Communications Network Design

lecture 19

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

March 2, 2009

Networks of networks

The Internet is a network of networks. Most of the problems we have considered up to this point concern a single network. There are many interesting problems when we consider how these networks interconnect.

The lecture extends our discussion of routing to allow for more complicated routing objectives such as policies, which are used for inter-domain routing.

the Internet

The Internet has

- many thousands of routers
- many millions of hosts

What does it look like

- can we do shortest path routing?
- should we do shortest path routing?

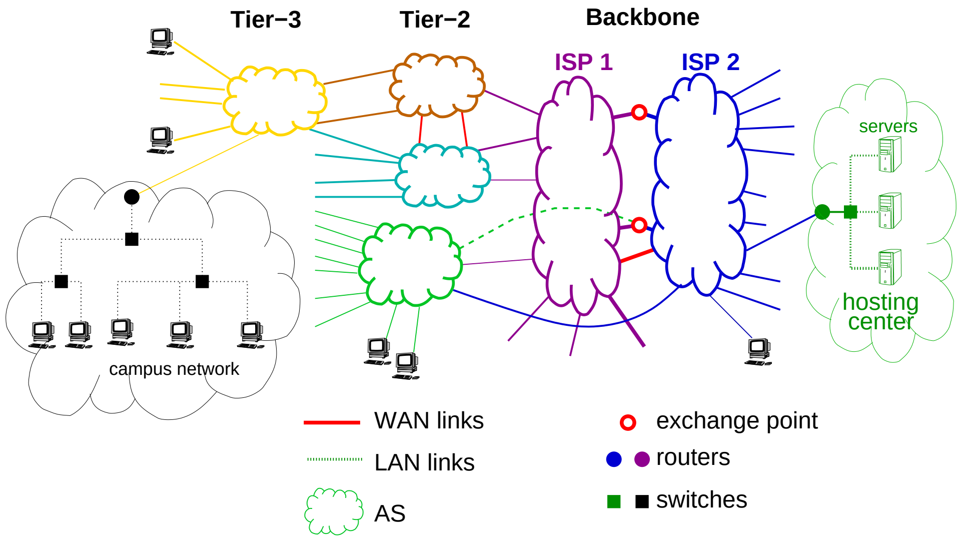
Obviously, we at least need some hierarchy?

Internet Topology

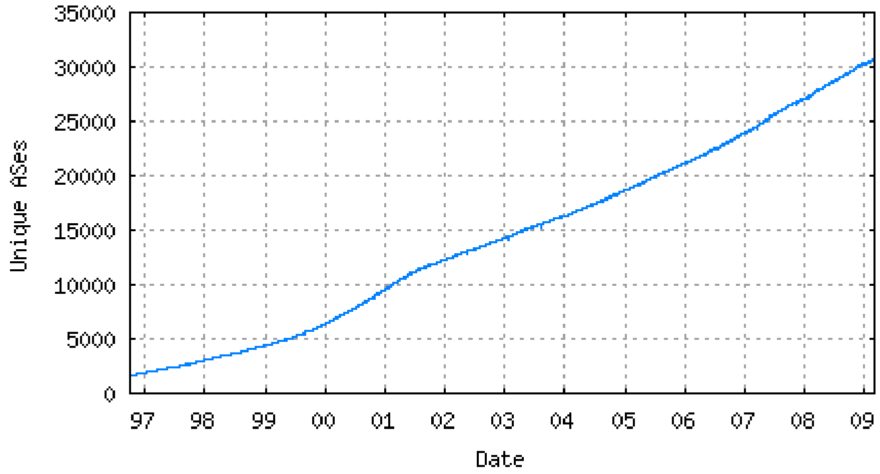
The Internet is broken into more than 10,000 Autonomous Systems (ASes)

- AS is a separately managed network
- within an AS may use different routing, technology, management, ...
- may be a LAN, WAN, or combination
- example ASes:
  - ISP (Internet Service Provider)
  - Campus network
  - Enterprise network
  - Hosting center
- see RFC 1009 for definition (obsoleted by 1812)

the Internet



Number of ASes



http://www.cidr-report.org/

maximum number is ∼ 65k

Communications Network Design: lecture 19 – p.5/32

Communications Network Design: lecture 19 – p.6/32

the Internet

- an AS is a network under one administrative domain
  - from the outside, we don't see the details
  - all we see are a set of subnetworks which are reachable via that AS
- subnets
  - either a group directly attached computers
  - or a group of customers' computers
- CIDR (Classless Inter-Domain Routing)
  - subnet = group of IP addresses with a common **prefix**
    - e.g. private addresses 192.168.0.0/16
      - all address with same first **16** bits 192.168
      - 192.168.0.0 – 192.168.255.255

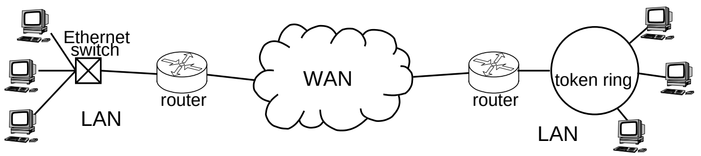
Different Flavours of Routing

Routing is different inside an AS from between ASes

- **intra-domain** (inside an AS)
  - called Interior Gateway Routing (**IGP**) protocols
  - examples: OSPF, RIP, EIGRP, IS-IS, ...
  - can use any one of these
  - can even use more than one at once!
- **inter-domain** (between ASes)
  - called Exterior Gateway Routing (**EGP**) protocols
  - one defacto standard BGPv4
    - Border Gateway Protocol
    - must talk internationally
      - can be only one

An Aside on Gateways

- router sometimes called gateway



  - RFC 1009 "Requirements for Internet Gateways" has definitions of such
  - better to use this term for gateway routers (that join two networks)
  - also for high level (e.g. network level) protocol conversion, e.g. IP to IPX
- but routing protocols still get called 'gateway' protocols

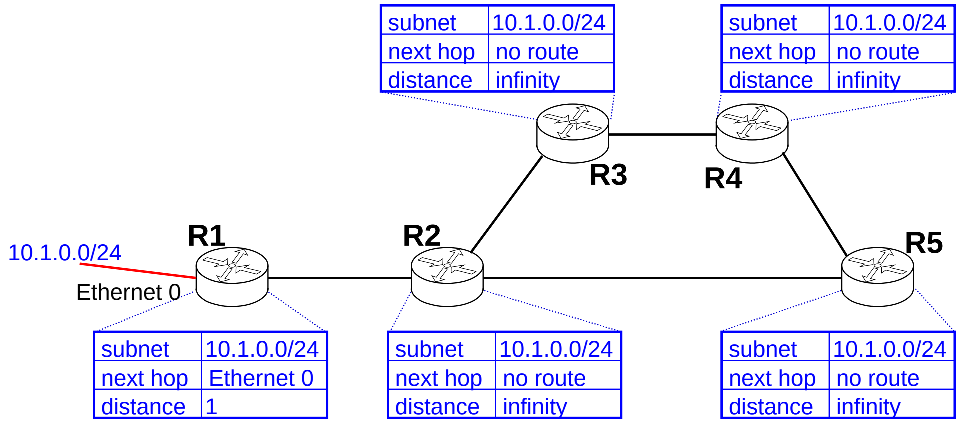
Link state vs Distance Vector

- We saw OSPF was a **link-state** routing protocol
  - floods topology (link states), and computes SPF
  - solves shortest path problem
- alternative is called **distance-vector** protocol
  - examples: RIP, IGRP, ...
  - originally also aimed to solve shortest paths
    - but nodes don't need to know complete topology
  - does BGP still do this?
    - BGP is a generalization called **path-vector** protocol

Distance Vector reminder

- Make a list of destinations you can reach and the distance to these destinations.
  - Store in routing table
- Share this list with your neighbours
- Add to routing table new information gained from adjacent routers about the destinations they can reach
  - remember to increment their distance
  - keep the source as the next hop
- If two paths to the same destination exists, keep the shortest distance path.
- Repeat periodically (in RIP every 30 seconds).

Distance Vector example



Internet structure/topology

- RIP still used, but only in small networks
  - IGRP similar to RIP, but few improvements to make it more scalable
  - I don't know how widely IGRP is used
- to really understand why distance-vector protocols are so important, we need to look at BGP
- BGP needs to support connectivity between ASes
- structure of AS graph is therefore important
  - tiering
    - customer-provider relationship
  - peering
  - routing **policy**

Tiering

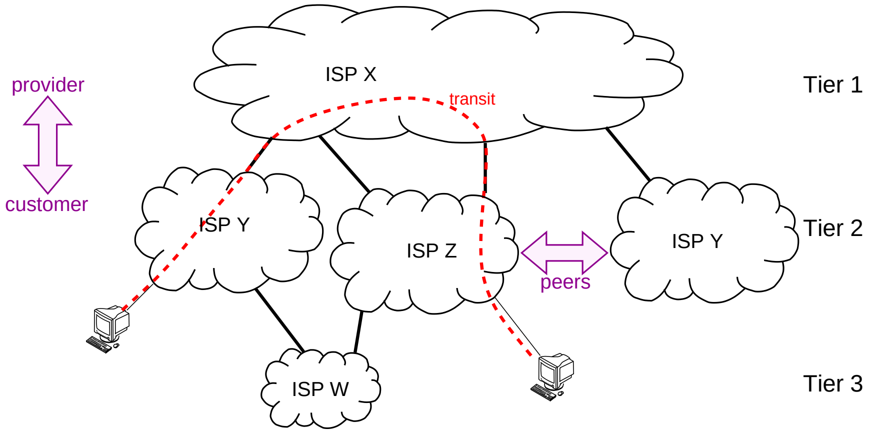
- no hard and fast rules, but
  - tier-1 ISP: international, or national backbone
    - provide transit
    - have at least some default-free routers
    - have connectivity over large geographic area
  - tier-2 ISP: regional ISP
    - provide transit within a geographic area
    - may have default-free routers
  - tier-3 ISP: local ISP
    - do not provide commercial transit services, although they may incidentally provide transit among their customers
  - tier-4 ISP: e.g. company network
    - Internet access through provider only

Communications Network Design: lecture 19 – p.13/32

Communications Network Design: lecture 19 – p.14/32

Tiering

Higher tiers provide **transit** for lower tiers



- Lower tiers are **customers** of higher tiers
- Higher tiers are **providers** for lower tiers

Tiering

Some "tier-1" ISPs (in no particular order)

- UUNET/WorldCom/MCI (AS 701)
- AT&T (AS 7018 - North American backbone)
- Verio (AS 2914)
- Sprint (AS 1239)
- Level 3 (AS 1)
- Cable & Wireless (AS 3561)
- Global Crossing (AS 3549)
- Qwest (AS 209)

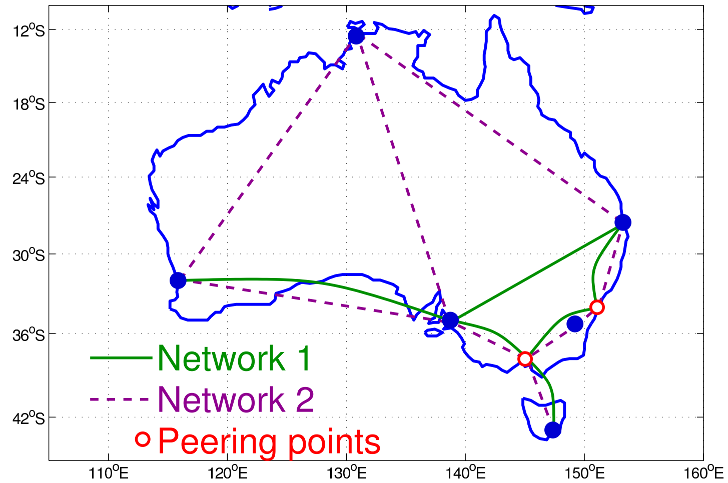
Note that some companies run more than one AS

Communications Network Design: lecture 19 – p.15/32

Communications Network Design: lecture 19 – p.16/32

Peering

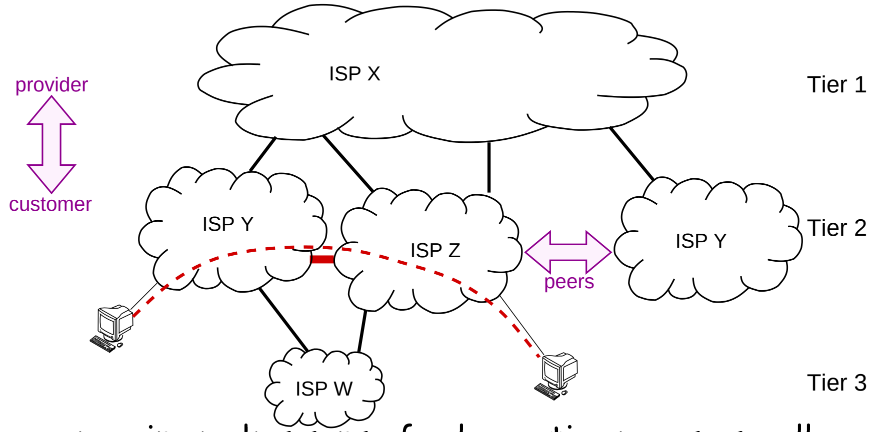
Two national networks.



Traffic has to get between them: **peering** links [1, 2].

Tiering and Peering

Peering between tier-1's is **needed**



- peering makes sense for lower tier peers as well
- avoid transit charges from providers

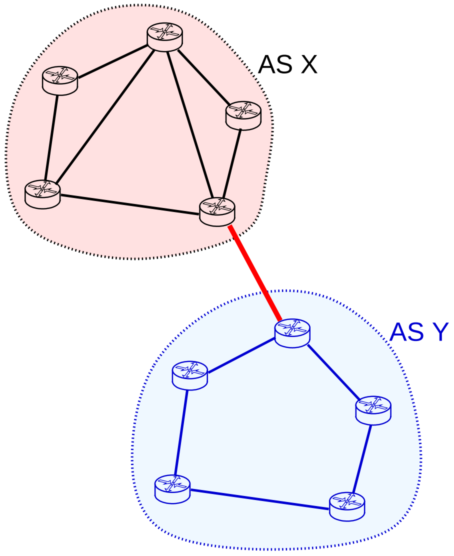
Peering Connections

What are the physical connections between ASes

- private peering
  - a point-to-point connection between a gateway router on each network
  - usually a WAN link
- Internet Exchange Point (IXP)
  - third party runs a router or switch or network
  - ISPs connect to the switch
  - similar concept Network Access Point (NAP)
- co-location facility
  - third party provides premises (and power etc)
  - multiple ISPs maintain routers in the premises
  - create local connection between their routers
  - e.g. carrier hotel

Private Peering Connections

- advantage:
  - high capacity
  - only two parties involved
- disadvantages:
  - not very flexible
  - e.g. can't change peers easily

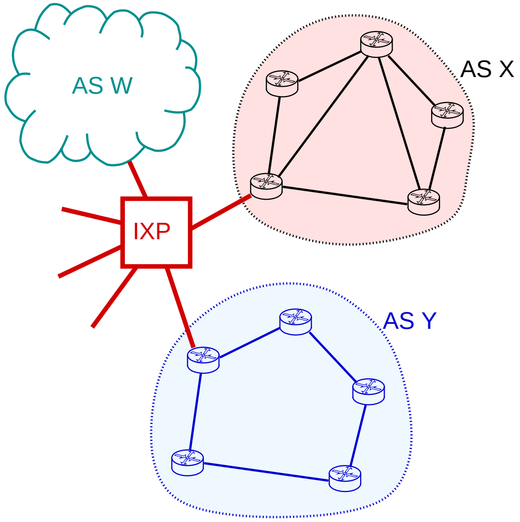


Communications Network Design: lecture 19 – p.19/32

Communications Network Design: lecture 19 – p.20/32

Exchange points

- advantage:
  - multiple parties
  - very flexible
- disadvantages:
  - connection to a single PoP
  - lower capacity
  - subject to a third party



Exchange points

Some US Exchange points

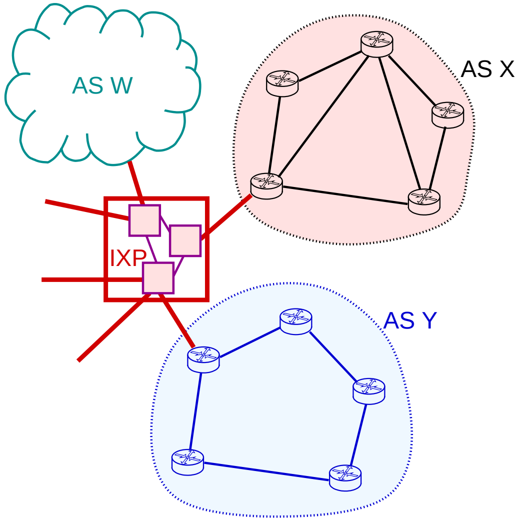
- MAE East (N VA)
- Sprint NAP (NJ)
- PAIX
- MAE West

Australian exchange points

- AUSIX.NET - Sydney
- Melbourne NAPette
- VIX - Victorian Internet Exchange (Melbourne)
- SAIX - Southern Australian Internet Exchange
- WAIX - Western Australia Internet Exchange

Distributed Exchange points

- advantage:
  - multiple parties
  - very flexible
- disadvantages:
  - subject to a third party



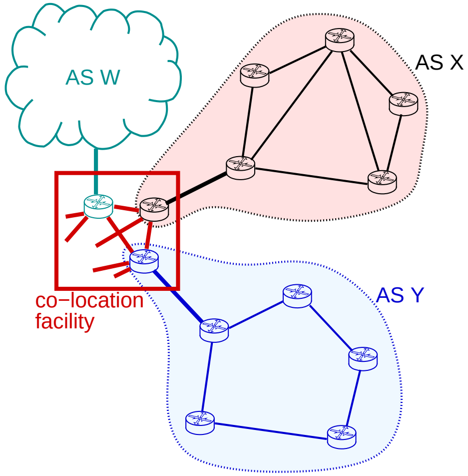
Distributed Exchange points

Some distributed exchange points

- LoNAP, London http://www.lonap.net/ see their peering matrix at http://stats.lonap.net/cgi-bin/matrix.cgi
- LYNX, London http://www.linx.net/ http://www.nanog.org/mtg-9901/ppt/linx/sld001.htm

Co-Location

- advantage:
  - best of private peering and NAPs
- disadvantages:
  - extra expense
- example:
  - Internap



http://www.internap.com/products/preferredcollo.html

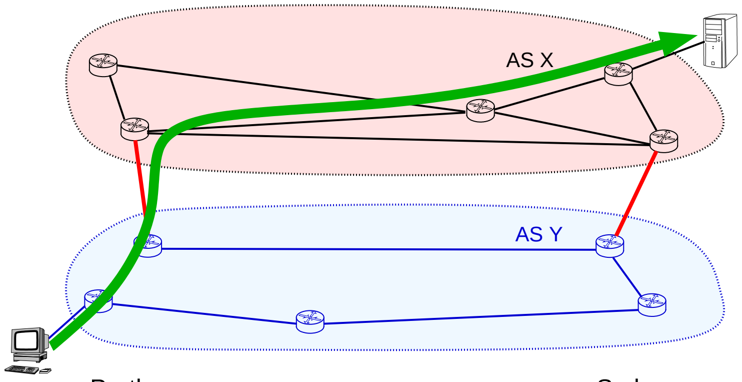
http://www.internap.com/products/locationmap.html

Routing Policy

- policy is a set of arbitrary rules for routing
- examples
  - we prefer to route to **peers** rather than **providers**
    - ⋆ providers charge us money
    - ⋆ traffic exchanged with peers for free
  - we prefer to route to route traffic with X
    - ⋆ maybe X provides better QoS
    - ⋆ maybe X's network is more secure
  - hot-potato routing
    - ⋆ reduce cost of carrying traffic on our network by dumping onto someone else's as soon as possible

Hot Potato routing

- dump traffic off your network as fast as possible

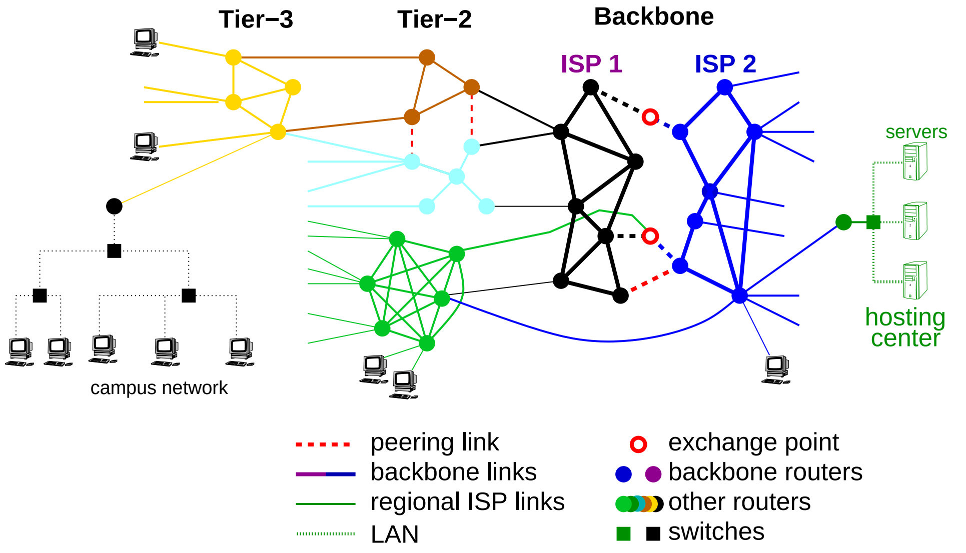


- results in **intrinsic** asymmetry in routing
- only fair if traffic is balanced

Complexities multiply

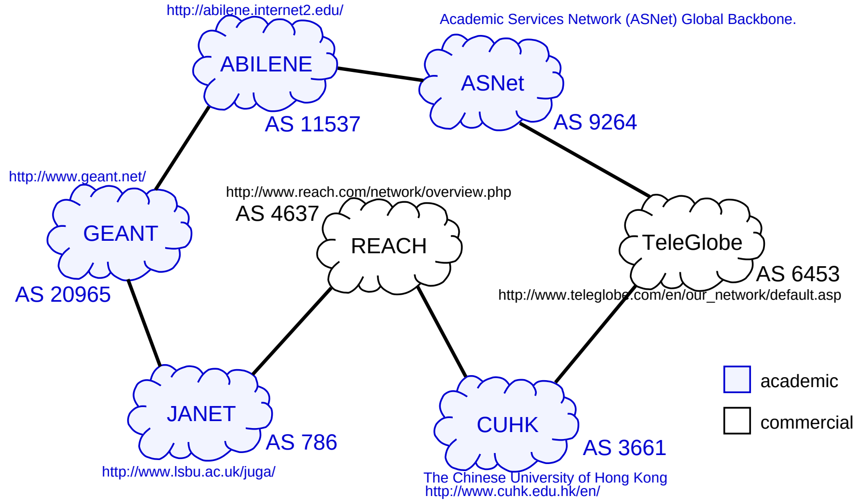
- no hard and fast rules about tiering
  - companies would like to be called tier-1
  - companies operate multiple networks
  - regional coverages overlap, but aren't equal
- peering between lower tiers to avoid transit fees
- relationships are more than just
  - customer-provider
  - peer-peer
- physical layers add complexity
  - two IP networks (layer 3)
  - relate as peers (so they are competing at level 3)
  - but both buy layer-1 physical transport from same company

A Picture of the Internet



A Real Example

A real example from [3]



Communications Network Design: lecture 19 – p.29/32

Communications Network Design: lecture 19 – p.30/32

Inter-domain optimization

- network design: e.g., where should peering links go?
- traffic engineering: balancing loads on peering links
- routing: optimize WRT policies (BGP)

To work with any of these, we need to know more about how BGP works.

References

[1] G. Huston, "Peering and settlements - part I," The Internet Protocol Journal, vol. 2, March 1999.

[2] G. Huston, "Peering and settlements - part II," The Internet Protocol Journal, vol. 2, June 1999.

[3] H. Zheng, E. K. Lua, M. Pias, and T. Griffin, "Internet routing policies and round-trip-times," in Passive and Active Measurements Workshop, (Boston, MA, USA), 2005.

[4] J. Stewart III, BGP4: Inter-domain Routing in the Internet. Addison-Wesley, Boston, 1999.

[5] T. Griffin, "Does BGP Solve the Shortest Paths Problem?," in The North American Network Operators' Group (NANOG) 18, (San Jose, CA, USA), February 2000. http://www.nanog.org/mtg-0002/ppt/griffin/.

[6] T. Griffin, F.Shepherd, and G.Wilfong, "The stable paths problem and interdomain routing," IEEE/ACM Transactions on Networking, vol. 10, no. 2, pp. 232–243, 2002.

Communications Network Design: lecture 19 – p.31/32