The Internet/Network Layer

IP Addresses and Routing Tables

Destination	Gateway	Genmask	Flags	MSS	Window	Irtt	Iface
138.38.96.0	0.0.0.0	255.255.248.0	U	0	0	0	eth0
127.0.0.0	0.0.0.0	255.0.0.0	U	0	0	0	lo
default	138.38.103.254	0.0.0.0	UG	0	0	0	eth0

Simple routing table as might be found in a host

- Send local traffic directly to the destination out on interface eth0
- Otherwise send to the default gateway 138.38.103.254, also on interface eth0

The Internet/Network Layer

IP Addresses and Routing Tables

Destination	Gateway	Genmask	Flags	MSS	Window	Irtt	Iface
138.38.96.0	0.0.0.0	255.255.248.0	U	0	0	0	eth0
127.0.0.0	0.0.0.0	255.0.0.0	U	0	0	0	lo
default	138.38.103.254	0.0.0.0	UG	0	0	0	eth0

The mask tells us how to divide the IP address into network and host parts (see later). Work down the table ANDing our address with the mask. If the result equals the Destination value, we use this row

The Internet/Network Layer

IP Addresses and Routing Tables

Destination	Gateway	Genmask	Flags	MSS	Window	Irtt	Iface
138.38.96.0	0.0.0.0	255.255.248.0	U	0	0	0	eth0
127.0.0.0	0.0.0.0	255.0.0.0	U	0	0	0	lo
default	138.38.103.254	0.0.0.0	UG	0	0	0	eth0

“Default” is actually destination 0.0.0.0, and so always matches any address

The Internet/Network Layer

IP Addresses and Routing Tables

Destination	Gateway	Genmask	Flags	MSS	Window	Irtt	Iface
138.38.96.0	0.0.0.0	255.255.248.0	U	0	0	0	eth0
127.0.0.0	0.0.0.0	255.0.0.0	U	0	0	0	lo
default	138.38.103.254	0.0.0.0	UG	0	0	0	eth0

There is also a *loopback* address for a virtual internal network connecting the machine to itself on (virtual) interface lo0. This is useful for many things, such as testing

The Internet/Network Layer

IP Addresses and Routing Tables

Destination	Gateway	Genmask	Flags	Metric	Ref	Use	Iface
213.121.147.69	*	255.255.255.255	UH	0	0	0	ppp0
172.18.0.0	*	255.255.0.0	U	0	0	0	eth0
172.17.0.0	*	255.255.0.0	U	0	0	0	eth1
127.0.0	*	255.0.0.0	U	0	0	0	lo
default	213.121.147.69	0.0.0.0	UG	0	0	0	ppp0

A host with three interfaces: ppp0, eth0, eth1 (and lo)

The Internet/Network Layer

IP Addresses and Routing Tables

Destination	Gateway	Genmask	Flags	Metric	Ref	Use	Iface
213.121.147.69	*	255.255.255.255	UH	0	0	0	ppp0
172.18.0.0	*	255.255.0.0	U	0	0	0	eth0
172.17.0.0	*	255.255.0.0	U	0	0	0	eth1
127.0.0	*	255.0.0.0	U	0	0	0	lo
default	213.121.147.69	0.0.0.0	UG	0	0	0	ppp0

- A packet with address 213.121.147.69 goes directly out on interface ppp0
- Packets with addresses in the network 172.18 go directly on interface eth0

The Internet/Network Layer

IP Addresses and Routing Tables

Destination	Gateway	Genmask	Flags	Metric	Ref	Use	Iface
213.121.147.69	*	255.255.255.255	UH	0	0	0	ppp0
172.18.0.0	*	255.255.0.0	U	0	0	0	eth0
172.17.0.0	*	255.255.0.0	U	0	0	0	eth1
127.0.0	*	255.0.0.0	U	0	0	0	lo
default	213.121.147.69	0.0.0.0	UG	0	0	0	ppp0

- Packets with addresses in the network 172.17 go directly on interface eth1
- Otherwise packets are routed to the gateway 213.121.147.69 on the interface ppp0

The Internet/Network Layer

IP Addresses and Routing Tables

Destination	Gateway	Genmask	Flags	Metric	Ref	Use	Iface
213.121.147.69	*	255.255.255.255	UH	0	0	0	ppp0
172.18.0.0	*	255.255.0.0	U	0	0	0	eth0
172.17.0.0	*	255.255.0.0	U	0	0	0	eth1
127.0.0	*	255.0.0.0	U	0	0	0	lo
default	213.121.147.69	0.0.0.0	UG	0	0	0	ppp0

- The first row of the table is actually redundant here
- Other information, in particular the flags, will be explained later

The Internet/Network Layer

Networks and IP Addresses

- Dividing the 32 bit address into the network and host parts
- 8 bits for network? Then 2^8=256 networks each with 2^{24}=16777216 possible hosts
  - Not enough networks
  - Too many hosts (for most installations)

The Internet/Network Layer

Networks and IP Addresses

- 24 bits for network? Then 16777216 networks each with 256 possible hosts
  - Plenty of networks
  - Not enough hosts per network

The Internet/Network Layer

Networks and IP Addresses

- 16 bits for network? Then 65536 networks each with 65536 possible hosts
  - Not really enough networks
  - Plenty of hosts per network

The Internet/Network Layer

Networks and IP Addresses

- Do all of them!
  - *Class A* networks. From 1.0.0.0 to 127.255.255.255 have 7 bits for network and 24 bits for host. This is 126 networks each with 16777214 hosts. The address x.y.z.w has x as network, y.z.w as host

The Internet/Network Layer

Networks and IP Addresses

- Do all of them!
  - *Class B* networks. From 128.0.0.0 to 191.255.255.255 have 14 bits for network and 16 bits for host. This is 16382 networks each with 65534 hosts. The address x.y.z.w has x.y as network, z.w as host

The Internet/Network Layer

Networks and IP Addresses

- Do all of them!
  - *Class C* networks. From 192.0.0.0 to 123.255.255.255 have 21 bits for network and 8 bits for host. This is 2097152 networks each with 254 hosts. The address x.y.z.w has x.y.z as network, w as host

The Internet/Network Layer

Networks and IP Addresses

- Many empty gaps in the allocation. Also have
  - *Class D* networks. From 224.0.0.0 to 239.255.255.255 are *multicast* addresses. Multicast is a means to send a single packet to multiple hosts. More on multicast later.

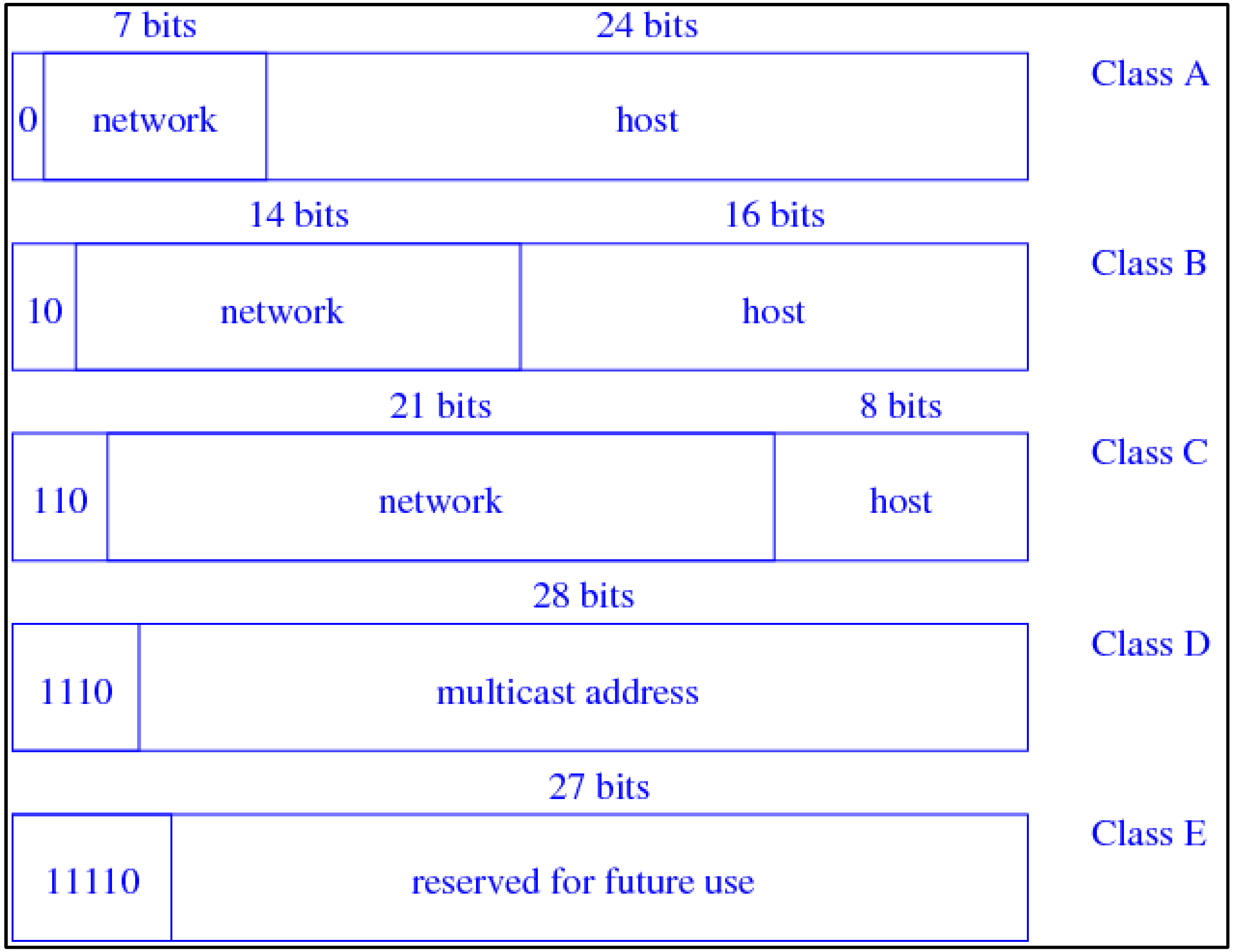
The Internet/Network Layer

Networks and IP Addresses

- Many empty gaps in the allocation. Also have
  - *Class E* networks. From 240.0.0.0 to 247.255.255.255 are reserved for experimental and future use

The Internet/Network Layer

Networks and IP Addresses



The Internet/Network Layer

Networks and IP Addresses

- Many empty gaps in the allocation. Also have
  - Host part all 0s. Refers back to the originating host. So, any host on the network 172.16 sending to 172.16.0.0 is sending to itself. Not commonly supported.

The Internet/Network Layer

Networks and IP Addresses

- Many empty gaps in the allocation. Also have
  - Host part all 1s. Refers to all hosts on that network. So, any host sending to 172.16.255.255 is broadcasting to all the hosts on that network

The Internet/Network Layer

Networks and IP Addresses

- Many empty gaps in the allocation. Also have
  - Network part all 0s. Refers to hosts on the current network. So, any host sending to 0.0.12.34 (on a Class B) is sending to the host 12.34 on its network. Not commonly supported

The Internet/Network Layer

Networks and IP Addresses

- Many empty gaps in the allocation. Also have
  - Network 127.0.0.0. The loopback network. Always implemented. The address 127.0.0.1 is commonly used by a host to send packets over its internal virtual network to itself

The Internet/Network Layer

Networks and IP Addresses

- The class scheme allows IANA to allocate large chunks of addresses to people who need them, and small chunks to those that only need a few
- This scheme has been historically very successful, but with the growth of the Internet has revealed several weaknesses. These days, a *classless* allocation is used.

The Internet/Network Layer

Subnetting

- Suppose you have been allocated class B network: 64 thousand addresses are very hard to manage
- Think of the broadcast traffic (e.g., ARP)
- Technical issues (limits on Ethernet)
- Political issues (traffic from one department must be separate from another department)

The Internet/Network Layer

Subnetting

- Use *subnetting* to split the network into smaller pieces
- Subnets can be administered by separate departments and are joined by routers
- Just like the Internet!
- And, also just like the Internet, we split the address into some bits for the subnetwork and the rest for the hosts

The Internet/Network Layer

Subnetting

- Routers will need to know which bits are the subnet part to be able to decide how to route packets
- Use a *subnet mask*

The Internet/Network Layer

Subnetting

- For example, the University of Bath has a class B, address 138.38
- Department of Mathematical Sciences has a subnet consisting of addresses 138.38.96.0 to 138.38.103.255 (4096 addresses)

The Internet/Network Layer

Subnetting

Network address	138.38.96.0	10001010 00100110 01100000 00000000
Broadcast address	138.38.103.255	10001010 00100110 01100111 11111111
Subnet mask	255.255.248.0	11111111 11111111 11111000 00000000

- A machine can tell if an address is on the local network if the address ANDed with the mask gives the network address

The Internet/Network Layer

Subnetting

Host address	138.38.100.20	10001010 00100110 01100100 00010100
mask	255.255.248.0	11111111 11111111 11111000 00000000
AND	138.38.96.0	10001010 00100110 01100000 00000000

- So 138.38.100.20 is on the subnetwork

The Internet/Network Layer

Subnetting

Host address	138.38.104.20	10001010 00100110 01101000 00010100
mask	255.255.248.0	11111111 11111111 11111000 00000000
AND	138.38.104.0	10001010 00100110 01101000 00000000

- But 138.38.104.20 is not on the subnetwork

The Internet/Network Layer

Subnetting

- Outside of 138.38 the subnetting is invisible so no changes to global routing tables are necessary if we rearrange our network
- Subnets can be further subnetted for exactly the same reason

The Internet/Network Layer

Subnetting

- The subnet is described as “138.38.96.0, subnet mask 255.255.248.0”
- More commonly “138.38.96.0/21”, where 21 is the number of 1 bits in the mask
- Don't have to use the top *n* bits for a mask, but overwhelmingly common to do so (/*n* notation is only for a top-*n*-bit mask)

The Internet/Network Layer

More and more Networks

- Everybody wants a class B as C is too small and A is too large
- Called the *Three Bears Problem*
- Not many class Bs left
- Can we link together some class Cs?
- Awkward as this leads to multiple networks, each needing separate routing

The Internet/Network Layer

More and more Networks

- The growth of the Internet meant that a new way of allocating addresses was needed
  1. Change the way classes are defined
  2. Use private addresses with *network address translation*
  3. Increase the number of addresses available by changing the IP
- We shall be looking at each of these

The Internet/Network Layer

Classless Networks

- *Classless Interdomain Routing*, or *CIDR*
- CIDR takes class C networks and joins them together in such a way that simplifies routing
- Blocks of addresses are allocated to regions
  - 194.0.0.0-195.255.255.255 Europe
  - 198.0.0.0-199.255.255.255 North America
  - 200.0.0.0-201.255.255.255 Central and S America
  - 202.0.0.0-203.255.255.255 Asia and the Pacific

The Internet/Network Layer

Classless Networks

- About 32 million addresses per region
- More addresses available for later allocation
- Easy routing: anything 194 or 195 goes to Europe
- Repeat the idea *within* each region: contiguous block of C networks are allocated to ISPs or organisations
- Keeps simple routing within the region

The Internet/Network Layer

Classless Networks

- E.g., 192.24.0 to 192.24.7, normally written 192.24.0.0/21, where 21 it the number of leading 1s in the mask: *exactly like subnetting*

192.24.0.0	11000000 00011000 00000000 00000000
192.24.7.0	11000000 00011000 00000111 00000000
255.255.255.248	11111111 11111111 11111000 00000000

- Any packet with address that has addr AND 255.255.255.258 = 192.24.0.0 should be routed to that ISP or organisation

The Internet/Network Layer

Classless Networks

- End hosts do not need to know about CIDR, destination networks are treated just as classed networks
- Classless networks can be subnetted
- CIDR has allowed the continuous growth of the Internet
- Can use class A networks similarly

The Internet/Network Layer

Classless Networks

- CIDR merges small networks into a larger one
- Subnetting divides a large network into smaller ones
- CIDR is sometimes called *supernetting*

The Internet/Network Layer

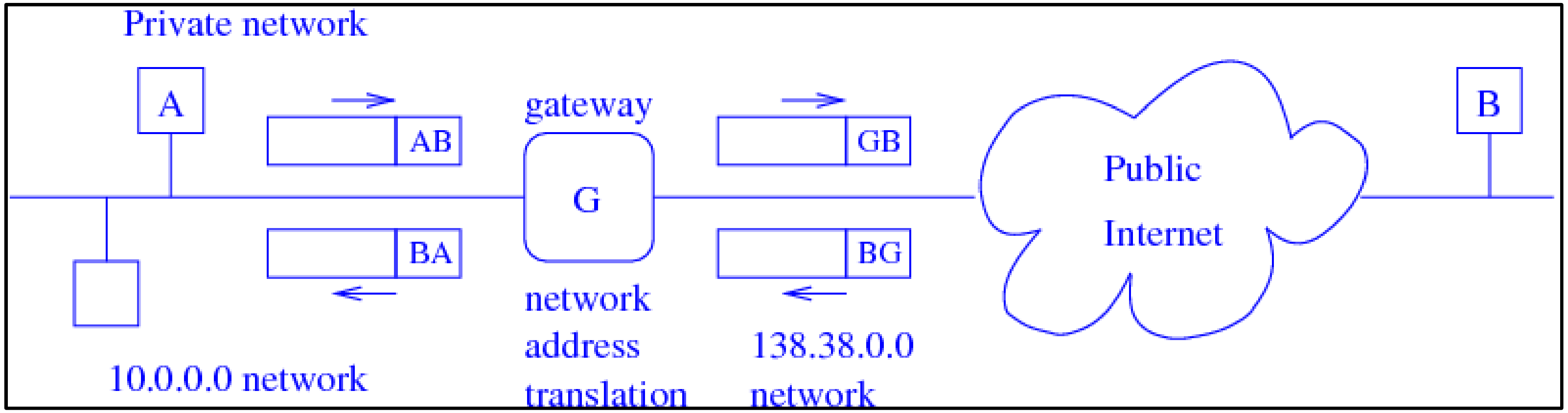
Network Address Translation

- Some IP addresses are reserved for private networks
  - Class A: 10.0.0.0-10.255.255.255
  - Class B: 172.16.0.0-172.31.255.255
  - Class C: 192.168.0.0-192.168.255.255

One class A, 16 class B and 256 class C networks are guaranteed never to be allocated for public use in the Internet

The Internet/Network Layer

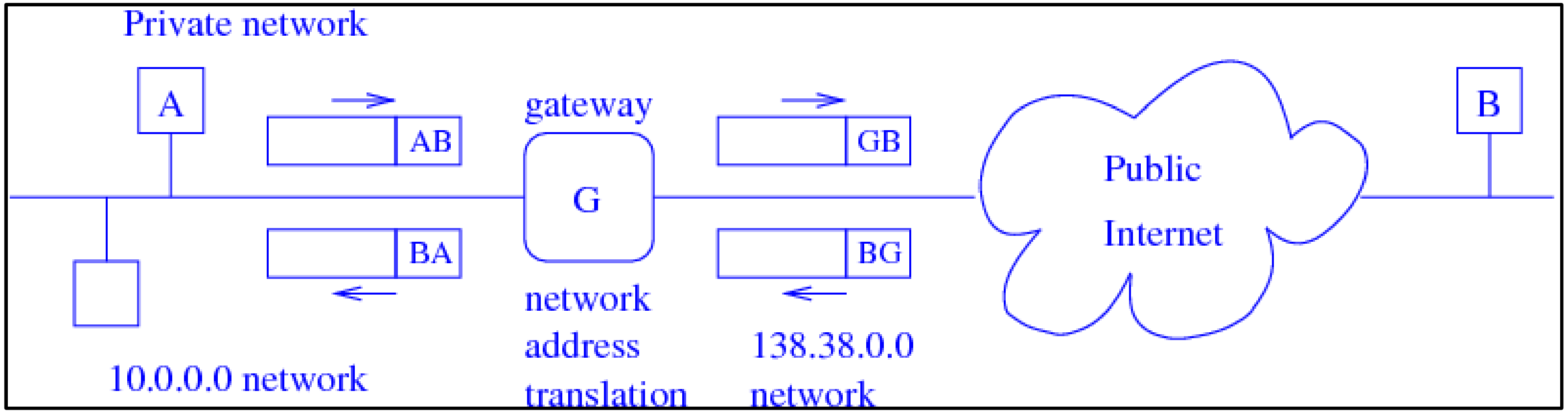
Network Address Translation



- A packet from 10.0.1.1 (A) is sent to 212.58.226.33 (B)
- The gateway overwrites the source address with its own address (G)

The Internet/Network Layer

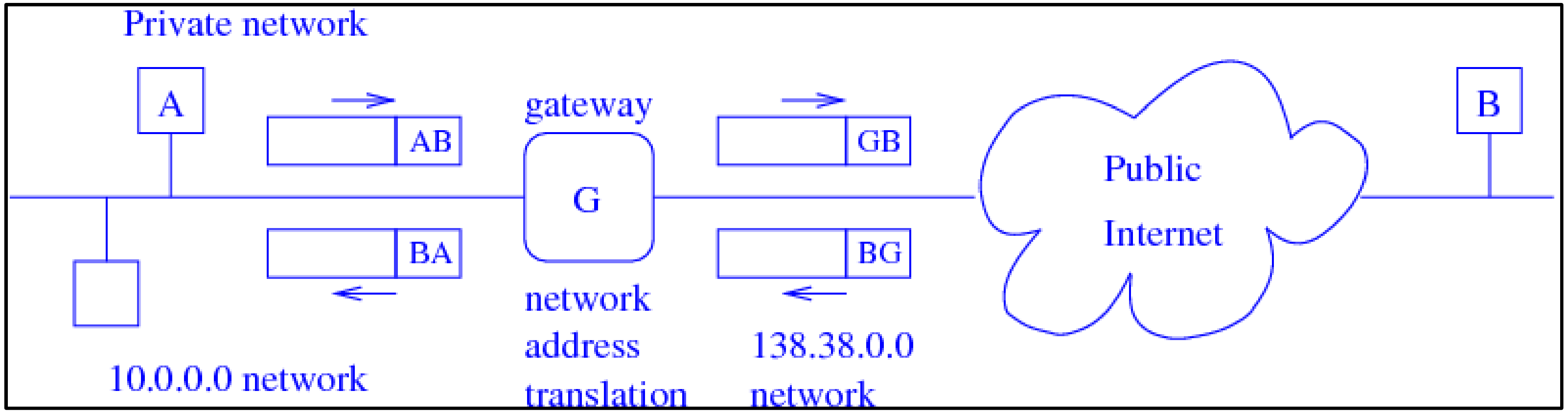
Network Address Translation



- B replies with a packet with destination address G
- The gateway recognises this packet as a reply to A and rewrites the destination address

The Internet/Network Layer

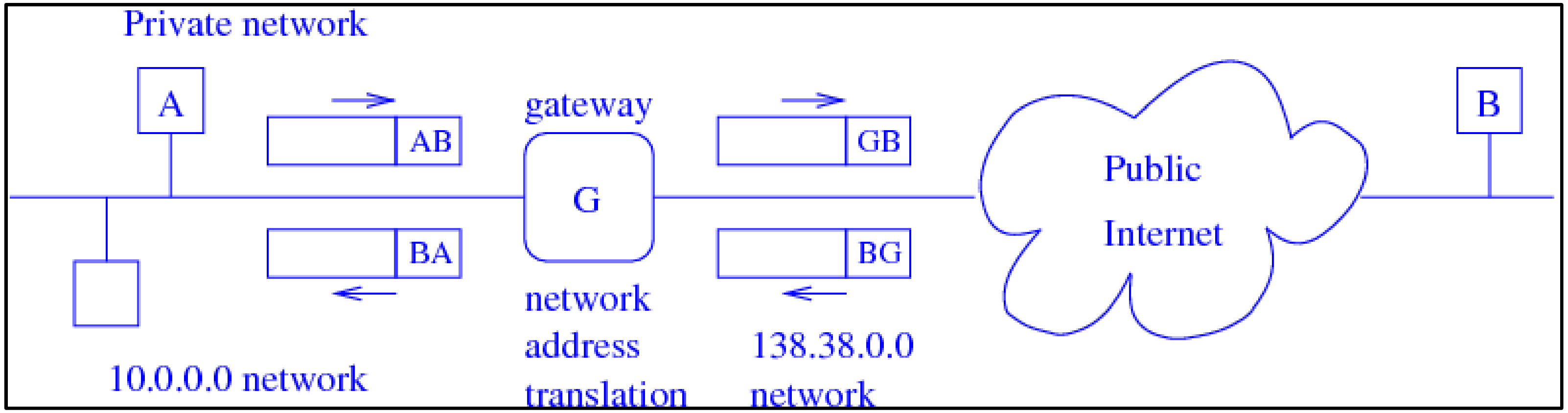
Network Address Translation



- A thinks it is connected to the public Internet, and B thinks data is coming from G
- Machines on the public Internet *cannot* initiate traffic to A as 10.0.1.1 is an unroutable address

The Internet/Network Layer

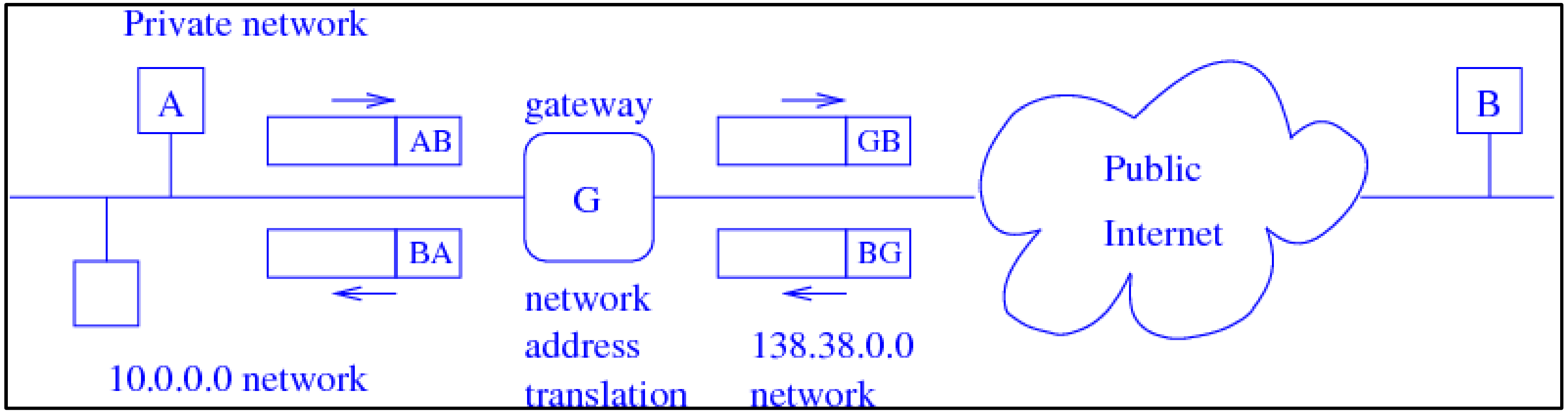
Network Address Translation



- G will only pass on replies to messages initiated by A
- This provides some measure of protection to A from external attack

The Internet/Network Layer

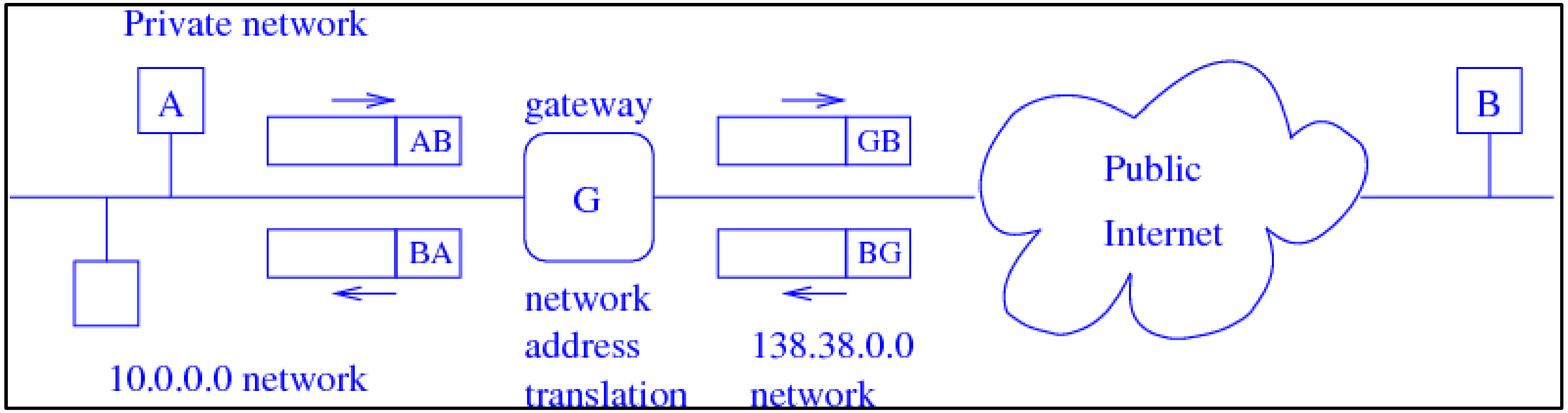
Network Address Translation



- Problems arise when the data in the packet contain IP addresses that, say, will be used to set up new connections. E.g., FTP, Quake

The Internet/Network Layer

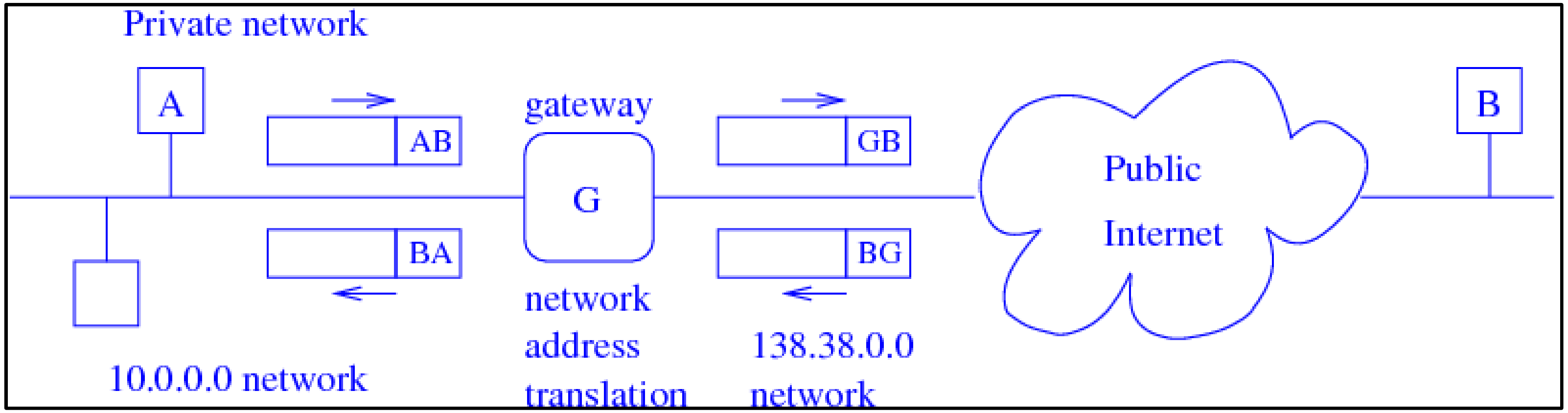
Network Address Translation



- Unless the gateway is intelligent enough to look inside the data and know where the IP addresses are to be found and rewrite them the addresses will remain untranslated and the protocol will fail

The Internet/Network Layer

Network Address Translation



- Not many protocols do this these days, but those that do must be treated specially by the NAT gateway
- Note this is a problem due to a violation of layering