oscillations