



Network Routing

Redes de Computadores

2021/22

Pedro Brandão

References

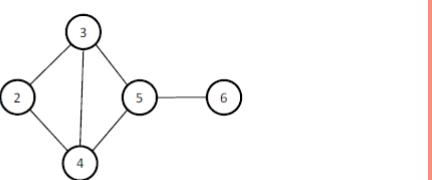
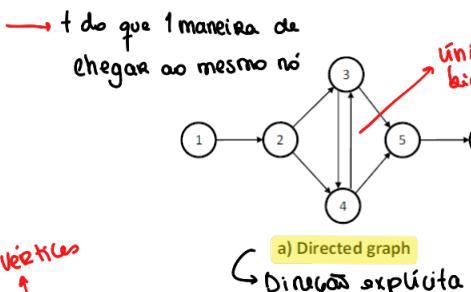
- These slides are from “Computer Networking: A Top Down Approach 5th edition. Jim Kurose, Keith Ross Addison-Wesley, April 2009”
- And “Computer Networks, 5th edition. Andrew S. Tanenbaum, David J. Wetherall, Prentice Hall, 2011”
- With adaptations/additions by Manuel Ricardo and Pedro Brandão

Driving questions...

- What is a graph?
- What is a spanning tree?
- What is a shortest path tree?
- How are paths defined in a network?
- How does the Dijkstra algorithm work?
- How does a link state routing protocol work?
- How does a node learn about neighbours?
- How does the Bellman-Ford algorithm work?
- How does a distance vector work?
- What are the limitations of the layer 2 network of switches?
- How does the IEEE spanning tree protocol work?
- What is the maximum capacity of a flow network?

3

Graph – Directed and Undirected



vertices

$$G = (V, E)$$

$$V = \{v_1, v_2, v_3, v_4, v_5, v_6\}, \quad |V| = 6$$

$$E = \{(v_1, v_2), (v_2, v_3), (v_2, v_4), (v_3, v_4), (v_4, v_3), (v_3, v_5), (v_4, v_5), (v_5, v_6)\}, \quad |E| = 8$$

edges

$$G = (V, E)$$

$$V = \{v_1, v_2, v_3, v_4, v_5, v_6\}, \quad |V| = 6$$

$$E = \{(v_1, v_2), (v_2, v_3), (v_2, v_4), (v_3, v_4), (v_3, v_5), (v_4, v_5), (v_5, v_6)\}, \quad |E| = 7$$

Aparece +1 devido à bidirecionalidade de 3 e 4

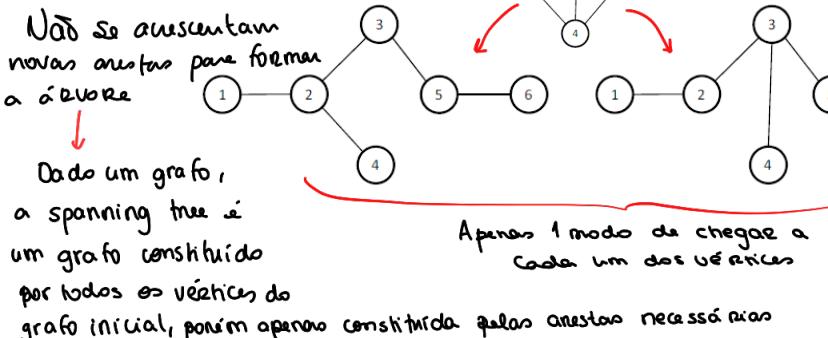
4

Tree

- Trees $T = (V, E)$
 - graph with no cycles
 - number of edges $|E| = |V| - 1$
 - any two vertices of the tree are connected by exactly one path

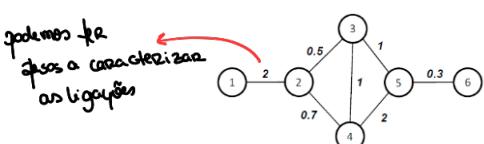
- A tree T is said to span a graph (spanning tree) if:

- Given graph $G = (V, E)$
- $T = (V, E')$ and $E' \subseteq E$



Shortest Path Trees

- Graphs and Trees can be weighted
 - $G = (V, E, w)$
 - $T = (V, E', w)$
 - $w: E \rightarrow \mathbb{R}$ → para uma dada aresta/ligação temos um valor associado
- Total cost of a tree $T \rightarrow C_{total}(T) = \sum_{i=1}^{|E|} w(e_i)$
- Minimum Spanning Tree $T^* \rightarrow C_{total}(T^*) = \min(C_{total}(T))$ → árvore com um custo igual ao menor custo total
 - algorithms used to compute MST: Prism, Kruskal
- Shortest Path Tree (SPT) Rooted at Vertex s
 - tree composed by the union of the shortest paths between s and each of other vertices of G
 - Algorithms used to compute SPT: Dijkstra, Bellman-Ford
- Computer networks use Shortest Path Trees



Routing in Layer 3 Networks

(Routing)

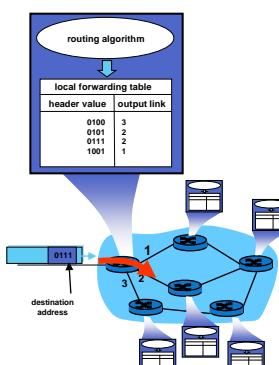
Párote destino X
 Verificar tabela e
 reencaminhar para
 rotear seguinte
 Construção da tabela
 de forwarding. Pode implicar
 saber qual o nó seguinte + favorável da
 conexão todos os caminhos e, portanto,
 saber qual o +
 favorável

Forwarding, Routing

- **Forwarding** → data plane
 - directing packet from input to output link
 - using a forwarding table

- **Routing** → control plane
 - computing paths the packets will follow
 - routers exchange messages
 - each router creates its forwarding table

Vai sendo adaptada à medida que se vão
 recebendo novas informações sobre o
 que são as ligações na rede onde se
 encontram os routers



Importance of Routing

- End-to-end performance

- path affects quality of service

escolha do caminho tem influência na qualidade relativamente à transmissão dos dados

- Use of network resources

- balance traffic over routers and links

Vai, indiretamente, implicar um melhoramento dos parâmetros da rede.

- avoiding congestion by directing traffic to less-loaded links

- Transient disruptions

- failures, maintenance

- limiting packet loss and delay during changes

Caso ocorram falhas e/ou existir uma rápida adaptação das rotas para outra rota redundante

9

Shortest-Path Routing

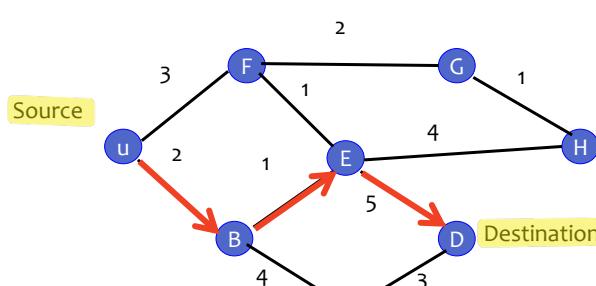
- Path-selection model

- Destination-based

Não tem adaptação ao que seja a largura de banda disponível, delay associado, perdas a acontecer, etc.

- Load-insensitive (e.g., static link weights)

- Minimum hop count or minimum sum of link weights



os outros têm todos um custo maior
shortest path:
 $2+1+5 = 8$

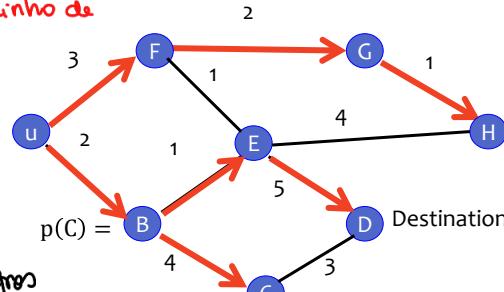
10

Shortest-Path Routing

- Given a network topology with link costs
 - $c(x, y)$ - link cost from node x to node y
 - Infinity (∞) if x and y are not direct neighbours
- Compute the least-cost paths from source u to all nodes
 - $p(v)$ - node predecessor of node v in the path to u

predecessor no caminho de menor custo

Source



Inicialmente, o nó u tem 2 ligações para os vizinhos que custam 3 e 2, para todos os outros tem um custo de 4 pois ele não sabe nenhuma rota ainda.

$$\begin{aligned} p(C) &= B \\ p(E) &= B \\ p(G) &= F \\ p(H) &= G \\ p(E) &= B \\ p(D) &= E \end{aligned}$$

11

Dijkstra's Shortest-Path Algorithm

- Iterative algorithm
 - After k iterations \rightarrow known least-cost paths to k nodes
- $S \rightarrow$ set of nodes for which least-cost path is known
 - Initially, $S=\{u\}$, where u is the source node
 - Add one node to S in each iteration
- $D(v) \rightarrow$ current cost of path from source to node v
 - Initially
 - $D(v)=c(u, v)$ for all nodes v adjacent to u
 - $D(v)=\infty$ for all other nodes v
 - Continually update $D(v)$ when shorter paths are learned

Conforme se vão adicionando nós a S

independente de estarem em S ou não só uma estimativa
inicialmente só conhecemos D nos seus vizinhos
setas as ligações que o nó tem e consegue

12

Dijkstra's Algorithm

```

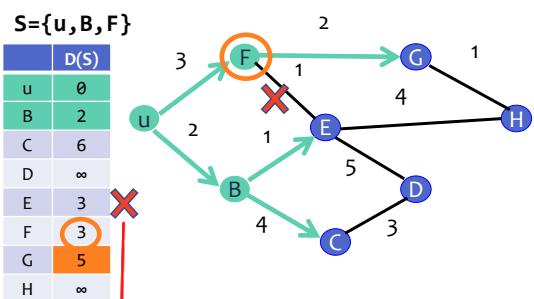
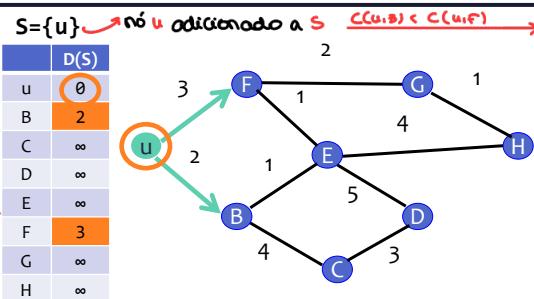
1 Initialization:
2   S = {u}
3   for all nodes v
4     if v adjacent to u {
5       D(v) = c(u,v)
6     }
7     else D(v) = ∞
8 Loop
9   find node w not in S with the smallest D(w)
10  add w to S
11  update D(y) for all v adjacent to w and not in S:
12    D(v) = min{D(v), D(w) + c(w,v)}
13 until all nodes in S
  
```

m^ínimo entre
 o valor que j^a l^e est^a
 . caminho + curto da u a w + c(w,v)

Inicialmente vamos usar
o vizinho com menor peso na ligação

13

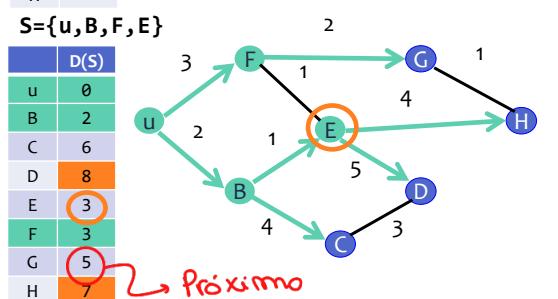
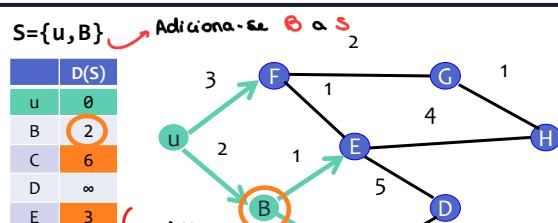
Dijkstra's Algorithm – Example – 1



N^{ão} s^eo vizinhos
a u

RCom 21/22 - Routing

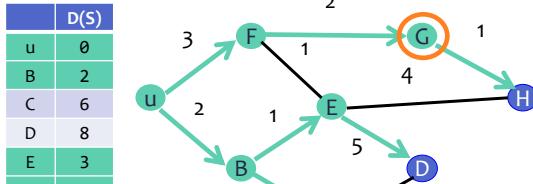
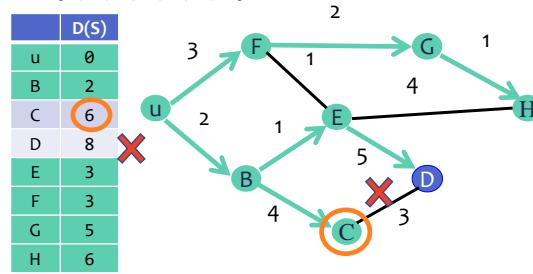
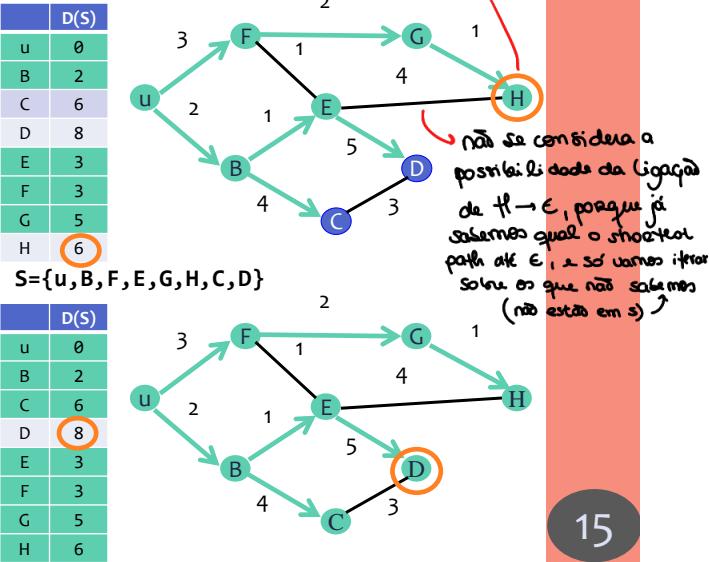
há renhuma melhoria
associada, o m^ínimo era o
que j^a conhecemos



Próximo
Vai ser atualizado o
iterar sobre G

14

Dijkstra's Algorithm – Example – 2

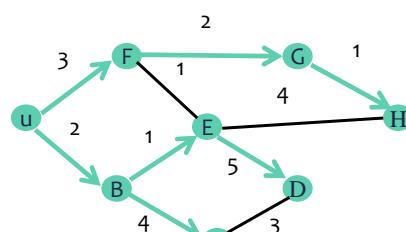
 $S = \{u, B, F, E, G\}$  $S = \{u, B, F, E, G, H, C\}$  $S = \{u, B, F, E, G, H\}$ 

15

Dijkstra's Algorithm – Example – 3

	D(S)	Next Hop
u	0	—
B	2	B
C	6	B
D	8	B
E	3	B
F	3	F
G	5	F
H	6	G

Forwarding table at u



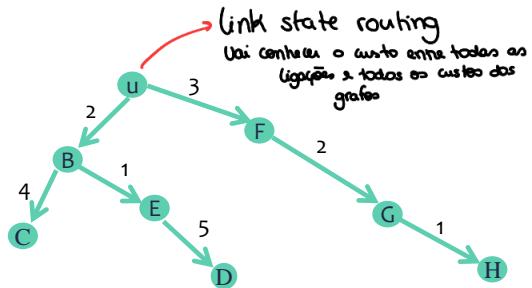
Shortest-path tree from u

16

Dijkstra's Algorithm – Example – 4

	D(S)	Next Hop
u	0	--
B	2	B
C	6	B
D	8	B
E	3	B
F	3	F
G	5	F
H	6	G

Forwarding table at u



Shortest-path tree from u

17

Link-State Routing

- Each router keeps track of its incident links
 - link up, link down
 - cost on the link → custo dinâmico em vez de estático
 - Each router broadcasts link state
 - every router gets a complete view of the graph
 - Each router runs Dijkstra's algorithm, to
 - compute the shortest paths
 - construct the forwarding table
 - Example protocols
 - Open Shortest Path First (OSPF), [STD 54](#), [RFC 2328](#), [OSPF Version 2](#)
 - Intermediate System – Intermediate System (IS-IS)
 - [ISO/IEC 10589:2002, Information technology — Telecommunications and information exchange between systems — Intermediate System to Intermediate System intra-domain routing information exchange protocol for use in conjunction with the protocol for providing the connectionless-mode network service \(ISO 8473\)](#)
 - [RFC 1195](#), Use of OSI IS-IS for routing in TCP/IP and dual environments
- Quando ocorre algum evento que possa ter influência nas rotas das tabelas de Redirecionamento, é preciso avisar todos os nós da rede (avisam os vizinhos que avisam os vizinhos, etc) para que saibam conhecer as alterações em toda a rede

18

Detection of Topology Changes

- Beacons generated by routers on links
 - Periodic “hello” messages in both directions
 - few missed “hellos” → link failure

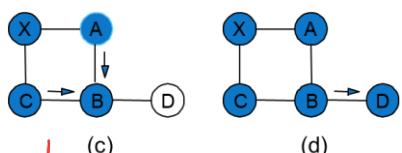
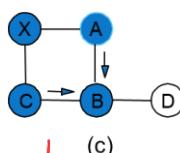
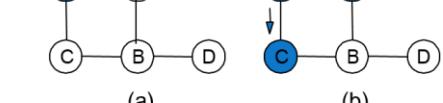
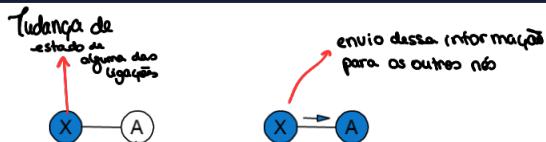
• São enviadas mensagens periodicamente, para que se consiga detectar falhas em determinadas ligações



19

Broadcasting the Link State

- How to Flood the link state?
 - every node sends link-state information through adjacent links
 - next nodes forward that info to all links
 - except the one where the information arrived



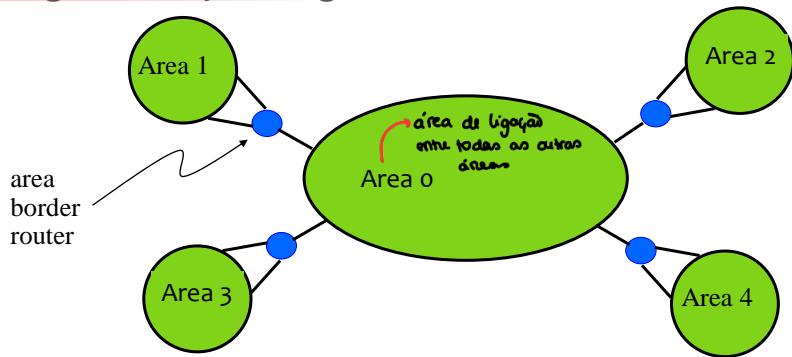
mesmo que não hajam mudanças
envio controlado para não ser devolvida a informação para quem enviou aquele nó:

→ Há medida que aumentamos o tamanho da rede, vai haver muita informação a ser transmitida periodicamente ou em caso de mudança, por toda a rede

20

Scaling Link-State Routing

- Overhead of link-state routing
 - flooding link-state packets throughout the network
 - running Dijkstra's shortest-path algorithm
- Introducing hierarchy through “areas”

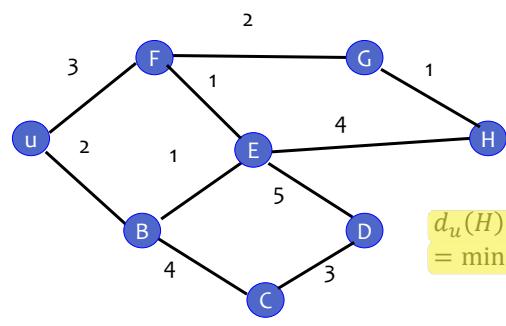


21

→ Criam-se **áreas**, onde os routers estão nas fronteiras e acumulam o conhecimento de um lado e do outro, mas só transmitem resumos, ou a possibilidade de chegar a determinadas redes a partir delas, e não todo o estado associado a internamente a cada uma das áreas (como os apss associados a todos os links dessas áreas)

Bellman-Ford Algorithm

- Define distances at each node x
 - $d_x(y) = \text{cost of least-cost path from } x \text{ to } y$
- Update distances based on neighbours
 - $d_x(y) = \min(c(x, n) + d_n(y)) \text{ over all neighbours } n \text{ of } x$



• não precisamos de saber o percurso, apenas precisamos de saber que há um distância entre o nosso vizinho e o valor que queremos colocar na nossa tabela.

• Distância até esse ponto:

Dist. até ao vizinho + Dist. vizinho até este nó

$$d_u(H) = \min(c(u, F) + d_F(H), c(u, B) + d_B(H))$$

$$\frac{3}{6} + \frac{3}{2} = \frac{5}{7}$$

22

Distance Vector Algorithm

- $c(x,n)$ = cost for direct link from x to n
 - node x maintains costs of direct links $c(x,n)$
 - n are x 's neighbours
- $D_x(y)$ = estimate of least cost from x to y
 - node x maintains distance vector $D_x = [D_x(y): y \in V]$
 - V is all nodes

↳ Vetor que o nó x mantém sobre os custos para chegar a todos os outros nós
- Node x maintains also its neighbours' distance vectors
 - for each neighbour n , x maintains $D_n = [D_n(y): y \in V]$

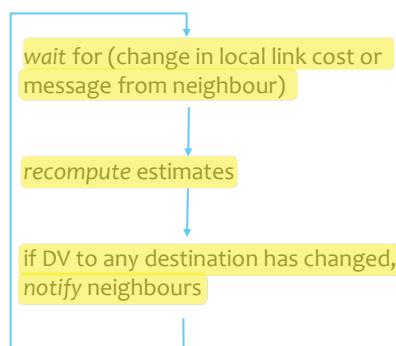
↳ para poder ir atualizar a medida que os vizinhos vão comunicando os custos para chegar aos outros nós
- Each node x periodically sends D_x to its neighbours
 - and neighbours update their own distance vectors
 - $D_z(y) \leftarrow \min_z \{c(z,x) + D_x(y)\}$ for each node $y \in V$
 - z is a neighbour of x

↳ Isto consegue chegar a todos nós com este custo
- Over time, the distance vector D_x converges

23

Distance Vector Algorithm

- A cada iteração podemos alterar vetor D
- Iterative, asynchronous
 - each local iteration caused by:
 - local link cost change
 - distance vector update message from neighbour
- Each node:



- Distributed
 - node notifies neighbours
 - only when its DV changes
- Neighbours then
 - notify their neighbours,
 - if necessary

↳ Se utilizavam aquela ligação para chegar ao nó destino

24

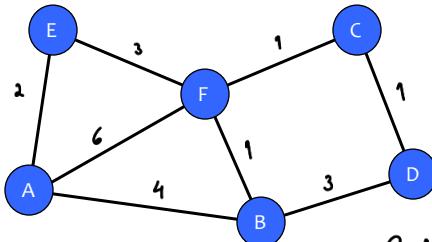
Distance Vector Example - Step 0

* Ainda não tem valores, pois não conhecemos

Os que não são vizinhos têm os
Distância para todos os outros

Table for A			Table for B		
Dst	Cst	Hop	Dst	Cst	Hop
A	0	A	A	4	A
B	4	B	B	0	B
C	∞	-	C	∞	-
D	∞	-	D	3	D
E	2	E	E	∞	-
F	6	F	F	1	F

Table for C			Table for D			Table for E			Table for F		
Dst	Cst	Hop									
A	∞	-	A	∞	-	A	2	A	A	6	A
B	∞	-	B	3	B	B	∞	-	B	1	B
C	0	C	C	1	C	C	∞	-	C	1	C
D	1	D	D	0	D	D	∞	-	D	∞	-
E	∞	-	E	∞	-	E	0	E	E	3	E
F	1	F	F	∞	-	F	3	F	F	0	F



→ Cada um dos nós envia informações para os vizinhos até onde é que consegue chegar e com que custos associados

25

q) $D(A, E) = D(E, A)$ no caso de um grafo não direcionado

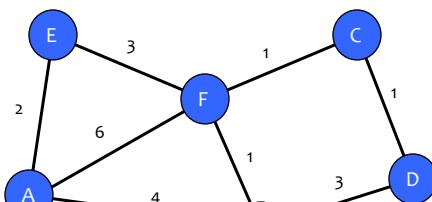
Um grafo direcionado pode ter custos diferentes

Distance Vector Example - Step 1

$$D(A, F) = 5 \rightsquigarrow A \rightarrow E \rightarrow F$$

Table for A			Table for B		
Dst	Cst	Hop	Dst	Cst	Hop
A	0	A	A	4	A
B	4	B	B	0	B
C	7	F	C	2	F
D	7	B	D	3	D
E	2	E	E	4	F
F	5	E	F	1	F

Table for C			Table for D			Table for E			Table for F		
Dst	Cst	Hop	Dst	Cst	Hop	Dst	Cst	Hop	Dst	Cst	Hop
A	7	F	A	7	B	A	2	A	A	5	B
B	2	F	B	3	B	B	4	F	B	1	B
C	0	C	C	1	C	C	4	F	C	1	C
D	1	D	D	0	D	D	∞	-	D	2	C
E	4	F	E	∞	-	E	0	E	E	3	E
F	1	F	F	2	C	F	3	F	F	0	F



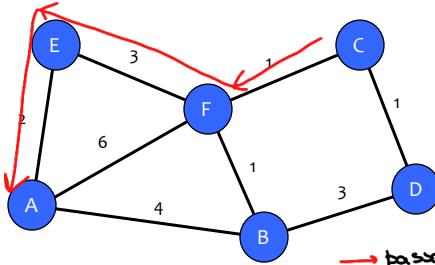
26

L) Alterações que cada nó faz ao seu vetor de distâncias

→ A alterou a sua rota para chegar a F, mas para chegar até C o valor que utilizou é o valor antigo que tinha para chegar a F

Distance Vector Example - Step 2

Table for A			Table for B			Table for C			Table for D			Table for E			Table for F		
Dst	Cst	Hop															
A	0	A	A	4	A	A	6	E	B	7	B	B	2	F	C	4	F
B	4	B	B	0	B	B	2	F	C	3	D	D	1	F	D	1	F
C	6	E	C	2	F	C	1	C	C	0	C	C	1	C	E	4	C
D	7	B	D	3	D	D	0	D	D	1	D	D	2	C	F	2	C
E	2	E	E	4	F	E	5	C	E	0	E	E	3	E	F	0	F
F	5	E	F	1	F	F	2	C	F	3	F	F	0	F	F	0	F



→ passadas algumas
iterações os valores
estabilizam até que ocorram
mudanças novamente

é na
mesma através
def, mas o
caminho depois
de F é alterado

altera custos

Custo para chegar a C vai ser recalculado

Propaga a informação

27

Routing Information Protocol (RIP)

- STD 56, RFC 2453, RIP Version 2
- Distance vector protocol
 - nodes send distance vectors every 30 seconds
 - or, when an update causes a change in routing
- RIP is limited to small networks

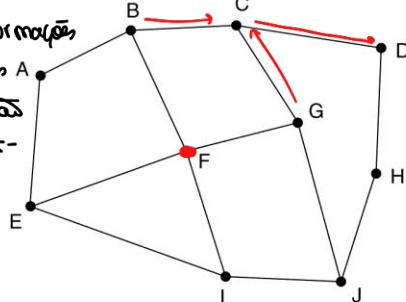
limitado a pequenas redes
visto que há uma atualiza-
ção um pouco mais constante

28

BGP – The Exterior Gateway Routing Protocol

↳ Numa rede interna vamos ter interior routing protocols

- Vão agregar as informações que possuem das redes A internas, portanto não passa toda a informação do que se está a passar. Só um sumário.



(a)

A set of BGP routers.

Information F receives from its neighbors about D

From B: "I use BCD"
From G: "I use GCD"
From I: "I use IFGCD"
From E: "I use EFGCD"

Vai receber a informação de como é que os outros chegam a D

(b)

Information sent to F

* B diz que passa pela rede B → rede C → rede D

→ Em muitos dos casos não é só o que vai ser feito em consideração, mas também certas **póliticas de propagação**.

→ Nó G consegue chegar a D através de J ou C. Anuncia só 1. Pode ter algum acordo com C que lhe figura + levanta chegar até D. Ele não anuncia que chega a J. Assim os outros nós não o vão utilizar como rede de passagem até J. Dependendo dos acordos que existem

29



FACULDADE DE CIÉNCIAS
UNIVERSIDADE DO PORTO

U.PORTO

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Unique Spanning Tree in Ethernet Networks

Routing

TTL → Impede que pacote fique pendendo na rede. Ao chegar a 0 é descartado.
 → Vai sempre diminuindo cada vez que o roteador encaminha o pacote

L2 Networking - Single Tree Required havendo ciclos

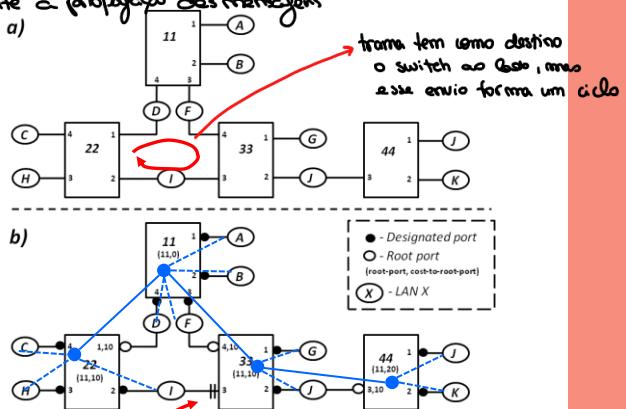
- Ethernet frame
 - No hop-count
 - Could loop forever
 - broadcast frame, mis-configuration

No caso da ethernet ao passar pelos switches, como não há nenhum controlo podia acontecer de uma trama ficar extamente na rede
 ↳ relativamente à propagação das mensagens

- Layer 2 network
 - Required to have tree topology
 - Single path between every pair of stations

- Spanning Tree Protocol (STP)
 - Running in bridges
 - Helps building the spanning tree
 - Blocks ports

Possíveis portas que levariam a loops
 Relativamente ao envio de determinadas tramas

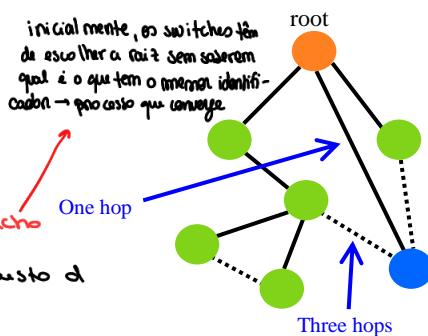


31

Constructing a Spanning Tree

- Distributed algorithm
 - switches need to elect a “root”
 - the switch with the smallest identifier
 - each switch identifies if its interface is on the shortest path from the root
 - messages (Y, d, X)
 - from node X
 - claiming Y is the root
 - and the distance is d

Há um envio de mensagens:
 - consegui chegar à raiz que acho
 que é Y desde de X com o custo d



32

Steps in Spanning Tree Algorithm

- Initially, each switch thinks it is the root
 - switch sends a message out every interface
 - identifying itself as the root with distance 0
 - example: switch X announces $(X, 0, X)$

X é um valor numérico: root ID
consegue chegar a X ate X com dsf
- Other switches update their view of the root
 - upon receiving a message, check the root id
 - if the new id is smaller, start viewing that switch as the root

Atualizam raiz = distância
Aparecem - se de quem é a raiz
- Switches compute their distance from the root
 - add 1 to the distance received from a neighbour
 - identify interfaces not on a shortest path to the root and exclude them from the spanning tree

a medida que vão recebendo informações vai atualizando

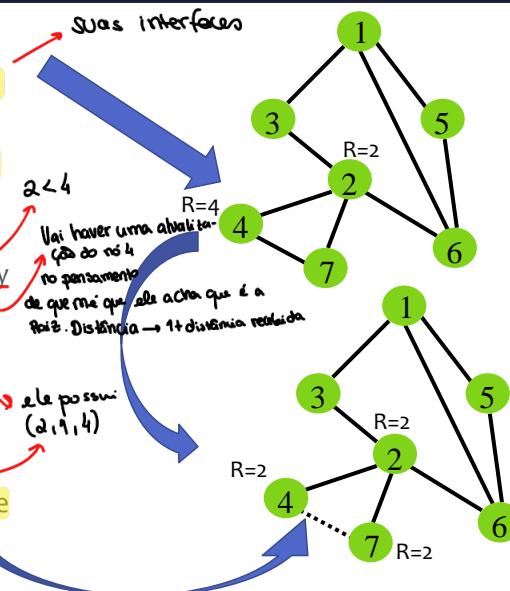
33

Example - Switch #4's Viewpoint

- Switch #4 thinks it is the root
 - sends $(4, 0, 4)$ message to 2 and 7
- Then, switch #4 hears from #2
 - receives $(2, 0, 2)$ message from 2
 - ... and thinks that #2 is the root
 - and realizes it is just one hop away

2 < 4
Vai haver uma atualização
que do rd 4 no pensamento
de que não que elle acha que é a
raiz. Distância → 1 + distância recebida
- Then, switch #4 hears from #7
 - receives $(2, 1, 7)$ from 7
 - and realizes this is a longer path
 - so, prefers its own one-hop path
 - and removes 4-7 link from the tree

ele possui
(2, 1, 4)
Não necessita desse caminho
para chegar a 2



34

3 de saídas que
a raiz seria 1 e que ele
chega lá com 1 hop

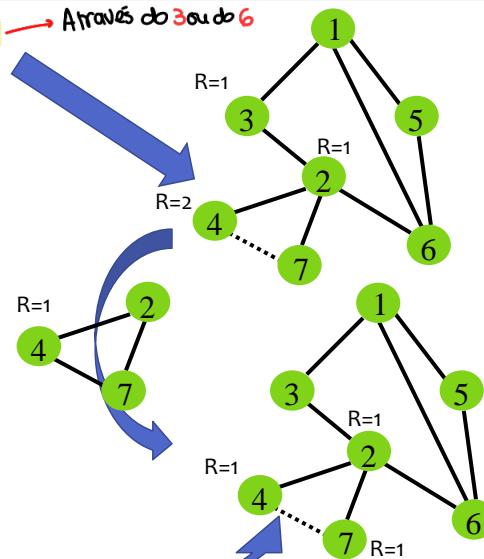
4 idem
mas se houver
2 saídas

Example - Switch #4's Viewpoint

- **Switch #2 hears about switch #1**
 - switch 2 hears $(1, 1, 3)$ from 3
 - switch 2 starts treating 1 as root
 - and sends $(1, 2, 2)$ to neighbours
- **Switch #4 hears from switch #2**
 - switch 4 starts treating 1 as root
 - and sends $(1, 3, 4)$ to neighbours
- **Switch #4 hears from switch #7**
 - switch 4 receives $(1, 3, 7)$ from 7
 - and realizes this is a longer path
 - so, prefers its own three-hop path
 - and removes 4-7 link from the tree

Assumindo que para o 1 podia ter um caminho +
curto pelo lado do 7. Não estava excluído, pois a raiz
mudou.

A descoberta de qual era o switch que era a raiz foi distribuída e som haber rada pré-definida
Vai convergir para uma árvore com raiz em 1. A árvore final é a árvore que determinada para a comunicação



35



FACULDADE DE CIÉNCIAS
UNIVERSIDADE DO PORTO

U.PORTO

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Maximum Flow of a Network

Routing

Flow Network Model

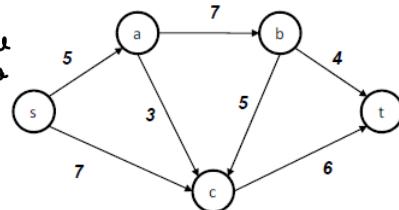
- Flow network

- source s
- sink t
- nodes a, b and c

- Edges are labelled with capacities

- (e.g. bit/s)

largura de banda



- Communication networks are not flow networks

- they are queue networks
- flow networks enable to determine limit capacity values

37

Maximum Capacity of a Flow Network

- Max-flow min-cut theorem

- maximum amount of flow transferable through a network
- equals minimum value among all simple cuts of the network

- Cut → split of the nodes V into two disjoint sets S and T

- $S \cup T = V$ → no momento da divisão não podemos excluir nós
- there are $2^{|V|-2}$ possible cuts

- Capacity of cut (S, T) :

$$c(S, T) = \sum_{(u,v) | u \in S, v \in T, (u,v) \in E} c(u, v)$$

tem de existir uma
aresta a ligar os
nós

→ Os nós do
mesmo cut têm de
ter ligações entre si

u num lado da divisão
 v no outro lado da divisão

38

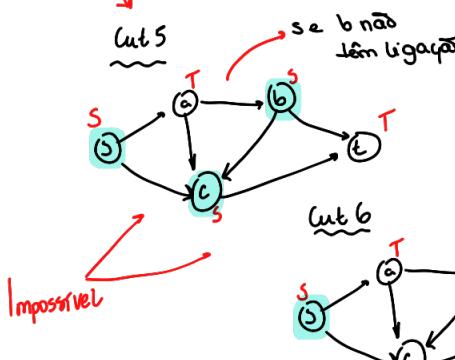
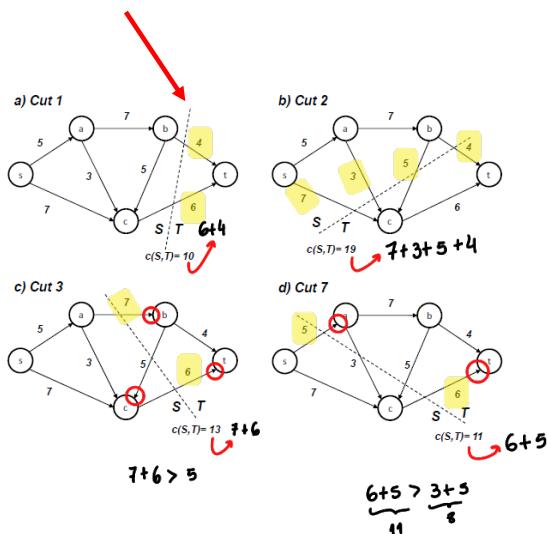
Max-flow Min-cut - Example

todas as possibilidades

$$2^{|V|-2} = 8 \text{ possible cuts}$$

Cut	Vertices					$c(S, T)$	Feasibility
	s	a	b	c	t		
1	S	S	S	S	T	10	✓
2	S	S	S	T	T	19	✓
3	S	S	T	S	T	13	✓
4	S	S	T	T	T	17	✓
5	S	T	S	S	T	-	✗
6	S	T	S	T	T	-	✗
7	S	T	T	S	T	11	✓
8	S	T	T	T	T	12	✓

Maximum flow = 10



Homework

1. Review slides

2. Read from Tanenbaum

- Section 5.2 – Routing algorithms
- Section 4.8.3 Spanning Tree Bridges

3. Answer questions at Moodle

39

40



FC

FACULDADE DE CIÉNCIAS
UNIVERSIDADE DO PORTO

U.PORTO

U.PORTO

FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

End of Routing