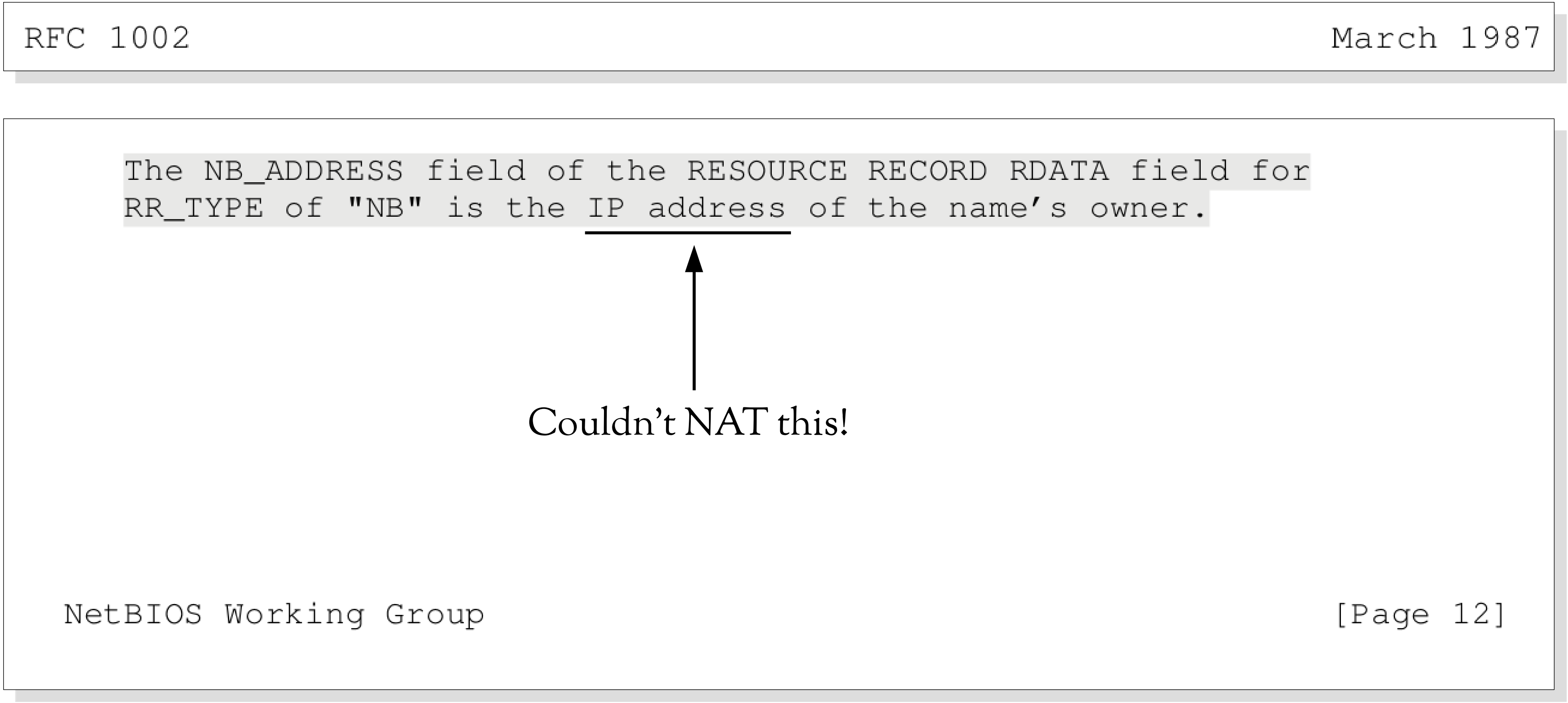
The Trouble with NAT

(Or why I care about IPv6)

AusNOG Conference
2016

Mark Smith
markzzzsmith@gmail.com
@markzzzsmith

A Sad Story from 1996



RFC 1002 March 1987

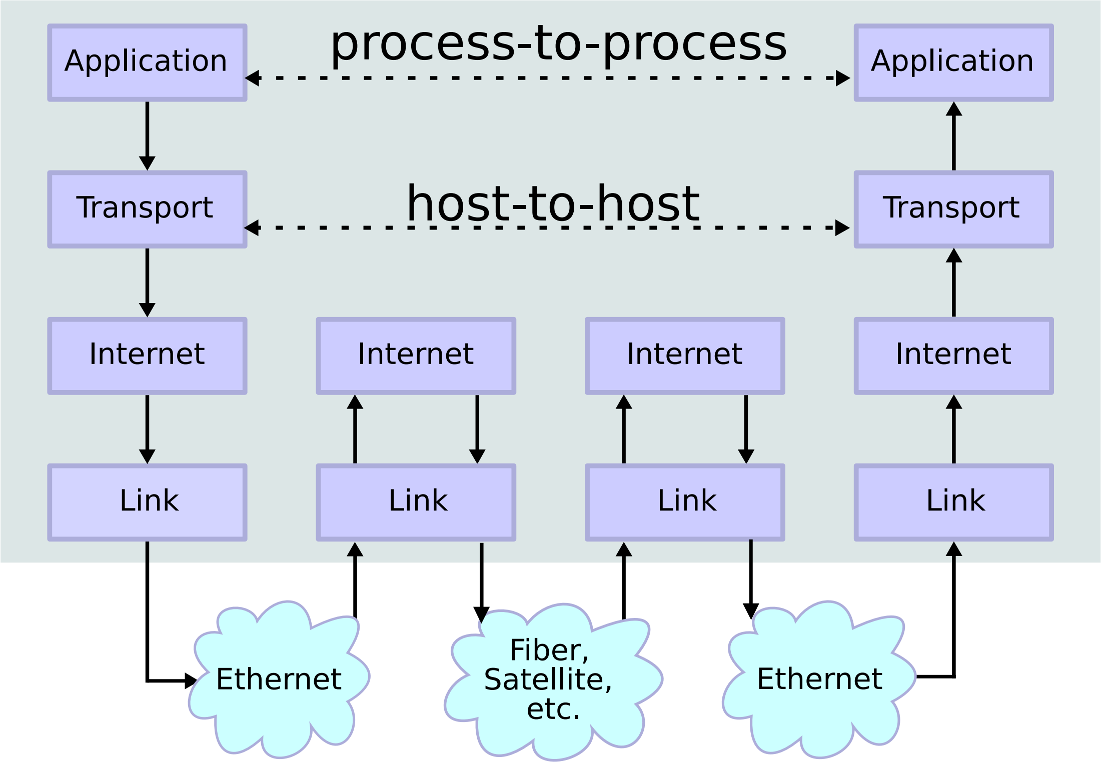
The NB_ADDRESS field of the RESOURCE RECORD RDATA field for
RR_TYPE of "NB" is the IP address of the name's owner.

Couldn't NAT this!

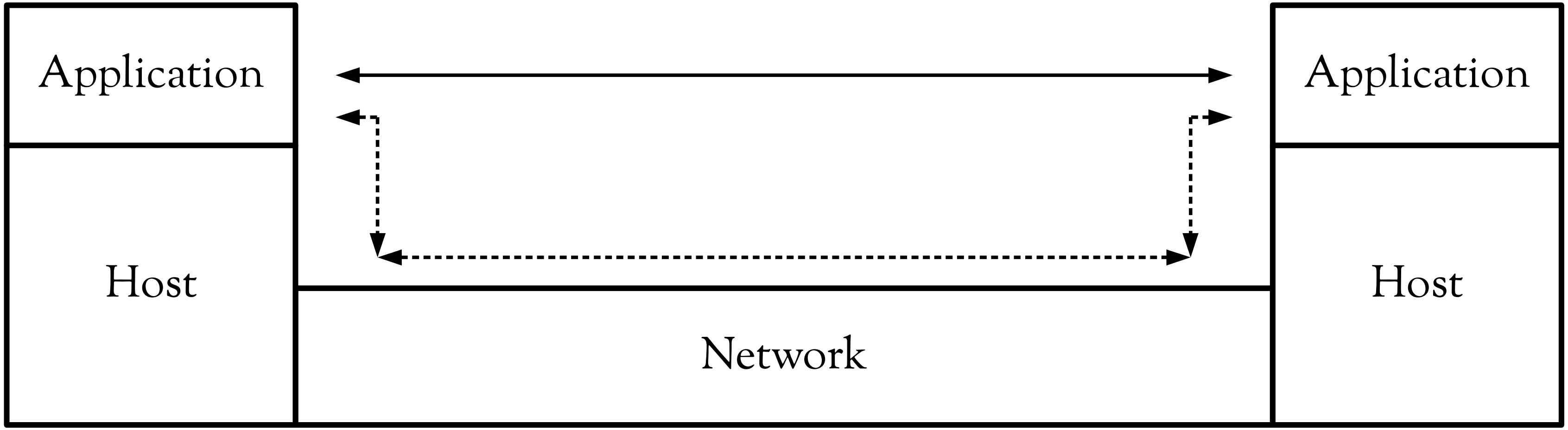
NetBIOS Working Group [Page 12]

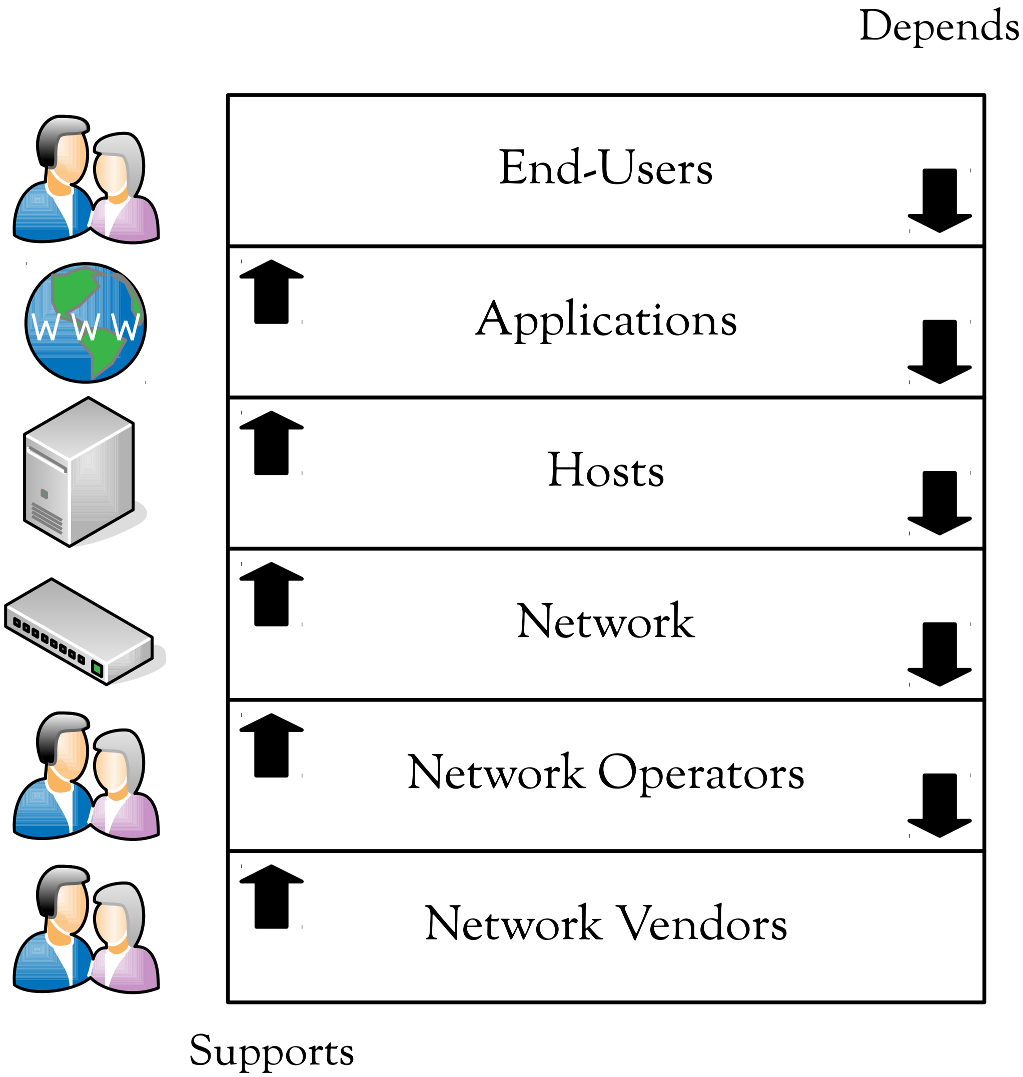
On Being a Network Operator

Data Flow



https://commons.wikimedia.org/wiki/File:IP_stack_connections.svg, Author: Kbrose, Permission: GDFL


Application
Application
Host
Host
Network



Depends

End-Users

Applications

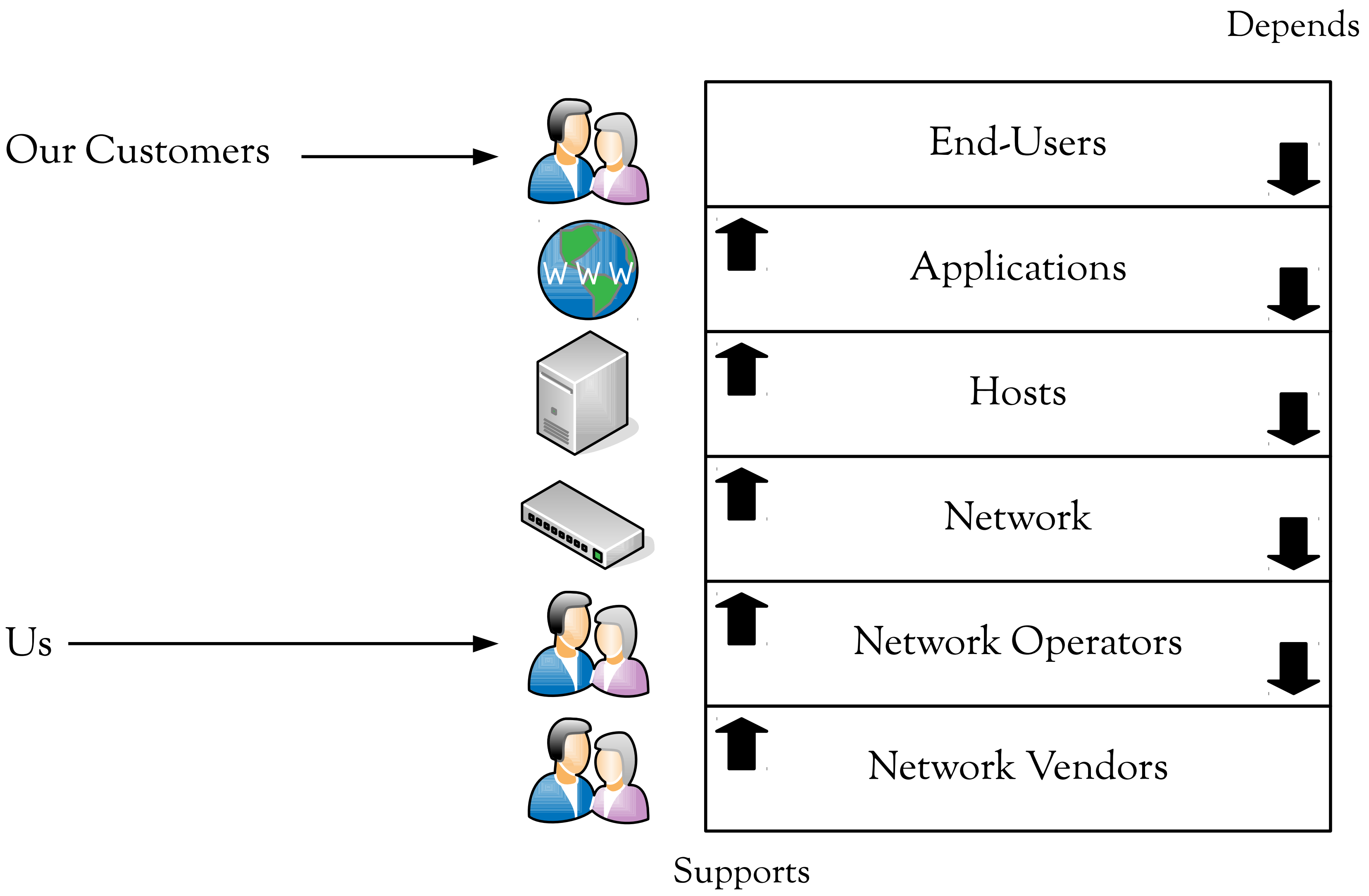
Hosts

Network

Network Operators

Network Vendors

Supports



Depends

Our Customers

End-Users

Applications

Hosts

Network

Us

Network Operators

Network Vendors

Supports

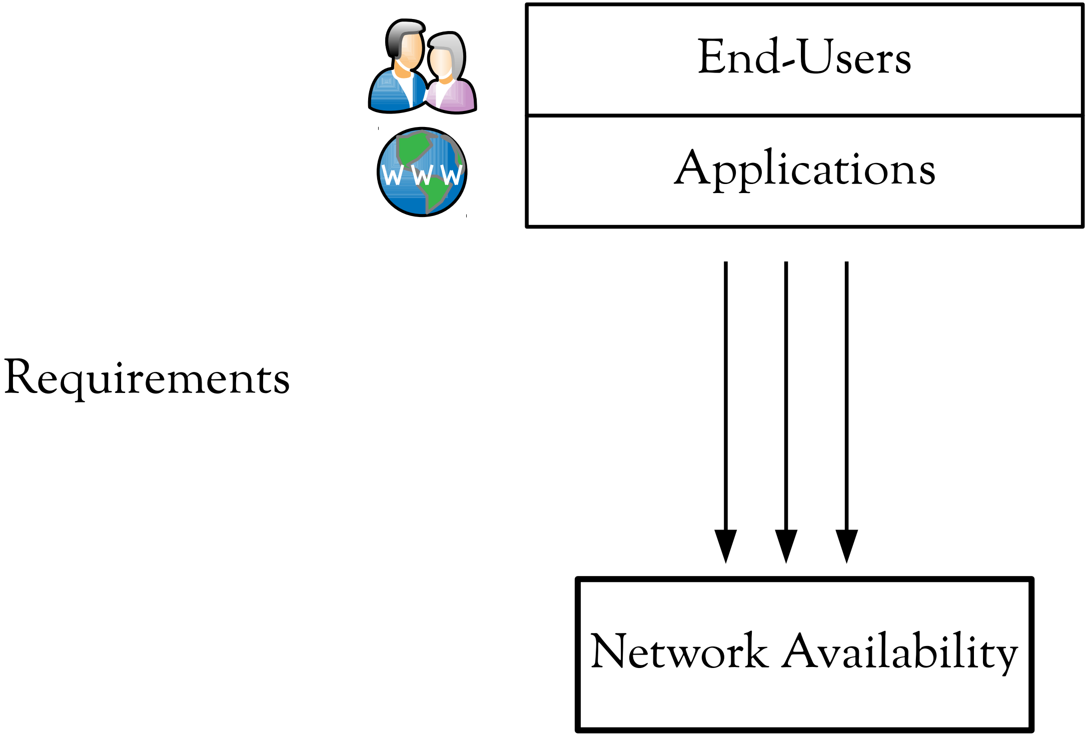
Our Mission

Or,

Network Critical Success Factors (NCSFs)

Available

https://flic.kr/p/81SRQG; Author: Daniel Weir; No Mods; Lic:CC BY-NC 2.0



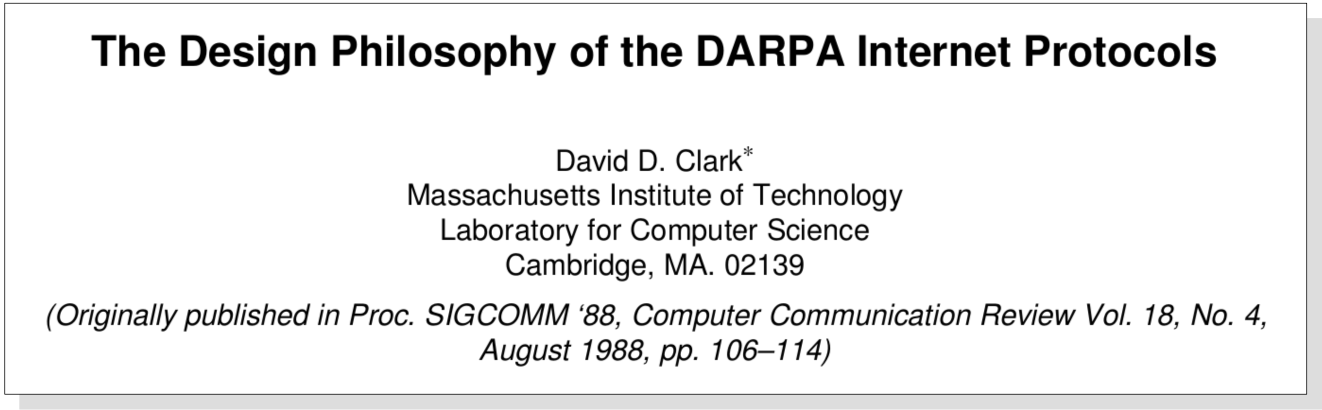
End-Users

Applications

Requirements

Network Availability

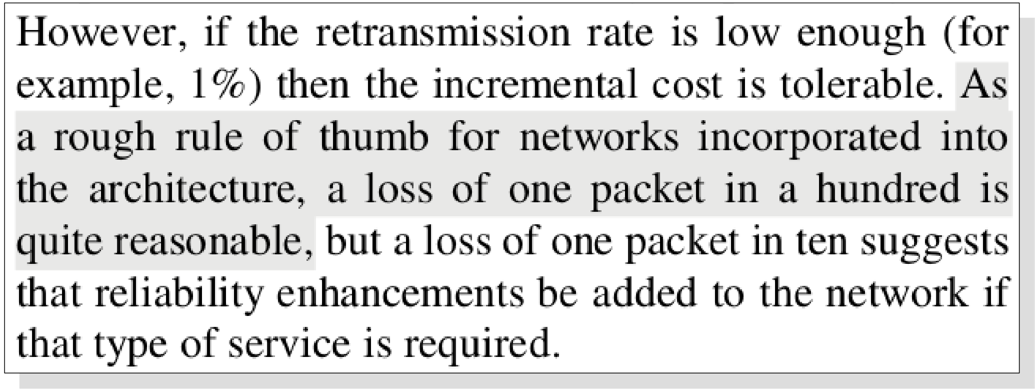
Packets sent by Hosts should have a good Probability of Arriving at the destination Host within an Acceptable Timeframe.



The Design Philosophy of the DARPA Internet Protocols

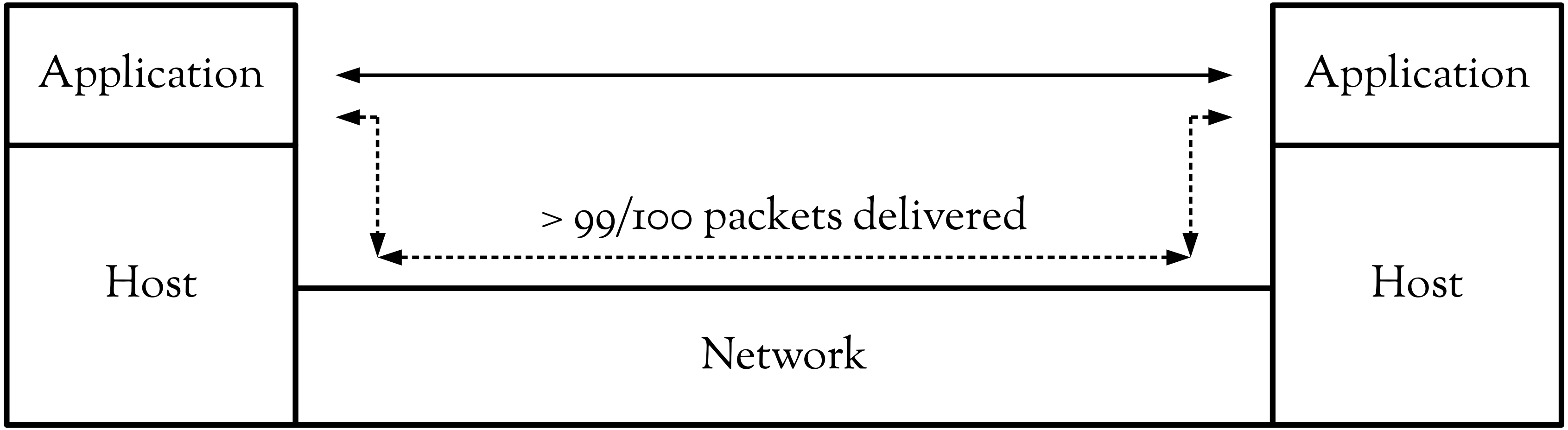
David D. Clark*
Massachusetts Institute of Technology
Laboratory for Computer Science
Cambridge, MA. 02139

(Originally published in Proc. SIGCOMM '88, Computer Communication Review Vol. 18, No. 4, August 1988, pp. 106–114)

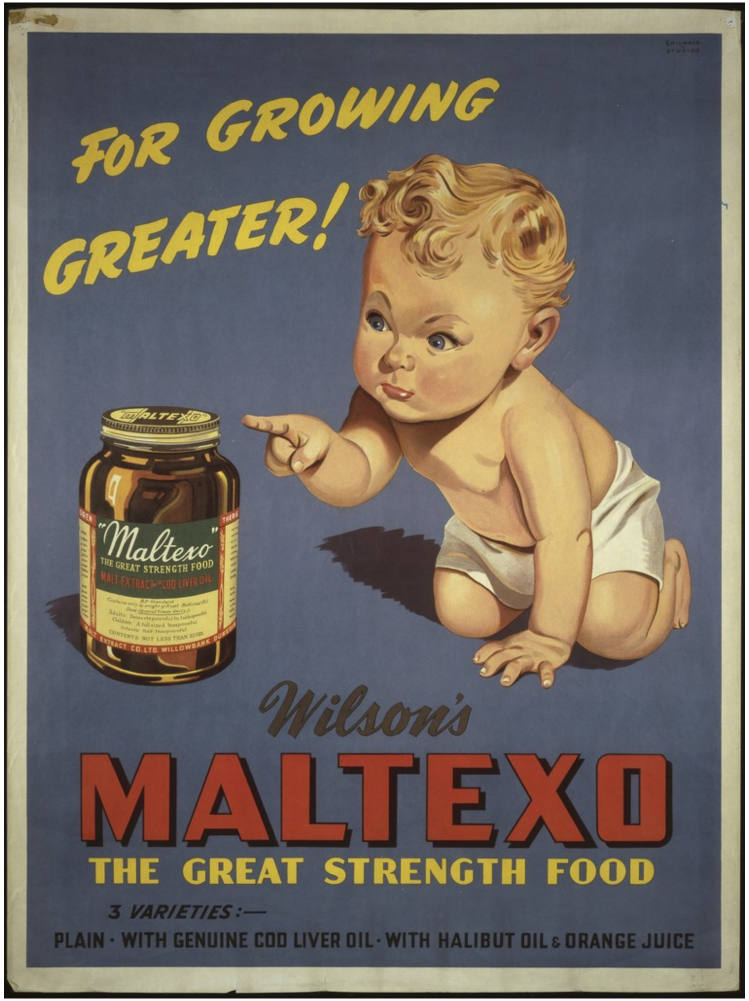


However, if the retransmission rate is low enough (for example, 1%) then the incremental cost is tolerable. As a rough rule of thumb for networks incorporated into the architecture, a loss of one packet in a hundred is quite reasonable, but a loss of one packet in ten suggests that reliability enhancements be added to the network if that type of service is required.

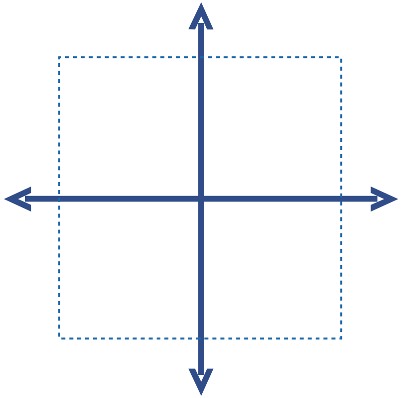
http://ccr.sigcomm.org/archive/1995/jan95/ccr-9501-clark.pdf


Application
Application
> 99/100 packets delivered
Host
Host
Network

Scalable


RAILWAYS STUDIOS
FOR GROWING GREATER!
"MALTEXO"
"Maltexo"
THE GREAT STRENGTH FOOD
MALT EXTRACT AND COD LIVER OIL
B.P. Standard
Contains 20% by weight of Finest Medicinal Oil
Dose (several times daily):
Adults: Dessertspoonful to tablespoonful
Children: A full sized teaspoonful
Infants: Half-teaspoonful
CONTENTS NOT LESS THAN 32 OZS.
MALT EXTRACT CO. LTD. WILLOWBANK DUNEDIN
Wilson's
MALTEXO
THE GREAT STRENGTH FOOD
3 VARIETIES:—
PLAIN · WITH GENUINE COD LIVER OIL · WITH HALIBUT OIL & ORANGE JUICE

https://flic.kr/p/aAu2Py



Scaling Dimensions

Network Elements (Routers / Switches)

Links

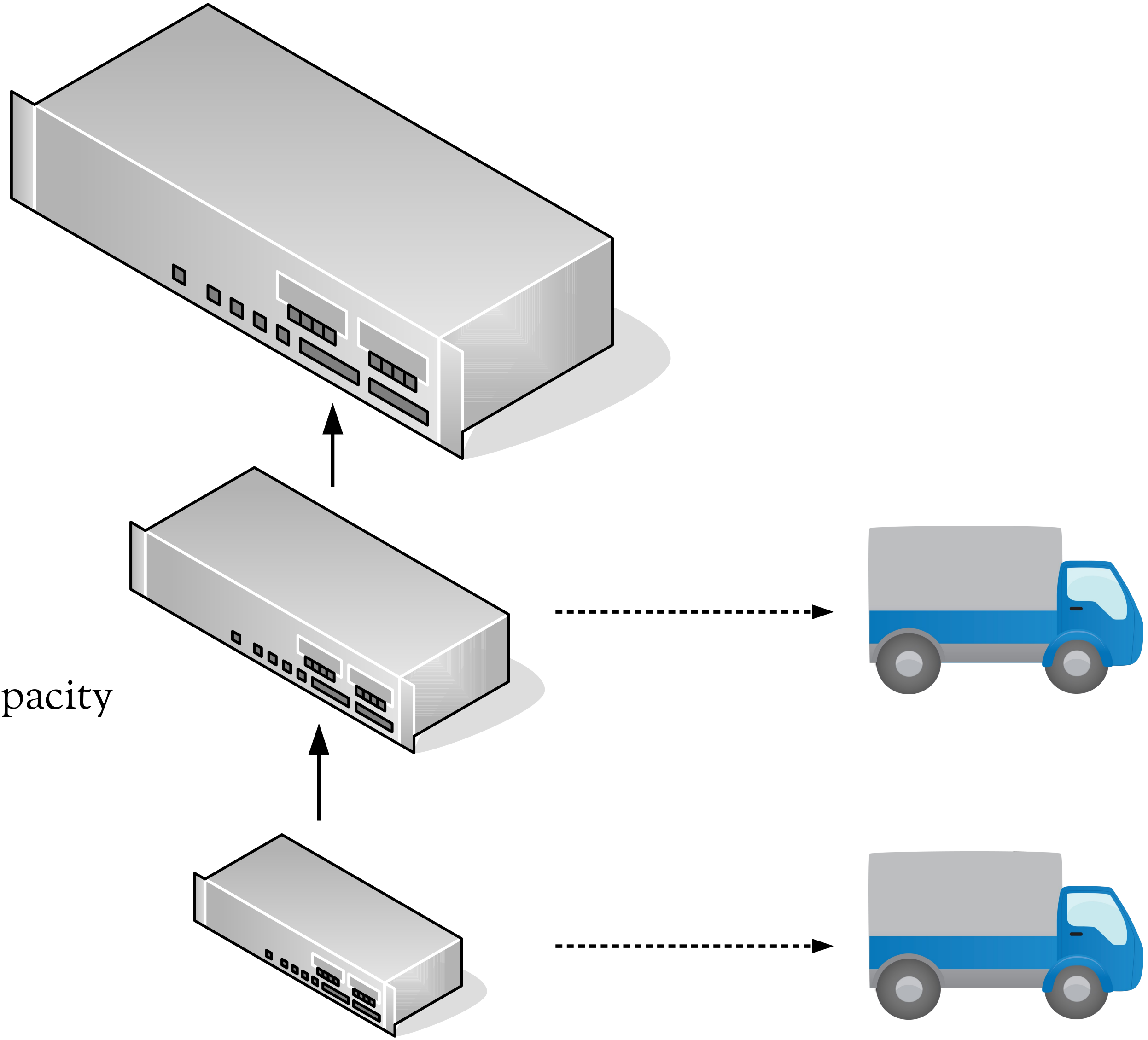
Network Capacity

Hosts

Geographic Sites

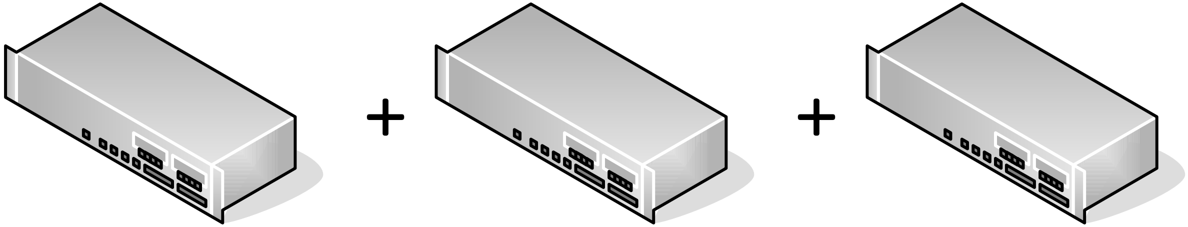
Vertical Scaling

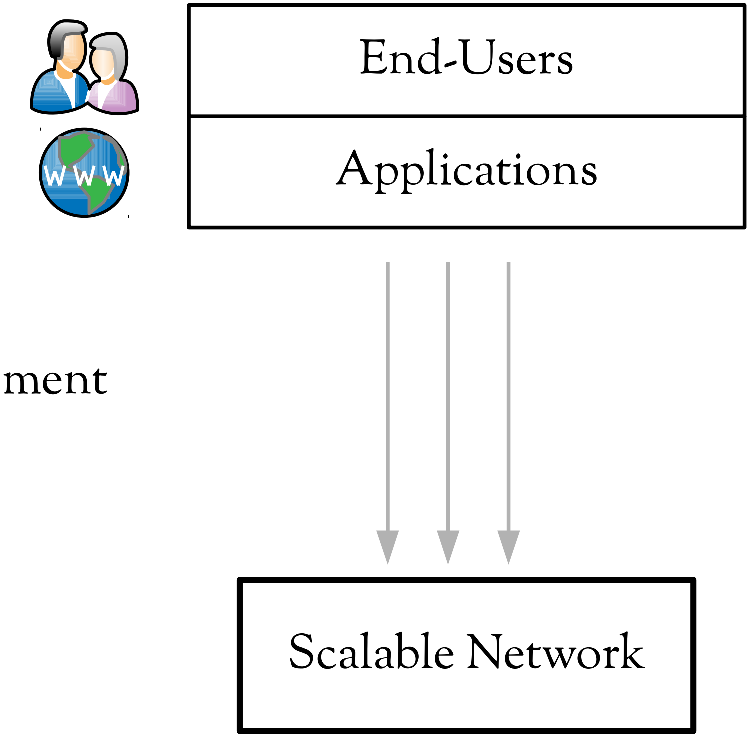
- “Scaling Up”
- Need to replace existing capacity while adding new capacity
- Using a bigger hammer!



Horizontal Scaling

- “Scaling Out”
- Adding new capacity to existing capacity
- No capacity replacement!
- Divide-and-conquer!





End-Users

Applications

Inherent requirement

Scalable Network

Adequately Performing

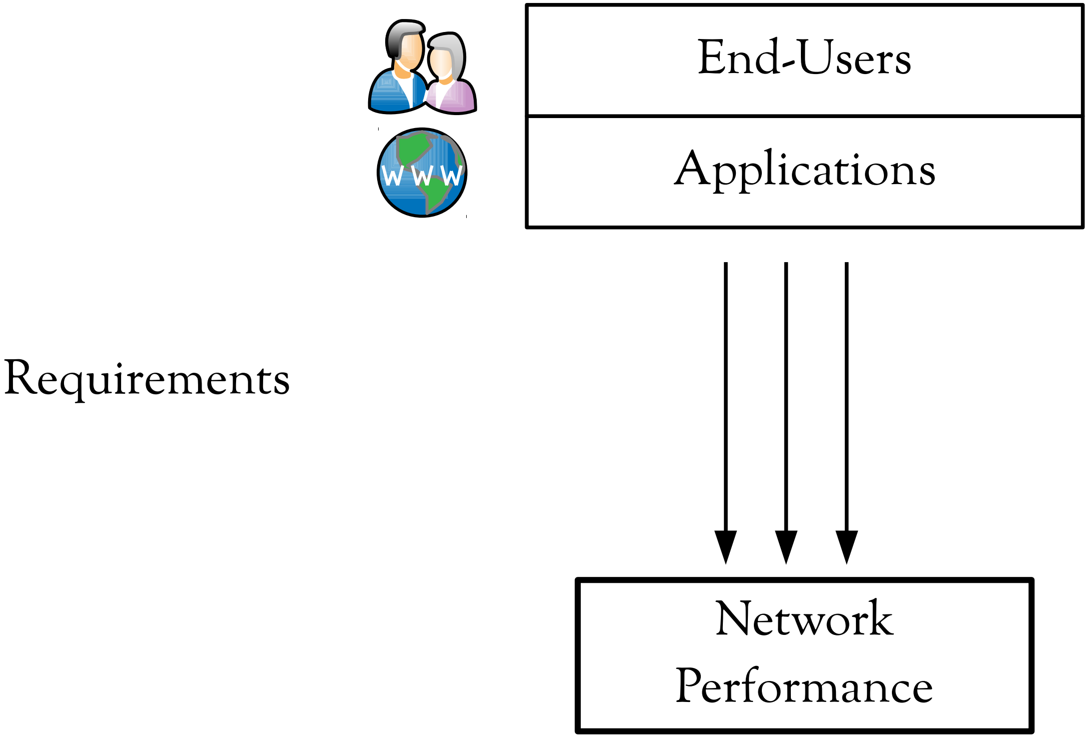
Adequate network:

Throughput

Latency

Packet Delivery Success

Packet Order



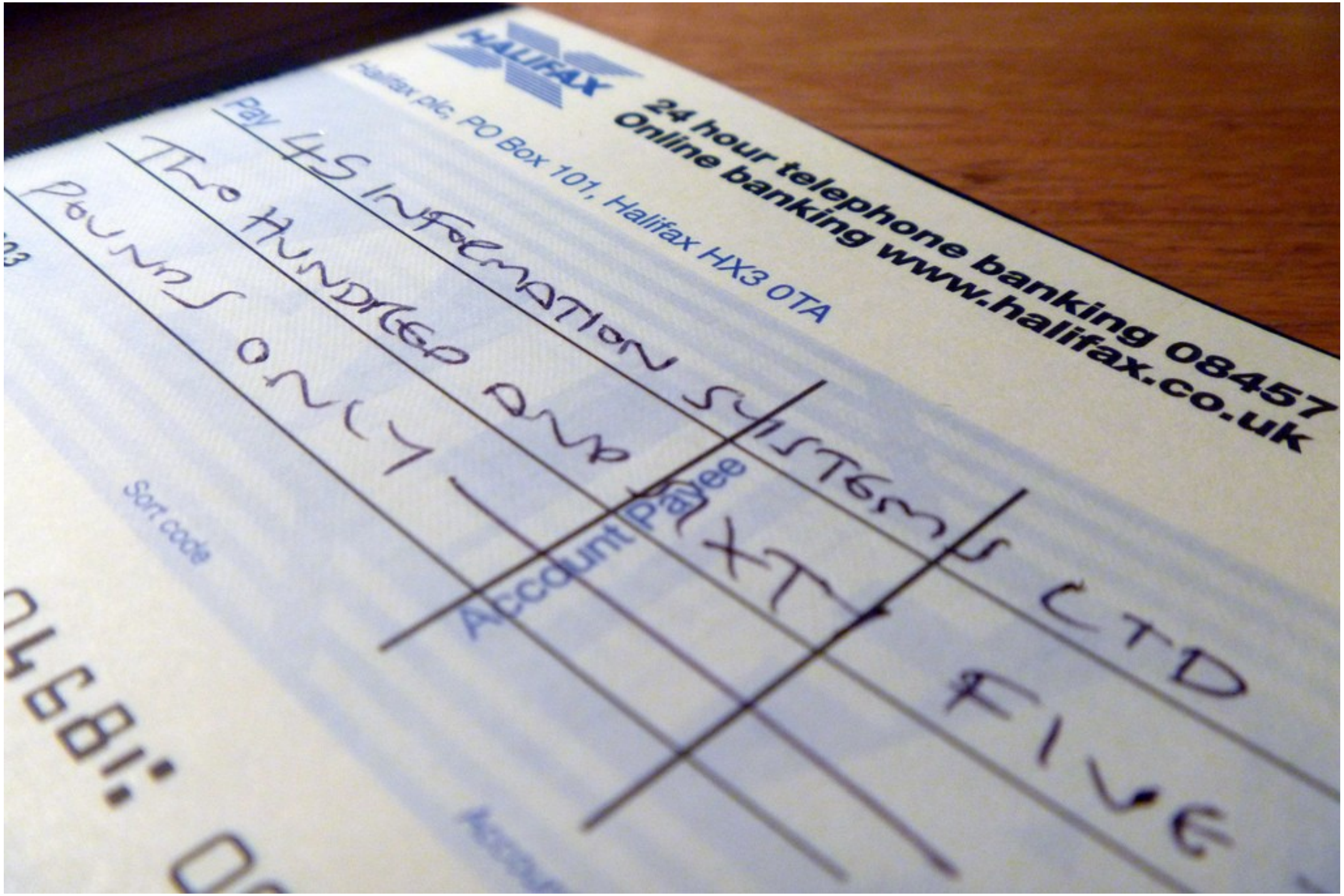
End-Users

Applications

Requirements

Network Performance

Constrained by Budget



HALIFAX

24 hour telephone banking 08457
Online banking www.halifax.co.uk

Halifax plc, PO Box 101, Halifax HX3 0TA

Pay 4S INFORMATION SYSTEMS LTD

Two HUNDRED AND SIXTY FIVE

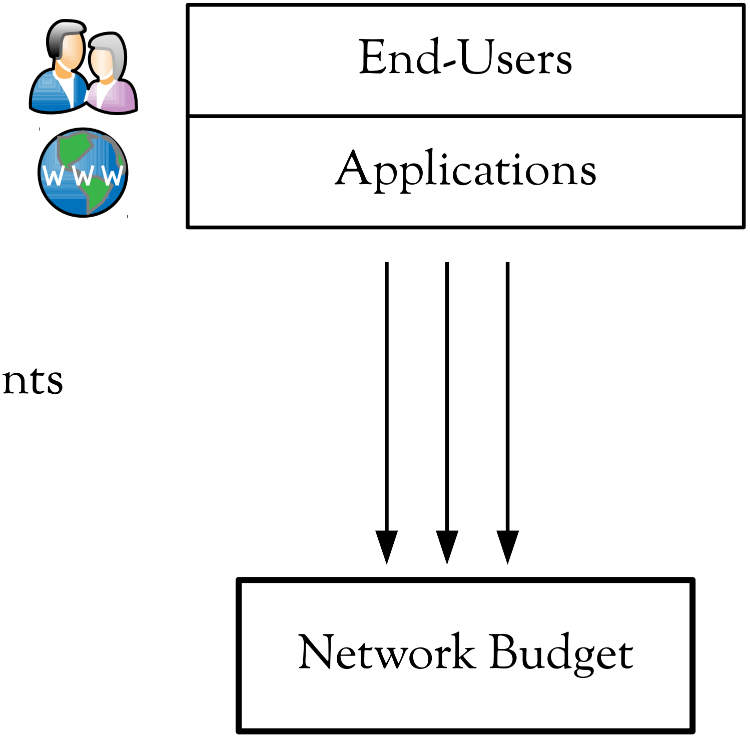
POUNDS ONLY —

Account Payee

Sort code

04681: 0

https://flic.kr/p/akxa03



End-Users

Applications

Network Budget

Budget Constraints

Available

Scalable

Adequately
Performing

Budget

https://flic.kr/p/8a9uyV; Author: ph-stop; No Mods; Lic: CC BY-SA 2.0

Available >

Scalable >

Performance?

Performance means nothing if you crash!

The Trouble with NAT

“NAT”

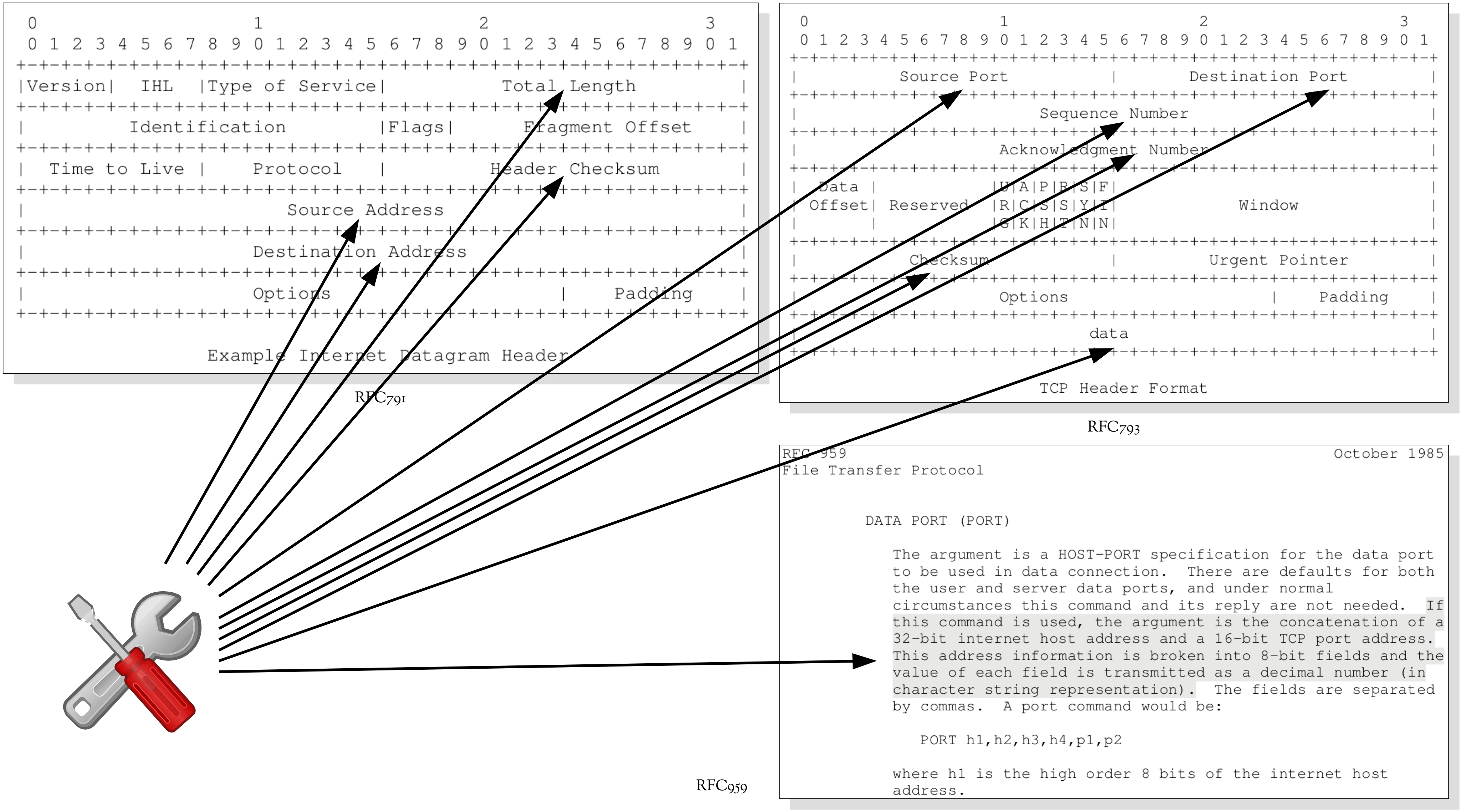
Basic NAT –

one:one address translation

Network Address Port Translation (NAPT) –

many:one address translation

RFC2663



```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Version|  IHL  |Type of Service|          Total Length         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Identification        |Flags|      Fragment Offset    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Time to Live |    Protocol   |         Header Checksum       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Source Address                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Destination Address                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Options                    |    Padding    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

                  Example Internet Datagram Header
```

RFC791

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Source Port          |       Destination Port        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Sequence Number                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Acknowledgment Number                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Data |           |U|A|P|R|S|F|                               |
| Offset| Reserved  |R|C|S|S|Y|I|            Window             |
|       |           |G|K|H|T|N|N|                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Checksum            |         Urgent Pointer        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Options                    |    Padding    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             data                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

                        TCP Header Format
```

RFC793

```
RFC 959                                                     October 1985
File Transfer Protocol


         DATA PORT (PORT)

            The argument is a HOST-PORT specification for the data port
            to be used in data connection.  There are defaults for both
            the user and server data ports, and under normal
            circumstances this command and its reply are not needed.  If
            this command is used, the argument is the concatenation of a
            32-bit internet host address and a 16-bit TCP port address.
            This address information is broken into 8-bit fields and the
            value of each field is transmitted as a decimal number (in
            character string representation).  The fields are separated
            by commas.  A port command would be:

               PORT h1,h2,h3,h4,p1,p2

            where h1 is the high order 8 bits of the internet host
            address.
```

RFC959

NAT Impact #1 – Packet Modification

- Fails to understand Transport Layer Protocol (TLP).

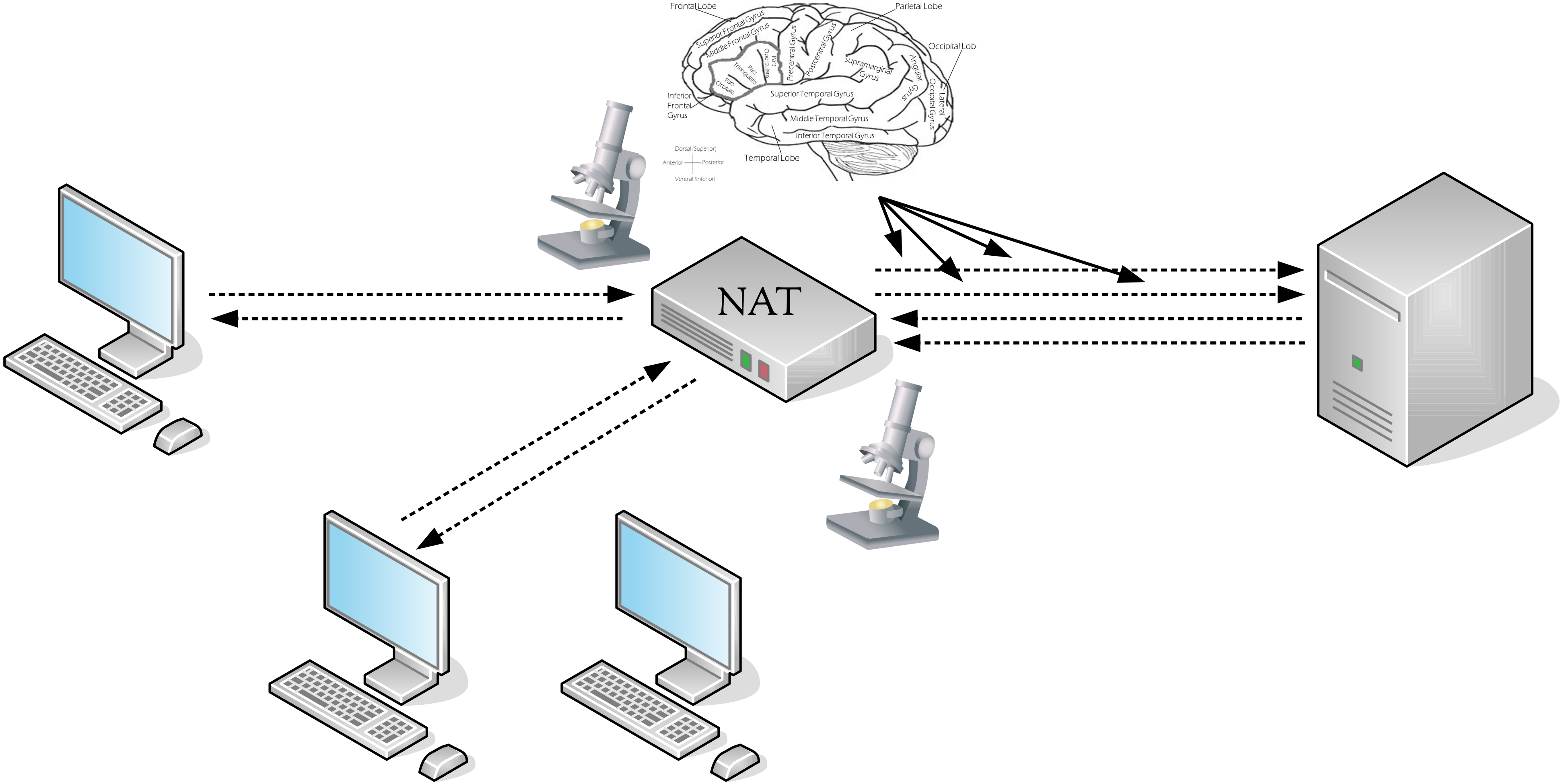
- Fails to understand Application Layer Protocol (ALP).

- Can’t see TLP and/or ALP due to encryption.

- Receiver considers modifications to be an MITM attack.

Any of Above may result in the Packet being Dropped.

NCSF: Availability IMPACT.


NAT
Frontal Lobe
Superior Frontal Gyrus
Middle Frontal Gyrus
Pars Opercularis
Pars Triangularis
Pars Orbitalis
Inferior Frontal Gyrus
Precentral Gyrus
Postcentral Gyrus
Parietal Lobe
Supramarginal Gyrus
Occipital Lob
Angular Gyrus
Occipital Gyrus
Lateral
Superior Temporal Gyrus
Middle Temporal Gyrus
Inferior Temporal Gyrus
Temporal Lobe
Dorsal (Superior)
Anterior
Posterior
Ventral (Inferior)

NAT Impact #2 – State / Loss of State

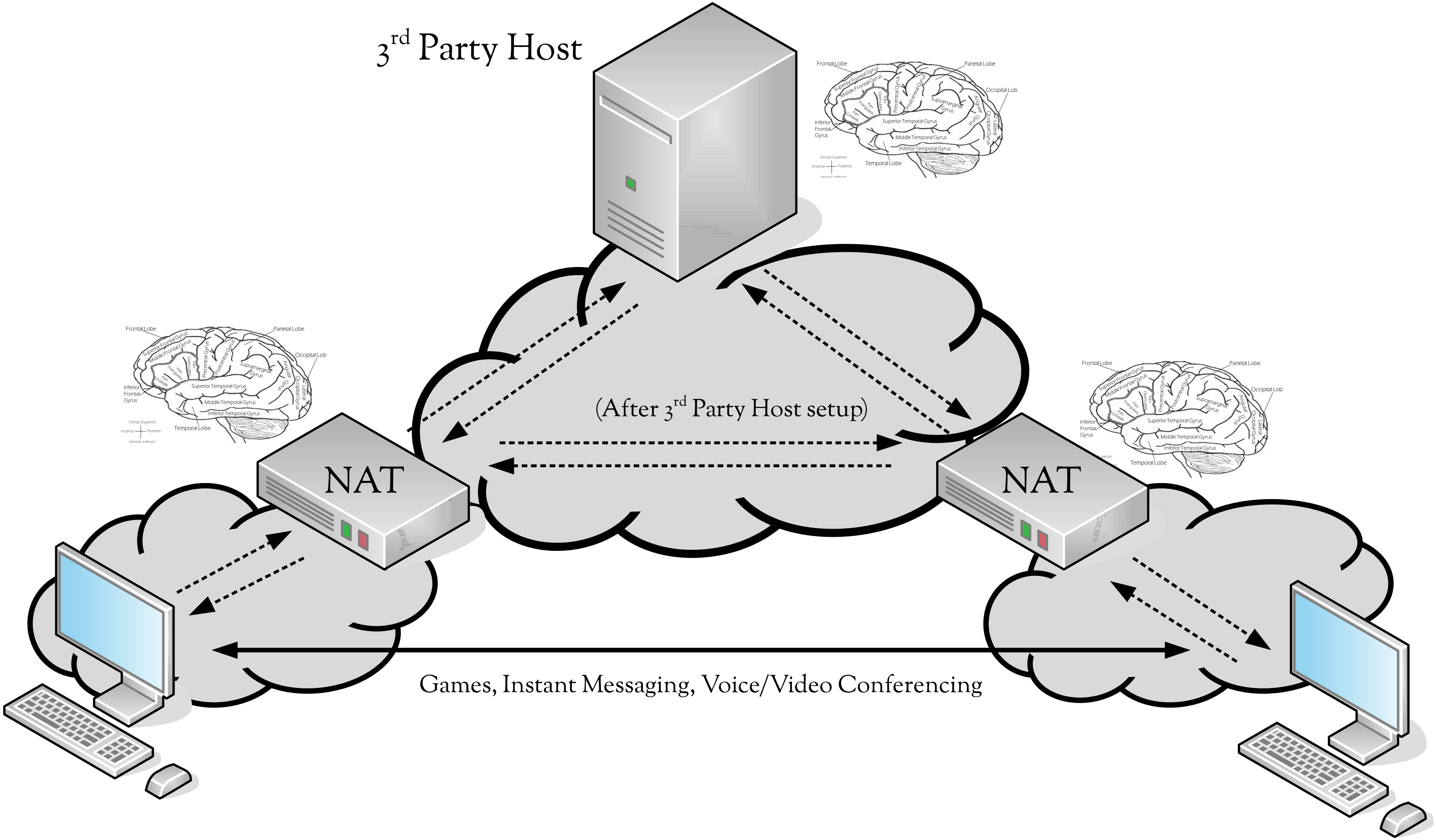
- Traffic driven state means vulnerable to State Exhaustion Denial of Service attack.

- Loss of State due to device Failure means Application Sessions can fail even if there is an alternate Network Path.

- State Synchronisation between redundant NAT devices can be Expensive if devices are Geographically Diverse e.g., different racks, different DCs

NCSF: Availability IMPACT.

NCSF: Budget IMPACT.



3rd Party Host

NAT

(After 3rd Party Host setup)

NAT

Games, Instant Messaging, Voice/Video Conferencing

NAT Impact #3 – 3rd Party Host Required

- Applications that suit Direct Communication are forced to use a 3rd Party Host

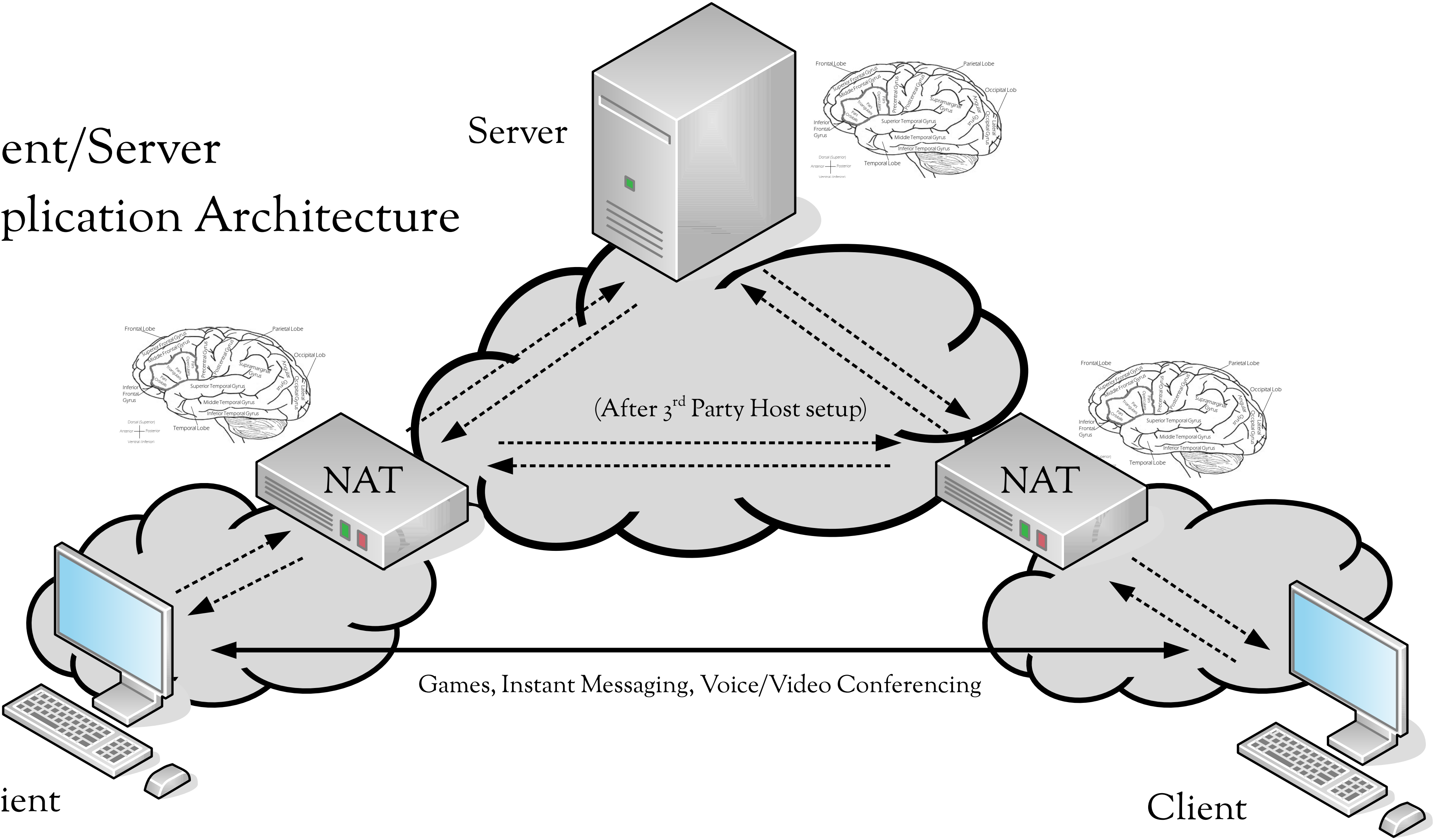
- 3rd Party Host acts as a Relay for All Traffic or is involved in setting up Direct NAT-to-NAT path.

- 3rd Party Host may be relied on (relay), perform well (relay) and must be Trusted.

NCSF: Availability IMPACT.

NCSF: Performance IMPACT.

Client/Server Application Architecture



Server

NAT

NAT

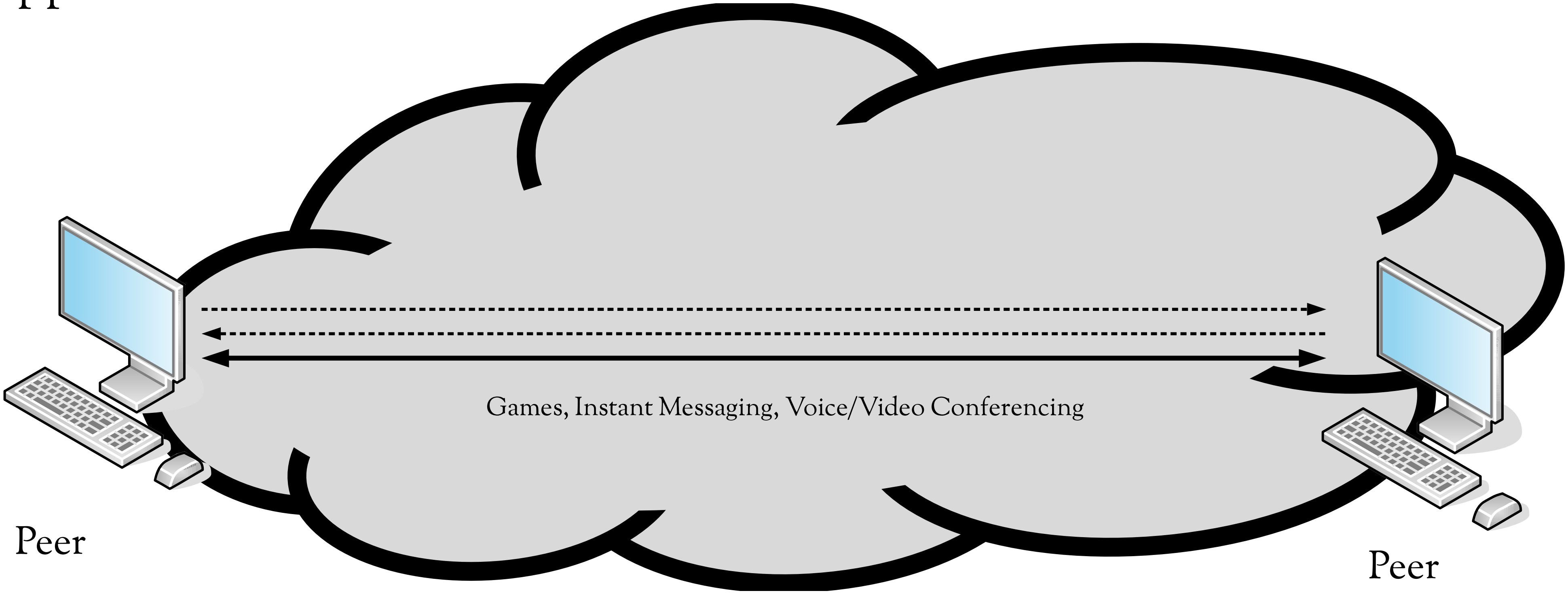
(After 3rd Party Host setup)

Games, Instant Messaging, Voice/Video Conferencing

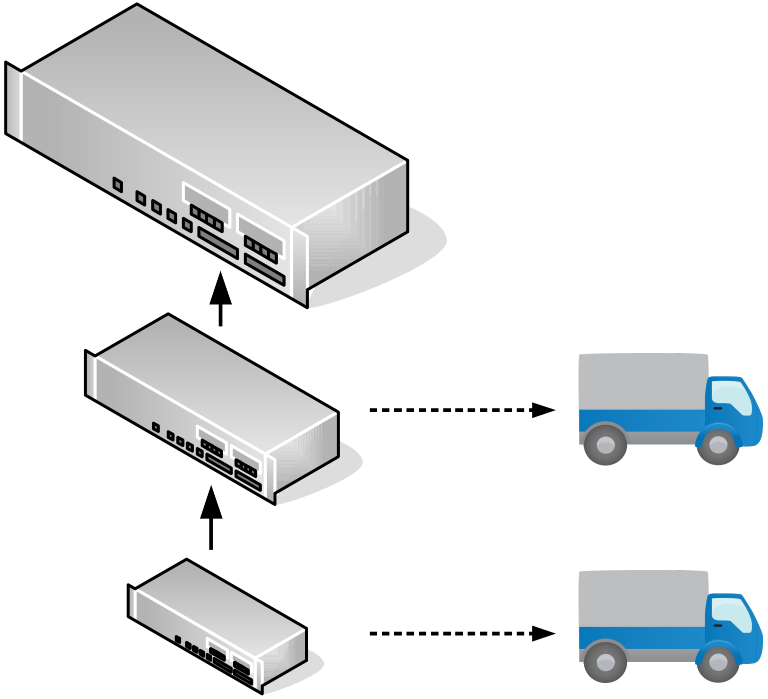
Client

Client

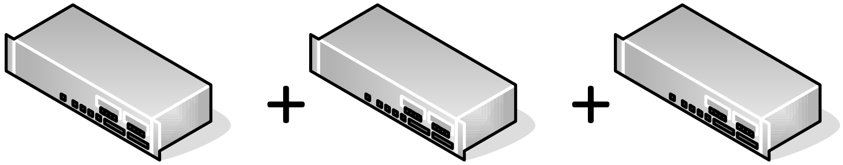
Peer-to-Peer
Application Architecture



Client/Server Architectures



Peer-to-Peer Architectures



What is the Nature of the Internet Protocols?

Client/Server?

IPv4

```
[mark@x13 RFCs]$ egrep -i ”(Client|Server)” rfc791.txt
[mark@x13 RFCs]$
```

IPv6

```
[mark@x13 RFCs]$ egrep -i ”(Client|Server)” rfc2460.txt
[mark@x13 RFCs]$
```

2. Terminology

```
   node          - a device that implements IPv6.

   router        - a node that forwards IPv6 packets not explicitly
                   addressed to itself.  [See Note below].

   host          - any node that is not a router.  [See Note below].
```

Peer-to-Peer, just like People!

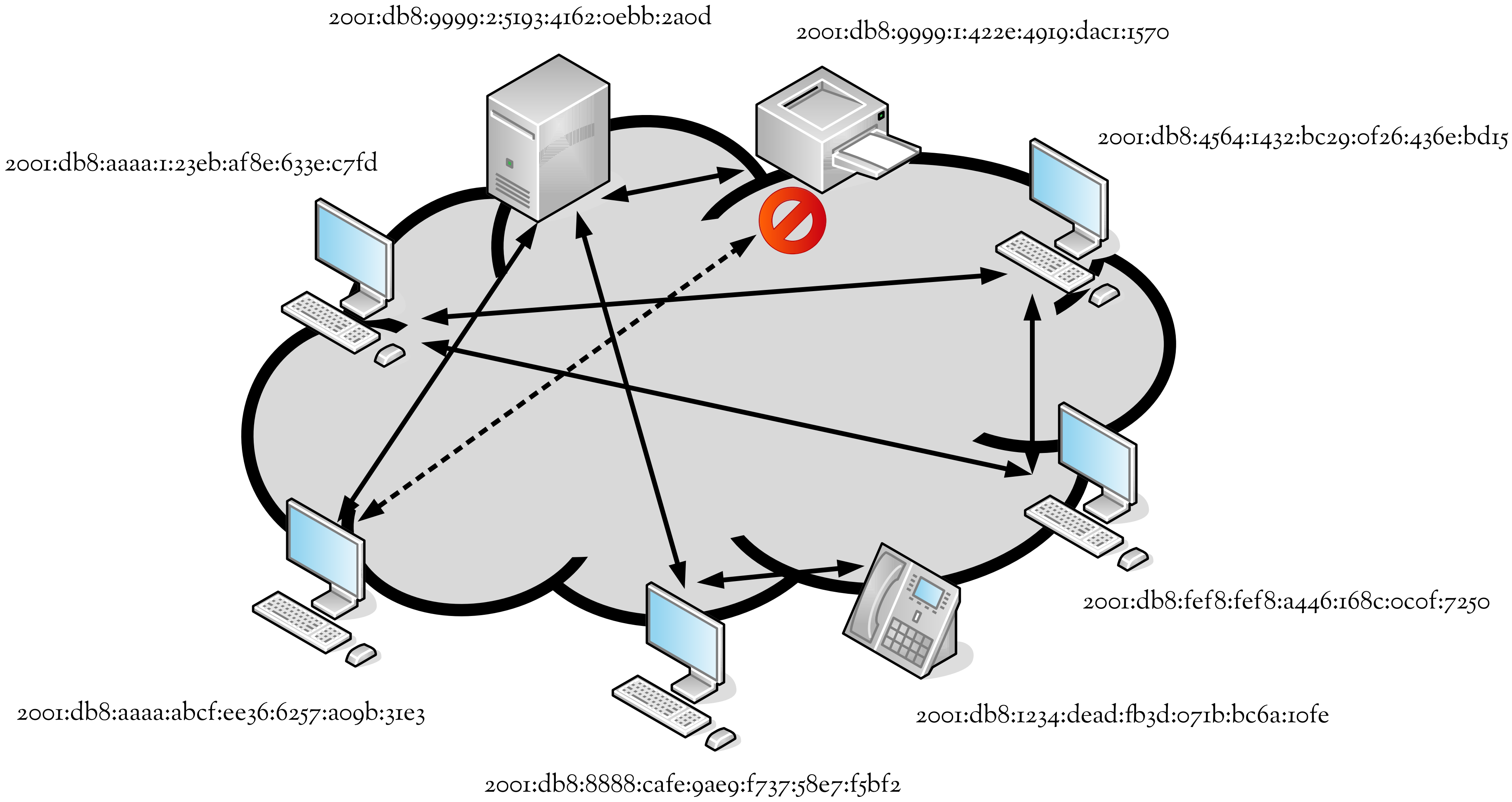
https://flic.kr/p/9PiRpm

Being a Peer

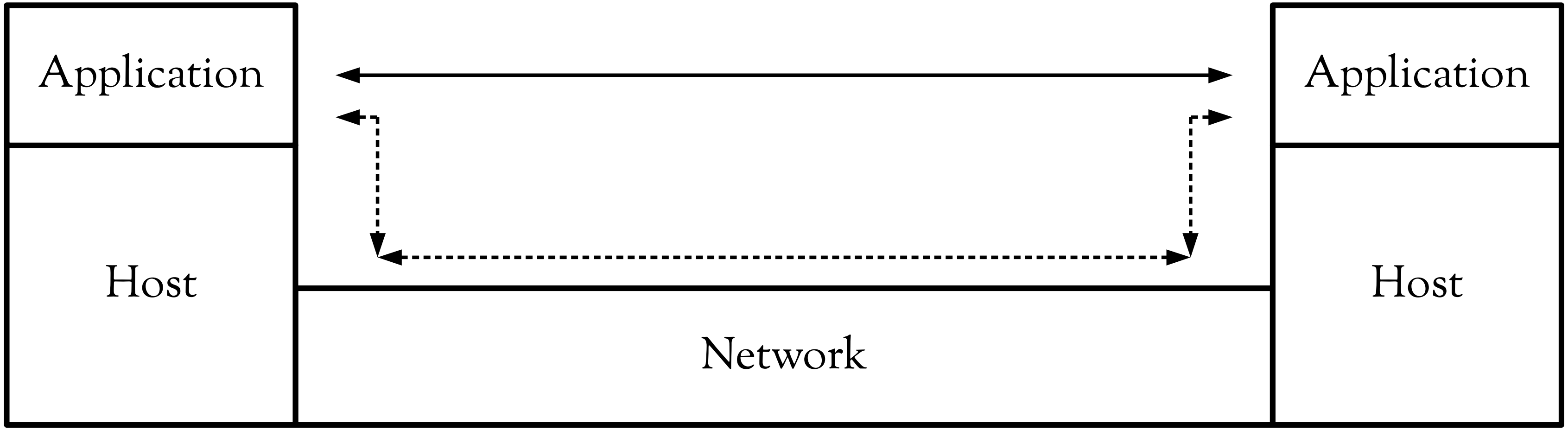
A device with an IP address should be able to:

Send Packets to and receive Packets from All other devices with IP addresses attached to the Same Network, Security Permitting.

Use its own IP address to Identify itself to Others when Referring to itself.

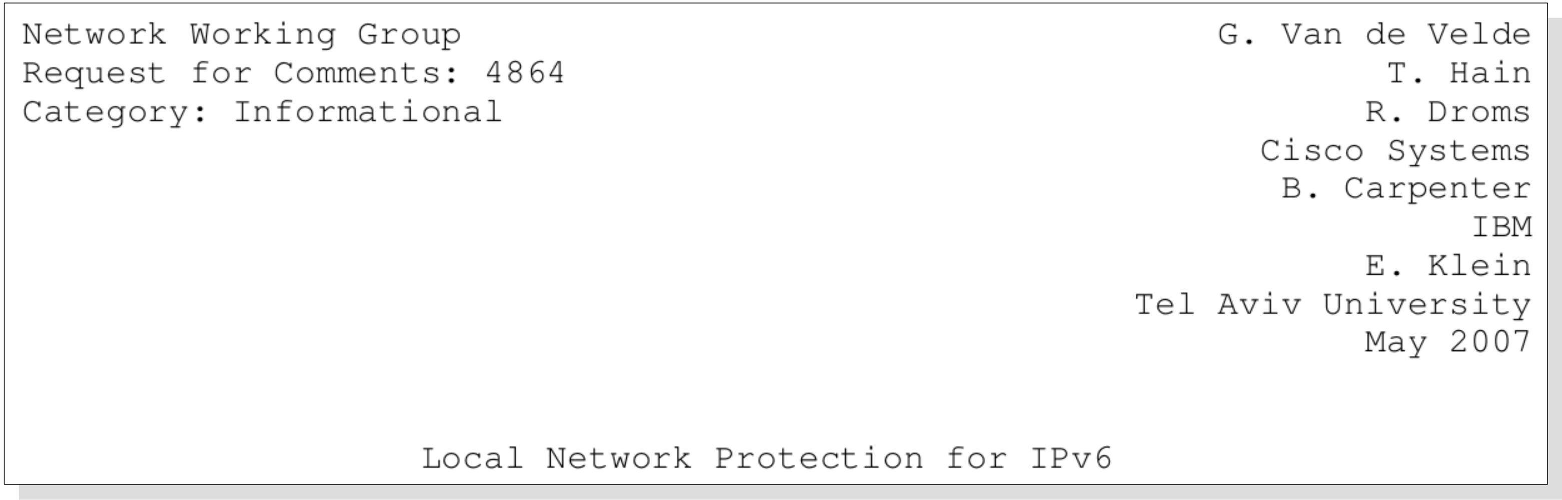

2001:db8:9999:2:5193:4162:0ebb:2a0d
2001:db8:9999:1:422e:4919:dac1:1570
2001:db8:4564:1432:bc29:0f26:436e:bd15
2001:db8:aaaa:1:23eb:af8e:633e:c7fd
2001:db8:fef8:fef8:a446:168c:0c0f:7250
2001:db8:aaaa:abcf:ee36:6257:a09b:31e3
2001:db8:1234:dead:fb3d:071b:bc6a:10fe
2001:db8:8888:cafe:9ae9:f737:58e7:f5bf2

Remember This?



The Fundamental Constraint of NAT is that it Prevents IP nodes attached to the same network from Acting as Peers of each Other.

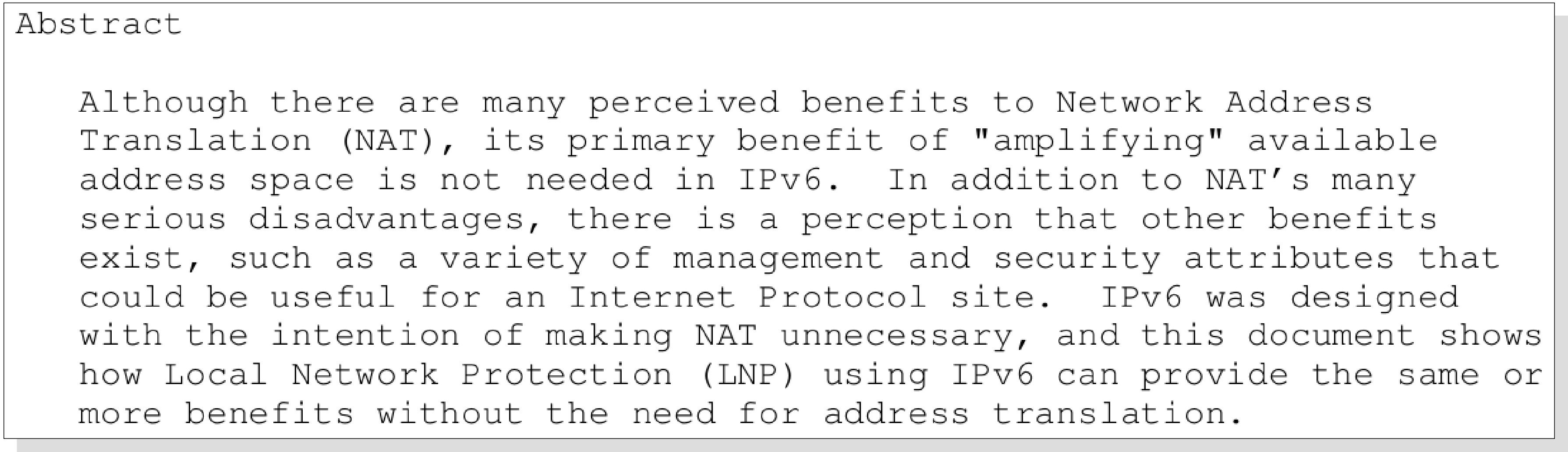
IPv6 without NAT



Network Working Group
Request for Comments: 4864
Category: Informational

G. Van de Velde
T. Hain
R. Droms
Cisco Systems
B. Carpenter
IBM
E. Klein
Tel Aviv University
May 2007

Local Network Protection for IPv6



Abstract

Although there are many perceived benefits to Network Address Translation (NAT), its primary benefit of "amplifying" available address space is not needed in IPv6. In addition to NAT's many serious disadvantages, there is a perception that other benefits exist, such as a variety of management and security attributes that could be useful for an Internet Protocol site. IPv6 was designed with the intention of making NAT unnecessary, and this document shows how Local Network Protection (LNP) using IPv6 can provide the same or more benefits without the need for address translation.

FAQ: Renumbering

IPv6 formally supports multiple concurrent addresses on each interface and addresses lifetimes.

Use Unique Local Addresses (RFC4193) for internal or local traffic, Global prefix(es) for external Internet access.

ULA prefix stays stable and in use during Global renumbering procedure.

Future: Multipath transport protocols e.g., MPTCP, Source Address Dependent Routing (SADR).

FAQ: NAT provides Stateful Firewalling

Stateful Firewalling property of NAT is a side effect of what is necessary to do to perform address translation.

Stateful Firewalling can be performed without address translation (and is, see Linux kernel ’ip6tables’ as an example).

FAQ: NAT hides devices

People are really saying, “NAT hides devices from unsolicited inbound address probes”.

Devices are not hidden from other forms of discovery such as HTTP cookies, or addresses and other identifiers that are leaked in other places in protocols.

Network or host stateful or stateless inbound filters can “hide” IPv6 devices, as well as addressing schemes such as IPv6 Temporary/Privacy Addresses and hard to find using probing Stable Opaque (RFC7217) Addresses.

FAQ: NAT Internal Topology Hiding

RFC4864 mentions using host routes for small scale sites and Mobile IPv6 larger ones.

Another option is various forms of tunnelling over IPv4 to make an IPv6 device appear where the tunnelling concentrator is located.

For example, ISATAP (Intra-Site Automatic Tunnel Addressing Protocol) makes IPv6 devices attached to an IPv4 network appear to all come from the same single /64.

Convinced?

Some Further Reading

RFC1627 - “Network 10 Considered Harmful (Some Practices Shouldn’t be Codified)”

RFC1958 - “Architectural Principles of the Internet”

RFC2775 - “Internet Transparency”

RFC2993 - “Architectural Implications of NAT”

RFC3439 - “Some Internet Architectural Guidelines and Philosophy”

RFC3879 - “Deprecating Site Local Addresses”

RFC4924 - “Reflections on Internet Transparency”

RFC5902 - “IAB Thoughts on IPv6 Network Address Translation”

Questions?

https://flic.kr/p/dA1sTY; Author:Shara Miller; No Mods; Lic: CC BY-NC-ND 2.0

Thanks for listening.