

100G D2D Tx

- Goal
- Method
- Ref
 - Ref spec for 100G DTN
 - 100G Experimental (~2018) DTN Design
 - (~2020) 40/50/100Gbps Capable DTN Design
 - DTN tuning

Goal

Test disk to disk transmission for 100Gbps throughput in local environment

Method

Ref

Ref spec for 100G DTN

100G Experimental (~2018) DTN Design

(captured at 2021) <https://fasterdata.es.net/science-dmz/DTN/100g-dtn/>

If you wish to build a 100G DTN, here are the important hardware considerations:

1. **100G NICs require PCIe Gen3 x16**. If you are assuming all transfers are 4 or more parallel streams, any CPU with a **clock rate greater than 2.5GHz** should be fast enough to push 25Gbps per flow with the standard tuning applied.
2. 100 Gbps (or 12.5 GBytes/sec) of disk IO is challenging.

Based on reports from colleagues, to get this much disk I/O you'll need either:

- **10 SSD NVME PCIe Gen 3 x4 drives** (e.g.: Samsung 950 Pro with U.2 to M.2 2.5" adapters), or 8 high-end NVME PCIe Gen 3 x4 drives (e.g.: Intel DC P3700 with U.2, 2.5" version that has high endurance).
- 24 SSD SATA drives (with two PCIe Gen 3 x8 RAID controllers, or one PCIe Gen 3 x16 controller)

Note that both of these configurations will require a special chassis that can hold that many drives, such as:

- for nvme drives: https://www.supermicro.com/products/system/2u/2028/SYS-2028U-TN24R4T_.cfm
- for SATA drives: <http://www.supermicro.com/products/chassis/2U/?chs=216>

Sample drives that have been used for 100G DTNs include:

- SSD NVME drives:
 - Samsung 950 Pro or Intel DC P3700
- SSD SATA drives:
 - Samsung PM863 or SM863

(~2020) 40/50/100Gbps Capable DTN Design

(captured at 2022) <https://fasterdata.es.net/DTN/reference-implementation>

The total cost of this server was around \$21K in mid 2019. This system is being tested currently, and we be deployed to ESnet in mid/late 2020. Please note that specifics on configuration will be available after full evaluation. Note that this server uses VROC, and requires the purchase of a premium license.

- Base System: Gigabyte R281-NO0 dual socket P 2U server
 - Onboard: VGA, 2 x GbE RJ45 Intel i350, IPMI dedicated LAN
 - 24 x front access U.2 hotswap bays
 - 2 x rear access 2.5" SATA hotswap bays
 - Dual redundant hotswap 1600W PSU
- 2 x Intel Cascade lake **Xeon Gold 6246**
 - 12 cores each
 - **3.3GHz** 165W TDP processor
- 12 x 16G DDR4 2933 ECC RDIMM (**192G total**)
- **10 x Intel P4610 1.6TB U.2/2.5"** PCIe NVMe 3.0 x4 Drives (connect directly to CPU for **VROC**)

- 2 x Enterprise 960G 2.5" SATA SSD (OS, onboard Intel SATA Raid 1)
- VIntel® Virtual RAID On CPU (VROC) RAID 0, 1, 10, 5
- Mellanox **ConnectX-5** EN MCX516A-CCAT 40/50/100GbE dual-port QSFP28 NIC

DTN tuning

<https://fasterdata.es.net/DTN/tuning/>