Peer-to-Peer Communication Across Network Address Translators

Bryan Ford – M.I.T.
Pyda Srisuresh – Caymas Systems
Dan Kegel – Ixia Communications

J'fais des trous, des petits trous...
toujours des petits trous
– S. Gainsbourg

USENIX – April 14, 2005
Network Address Translation (NAT)
Network Address Translation (NAT)

Public Internet

Server S
Public Endpoint (S<sup>IP</sup>:S<sup>port</sup>)

Client A
Public Endpoint (A<sup>IP</sup>:A<sup>port</sup>)

Home NAT Router
Home Network
Client B
Network Address Translation (NAT)

Server S

Public Endpoint
(SIP:S\text{port})

Public Internet

Home NAT Router

Home Network

Private Endpoint
(BIP:B\text{port})

Client A

Client B

Public Endpoint
(AIP:A\text{port})
Network Address Translation (NAT)

Public Endpoint (A<sup>IP</sup>:A<sup>port</sup>)

Server S

Public Internet

Public Endpoint (S<sup>IP</sup>:S<sup>port</sup>)

Home NAT Router

Home Network

Private Endpoint (B<sup>IP</sup>:B<sup>port</sup>)

Client A

Client B
Network Address Translation (NAT)

Public Endpoint (A<sub>IP</sub>:A<sub>port</sub>)

Server S

Temporary Public Endpoint (TB<sub>IP</sub>:TB<sub>port</sub>)

Public Internet

Home Network

Home NAT Router

Private Endpoint (B<sub>IP</sub>:B<sub>port</sub>)

Client A

Client B
Network Address Translation (NAT)

Public Endpoint (A<sup>IP</sup>:A<sup>port</sup>)

Server S

Public Endpoint (S<sup>IP</sup>:S<sup>port</sup>)

Public Internet

Client A

Home Network

Client B

Home NAT Router
Network Address Translation (NAT)

Public Endpoint (A^{IP}:A^{port})

Server S

Public Endpoint (S^{IP}:S^{port})

Public Internet

Home Network

Private Endpoint (B^{IP}:B^{port})

Home NAT Router

Clients:
- Client A
- Client B
Network Address Translation (NAT)

Public Endpoint (A<sub>IP</sub>:A<sub>port</sub>)

Public Endpoint (S<sub>IP</sub>:S<sub>port</sub>)

Server S

Public Internet

Home NAT Router

Home Network

Private Endpoint (B<sub>IP</sub>:B<sub>port</sub>)

Client A

Client B
Network Address Translation (NAT)

Public IP Addresses

Public Internet

ISP-deployed NAT

ISP-Private Network

Home NAT Router

Home Network

Private IP Addresses
Network Address Translation (NAT)
Network Address Translation (NAT)
Demand for P2P Communication

Many compelling apps need P2P communication, not just “P2P apps”:

- Teleconferencing, Voice over IP (VoIP)
- Multiplayer on-line games
- Remote access/administration (e.g., ssh)
Outline

- The NAT Traversal Problem
- UDP Hole Punching (not new)
- TCP Hole Punching (quite new)
- Multi-Level NAT Scenarios
- NAT Compatibility with Hole Punching
- Related Work
UDP Hole Punching

Usage model assumptions:

• Clients register with public “rendezvous server” to become accessible to other clients

• Application implements notion of “identity”
  – Username, public key [HIP], etc.

• Rendezvous server facilitates P2P session setup, but does not participate in resulting P2P sessions
UDP Hole Punching

Rendezvous Server S

Public Internet

Client A

Client B
UDP Hole Punching

Rendezvous Server $S$ ($S_{IP}$, $S_{port}$)

Public Internet

NAT

Home Network

Client A

NAT

Home Network

Client B
UDP Hole Punching

Rendezvous Server S
\((S^{IP}, S^{port})\)

Public Internet

NAT

Home Network

(A^{IP}, A^{port})

Client A

NAT

Home Network

(B^{IP}, B^{port})

Client B
UDP Hole Punching

Rendezvous Server S
$(S^{IP}, S^{port})$

Client A
$(A^{IP},A^{port})$

Client B
$(B^{IP},B^{port})$

NAT

NAT

$\text{(TA}^{IP}:\text{TA}^{port})$

$\text{(TB}^{IP}:\text{TB}^{port})$
UDP Hole Punching

Rendezvous Server S

Session A-S

Session B-S

Client A

Client B

Session A-S

Session B-S
UDP Hole Punching

Rendezvous Server S
\( (S^{IP}, S^{port}) \)

Client A
\( (A^{IP}, A^{port}) \)

Client B
\( (B^{IP}, B^{port}) \)

NAT

“Help me reach B”
UDP Hole Punching

Rendezvous Server $S$ $(S^{IP}, S^{port})$

“B is at $(T^{IP}, T^{port})$”
UDP Hole Punching

Rendezvous Server S

(S<sup>IP</sup>:S<sup>port</sup>)

(B<sup>IP</sup>:B<sup>port</sup>)

“B is at (TB<sup>IP</sup>:TB<sup>port</sup>)”

(A<sup>IP</sup>:A<sup>port</sup>)

Client A

Client B
UDP Hole Punching

Rendezvous Server \( S \) (\( S^{IP}:S^{port} \))

Client A

\( (A^{IP}:A^{port}) \)

NAT

Client B

\( (B^{IP}:B^{port}) \)

NAT

"B is at \( (TB^{IP}:TB^{port}) \)"

"A is at \( (TA^{IP}:TA^{port}) \)"
UDP Hole Punching

Rendezvous Server S
(S IP:S port)

NAT Client A

NAT

(TA IP:TA port)

Client B

(TB IP:TB port)

(A IP:A port)

B IP:B port

UDP Hole Punching
UDP Hole Punching

Rendezvous Server $S$ ($S^{IP}:S^{port}$)

Session A-B

($A^{IP}:A^{port}$) $\leftrightarrow$ ($TB^{IP}:TB^{port}$)

Client A

Client B
UDP Hole Punching

Rendezvous Server S

\((S^{IP}:S^{port})\)

Session A-B

\((TA^{IP}:TA^{port}) \Leftrightarrow (TB^{IP}:TB^{port})\)

NAT

Client A

\((A^{IP}:A^{port})\)

Session A-B

\((A^{IP}:A^{port}) \Leftrightarrow (TB^{IP}:TB^{port})\)

NAT

Client B

\((B^{IP}:B^{port})\)
UDP Hole Punching

Rendezvous Server S
\((S^{IP}:S^{port})\)

Session A-B
\((TA^{IP}:TA^{port}) \leftrightarrow (TB^{IP}:TB^{port})\)

NAT

Client A
\((A^{IP}:A^{port})\)

NAT

Client B
\((B^{IP}:B^{port})\)
UDP Hole Punching

Rendezvous Server S
(SIP:Sport)

Session A-B
(TAIP:Tport) ⇔ (TBIP:Tport)

NAT
Client A
(AIP:Aport)

NAT
Client B
(BIP:Bport)
UDP Hole Punching

Rendezvous Server S
\( (S^{IP},S^{port}) \)

Session A-B
\((TA^{IP},TA^{port}) \Leftrightarrow (TB^{IP},TB^{port})\)

NAT

Client A
\( (A^{IP},A^{port}) \)

NAT

Client B
\( (B^{IP},B^{port}) \)
UDP Hole Punching

Rendezvous Server S
(S^{IP}:S^{port})

Session A-B
(TA^{IP}:TA^{port}) ⇔ (TB^{IP}:TB^{port})

Session A-B
(A^{IP}:A^{port}) ⇔ (TB^{IP}:TB^{port})

Session A-B
(B^{IP}:B^{port}) ⇔ (TA^{IP}:TA^{port})

Client A

Client B

NAT
UDP Hole Punching

Rendezvous Server S

\( (S^{IP}.S^{port}) \)

Session A-B

\( (T^{A^{IP}}.T^{A^{port}}) \leftrightarrow (T^{B^{IP}}.T^{B^{port}}) \)

Session B-A

\( (T^{B^{IP}}.T^{B^{port}}) \leftrightarrow (T^{A^{IP}}.T^{A^{port}}) \)

Client A

\( (A^{IP}.A^{port}) \)

Client B

\( (B^{IP}.B^{port}) \)

NAT

UDP Hole Punching

\( (A^{IP}.A^{port}) \leftrightarrow (T^{B^{IP}}.T^{B^{port}}) \)

\( (B^{IP}.B^{port}) \leftrightarrow (T^{A^{IP}}.T^{A^{port}}) \)
UDP Hole Punching

Rendezvous Server S
(S^IP:S^port)

Session A-B
(TA^IP:TA^port) ⇔ (TB^IP:TB^port)

NAT

Client A

Session A-B
(A^IP:A^port) ⇔ (TB^IP:TB^port)

NAT

Client B

Session A-B
(B^IP:B^port) ⇔ (TA^IP:TA^port)
UDP Hole Punching Gone Wrong

Session A-S
$(A^{IP}:A^{port}) \leftrightarrow (S^{IP}:S^{port})$

Session A-S
$(TA^{IP}:TA^{port}) \leftrightarrow (S^{IP}:S^{port})$

Rendezvous Server S
$(S^{IP}:S^{port})$

Client A
$(A^{IP}:A^{port})$

Client B
$(B^{IP}:B^{port})$

NAT
$(TA^{IP}:TA^{port})$

NAT
$(TB^{IP}:TB^{port})$
UDP Hole Punching Gone Wrong

Session A-S
\((T_A^{IP}:T_A^{port}) \leftrightarrow (S^{IP}:S^{port})\)

Session A-B
\((T_A^{IP}:T_A^{port}) \leftrightarrow (T_B^{IP}:T_B^{port})\)

Session A-S
\((A^{IP}:A^{port}) \leftrightarrow (S^{IP}:S^{port})\)

Session A-B
\((A^{IP}:A^{port}) \leftrightarrow (T_B^{IP}:T_B^{port})\)

Client A

Client B

Rendezvous Server S
\((S^{IP}:S^{port})\)

NAT
UDP Hole Punching Gone Wrong

Session A-S
(TA\textsuperscript{IP}:TA_{\text{port}}) \leftrightarrow (S\textsuperscript{IP}:S_{\text{port}})

Session A-B
(TA2\textsuperscript{IP}:TA_{2\text{port}}) \Rightarrow (TB\textsuperscript{IP}:TB_{\text{port}})

Session A-S
(A\textsuperscript{IP}:A_{\text{port}}) \leftrightarrow (S\textsuperscript{IP}:S_{\text{port}})

Session A-B
(A\textsuperscript{IP}:A_{\text{port}}) \leftrightarrow (TB\textsuperscript{IP}:TB_{\text{port}})
UDP Hole Punching Gone Wrong

Session A-S
(TA\textsuperscript{IP}:TA\textsuperscript{port}) ⇔ (S\textsuperscript{IP}:S\textsuperscript{port})

Session A-B
(TA2\textsuperscript{IP}:TA2\textsuperscript{port}) ⇔ (TB\textsuperscript{IP}:TB\textsuperscript{port})

Session A-S
(A\textsuperscript{IP}:A\textsuperscript{port}) ⇔ (S\textsuperscript{IP}:S\textsuperscript{port})

Session A-B
(A\textsuperscript{IP}:A\textsuperscript{port}) ⇔ (TB\textsuperscript{IP}:TB\textsuperscript{port})
TCP Hole Punching

TCP has always supported crucial feature

- “Simultaneous TCP Open” [RFC 793]

Difficulties:

- More ways for NATs to behave poorly
- TCP sockets API oriented toward client/server
TCP Hole Punching

Rendezvous Server $S$ 
($S_{IP}:S_{port}$)

Client A

NAT

Client B

NAT
TCP Hole Punching

Rendezvous Server $S$ ($S^{IP}:S^{port}$)

Client $A$ ($A^{IP}:A^{port}$)
Connect Socket to $S$

NAT

Client $B$ ($B^{IP}:B^{port}$)
Connect Socket to $S$
TCP Hole Punching

Rendezvous Server S

(A\text{IP}:A^{\text{port}}) \rightarrow \text{NAT} \rightarrow \text{S}\text{IP}:	ext{S}^{\text{port}} \rightarrow \text{NAT} \rightarrow (B\text{IP}:B^{\text{port}})

Connect Socket to S

Client A

Client B
TCP Hole Punching

Rendezvous Server S
(S^IP : S^port)

Client A
(A^IP : A^port)

Connect Socket to S

NAT

"Help me reach B"

Client B
(B^IP : B^port)

Connect Socket to S
TCP Hole Punching

Rendezvous Server $S$ ($S^{IP}:S^{port}$)

Client A

$A^{IP}:A^{port}$

Connect Socket to S

Client B

$B^{IP}:B^{port}$

Connect Socket to S

```
TCP Hole Punching

(Rendezvous Server $S$ ($S^{IP}:S^{port}$))

Client A

$(A^{IP}:A^{port})$

Connect Socket to S

Client B

$(B^{IP}:B^{port})$

“B is at $(TB^{IP}:TB^{port})$”

“$A$ is at $(TA^{IP}:TA^{port})$”
```
TCP Hole Punching

Rendezvous Server S

Client A

Client B

Connect Socket to S

Connect Socket to B
TCP Hole Punching

Rendezvous Server S

\((S^{IP}:S^{port})\)

Client A

Connect Socket to S
Connect Socket to B

Client B

Connect Socket to S

Connect Socket to A
TCP Hole Punching

Rendezvous Server S

(A_{IP}:A_{port})

Connect Socket to S
Connect Socket to B

Client A

(S_{IP}:S_{port})

Connect Socket to S
Connect Socket to A

Client B

(TB_{IP}:TB_{port})

(TA_{IP}:TA_{port})
TCP Hole Punching

Rendezvous Server S

Client A

Client B

SYN

Connect Socket to S
Connect Socket to B

Connect Socket to S
Connect Socket to A

(AIP:A^port)

(TAIP:TA^port)

(TBIP:TB^port)

(BIP:B^port)
TCP Hole Punching

Rendezvous Server $S$ 

SYN

NAT

SYN

NAT

SYN

The Magic Socket Option: **SO_REUSEADDR**
TCP Hole Punching

Rendezvous Server S

(NAT): (S\text{IP}:S\text{port})

(A\text{IP}:A\text{port})

Connect Socket to S

Connect Socket to B

(A\text{IP}:A\text{port})

“Simultaneous TCP Open”

SYN

SYN

ACK

ACK

(B\text{IP}:B\text{port})

Connect Socket to B

Connect Socket to A

Client A

NAT

Client B

NAT
TCP Hole Punching

Rendezvous Server S

Client A

NAT

Client B

Connect Socket to S

Connect Socket to B

Connect Socket to A

“Simultaneous TCP Open”

SYN

ACK

SYN

ACK

ACK

ACK

(A\text{IP:}A\text{port})

(B\text{IP:Bport})

(TA\text{IP:TAport})

(TB\text{IP:TBport})

(S\text{IP:Sport})

TB:port

TA:port

(A:port)

(B:port)
TCP Hole Punching

Rendezvous Server S
(S_Ip,S_port)

Client A
(A_Ip:A_port)
Connect Socket to S
Connect Socket to B

Client B
(B_Ip:B_port)

NAT

“Simultaneous TCP Open”

SYN

ACK

SYN

ACK

Connect Socket to S
Connect Socket to A

Simultaneous TCP Open
Timing Caveat

Rendezvous Server S

Client A

Client B

Connect Socket to S

Connect Socket to B
Timing Caveat

Rendezvous Server S

(TA\text{IP}:TA_{\text{port}}) 

SYN

(A\text{IP}:A_{\text{port}}) 

Connect Socket to S

Connect Socket to B

Client A

(TB\text{IP}:TB_{\text{port}}) 

SYN

(B\text{IP}:B_{\text{port}}) 

Connect Socket to S

Client B
Timing Caveat

Rendezvous Server S
(S<sup>IP</sup>:S<sup>port</sup>)

Client A

Client B

SYN

Connect Socket to S

Connect Socket to B

RST

(TA<sup>IP</sup>:TA<sup>port</sup>)

NAT

NAT

(TB<sup>IP</sup>:TB<sup>port</sup>)

Connect Socket to S
Timing Solution

Rendezvous Server S

Client A: (A<sup>IP</sup>:A<sup>port</sup>)
- Connect Socket to S
- Listen Socket

Client B: (B<sup>IP</sup>:B<sup>port</sup>)
- Connect Socket to S
- Listen Socket

NAT

(S<sup>IP</sup>:S<sup>port</sup>)
Timing Solution

Rendezvous Server $S$ ($S_{IP}:S_{port}$)

Client A
- Listen Socket
- Connect Socket to S
- Connect Socket to B

Client B
- Listen Socket
- Connect Socket to S

NAT

SYN
Timing Solution

Rendezvous Server S
(S^P:S^p)

Client A

(A^P:A^p)

Listen Socket

Connect Socket to S

Connect Socket to B

Client B

(B^P:B^p)

Listen Socket

Connect Socket to S

"Normal TCP Open"

SYN

SYN-ACK

ACK
TCP Hole Punching Gone Wrong

Potential problems:

• Inconsistent endpoint translation
  – Same as for UDP

• NAT could reject “unsolicited” incoming SYN with RSTs or ICMP errs instead of just dropping
  – Connection failures, retry oscillation

• Buggy TCP state machine in host OS
  – Windows before XP SP2
Multi-Level NAT

Public Internet

ISP-deployed NAT

ISP-Private Network

Home Network

ISP-deployed NAT

Home Network
Multi-Level NAT

Public Internet

ISP-deployed NAT

ISP-Private Network

Home Network

Home NAT

Home Network

Home NAT
Multi-Level NAT

Public Internet

ISP-deployed NAT

ISP-Private Network

Home Network

Home NAT

Home Network

Home NAT
Multi-Level NAT

Public Internet

ISP-deployed NAT

ISP-Private Network

Home Network

Hairpin Translation

Home NAT

Home Network

Home NAT
NAT Check

Tests hole punching “end-to-end” from user's host

- Results reflect composition of all NAT(s) in path
- No isolation of contention-related “bad” behaviors
- No tests for “bad but semi-predictable” behaviors

More detailed tests of specific NATs elsewhere [Jennings–STUN, Guha–STUNT]

http://midcom-p2p.sourceforge.net/
Data Collection

Results submitted over Web by (self-selecting) community of volunteers

- UDP: 380 data points
- TCP: 286 data points

Covers

- NAT router hardware from 68 vendors
- NAT support in 8 popular operating systems

(Breakdown by vendor in paper)
Testing Results

 UDP Hole Punching
- 82% of NATs support
- Most common NATs:
  - Linksys  98% (45/46)
  - Netgear  84% (31/37)
  - Windows 94% (31/33)
  - Linux  81% (26/32)
- Hairpin: 24%

 TCP Hole Punching
- 64% of NATs support
- Most common NATs:
  - Linksys  87% (33/38)
  - Windows 52% (16/31)
  - Netgear 63% (19/30)
  - Linux  67% (16/24)
- Hairpin: 13%
Related Work

• UDP hole punching: [Kegel 1999]
  – Voice over IP: SIP/ICE [Rosenberg 2003]

• Asymmetric TCP hole punching
  – NUTSS, NATBLASTER, NatTrav
  – Sometimes compensate for bad NAT behaviors, but more complex, timing-sensitive

• Proxy protocols
  – SOCKS, RSIP, MIDCOM, UpnP
    require explicit NAT support, user setup
Conclusion

- NAT is evil, but is here to stay
- Hole punching enables practical, automatic traversal of majority of existing NATs
- Compatibility good for UDP, tolerable for TCP, increasing with NAT vendor awareness (hint, hint)