

## **I think I need a 100G perfSONAR Node!**

### **EPOC advice on evaluating if this is a good idea or not**

EPOC Contact Point: Jason Zurawski (zurawski@es.net)

Last Edit: December 7, 2020

### **The Request**

As part of the EPOC consultation process, we are often contacted for advice in using or deploying 100G perfSONAR nodes. Our response generally begins by suggesting that the folks take a step back and first identify their use cases versus making a choice based on technology availability, and to take a closer look at the networks they're trying to test. One mis-step that sites make when looking at 100G perfSONAR nodes is being tempted to build the biggest they can, with the belief that bigger is always better. However, our experience has shown this approach is often counterproductive for several reasons that are apparent when you examine the environment in more detail.

### **Why 100G Isn't the Right Solution**

Regular 100G testing is still not a very common thing. It is true people have capability, but two things happen when a 100G node is available:

- A. People use the perfSONAR lookup service and test to it (because they can), which can generate 100G traffic bursts at a campus, through a regional/through a backbone, or across an exchange point.
- B. People may not know a resource is 100G capable and test against it from their side with a smaller tester (100M/1G/10G/40G). When this happens, the results include microburst loss events, errors, and measured results that are unclear and not indicative of true performance

A perfSONAR node that can run a 100G test will also be significantly more expensive than one that only runs 1 or 10G tests,

### **Our Current Recommendations**

We advocate, as does the perfSONAR consortium (<http://perfsonar.net>), to build more smaller (e.g. 10G) perfSONAR nodes that can be deployed in multiple places as opposed to one giant node that is only in one place. Often having more testers available is far more valuable in practice.

In terms of the hardware, we recommend construction that can scale with the use case:

- CPUs that are high clock rate, so . 3.6Ghz or above, with a moderate core count. A fast CPU is critical to TCP working effectively, and multiple cores will ensure that one core is pinned to testing, which is still a serial operation, and then the others can handle needed system operation (I/O, graphical displays like maddash, etc).
- Maximize the use of the RAM slots when applicable. This doesn't mean getting the largest capacity for each DIMM, but it does mean filling all available RAM sockets with equal size DIMMs. The equal size is important as it greatly helps the machine to balance memory load.

- Main storage can be conservative. For perfSONAR nodes, regular spinning disks for the OS and measurement storage are fine. Alternatively, it is possible to configure a central collector and not worry about local storage at all. Unlike when building a Data Transfer Node (DTN), storage doesn't need to be NVMe/VROC or any more performant variety.
- Use a network card (or cards) that facilitate swappable optics and multiple Ethernet standards. For example, several available options can facilitate 10/40/50/100G operation, depending on the optic or cable used in the QSFP28 port. This way it is possible to have 100G capability if absolutely needed on occasion, but it can be left at a lower more friendly speed at other times.

The advantage of having multiple smaller perfSONAR test points at different locations in the network cannot be overstated.

For more information on how ESnet and the perfSONAR consortium design the perfSONAR test points, we recommend:

- perfSONAR Hardware Requirements: [https://docs.perfsonar.net/install\\_hardware.html](https://docs.perfsonar.net/install_hardware.html)
- General information about perfSONAR: <https://www.perfsonar.net/>