Gateways & Routing

ISP

Virgin Media Fiber:

- 1Gbit/s down
- 100Mbit/s up
- XGS-PON (10-Gigabit-capable passive optical network; 10 Gbit/s shared symmetric capacity)
- MTU: ~~1468~~ `1460` for IPv4 (IPv4 in IPv6),
- no IPv4 on router - IPv6 DS-Lite (IPv4 tunneled in IPv6 to DS-Lite carrier-grade NAT),
- no router bridge mode.

ISP gateway MTU

From VPS:

```
ping -s 1472 8.8.8.8
PING 8.8.8.8 (8.8.8.8) 1472(1500) bytes of data.
1480 bytes from 8.8.8.8: icmp_seq=1 ttl=110 time=1.34 ms
```

From router itself max ping size `1440` (IP packet size: `1468`). For IPv6 it is `1500` (`1452 + 40 + 8`).

From network (via Caroline/no VPN):

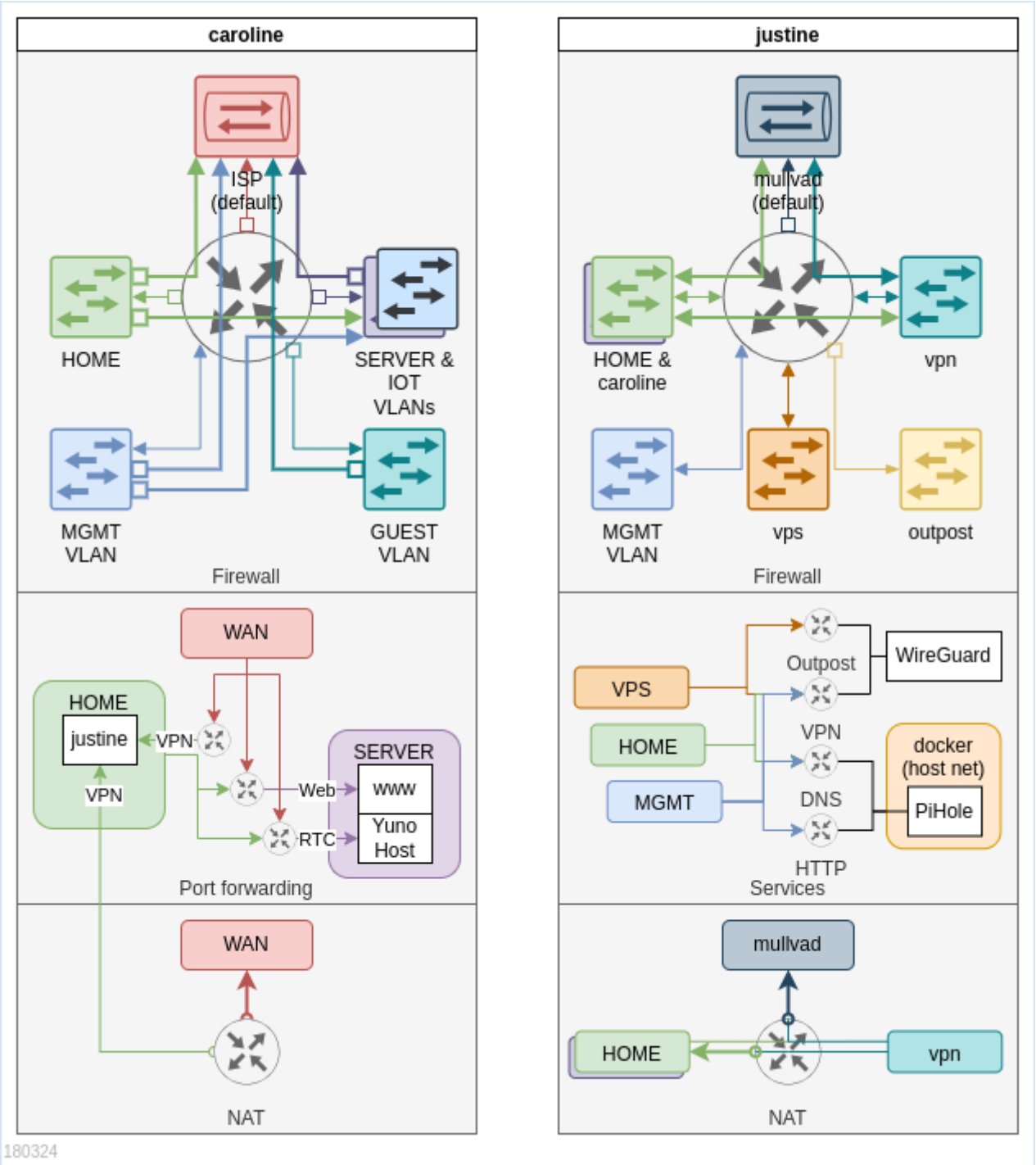
```
ping -s 1440 57.128.183.232
PING 57.128.183.232 (57.128.183.232) 1440(1468) bytes of data.
1448 bytes from 57.128.183.232: icmp_seq=1 ttl=48 time=25.1 ms
```

Router configuration:

- `Gateway MTU size    2000 (1280-1500)` ???
  - When set to `1500` MTU drops to `1460` and I cannot go back! They took 8 byte WTF is this?!?!?!
    - Looks like I was getting extra 8 bytes (`1508`) with `2000` setting.

IPv6 header occupies 40 bytes so IPv4 in IPv6 gets `1500 - 40 = 1460` MTU.

Two gateways



There are two gateways on the network:

1. **caroline** - exposed to the internet, provides access to internet and forwards connections to servers in SERVER VLAN
2. justine - VPN G/W that connects to Mullvad and terminates incoming WireGuard VPN connections

Clients use **caroline** as G/W for direct internet access and **justine** as G/W for Mullvad protected internet access. Additionally **caroline** runs DNS server that uses the ISP DNS server, while **justine**

will use PiHole and Mullvad's DNS server.

Routing with two gateways

Things get very complicated with two gateways setup. Clients need to be able to direct traffic to correct gateway in response to connections coming from one or the other gateway.

Gateway forwarded connections:

1. **caroline** forwards from the internet to access internal network to:
    1. public SERVER network services from outside: blog, younohost etc.
    2. **justine** WireGuard VPN
2. **justine** forwards from internet VPN connected devices to:
    1. HOME network
    2. to **caroline** for SERVER network

This creates the challenge where devices can be configured with any G/W and need to be able to forward the traffic to the other G/W in some cases:

1. local IP & bridge - VPN clients could be bridged directly and assigned bridged network IP
2. NAT - packets coming into the network are MASQUERADE'd to G/W IP address (how it is done currently)
3. static route - push static routes to all clients so response to packets coming from G/W terminated IPs (e.g. VPN) are forwarded back to correct G/W
4. ICMP redirect - both G/W could be configured to inform clients on the correct G/W to use for packets destination

Problems:

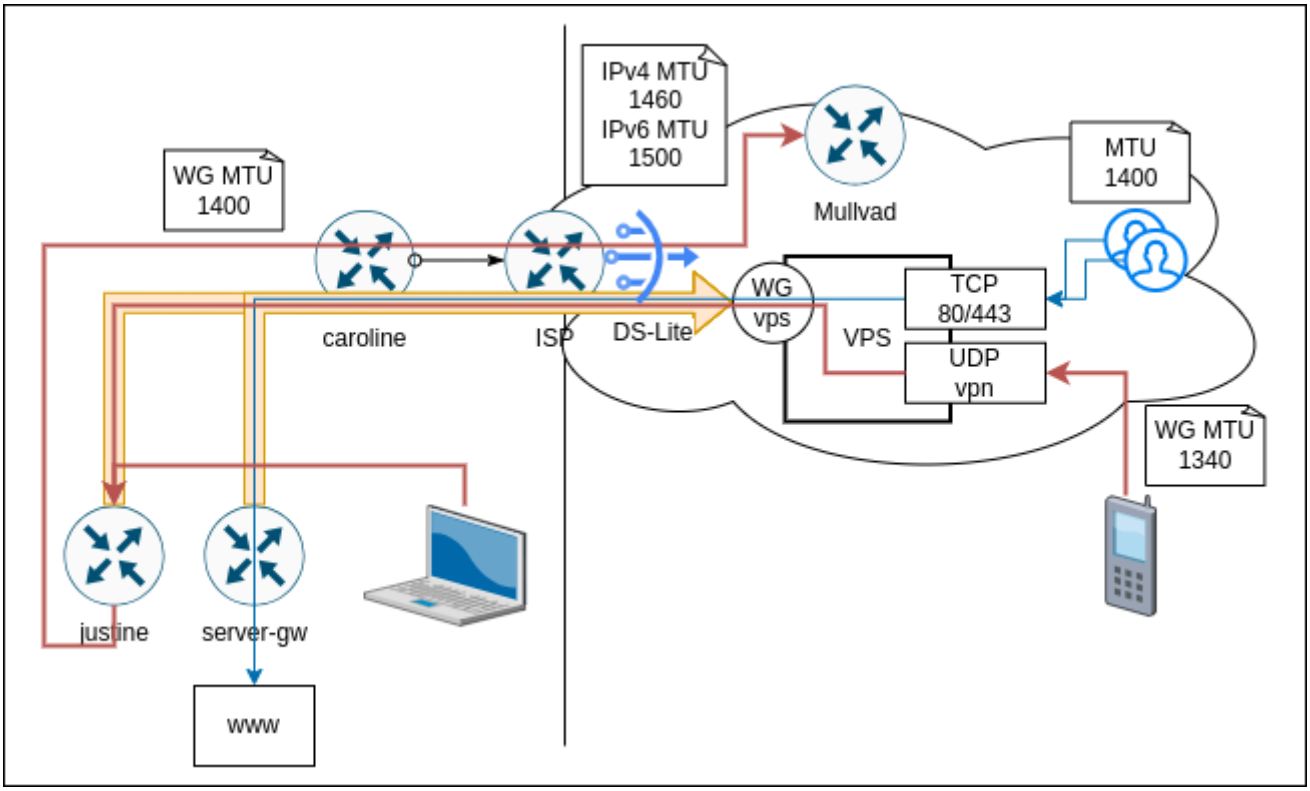
- NAT will obscures the source IP address making troubleshooting, monitoring and accounting more difficult.
- Static routes or redirects will work if G/W can be deduced from destination IP address.
- ICMP redirects many not work reliably, will probably drop first packet?
- Pushing routes to clients requires client support, NAT makes things transparent to clients.

Inbound connectivity

Since ISP does not provide direct IPv4 (DS-Lite - only IPv6 inbound connection forwarding is supported) in I use a VPS service and WireGuard VPN to establish inbound channels.

Vps server uses firewall rules to SNAT/DNAT incoming connections over incoming `vps` WireGuard tunnels to:

- Justine: for inbound VPN connections for roaming,
- Server-gw: for inbound HTTP connections for this wiki and other services.



MTU

Given this layered approach calculating correct MTU for WireGuard endpoints becomes tricky. Default WireGuard MTU of `1420` assumes IPv6 (as worst case) connection to WireGuard server over full `1500` MTU link.

Protocol overhead:

- IPv4 - 20 bytes
- IPv6 - 40 bytes
- UDP + WireGuard - 40 bytes

Link type	Link MTU	IPv4 max payload	IPv6 max payload	WireGuard MTU (IPv4)
Ethernet/Wi-Fi (LAN connectivity)	`1500`	`1480`	`1460`	`1440`
ISP link (DS-Lite)	`1500`	`1440` (IPv4 in IPv6)	`1460`	`1400`
vps VPN (WG IPv4) over ISP link	`1400`	`1380`	`1360`	`1340`

WireGuard MTU settings

Source	Destination	Bottleneck Link	WireGuard MTU
Justine	Mullvad (IPv4)	ISP DS-Lite	`1400`
Justing & Server-gw	Vps (IPv4)	ISP DS-Lite	`1400`
In-LAN (Laptop)	Mullvad via Justine (IPv4)	ISP DS-Lite	`1400`
Roaming (Phone, Laptop)	Mullvad via Justine via Vps (IPv4)	Vps VPN over ISP DS-Lite	`1340`

Server VLAN

Uses dedicated VM `server-gw` that uses WireGuard VPN to connect to Vps server. It acts as default G/W for all SERVER VLAN hosts and routes traffic out via Vps over the VPN connection. This way all servers have Vps public IP as their outgoing IP.

Incoming traffic is forwarded by Vps over same VPN connection to `server-gw` and from there to `www` for HTTP(S) termination and also to `younohost` service for Jitsi meet streams.

Sanbox VM network

Igor runs dedicated network (vnet) with `sanbox-gw` instance acting as default G/W for VMs connected to it. It runs Mullvad VPN and this way provides private connectivity out to the internet. There is no port forwarding into the network. The network is isolated from all other networks.

Revision #55
Created 2023-05-13 11:01:48 IST by hxd
Updated 2024-12-19 19:08:35 GMT by hxd