Network Overlays.

Ken Birman

Cornell University. CS5410 Fall 2008.

Network Overlays

- Consider the Internet
  - It creates the illusion of a fully connected *n x n* world of addressable endpoints
  - In reality, packets must route through a complex infrastructure, but the end user doesn't see that infrastructure
- Overlay concept takes this one step further
  - We focus on some application… and create a dedicated personal internet just for it
  - The dedicated network might have special properties

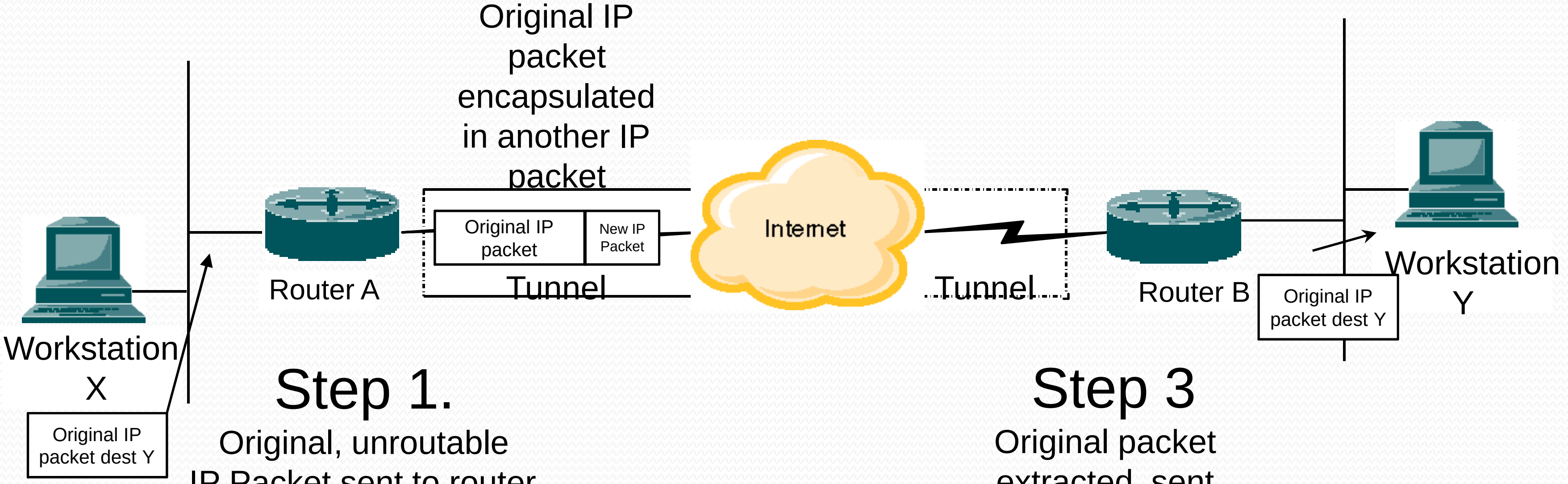
Uses of overlays

- Load balancing, other forms of quality of service
- Distributing files or data down some form of tree structure (allows massive fanouts without forcing any single node to send huge numbers of copies)
- Route around congestion
- *Content routing:* packets routed on the basis of the data inside them (could look at fields, or might do a whole xquery)
- *Publish subscribe:* packets route on the basis of topic
- *DHT*: In fact, even a DHT is an overlay!

Early Overlays

- The first overlays were really Internet "tunnels"
  - Idea was to encapsulate IP packets in some other network standard
  - … then route them over a link that used non-IP technology
  - … then unpack them and drop them back into IP-land
- Then we started to see fancier tunnels
  - IP multicast over TCP
  - IPv6 over IPv4

Tunneling Illustrated



Step 2
Original IP packet encapsulated in another IP packet

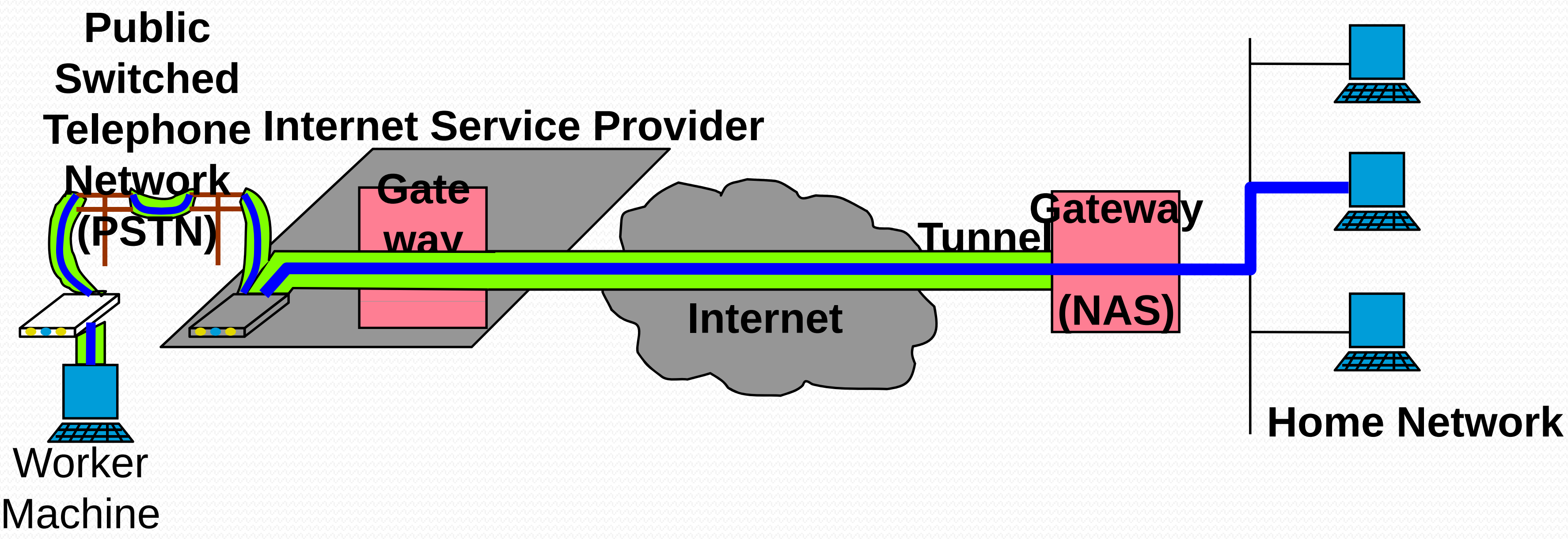
Step 1.
Original, unroutable IP Packet sent to router

Step 3
Original packet extracted, sent to destination

Widely known overlays

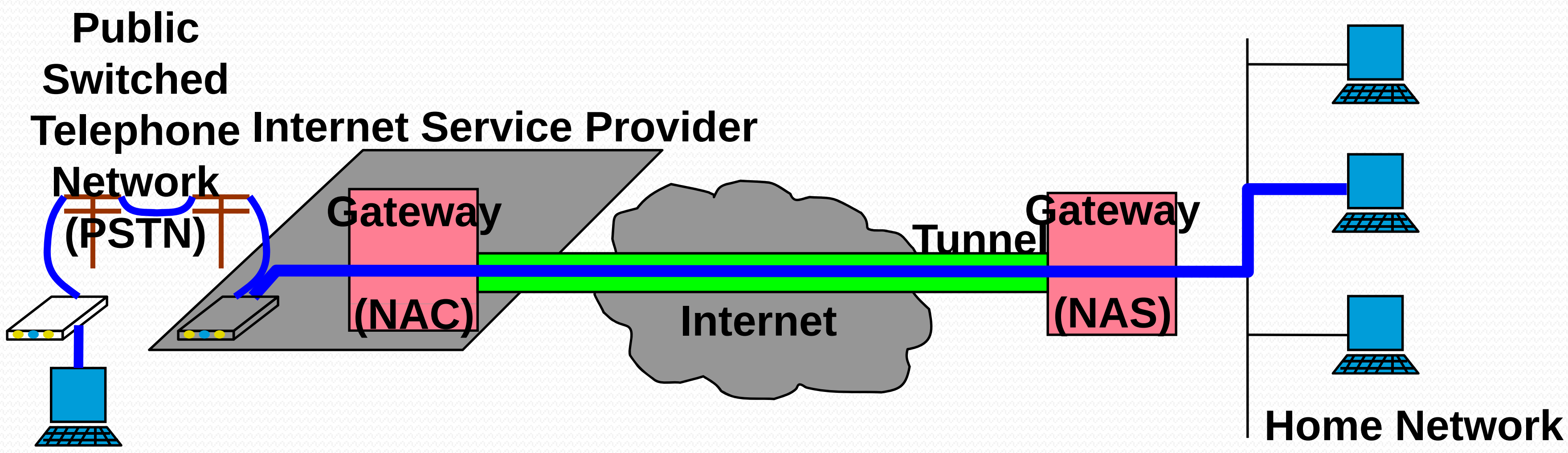
- Virtual private networks
  - End point computers need to have some form of certificate that they use to identify themselves
    - Typically: each machine has a private key and a public key
    - With this can send “unforgeable” encrypted data
    - So: edge machine authenticates itself to the VPN server, which sends back the current secret key of the VPN (a symmetric key)
    - The edge machine tunnels traffic encrypted with the VPN key via the VPN server, which acts as a router

Virtual Dial-up Example (1)



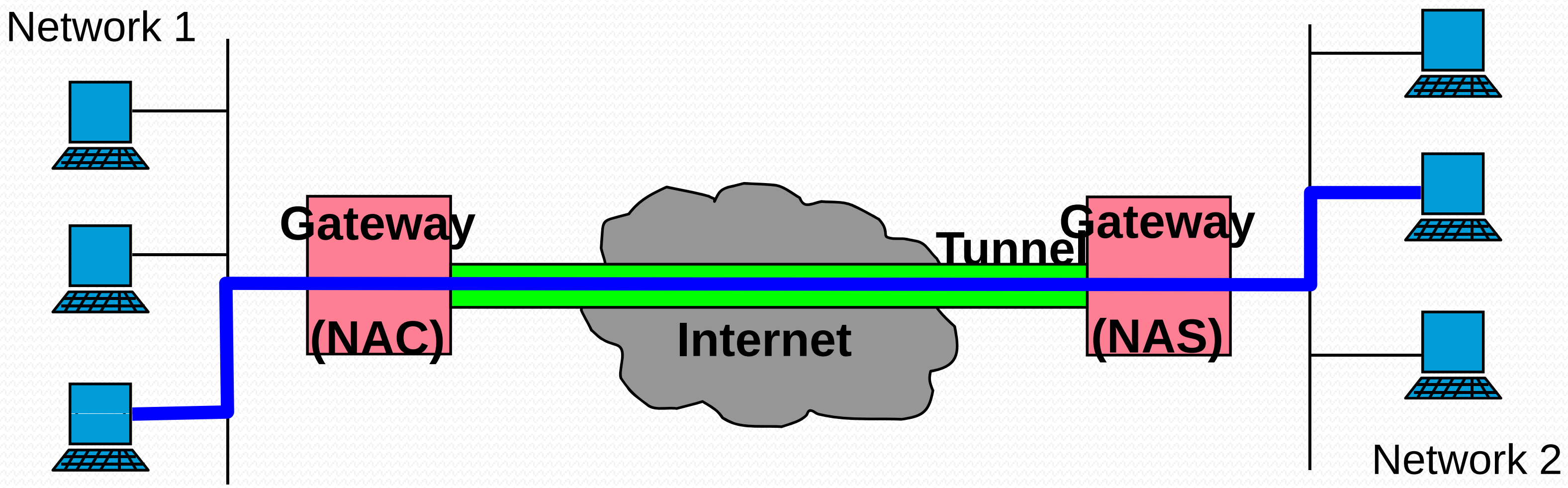
- Worker dials ISP to get basic IP service
- Worker creates his own tunnel to Home Network

Virtual Dial-up Example (2)


Public Switched Telephone Network (PSTN)
Internet Service Provider
Gateway
(NAC)
Internet
Tunnel
Gateway
(NAS)
Home Network

- Remote worker connects to Home Network through ISP created tunnel
- Allows wholesale dial-up

Logical Network Creation



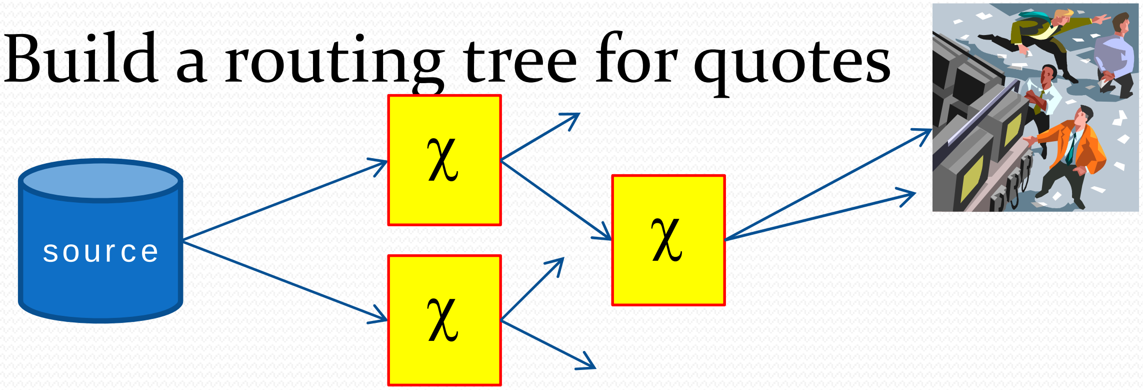
- Remote networks 1 and 2 create a logical network
- Secure communication at lowest level

Other uses for overlays

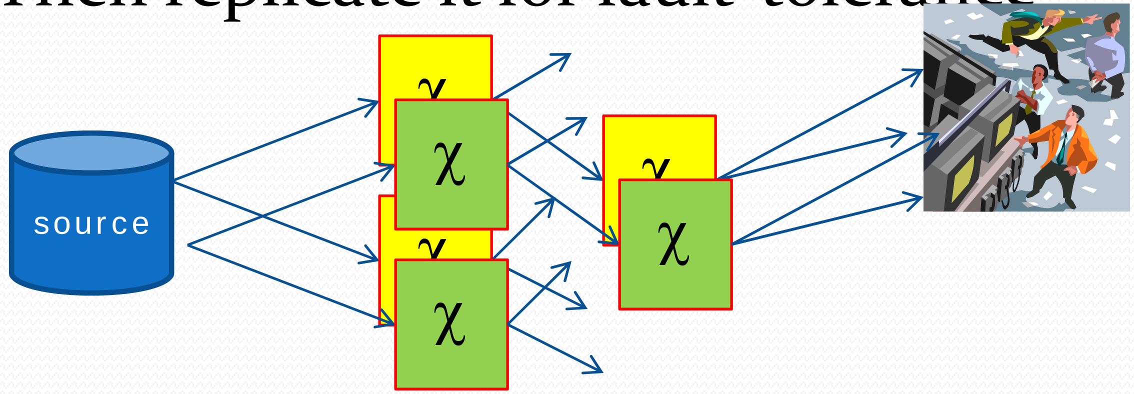
- New York Stock Exchange Quote Distribution System
  - Built around 1995
  - Issue: needed a customizable way to route quotes to overhead displays over internal network
  - Required fault-tolerance
  - Content sources ran at higher speeds than most display end systems could sustain

Basic idea…

- Build a routing tree for quotes



- Then replicate it for fault-tolerance

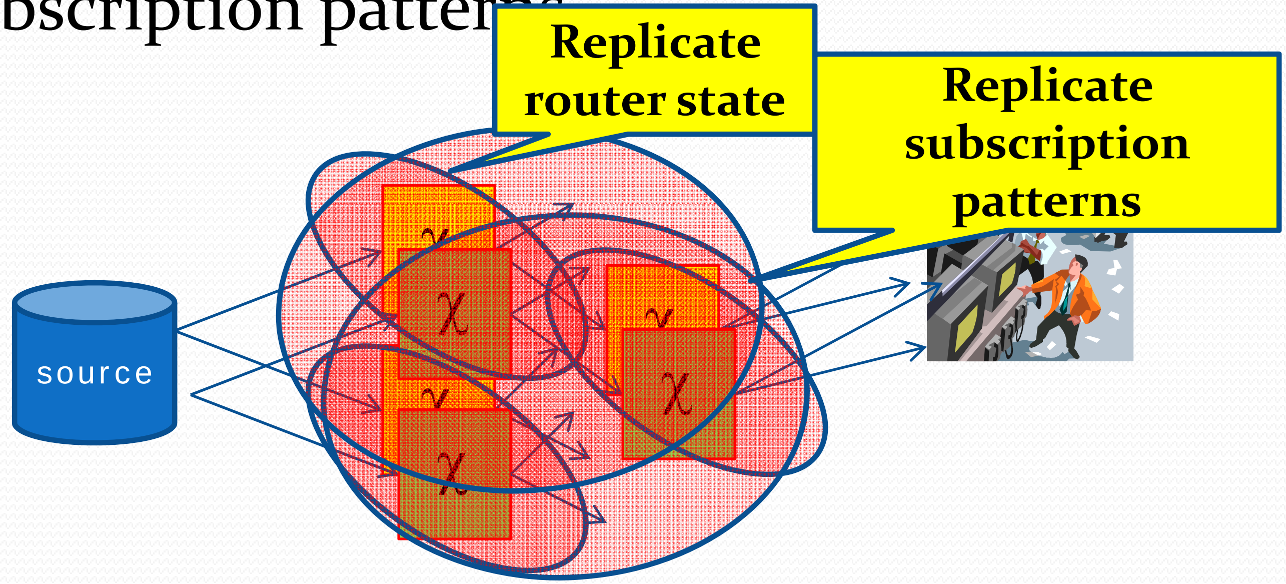


Components

- The source systems were the five or six “clearing” machines used by the NYSE to capture trades, bids, offered prices
- The routers were inexpensive dedicated computers with dual ethernet cards, one for each network
- Each network was a separate ethernet with distinct IP addresses and no automated routing
- The overhead displays were basically workstations

Fault-tolerance

- They used a virtual synchrony package (Isis) to replicate state within router pairs, and to track subscription patterns



Replicate router state

Replicate subscription patterns

- … lots of groups

Why an overlay?

- Isis wasn’t capable of supporting very large groups with very high data rates
  - So sending the actual trades/quotes wasn’t feasible
- Total number of routers was about 75… serving 1000 or more display systems

- By building a TCP-based overlay and using the Isis groups “out of band”, Isis wasn’t on the critical path
- Isis knew about the dual IP network… TCP didn’t.

Outcome?

- The solution was completely robust and was used from 1995 until mid 2006
  - During that decade there were many failures and even entire network outages
  - But the NYSE “rode them all out” absolutely unperturbed: traders saw no glitches at all

- So here the overlay plays two roles
  - Overlay carries the heavy communication burden
  - One overlay for each IP network

Resilient Overlay Networks

Ron Slides

http://nms.lcs.mit.edu/ron/

Final example for today: P6P

- Research by Li Dong Zhou and Van Renesse
- Issue addressed by this work
  - People want to use IPv6
  - But the Internet itself is locked into IPv4
- So idea is to support IPv6 as an overlay

- Features of IPv6?
  - Very long addresses (64 bits)
  - Address doesn’t reveal location (unlike IPv4)

How P6P works

- Assumes two worlds
  - An IPv6 world, invisible to them
  - An IPv4 world, where P6P lives
- Some IPv6 nodes live in both, call them “internal gateway nodes”
  - These have both an IPv6 and an IPv4 address
  - P6P itself implemented by what they call “external gateway” nodes that run in the IPv4 network

How P6P works

- They designed a DHT based on Chord
- Each IPv6 node must have an associated IG
  - So treat the (IPv6,IPv4) tuple as a (key,value) pair!
- IPv6 address is an index into Chord
  - New IPv6 node would create a new (key,value) pair
  - To send an IPv6 packet, look up the IPv4 helper node, then forward the IPv6 packet to the helper
  - Cache information for reuse
  - Plus many optimizations, and a security architecture…

How well does it work?

- They designed a detailed simulation and looked at random traffic (perhaps unrealistic…)
- In this model, P6P performed extremely well
  - Rapid routing
  - Fairly quick response when mobile nodes changed their associated IG node
    - Some false routing, but then automatically recovers
- Seems to be a very practical way to roll IPv6 out…

Summary: Overlays

- We’ve seen a few examples

- VPNs very widely used, origin of the whole idea
- RON is perhaps the most debated
  - Is RON “contrary to the end-to-end spirit of Internet”?
  - If RON becomes popular, will it break down?

- P6P illustrates how overlays can work-around a huge political question (“should we move to IPv6”?)