

ICMP redirects are ba'ad, mkay?

Version 1.1, 07/23/2002

Stephen Gill
E-mail: gillsr@cymru.com
Published: 06/29/2002

Contents

Introduction	2
Review	2
Rules of the Road	3
Raw Data	4
The Bad News	5
The Good News	8
Conclusion	9
References	9

Introduction

In a lecture to the third grade in a South Park cartoon, Mr. Mackey the counselor states that drugs, alcohol, smoking, and marijuana are all bad. He simply hopes that the children will accept his statements as truth on what to stay away from. As viewers, we know he is teaching the right thing, albeit without any supportive arguments.

In the IP world, networks with ICMP redirects are also bad. Yet, in this paper we present several supportive arguments as to why this is so. Contrary to what the title might imply, ICMP redirects can be useful because they point out bad network designs. However, a well designed network should never lend itself to the reliance on or desire for ICMP redirects.

Review

An ICMP redirect is an error message sent by a router to the sender of an IP packet . Redirects are used when a router believes a packet is being routed sub optimally and it would like to inform the sending host that it should forward subsequent packets to that same destination through a different gateway. In theory a host with multiple gateways could have one default route and learn more optimal specific routes over time by way of ICMP redirects. Consider the following diagram.

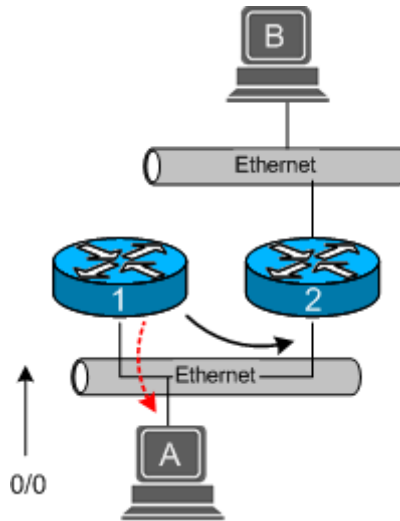


Figure 1 - ICMP Redirect from Router 1.

In Figure 1, host A wishes to send a packet to host B. The routing table on host A consists of a default route through router R1. When R1 receives a packet destined for host B, it looks at its routing table and notices that it has a route (not default) to R2 for packets destined to host B. Therefore, it forwards the datagram to R2, who in turn passes it along to host B. Router 1 also notices that the outgoing interface and network for the packet was the same as the interface it arrived on. Since it is configured to send ICMP redirects, it sends an error message to host A informing it that all packets destined for host B should be forwarded to R2.

It makes sense for router 1 to send a redirect because host A could have just as easily send packets through R2 to start with. Sending packets through R2 directly would decrease the number of routing hops for host A, and increase performance on R1's link. Obviously redirects would not be involved if host A contained a static route to R2 for packets to host B, or if R1 could forward packets to host B through a better path such as a directly connected interface.

Rules of the Road

There are several rules governing the use of ICMP redirects and their acceptance into a host's routing table. Before an ICMP redirect is generated, the following restrictions are typically adhered to on most of today's routers:

1. The outgoing and incoming interface of the packet must be the same.
2. The IP source address in the packet is on the same logical IP network as the next-hop IP address.
3. The route used for the outgoing packet must not be an ICMP redirect or a default route.

4. The packet does not contain an IP source route option.
5. The gateway must be configured to send redirects.

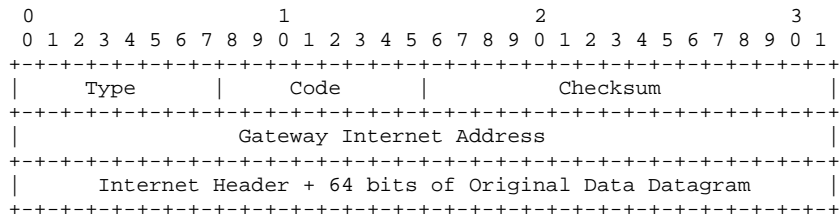
RFC 1122 [2] states that hosts should NOT send ICMP redirects. However, just as gateways follow a set of guidelines as to whether they will send redirects, hosts also have a built-in set of protections to prevent inadvertent additions to their routing table. The following checks are typically performed at the host, not all of which are listed in RFC 1122 before adding a redirect route to the routing table:

1. The new router must be reachable through a directly connected network.
2. The new route must be from the existing router for the destination in question.
3. The new route cannot specify the host as the next-hop address.
4. The redirect route must not point to an address on the local subnet.
5. The host must be configured to accept redirects.

In practice, some operating systems perform more security checks than others when adding redirect routes.

Raw Data

The ICMP redirect error message is consistent with the format of all other ICMP error messages. Following is a breakdown of what the ICMP redirect packet looks like according to RFC 792 [1]:



The redirect message contains an ICMP type of 5, and a code between 0 and 3. Next come the checksum and IP address of the next-hop router that should be used. Finally, the IP header and the first 8 bytes of the offending packet are included as part of the ICMP data. The code field allows for the following types of redirect messages to be sent:

- 0 = Redirect datagrams for the Network.
- 1 = Redirect datagrams for the Host.
- 2 = Redirect datagrams for the Type of Service and Network.
- 3 = Redirect datagrams for the Type of Service and Host.

Notice that there is no place in the packet to delineate a network mask. In practice, only codes 1 and 3 are useable because a host can not accurately guess the network mask in which the destination falls. Network

redirects were designed in the days of classful routing when masks were implied, but very few networks are still deployed this way today.

The Bad News

Based on the information presented so far it would seem that ICMP redirects serve quite a useful purpose. In fact, they do! Yet there are several reasons to steer clear from redirects, all of them relating to network performance, consistency, reliability, and security.

Networks that qualify for ICMP redirects are inefficient. This inefficiency stems from several reasons. Firstly, packets that enter and exit the same interface and IP network towards their final destination are taking a sub optimal routing path. These packets should be forwarded through a different gateway in order to reduce the number of routing hops and in turn decrease latency, and increase network throughput. Packets that enter and exit the same interface of a router are effectively doubling the load on that link!

Secondly, as we noted earlier, ICMP redirects are only host based because it is not possible for a host to guess for which networks a packet is being sub optimally routed. Consider the following diagram.

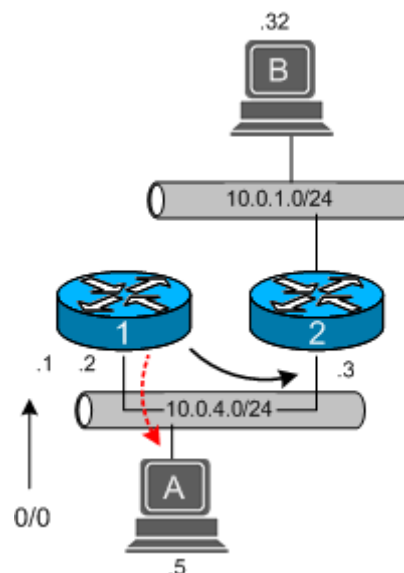


Figure 2 - Host Redirect

Once again, host A would like to send a packet to host B. The routing table on host A consists of only one default gateway and no other static

routes. Following is the routing table output from host A, a FreeBSD 4.6 system.

```
# netstat -rn -f inet
Routing tables

Internet:
Destination      Gateway          Flags    Refs      Use  Netif  Expire
default          10.0.4.1        UGSc     1         22   fxp0
10.0.4.0/24     link#1         UC       3         0   fxp0
10.0.4.1        00:00:5e:00:01:01 UHLW    0         0   fxp0    800
127.0.0.1       127.0.0.1      UH       0         234  lo0
```

To see if host B is alive, host A sends two ICMP Echo-Requests to host B as follows:

```
# ping 10.0.1.32
PING 10.0.1.32 (10.0.1.32): 56 data bytes
36 bytes from 10.0.4.1: Redirect Host(New addr: 10.0.4.3)
Vr HL TOS Len  ID Flg  off TTL Pro  cks      Src      Dst
 4  5  00 0054 012b  0 0000  40  01 6068 10.0.4.5 10.0.1.32

64 bytes from 10.0.1.32: icmp_seq=2 ttl=255 time=0.392 ms
36 bytes from 10.0.4.1: Redirect Host(New addr: 10.0.4.3)
Vr HL TOS Len  ID Flg  off TTL Pro  cks      Src      Dst
 4  5  00 0054 0146  0 0000  40  01 604d 10.0.4.5 10.0.1.32

64 bytes from 10.0.1.32: icmp_seq=2 ttl=255 time=0.407 ms
^C
--- 10.0.1.32 ping statistics ---
 2 packets transmitted, 2 packets received, 0% packet loss
round-trip min/avg/max/stddev = 0.392/0.407/0.437/0.018 ms
```

Notice that host A did receive an Echo-Reply from host B, but it also received something it may not have bargained for: one ICMP Redirect for every packet destined to host B! This would have continued indefinitely until host A canceled its pings as we have done above. Even though host A received an ICMP redirect for every packet, why didn't it take a new path right away? If we examine the routing table on host A once again we notice that a host based route was indeed installed for IP address 10.0.1.32 of host B.

```
# netstat -nr -f inet
Routing tables

Internet:
Destination      Gateway          Flags    Refs      Use  Netif  Expire
default          10.0.4.1        UGSc     1         26   fxp0
10.0.1.32       10.0.4.3      UGHD    0         0   fxp0
10.0.4.0/24     link#1         UC       3         0   fxp0
10.0.4.1        00:00:5e:00:01:01 UHLW    0         0   fxp0    779
10.0.4.2        00:a0:c9:69:c6:0d UHLW    2         0   fxp0    905
10.0.4.3        00:a0:c9:77:a2:52 UHLW    1         0   fxp0    903
127.0.0.1       127.0.0.1      UH       0         266  lo0
```

It would appear as if the ping program does not re-evaluate the routing table during runtime. The reason for this is explained quite elegantly by Richard Stevens [4] who points out that this happens because there is no control input function for the raw IP protocol, only for UDP and TCP. Since ICMP lies just above the IP layer, the redirects appear to be ignored. Thus, changes in the routing table are not taken into account until we stop and rerun the program like so:

```
# ping 10.0.1.32
PING 10.0.1.32 (10.0.1.32): 56 data bytes
64 bytes from 10.0.1.32: icmp_seq=0 ttl=255 time=0.265 ms
```

```

64 bytes from 10.0.1.32: icmp_seq=1 ttl=255 time=0.289 ms
64 bytes from 10.0.1.32: icmp_seq=2 ttl=255 time=0.233 ms
^C
--- 10.0.1.32 ping statistics ---
3 packets transmitted, 3 packets received, 0% packet loss
round-trip min/avg/max/stddev = 0.233/0.262/0.289/0.023 ms

```

This time our packets have taken their expected path and travel immediately through R2 rather than first traversing the default gateway R1. For every destination that host A receives a redirect it will install a redirect host based route in its routing table. In this small example, it could potentially install up to 255 routes with 32 bit masks. Consider how many potential host based routes this could generate for networks with more inclusive prefixes! For example, a single route with a class “B” mask could potentially add up to 65,535 redirects. Certainly it would be better for host A to be fitted with some better method of discovering what traffic to send through R2 instead of relying on ICMP redirects.

Depending on OS implementation, these host based routes can have the characteristic of being short lived. Solaris implements aggressive aging where redirects will only last for a specified amount of time. Each redirect is given a short lifetime value and will be automatically removed from the routing table when the timer expires. This is to ensure that redirects do not remain the routing table indefinitely at the risk of further sub optimal routing. In BSD networking code, redirects will be removed if they are being used by TCP and only after the fourth consecutive retransmission attempt. Routed and Gated perform similar forms of redirect expiration. Since ICMP redirects are dynamic and may not reflect the most current topology it is nice to have an automated method of redirect expiration.

Note that ICMP redirects can also cause problems in firewalled environments where flow traffic patterns are non-deterministic. Consider the following topology.

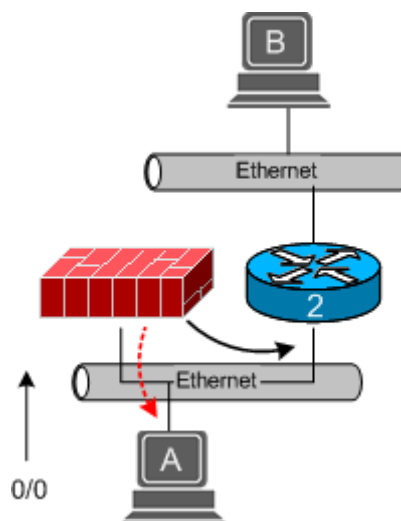


Figure 3 - ICMP Redirect from Firewall

Figure 3 looks quite similar to the previous diagrams, but suppose that host A's default gateway is now a firewall instead of a router. If host A were to receive an ICMP redirect from its default gateway, probably only part of the traffic flow from host A to host B would be seen by firewall.

Asynchronous flows pose particular problems for stateful firewalls because they need to see the entire flow cycle. This is especially the case for TCP traffic where individual packets can point to state transitions such as a beginning (SYN, SYN-ACK), middle (ACK, PSH), or end (FIN, FIN-ACK, RST). If the firewall were only to see the beginning of the flow or part of it, it may take an extra long time to time out entries from the session table because it was not able to see the middle and end of each flow.

The problem is exacerbated if an ICMP redirect only gets acknowledged by a host after the initial TCP handshake has completed. Firewalls rely on timeouts to age out old entries in cases where TCP control traffic is not available. Although embryonic TCP connection timeouts are generally an order of magnitude smaller than timeouts for established TCP connections, they can still lead to firewall session table entry exhaustion [5].

A common design flaw is that the firewall security in this topology could be bypassed altogether if host A were to simply create a static route to host B instead of using its default gateway. This is probably not what the network architect intended when placing a firewall in front of host A if he/she was looking to filter traffic properly from host A to host B.

Finally, redirects could be used by an attacker that has some knowledge of the network topology to inject malicious routes into a host's routing table. This could cause a denial of service against the host, or specific destinations it is trying to reach. One example of such a program is 'icgen' [3], a utility developed by the author simply for research that is capable of sending any type of ICMP unreachable message. Older Microsoft variants are much more susceptible to this form of attack since less security precautions are put in place. Newer operating systems make this more difficult to spoof, but not altogether impossible.

The Good News

Fortunately there are effective tactics when dealing with networks that wish to send redirects aside from disabling them altogether. Generally, these types of networks can be re-architected through one of the following means:

- Changing the network configuration to remove the possibility of multiple non-redundant gateways for a host.

- Changing the network configuration to make the gateways redundant.
- Adding optimum static routes to the hosts for the networks in question.
- Adding the capability for learning dynamic routes to a host.

Conclusion

Though ICMP redirects serve to point out issues with sub optimal routing, network re-architecting should be favored over their use. Well designed networks should never lend themselves to the reliance on or desire for ICMP redirects for reasons of performance, consistency, reliability, and security. ICMP redirects are ba'ad mkay?

References

- [1] Postel, John, "Internet Control Message Protocol", RFC 792, September 1981.
<http://www.ietf.org/rfc/rfc792.txt>
- [2] R. Braden, "Requirements for Internet Hosts – Communications Layers", RFC 1122, October 1989.
<http://www.ietf.org/rfc/rfc1122.txt>
- [3] Gill, Stephen, "ICMP Error Message Generator", icgen.c, June 2001.
<http://www.cymru.com/gillsr/code/icgen-1.2.tar.gz>
- [4] Stevens, Richard, "TCP/IP Illustrated Volume 2." Addison-Wesley. August, 1999.
- [5] Gill, Stephen, "Maximizing Firewall Availability", May 2002.
<http://www.cymru.com/gillsr/documents/maximizing-firewall-availability.pdf>