Linux – BRAS with accel-ppp

coolideas

INTERNET SERVICE PROVIDER

cool ideas
INTERNET SERVICE PROVIDER

Why ?

- Mikrotik is no longer fit for purpose within our org. Licensing on big brands is crazy.
  - High volumes of traffic, high volumes of users, IPv6 (dual stack), queues for rate limits.
  - Horizontal scaling becomes inefficient.
  - The CCR 1xxx is EOL and the 2xxx is not a drop in solution. They cannot cope with the volumes and PPPoE cannot be done in L3HW, including queues.
  - Currently we running 9000 subscribers in one DC and 4500 in the other, excludes CT and KZN
  - One server, tested, can handle 9000 subs, currently we seeing 1 server at 5064 subs doing 16.5G and the other 3900 subs doing 12.5G.

coolideas
INTERNET SERVICE PROVIDER

What

- We started looking at Accel-ppp and netElastic. We quickly dropped netElastic as it is a paid model.
- We started on a small intel server, dual 10G card, 16G ram, single CPU 4 core, no hyperthreading (3GHz).
- After some testing we added it to the PPPoE farm and had it pickup load, This proved successful in that it handled load, queues and through-put with consistent CPU.
- The next step was to decide on a production platform.
  1. We chose Intel S2600 2U 4 node chassis, this worked but lacked CPU
  2. Chassis is also an older model and at 2500 users they ran out of steam
  3. Swapping node was an issue due to Teraco cab design in that the PDU blocked the extraction path.
  4. Design of the sled put the CPU one behind the other on airflow path and the 2nd CPU ran hotter.

cooldeas

INTERNET SERVICE PROVIDER

coolideas
INTERNET SERVICE PROVIDER

What

- We then moved on to 2 test Gigabyte servers, Penguins, we tested them with the defaults they came with namely, Dual XEON CPU 2.2GHz and 64G ram.
- We also upgraded the CPU's to Intel(R) Xeon(R) Gold 6154 CPU @ 3.00GHz. This gave us a 10% improvement in performance.

coolideas
INTERNET SERVICE PROVIDER

What: part 2

- Power: at a very light calculation we found the:
  - Intels 230w per 7G
  - Penguins 270w per 15.3 and 200W per 3G
- So the Penguins are a little move power efficient for the bang.

- GUI: we found a lightweight GUI interface which gave our NOC the required functionality they are used to in troubleshooting client issues. The same as they were using on the Mikrotiks

- Support: We found the volunteer support group was very helpful and turn around time was reasonably good.

coolideas
INTERNET SERVICE PROVIDER

How

- Started small with PoC server running Devuan, no system standing in the way w.r.t udev.
  - This bypasses the need to do BRAS tuning system udev optimisations.
  - https://docs.accel-ppp.org/guides/BRAS_tuning.html#systemd-udev-optimizations

- Accel-ppp is compiled on server due to the vlan-mon kernel module.
  - Dynamic Kernel Module System is still w.i.p.

- Custom package created with default configs and settings for easy distribution of settings and faster deployment of a new server. Simplification so the NOC engineers can also do this ;)

- Most of our issues encountered stemmed from the L2TP module, which the devs have confirmed is not under active development. The devs were quick to implement fixes for our issues experienced.

- For the servers we have found that a higher clock-cycle is better than a higher core count and that the servers tend to respond with a PADI much quicker than a MikroTik, even at twice the active sessions. Memory and disk usage is negligible, currently running about 4500 users on a single server, with the cpu sitting around 50% and using 6G memory. Disk usage is mainly logging with per-username logs available for easier debugging of issues.

TER-JB1-AR1-OS Info

	Uptime
CPU	9.06%	3.17 weeks
Memory used	5.63%	3.17 weeks
Time	2025-04-04 14:50:00	

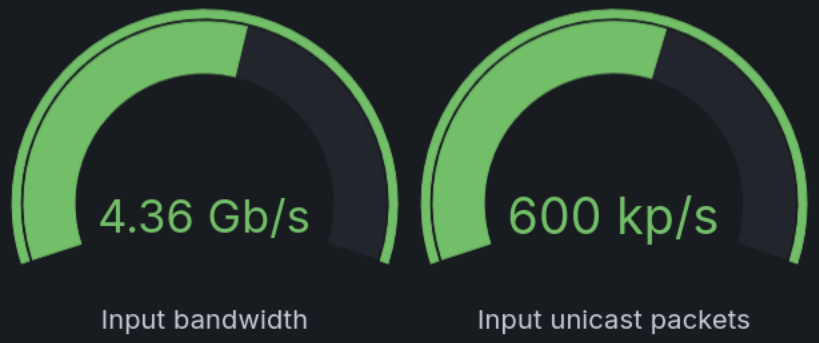
TER-JB3-AR1-OS Info

	Uptime
CPU	31.3%	6.21 weeks
Memory used	6.78%	6.21 weeks
Time	2025-04-04 14:50:00	

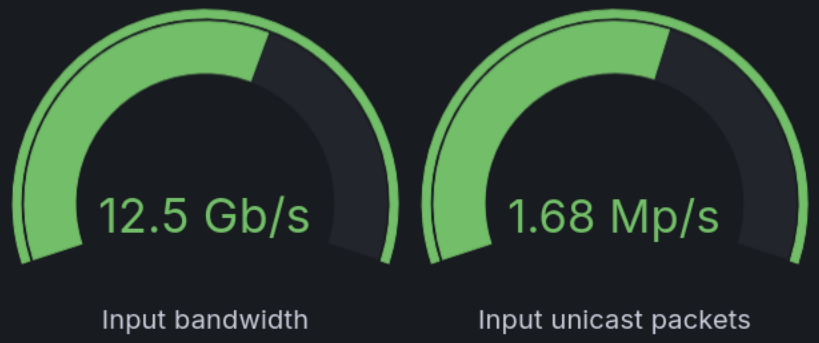
TER-JB3-AR2-OS Info

	Uptime
CPU	35.9%	9.20 weeks
Memory used	7.46%	7.00 weeks
Time	2025-04-04 14:50:00	

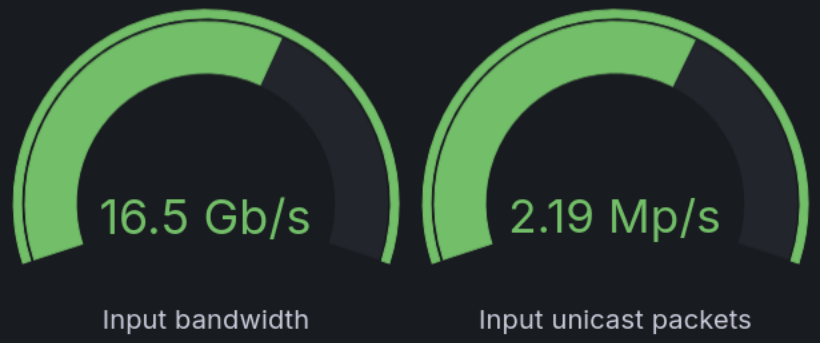
TER-JB1-AR1-Throughput and Packets



TER-JB3-AR1-Throughput and Packets



TER-JB3-AR2-Throughput and Packets



TER-JB1-AR1-Sessions

TER-JB3-AR1-Sessions

TER-JB3-AR2-Sessions

TER-JB1-AR1-Pools

TER-JB3-AR1-Pools

TER-JB3-AR2-Pools

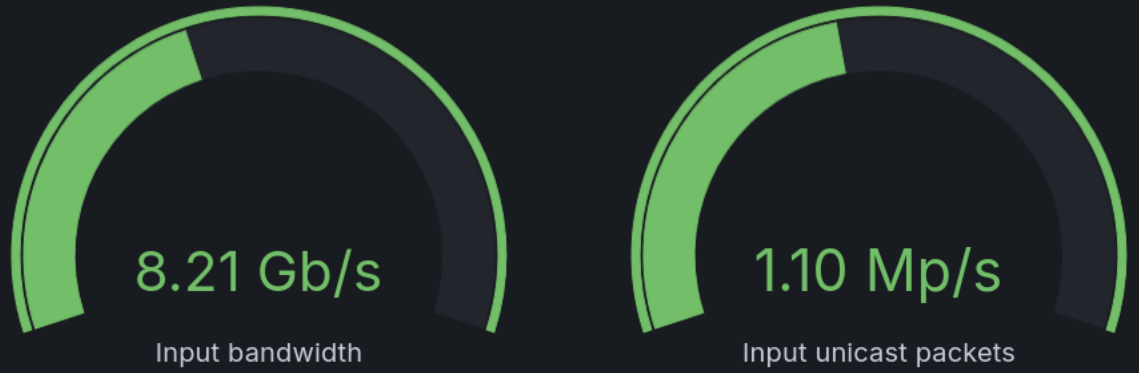
TER-F02-AR1-OS Info

	Uptime
CPU	34.0%	9.58 weeks
Memory used	3.33%	8.42 weeks
Time	2025-04-04 14:50:00	

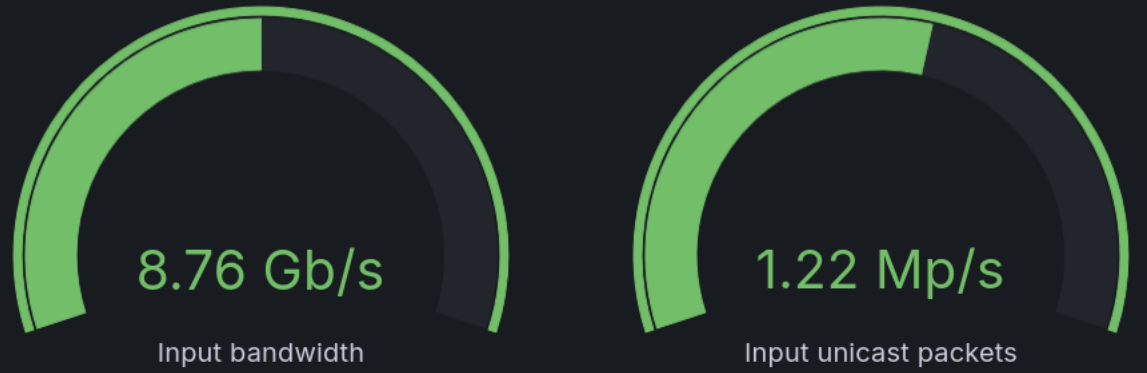
TER-F02-AR2-OS Info

	Uptime
CPU	36.5%	9.58 weeks
Memory used	3.28%	8.42 weeks
Time	2025-04-04 14:50:00	

TER-F02-AR1-Throughput and Packets



TER-F02-AR2-Throughput and Packets

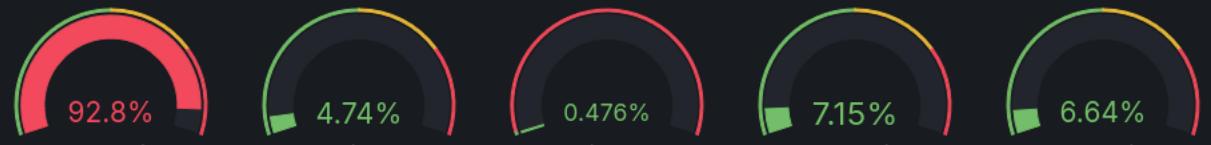


TER-F02-AR1-Sessions

TER-F02-AR2-Sessions

TER-F02-AR1-Pools

TER-F02-AR2-Pools



cooldeas
INTERNET SERVICE PROVIDER

How

- One thing we have found lacking is that some settings requires a restart of the service, including changes to the IP pools. Most of the common settings only requires a config reload.

- Jaco Kroon submitted a patch to the linux kernel to increase the hash bucket size. Testing in a lab environment improved throughput by about 100kpps, still needs to be tested in a live environment

- Also something that bit us is that since there is not a lot of entropy in PPPoE packets (ie dstmac on an incoming frame is constant) the traffic tends to be mapped to a single queue from the NIC which then get processed by a single core. Setting up RPS to distribute the load across all available core on the NUMA node that the NIC is connected too resolves the issue.

- Current estimation is that one of these GigaByte servers can handle about 9000 users, barring any sort of additional processing like NAT. With NAT (nftables) we have had as many as 6000 users connected (2000 natted) on a server without user complaints, with a hard ceiling of 36Gbps on the dual 40G NIC's we have installed. The 36G is imposed by the PCIe bus running 8 lanes at PCI 3.0 (LnkSta: Speed 8GT/s, Width x8).

- Foot notes:
  - https://finance.yahoo.com/news/vyos-networks-announces-acquisition-accel-100000718.html
  - https://accel-ppp.org/

Questions ?