Rethinking Internet Design

What's changed?

- operation in untrustworthy world
  - endpoints can be malicious
  - If endpoint not trustworthy, but want trustworthy network → more mechanism in network core
- more demanding applications
  - end-to-end best effort service not enough
  - new service models in network (Intserv, diffserv)
  - new application-level service architecture built on top of network core (e.g., CDN, p2p)

Rethinking Internet Design

What's changed?

- ISP service differentiation
  - ISP doing more (than other ISPs) in core is competitive advantage
- Rise of third party involvement
  - interposed between endpoints (even against will)
  - e.g., US recording industry
- less sophisticated users

All five changes motivate shift away from end-end!

What's at stake?

"At issue is the conventional understanding of the `Internet philosophy'

- freedom of action
- user empowerment
- end-user responsibility for actions taken
- lack of control "in" the net that limit or regulate what users can do

The end-to-end argument fostered that philosophy because it enables the freedom to innovate, install new software at will, and run applications of the users choice"

[Blumenthal and Clark, 2001]

Technical response to changes

- Trust: emerging distinction between what is "in" network (us, trusted) and what is not (them, untrusted).
  - ingress filtering
- Modify endpoints
  - Harden endpoints against attack
  - Endpoints/routers do content filtering: Net-nanny
  - CDN, ASPs: rise of structured, distributed applications in response to inability to send content (e.g., multimedia, high bw) at high quality

Technical response to changes

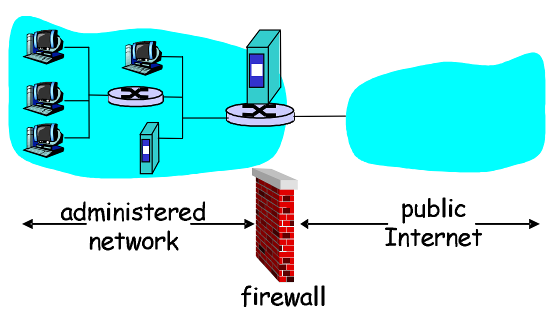
- Add functions to the network core:
  - filtering firewalls
  - application-level firewalls
  - NAT boxes
  - active networking

… All operate within network, making use of application-level information
  - which addresses can do what at application level?
  - If addresses have meaning to applications, NAT must "understand" that meaning

Firewalls

firewall

isolates organization's internal net from larger Internet, allowing some packets to pass, blocking others.



Firewalls: Why

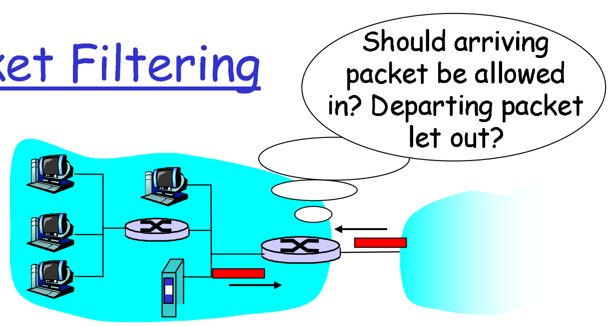
prevent denial of service attacks:
- SYN flooding: attacker establishes many bogus TCP connections, no resources left for "real" connections.

prevent illegal modification/access of internal data.
- e.g., attacker replaces CIA's homepage with something else

allow only authorized access to inside network (set of authenticated users/hosts)

two types of firewalls:
- packet-filtering
- application-level

Packet Filtering

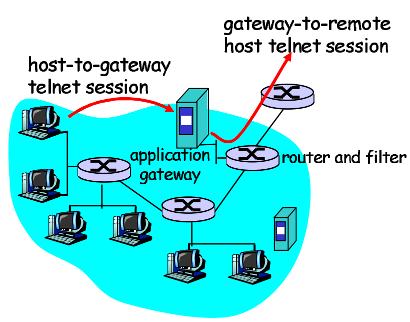


- internal network connected to Internet via router firewall
- router filters packet-by-packet, decision to forward/drop packet based on:
  - source IP address, destination IP address
  - TCP/UDP source and destination port numbers
  - ICMP message type
  - TCP SYN and ACK bits

Packet Filtering

- Example 1: block incoming and outgoing datagrams with IP protocol field = 17 and with either source or dest port = 23.
  - all incoming and outgoing UDP flows and telnet connections are blocked.
- Example 2: Block inbound TCP segments with ACK=0.
  - prevents external clients from making TCP connections with internal clients, but allows internal clients to connect to outside.

Application gateways



- Filters packets on application data as well as on IP/TCP/UDP fields.
- Example: allow select internal users to telnet outside.

1. Require all telnet users to telnet through gateway.
2. For authorized users, gateway sets up telnet connection to dest host. Gateway relays data between 2 connections
3. Router filter blocks all telnet connections not originating from gateway.

NAT: Network Address Translation

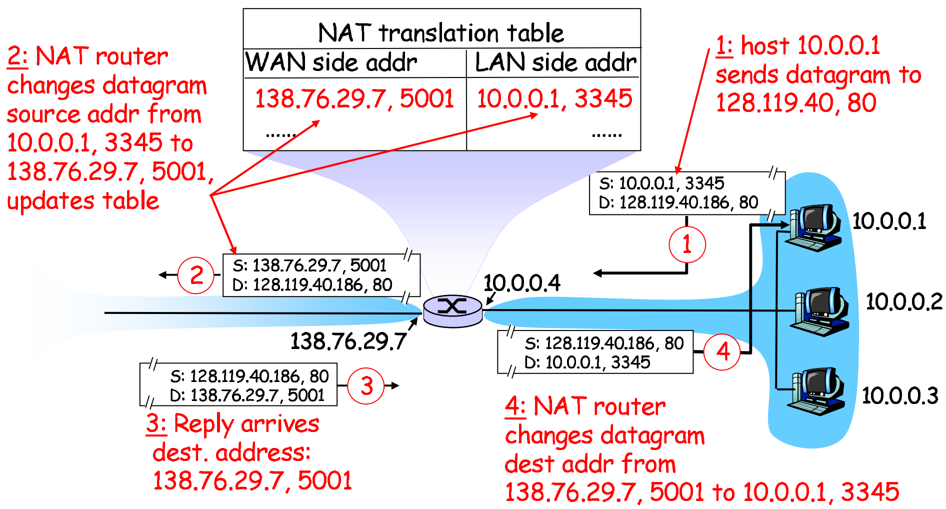
- Motivation: local network uses just one IP address as far as outside word is concerned:
  - no need to be allocated range of addresses from ISP: just one IP address is used for all devices
  - can change addresses of devices in local network without notifying outside world
  - can change ISP without changing addresses of devices in local network
  - devices inside local net not explicitly addressable, visible by outside world (a security plus).

NAT: Network Address Translation

Implementation: NAT router must:

- *outgoing datagrams: replace* (source IP address, port #) of every outgoing datagram to (NAT IP address, new port #)
  . . . remote clients/servers will respond using (NAT IP address, new port #) as destination addr.
- *remember (in NAT translation table)* every (source IP address, port #) to (NAT IP address, new port #) translation pair
- *incoming datagrams: replace* (NAT IP address, new port #) in dest fields of every incoming datagram with corresponding (source IP address, port #) stored in NAT table

NAT: Network Address Translation



NAT: Network Address Translation

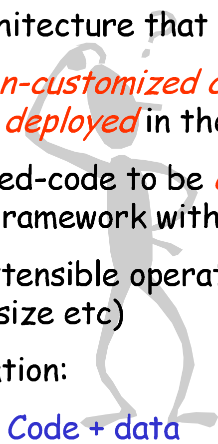
- 16-bit port-number field:
  - 60,000 simultaneous connections with a single LAN-side address!
- NAT is controversial:
  - routers should only process up to layer 3
  - violates end-to-end argument
    - NAT possibility must be taken into account by app designers, eg, P2P applications
  - address shortage should instead be solved by IPv6

What is an Active Network ?

- Depends on who you ask!
- active services: application-level services exploiting position within the network to provide enhanced service
  - CDN
  - streaming media caches
- capsule approach: packets carry programs, active node executes program when code-carrying packet arrives to active node
  - code may determine what to do with packet
  - may implement other service: e.g., network management, reliable multicast

Part 3 3-41

The capsule approach to active networks



- Network architecture that allows :
  - *application-customized code to be dynamically deployed* in the network
  - customized-code to be *executed* in controlled framework within network
- Similar to extensible operating systems (SPIN, Synthesize etc)
- the new equation:

  Packet = Code + data

  sort of like postscript

Part 3 3-42

Capsules

ANTS Header (spans Version through Dep fields)

IP header	Version	Type	Previous Address	Dep fields	Payload

- **Type**
  - Identifier for the forwarding routine to be executed (carries code by reference)
- **Previous address**
  - Where to get the forwarding routine from if it is not available in the present node (***Code Distribution)***
- **Dependent Fields**
  - Parameters for the forwarding code
- **Payload**
  - Header + data of higher layers

Active networking and End-to-End Arguments

- End-to-end principle: lower layers should have minimum functionality, but support widest variety of applications possible
  - active networking: support *all* higher-level applications
  - minimum common functionality: ability to execute code: *programmable versus pre-programmed* low layer functionality

Active networking: transparency?

- Transparency: use of network by others not very visible (can more or less predict behavior of network)
- Active networking: transparency difficult
  - constrain interactions among programmable entities in router (who knows *what* they will try to do)
  - like OS trying to constrain interaction among processes!

KISS

- success of LAN protocols, RISC architecture: KISS!
- "building complex functions into network optimizes network for small number of services, while substantially increasing cost for uses unknown at design time"
- "end-to-end argument does not oppose active networks per se but instead strongly suggests that enthusiasm for the benefits of optimizing current application needs by making the network more complex may be misplaced"

Epilogue: will IP take over the world?

Reasons for success of IP:

- *reachability:* reach every host, adapts topology when links fail.
- *heterogeneity:* single service abstraction (best effort) regardless of physical link topology

Note: our grading for these: A-, A

many other claimed (or commonly accepted) reasons for IP's success may not be true

…. let's take a closer look

1. IP already dominates global communications?

- business revenues:
  - ISPs: 13B
  - Broadcast TV: 29B
  - Cable TV: 29.8B
  - Radio broadcast: 10.6B
  - Phone industry: 268B

Q: IP equipment cheaper?
Economies of scale?
(lots of routers?)

Q: per-device, IP is cheaper
(one line into house, multiple devices)

Q: # bits carried in each network?

Q; Internet, more traffic and congestion is spread among all users (bad?)

- Router/telco switch markets:
  - Core router: 1.7B; edge routers: 2.4B
  - SONET/SDH/WDM: 28B, Telecom MSS: 4.5B

2. IP is more efficient?

- Statistical multiplexing versus circuit switching
- Link utilization:
  - Avg. link utilization in Internet core: 3% to 30% (ISPs: never run above 50%)
  - Avg. utilization of Ethernet is currently 1%
  - Avg. link utilization of long distance phone lines: 33%
- low IP link utilization: purposeful!
  - predictability, stability, low delay, resilience to failure
- At low utilization, we *forfeit* benefits of statistical multiplexing!

3. IP is more robust?

- Median IP network availability: downtime: 471 min/yr
- Avg. phone network downtime: 5 min/yr
- Convergence time with link failures:
  - BGP: 3 – 15 minutes
    - but ~ 1 sec with intra-domain, e.g., OSPF
  - SONET: 50 ms
- Inconsistent routing state
  - human misconfigurations
  - in-band signaling (signaling and data share same network)
  - routing computation "complex"

4. IP is simpler?

- Intelligence at edge, simplicity in core
  - Cisco IOS: 8M lines of code
  - Telephone switch: 3M lines of code
- Line-card complexity:
  - Router: 30M gates in ASICs, 1 CPU, 300M packet buffers
  - Switch: 25% of gates, no CPU, no packet buffers

5. Support of real-time app's telephony over IP

- Not yet

Discussion: benefits of IP?

- IP supports many different types of data applications at a wide range of data rates
  - phone network: 1 or many services (voice, fax, touch-tone service, 800 numbers, teletype, hearing impaired services, lots of enhanced voice services, voicemail…)
- what about ATM (versus IP versus phone)
- IP traffic, services more diverse (?). IP works at higher bandwidths (factually true for end applications, but cores are both high speed)
- Claim: IP supports short bursty connections "better" (implicit: less setup cost, less resources used – not that important given utilization figures)
- IP has 1 rtt transaction times, phone network is at least 2 rtt (setup plus transaction)