

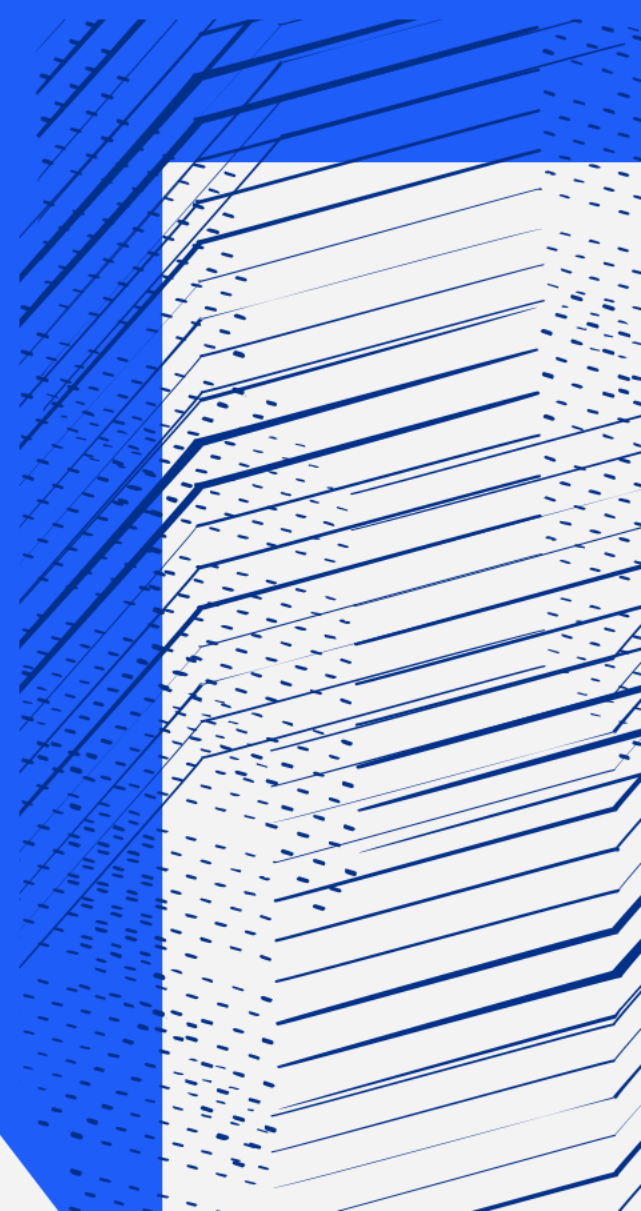


Science and
Technology
Facilities Council

Scientific Computing

IPv6 on RAL Tier1's Containerised WNs

HEPiX Spring 2024, Paris
Martin Bly / Tom Birkett
18th April 2024



Outline

- WNs
- Network
- IPv4
- IPv6

WNs at RAL Tier1

- 5 generations of WNs
 - 40 WNs with 48 slots (Intel Xeon)
 - 280 with 128 or 256 slots (AMD EPYC)
- 10 or 25GbE NIC, single connection to switch
- ~50k job slots (virtual cores)
- Each job runs in its own container, 1 per slot.
 - Multi-core jobs allowed on designated hosts, take up more slots
 - Two extra containers for CVMFS and local Ceph gateways
- Batch system is HTCondor
- BMC and administration (host provisioning) – IPV4
 - dedicated private network

Network overview

- Leaf/Spine(/SuperSpine) network
 - Three-layer CLOS network
 - L3, each Leaf has its own subnet for IPv4 and IPv6
 - BGP internally
 - Most Storage and WNs reside here
- Legacy Network
 - Tiered, resilient, L2
 - Some older sets of Storage and WNs
 - Rapidly reducing as kit ages out of use
 - Here be Dragons
- Joined via an adapter ‘leaf’.
- Hosts everywhere have static addresses including IPv6 addresses if assigned.

~~IPv6~~ IPv4

- Each host as an IPv4 address (provisioning, administration batch management etc.)
- Jobs run individually in a Docker container
- Each container has a private address NATed via the host IPv4 address
- Data in/out of container works fine, it's all on the host
- Doing something similar for IPv6...

IPv6

- Need to enable IPv6 in Docker, and the experimental features
 - It is a rapidly evolving technology

```
/etc/docker/daemon.json
{
  "experimental": true,
  "ipv6tables": true
}
```

- Mimic what we do for IPv4:
 - Each user defined Docker network has a IPv6 subnet derived from the host's IPV6 address.
 - Each user defined Docker network has a IPv6 subnet derived from the host's IPv6 address.
 - Each one of these subnets can be a user defined Docker network.

For example:

- Host IPv6 address: 2001:630:54:3:82f6:dd82::
- Leaf subnet: 2001:630:54:3:/64
- This bit: 82f6:dd82 - is the host's IPv4 address in Hex
- Docker user network:
 - --ipv6 --subnet 2001:630:54:3:82f6:dd82:1000::/112
- All containers on host will get an address in the range 2001:630:54:3:82f6:dd82:1000::/112
- Example
 - "GlobalIPv6Address": "2001:630:54:3:82f6:dd82:1000:b"
 - "IPv6Gateway": "2001:630:54:3:82f6:dd82:1000:1"

Pros and Cons

- The benefit of implementing IPv6 in this way is we get network isolation between different user defined Docker networks
- A drawback of this method is we are effectively NAT'ing IPv6 which isn't ideal, potentially lose out on a lot of IPv6 benefits



Science and
Technology
Facilities Council

Scientific Computing

A decorative graphic element consisting of numerous thin, blue, jagged lines that resemble a circuit board or a stylized 'L' shape. These lines are layered over a solid blue rectangular area that is positioned in the lower half of the slide. The background of the slide is split into an orange top half and a blue bottom half.

Questions?