P561: Network Systems
Week 4: Internetworking II

Tom Anderson
Ratul Mahajan

TA: Colin Dixon

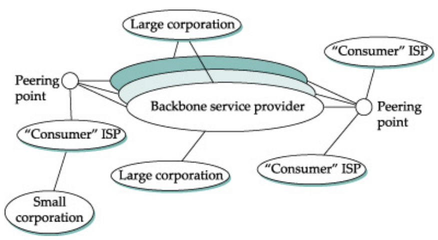
Today

Internet routing (BGP)
Tunneling and MPLS

Wireless routing
Wireless handoffs

2

Internet today



3

Key goals for Internet routing

Scalability

Support arbitrary policies
- Finding "optimal" paths was less important

(Supporting arbitrary topologies)

4

Internet routing overview

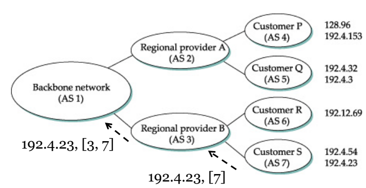
Two-level hierarchy for scalability
- Intra-domain: within an ISP (OSPF, MPLS)
- Inter-domain: across ISPs (BGP)

Path vector protocol between Ases
- Can support many policies
- Fewer messages in response to small changes
  - Only impacted routers are informed

5

Path vector routing

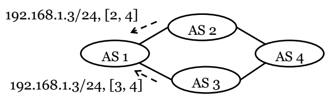
Similar to distance vector routing info includes entire paths



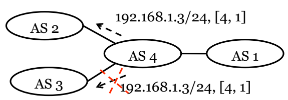
6

Policy knobs

1. Selecting one of the multiple offered paths

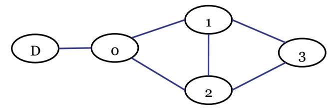


2. Deciding who to offer paths



7

Path vector vs. link state vis-à-vis policy



AS3 preferences
[310]
[320]
[3210]
~~[3120]~~

With path vector, implementing the policy above requires only local knowledge at AS3

With link state, AS3 would need to know the policies of other ASes as well

8

Typical routing policies

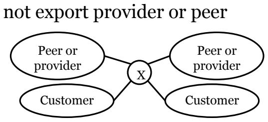
Driven by business considerations

Two common types of relationships between ASes

- Customer-provider: customer pays provider
- Peering: no monetary exchange

When selecting routes: customer > peer > provider

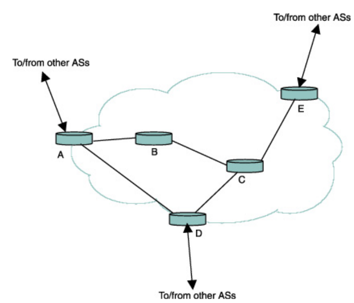
When exporting routes: do not export provider or peer routes to other providers and peers



Prefer routes with shorter AS paths

9

BGP at router level



10

BGP limitations

Path quality

Scale

Convergence

Security

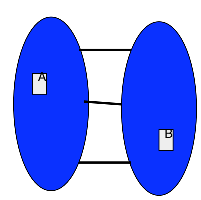
11

Path quality with BGP

Combination of local policies may not be globally good

- Longer paths, asymmetric paths
- Shorter "detours" are often available

Example:
hot potato routing

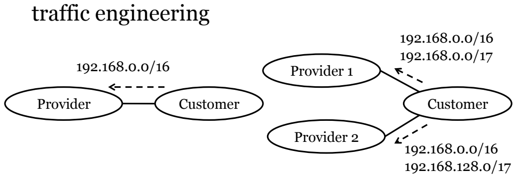


12

Scaling pressures on BGP

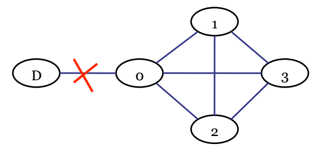
Too many prefixes (currently ~280K)

Major factors behind growth: multi-homing and traffic engineering



13

BGP convergence (1/4)

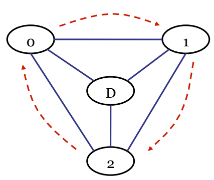


Temporary loops during path exploration

Differentiating between failure and policy-based retraction can help but not completely

14

BGP convergence (2/4)



To get to D, X prefers
- [X, (X+1) mod 3]
- [X]
- Others

Persistent loops can also form in BGP

Fundamentally, the combination of local policies may not have a unique global solution

15

BGP convergence (3/4)

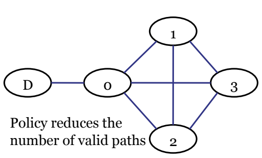
Several other issues have been uncovered
- Interaction with intra-domain routing
- Interaction with traffic engineering extensions
- Interaction with scalability extensions

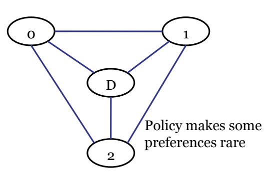
16

BGP convergence (4/4)

Q: What saves us in practice?

A: Policy! (No guarantees, however)





17

BGP security

Extreme vulnerability to attacks and misconfigurations
- An AS can announce reachability to any prefix
- An AS can announce connectivity to other Ases

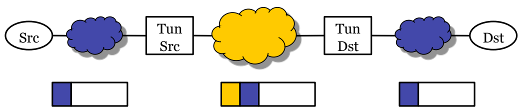
Many known incidents
- AS7007 brought down the whole internet in 1997
- 75% of new route adverts are due to misconfigs [SIGCOMM 2002]
- Commonly used for spamming

Technical solutions exist but none even close to deployment
- Incentives and deployability (Week 10)

18

Tunneling

Encapsulating one protocol within another



The blue sources, destinations, networks are oblivious to tunneling

The yellow network does not care if it carries blue (or green) packets

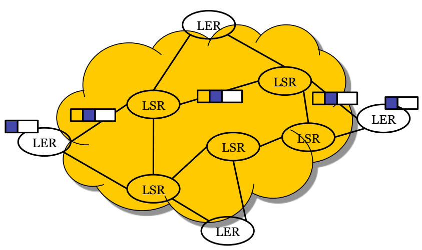
Tunneling is broadly useful technique

Used widely today

- Secure access to remote networks (VPNs)
  - Your laptop to corporate networks
  - Between different sites of a company
- MPLS
- 6to4
- GRE
- SSH tunnels
- ....

Think of it as a generalization of traditional layering

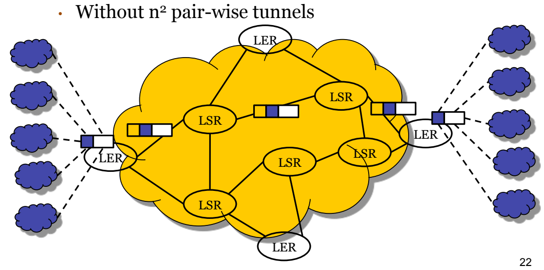
MPLS



Benefits of MPLS (1/3)

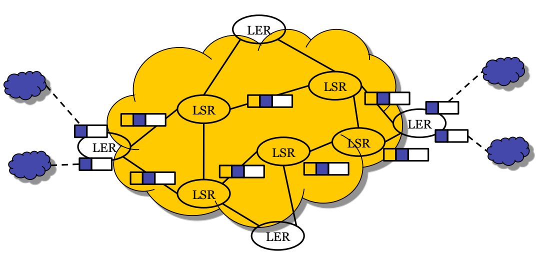
LSRs do not understand or maintain state for IP

- Can yield higher performance
- Without n^2 pair-wise tunnels



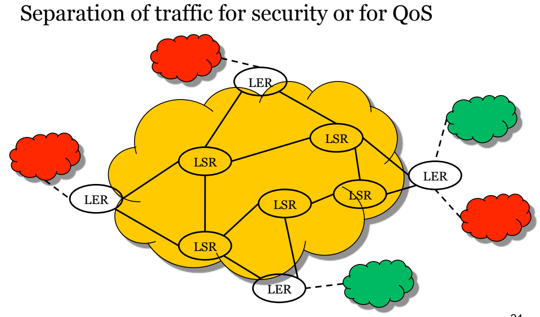
Benefits of MPLS (2/3)

Traffic engineering (load balancing)



Benefits of MPLS (3/3)

Separation of traffic for security or for QoS



Downsides of MPLS

Unnecessary overhead
- If all you want is IP forwarding
- If link state routing can provide effective traffic engineering

Robustness to failures
- Setting up a complete virtual circuit takes time
- Fast reroute works only for a handful for failures

Opacity
- Traditional diagnosis tools do not work

Complexity
- Requires more configuration at routers

25

MPLS adoption

Pretty widespread
- Almost all tier-1 ISPs have deployed MPLS

It offers tools that network admins badly need
- Practical concerns trumped purist views

26

Why is wireless routing different?

Mobility and fast changing conditions

Packet losses

Interference

27

First generation of protocols

Focus on mobility and changing conditions
- Used hop count as the quality metric
- Reactive route computation was more popular
  - To avoid unnecessary topology maintenance overhead

Examples: DSR, AODV

28

Hop count limitations

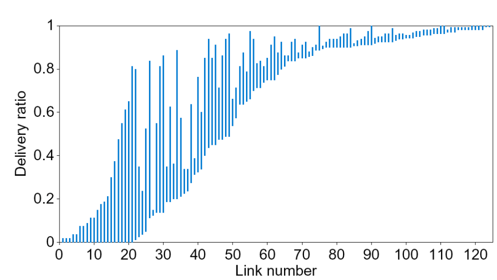
It minimizes the number of hops and thus prefers longer links

But longer links tend to have more loss
- Need more retransmissions for successful reception

Retransmissions can consume more spectrum resources than using shorter hops
- Need to balance hops and losses

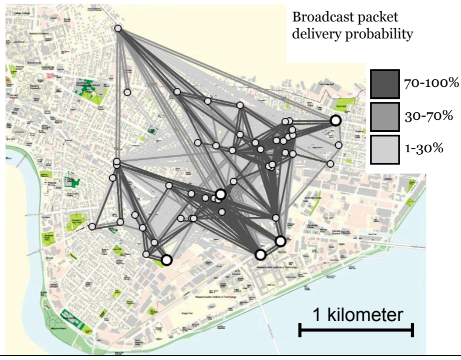
29

All links are not the same

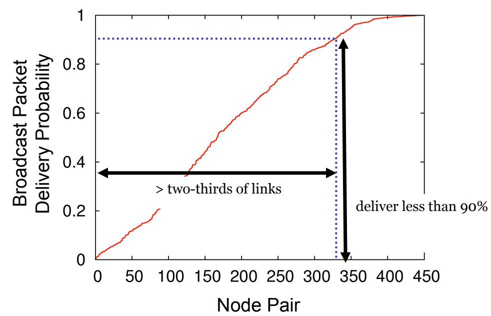


MIT's indoor testbed

Link qualities in Roofnet



Delivery probabilities in Roofnet



ETX: Expected transmissions

Estimate number of times a packet has to be retransmitted on each hop

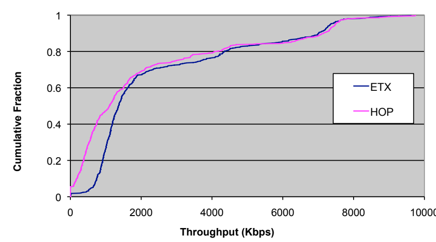
- Use probes to calculate forward and reverse loss rate to each neighbor

$$ETX = \frac{1}{(1 - P_f) * (1 - P_r)}$$

Select the path with least total ETX

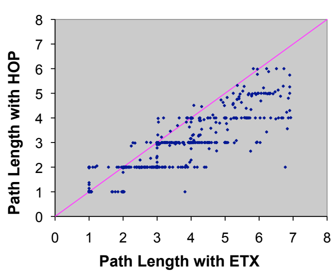
- Takes longer paths only when they are better

ETX avoids low-throughput paths



Results on MSR testbed (courtesy Jitu Padhye)

ETX paths are generally longer



ETX shortcomings

Assumes that all transmissions are equal

- In reality, different transmissions use different amount of spectrum

Assumes a simplistic interference model

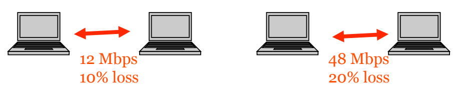
- Cross-flow interference not directly accounted
- Worst-cast self-interference

Ignores the broadcast nature of wireless

ETT: Expected transmission time

Generalizes ETX to the case of multiple bit rates
Directly measures spectrum resources used
On a link with loss rate *p*, bitrate *B*, packet size *S*

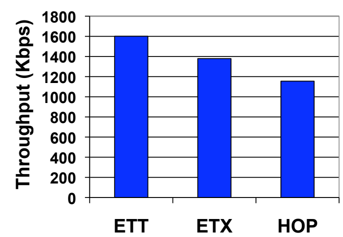
$$ETX = \frac{1}{1-p} \qquad ETT = \left(\frac{S}{B}\right) * ETX$$



[Routing in Multi-radio, Multi-hop Wireless Mesh Network, MOBICOM 2004]

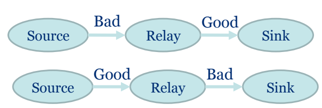
37

ETT performance



38

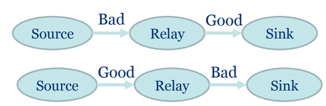
Is one path better than the other?

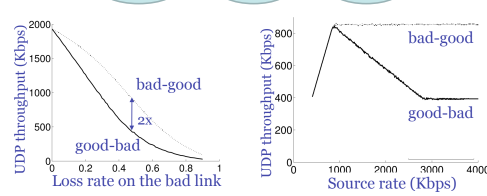


Hint: ETT (or ETX) of both is same

39

Unpredictable wireless performance





40

Predictable performance optimization

Measure the RF profile of the network
↓
Constraints on sending rate and loss rate of each link
↓
Find compliant source rates that meet the objective

[Predictable performance optimization for wireless networks, SIGCOMM 2008]

41

Measurements

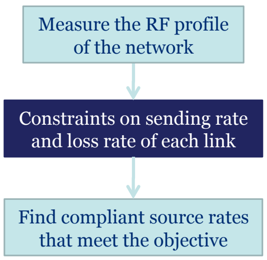
Measure the RF profile of the network
↓
Constraints on sending rate and loss rate of each link
↓
Find compliant source rates that meet the objective

One or two nodes broadcast at a time

Yields information on loss and deferral probabilities

42

Modeling

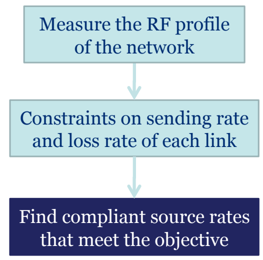


$O(n^2)$ constraints

1. Link throughput is a function of loss rate and tx probability
2. Link tx probability is a function of tx probability of other links and deferral probability
3. Link loss rate depends on tx probability of other links
4. Tx probability is bounded by a function of loss rate

43

Optimization

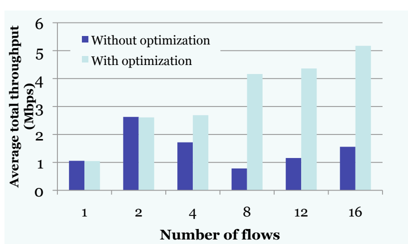


Inputs:
- Traffic matrix
- Routing matrix
- Optimization objective

Output:
- Per-flow source rate

44

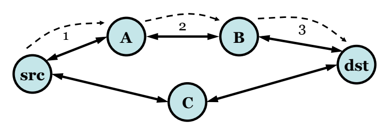
Benefit of predictable optimization



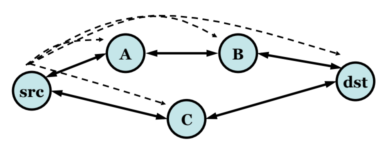
45

Leveraging wireless broadcast

Traditional routing



An opportunity



46

ExOR: Extremely opportunistic routing

Source identifies and prioritized list of relays

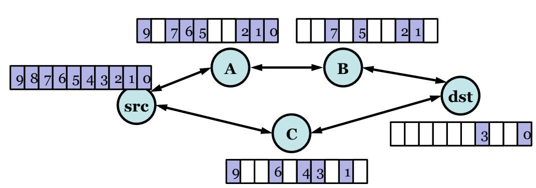
Source groups packets into a batch and transmits

Nodes run an agreement protocol
- The highest priority relay announces what it received
- The next relay transmits packets not received by higher priority relays
- Finally, the source retransmits what nobody got

[Opportunistic Routing in Multi-Hop Wireless Networks, SIGCOMM 2005]

47

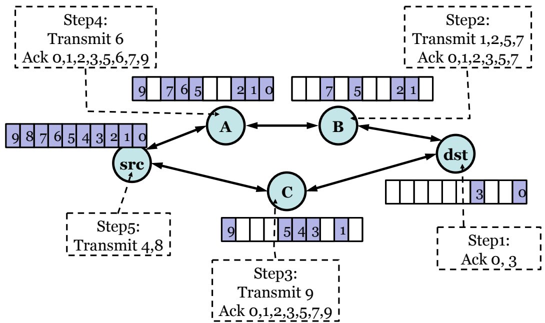
ExOR: Initial transmission



Source transmits the entire batch

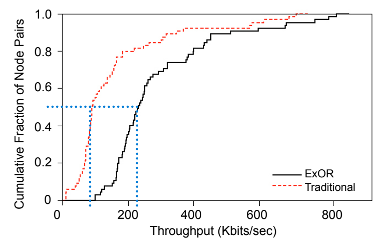
48

ExOR: Agreement protocol



49

ExOR performance improvement



50

Question

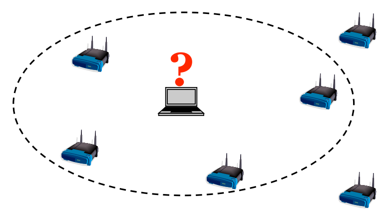
Is ExOR a forwarding or a routing protocol? (Or, is it a MAC-layer protocol?)

51

Wireless handoff

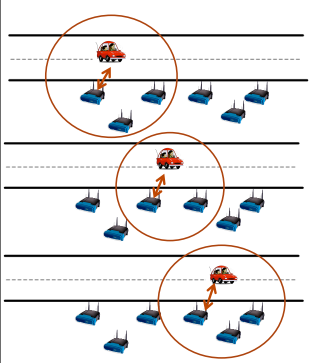
A special case of routing, with one mobile node

- How to provide connectivity to the node as it moves?



52

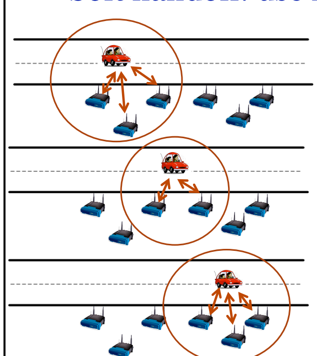
Hard handoff: one BS at a time



Select BS based on
- Signal strength (SNR)
- Loss rate
- Combination of factors

53

Soft handoff: use multiple BSes

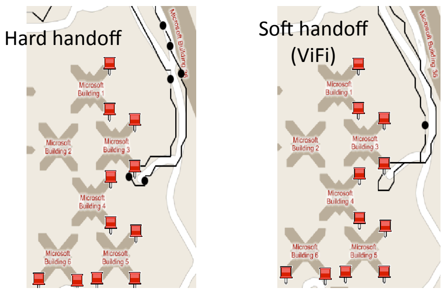


Adds reliability by leveraging diversity

The mobile node does not depend on only one BS

54

The handoff methods in a real setting



55

Soft handoff questions

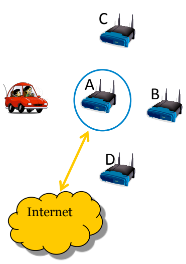
How to pick multiple BSes?
- A generalization of picking one
- Usually, two or three BSes suffice

What to do when multiple BSes hear a packet from the mobile?
- The BS backplane may be bandwidth-limited

How do multiple BSes send packet to the mobile?
- Simultaneous transmissions may collide

56

ViFi overview



Designed with the vehicular setting in mind but the underlying problem is more general

Vehicle chooses *anchor* BS
- Anchor responsible for vehicle's packets

Vehicle chooses a set of BSes in range to be *auxiliaries*
- ViFi leverages packets overheard by auxiliaries

57

ViFi protocol

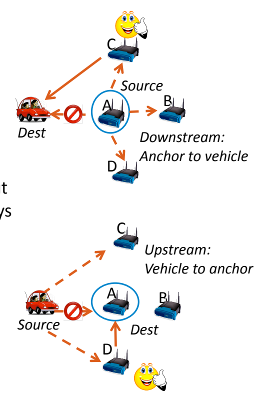
(1) Source transmits a packet

(2) If destination receives, it transmits an ack

(3) If auxiliary overhears packet but not ack, it *probabilistically* relays to destination

(4) If destination received relay, it transmits an ack

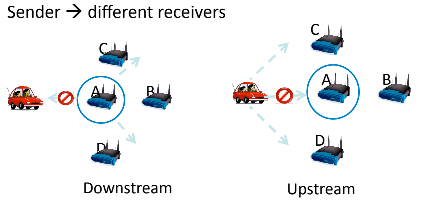
(5) If no ack within retransmission interval, source retransmits



58

Why is relaying effective?

- Losses are bursty
- Losses are independent
  - Different senders → receiver
  - Sender → different receivers



59

Determining relaying probability

Goal: Compute relaying probability R_B of auxiliary B such that
- Total # of relays are limited
- Prefer auxiliaries with better connectivity to destinations
- Avoid per-packet coordination

1: The probability that auxiliary B is *considering relaying*
C_B = P(B heard the packet) . P(B did not hear ack

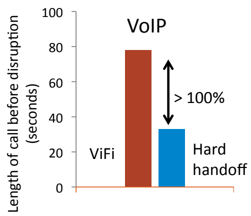
2: The expected number of relays by B is $E(B) = C_B \times R_B$

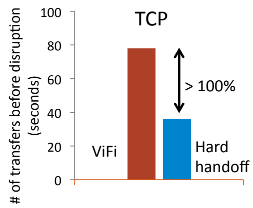
3: Formulate ViFi probability equation, $\sum E(x) = 1$
To solve uniquely, set R_B *proportional to P(destination hears B)*

4: B estimates P(auxiliary considering relaying) and P(destination heard auxiliary) for each auxiliary

60

ViFi reduces disruptions to apps





61

Other handoff challenges

Finding nearby BSes

- Scanning can be time consuming

Overhead of switching BSes

- Can involve association, authentication, DHCP, etc.

62

Summary

Forwarding and routing protocols enable you to construct bigger networks

- Seemingly simple protocols but complex dynamics

The wireless medium bring challenges of its own that forwarding and routing must address

Next week (tom @ UW):
how to communicate reliably and efficiently when so much can go wrong inside a network?

63