

RAL Tier1 Network – IPv6

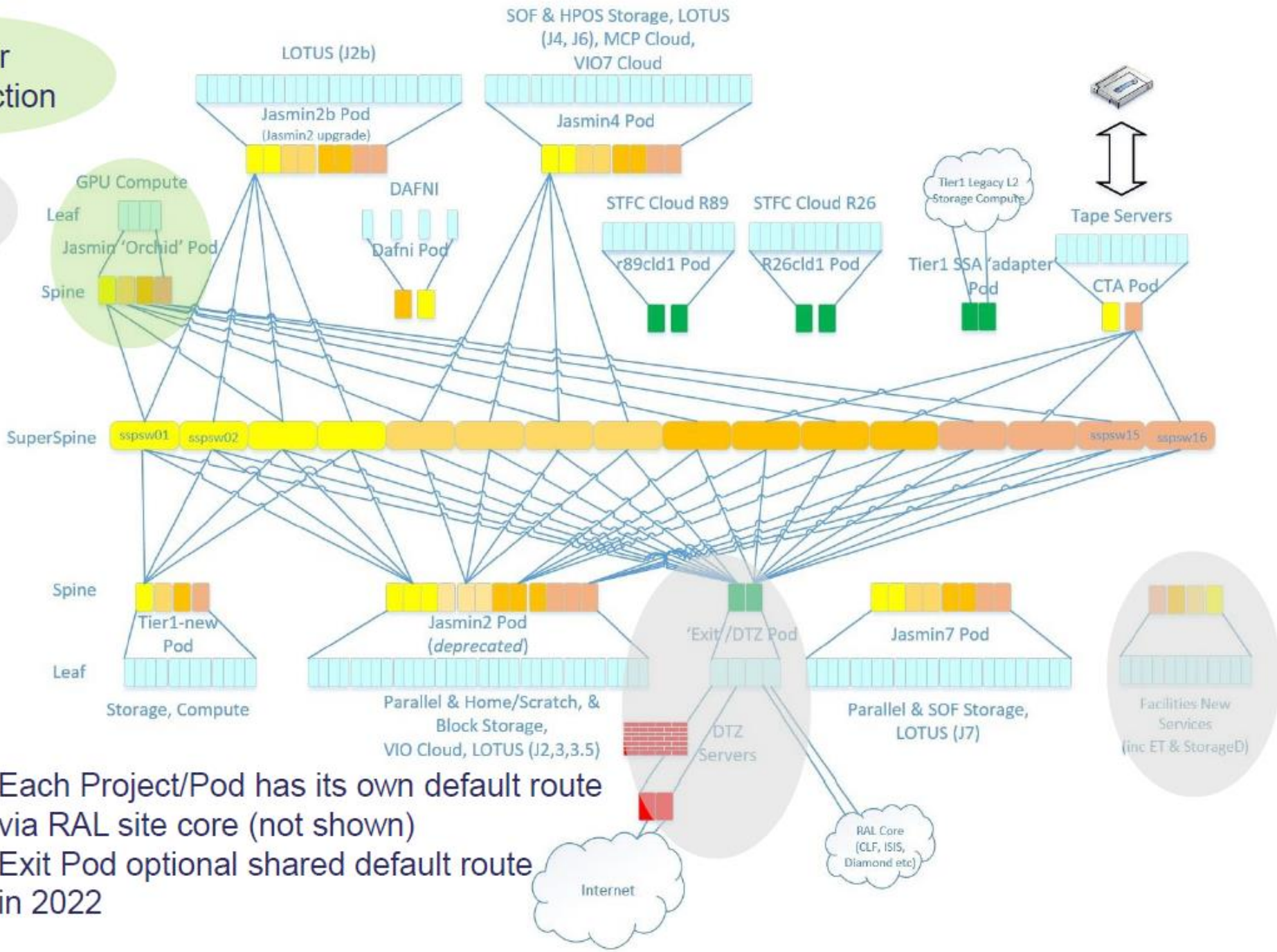
Martin Bly, STFC RAL Tier1

T1 Network

- Complex
- L2 Legacy network
 - Resilient core with links to...
 - ToR to CPU, Storage, various other stuff
 - No Vlans, or separation traffic of any kind – so blast-radius is large
 - Core to ToR links mostly paired or quad 40G with a few paired 100G links
- L3 Leaf/Spine future network
 - 4 x 64port @ 100G Spines
 - Increasing number of leaf switches (ToR) with 4 x 100G uplinks, 1 to each Spine
 - Storage, CPU and eventually everything else
 - Each leaf has its own subnet
- L3 SCD SuperSpine
 - 16 x 32 x 100G switches
 - Provides connections between Spine pods and n x 100G depending on need
 - Tier1 has 16 x 100G uplink to SS, 4 links from each Spine, one to each of 4 SS switches
- So a 3 or 5 layer CLOS depending from which angle you're looking

Under Construction

Planned



- Each Project/Pod has its own default route via RAL site core (not shown)
- Exit Pod optional shared default route in 2022

Routing

- Tier1 IPv4 IP space subnetted, each ToR in L/S space has own subnet
- Some subnet allocation in Legacy network ahead of re-provisioning on L/S network (but still L2)
- Routing between Legacy and L/S via an adapter router in the Legacy network between it and the SuperSpine.
- Exit points all on the Legacy side, so L/S traffic currently goes over the SS to Legacy exit to LCHOPEN, RAL site , JANET
- Current IPv6 allocation matches this since it's all in the Legacy side
 - OPN and non-OPN ranges, all handled in L2.
- But...

...IPv6 is not so simple

- RAL has a /48 allocation, with an addressing scheme applied
- But we need a /64 for each ToR
- The SCD Cloud needs many more /64s than there are in the IPv6 allocation scheme to allow all the stuff they need to do for their users
- We don't have enough subnet space to do things sensibly
 - Despite being almost the only IPv6 users on site
- We applied for and been allocated two adjacent /48s for SCD
 - That should be enough, right? :-)
- So now we have 5 x /48 for various projects
 - JASMIN, CLOUD, Tier1, RAL, ...
- Get a bigger-still allocation (/32) you ask?
 - It's complicated but under discussion.

Not only, but also:

- T1 Exit point will move to L/S network fairly soon (increase in bandwidth)
- Dual stack for the batch system has been delayed awaiting IPv6 support then upgrade
- Must provide Dual Stack (storage, tape) for some service users on-site because they don't do IPv6
- T1 network is in transition but even planned service interruptions not so welcome
- CTA/Antares (Tape Service), SCD Cloud, other SCD services have their own Spine pods with which we need to maintain connectivity
- We don't run the site network core or border routers or firewall, we need to cooperate with the site networking team to get some site access stuff done/enabled