Aggregate 55+ Gigabits per Second (Gbps) Transmits, 52+ Gbps Receives and 75+ Gbps Bi-Directional Transmissions Achieved Via A Single Workstation With a Single 6x10-Gigabit Ethernet Network Interface Card

- As part of plans to assess the throughput performance of wide-area file transfer applications, GSFC's High End Computer Network (HECN) Team tested a HotLava six-port 10-Gigabit Ethernet network interface card (NIC) in a HECN Team-assembled workstation that costs less than $ 6,800 with the NIC and achieved aggregate 55+ Gbps transmits, 52+ Gbps receives and 75+ Gbps bi-directional memory-to-memory data transmissions.

- The workstation consists of a 3.2-GHz single-processor (quad core) Intel Core i7 (Nehalem) with one HotLava NIC in one PCIe Gen2 x16 slot of an Asus P6T6 WS Revolution motherboard.

- Transmissions between the above workstation and two other HECN Team-assembled Intel Core i7 workstations with other NICs were measured using the nuttcp (www.nuttcp.net) network-performance testing tool.

- Demonstrations of similar workstations supporting 100 Gbps network testing and near-40 Gbps file transfer applications are planned in the NASA research exhibit at the SC10 conference, New Orleans, LA, Nov. 15–18.
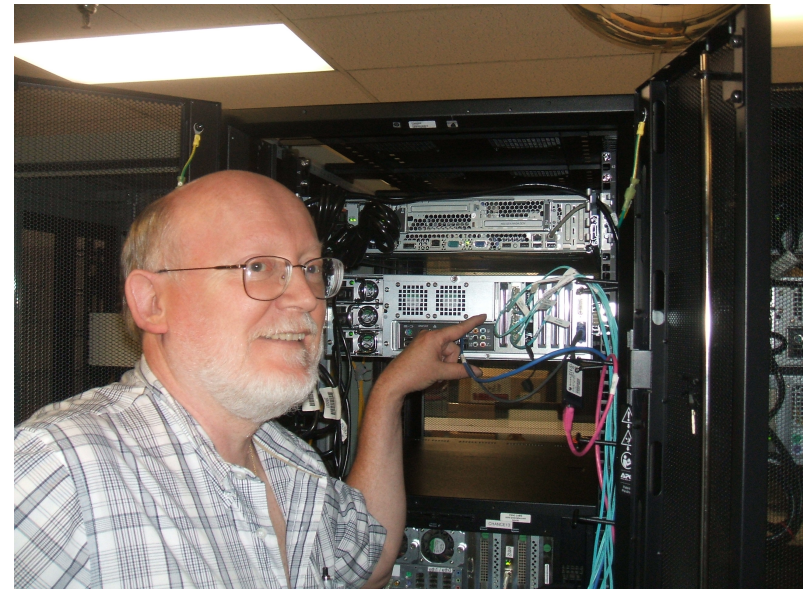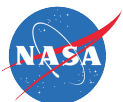

Figure: Bill Fink, author of nuttcp and the throughput performance tests, pointing to the 6x10GE HotLava NIC in the HECN Team's Intel Core i7 based workstation.

POC: Bill Fink, Bill.Fink@nasa.gov, GSFC Computational and Information Sciences and Technology Office

Nuttcp Sample Test Results With One HotLava Systems 6x10GE Tambora 120G6 NIC (1-of-8) [Source: Bill Fink/GSFC]

Configuration of Test Workstations

- Three HECN Team-assembled Intel core i7 server systems (each a 3.2-GHz single-processor (quad core) Intel Core i7 (Nehalem), over-clocked to 3.6 GHz, on an Asus P6T6 WS Revolution motherboard):

    – One using 1 HotLava 6x10-GigE NIC

    – One using 2 Myricom 2x10-GigE NICs

    – One using 1 Myricom 2x10-GigE NIC

Nuttcp Sample Test Results With One HotLava Systems 6x10GE Tambora 120G6 NIC (2-of-8) [Source: Bill Fink/GSFC]

- Theoretical maximum throughput (TMT) on a PCI-E 2.0 x16 card is nnn/(nnn+24)*8/10*16*5 Gbps, where:

  nnn = PCIe MaxPayload on the test systems
   24 = PCIe protocol overhead consisting of:
      1 Byte   Start of Packet (STP)
      2 Bytes   Sequence Number
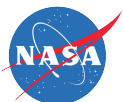   16 Bytes   Header (only 12 Bytes if < 4 GB address)
      4 Bytes   LCRC
      1 Byte   END
   8/10 = 8B10B signalling encoding
    16 = number of lanes
    5 = 5 Gbps per lane for PCIe 2.0

Nuttcp Sample Test Results With One HotLava Systems 6x10GE Tambora 120G6 NIC (3-of-8) [Source: Bill Fink/GSFC]

- With MaxPayload defaulted to 128,   the TMT = 53.8947 Gbps

- With MaxPayload increased* to 256**, the TMT = 58.5142 Gbps

  *Via the setpci command not only on the 6 10-GigE interfaces,
    but also on all the PCIe bridges and the Intel X58 I/O Hub in
    the data path, and with the PCIe MaxReadReq increased from
    512 to 4096 only on the 6 10-GigE interfaces

  **Not 512 because 256 is the maximum value supported by the
    Intel X58 I/O Hub
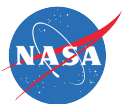
Nuttcp Sample Test Results With One HotLava Systems
6x10GE Tambora 120G6 NIC (4-of-8) [Source: Bill Fink/GSFC]

- 60 Second Transmit Test With MaxPayload = 128
  - n2: 8172.4651 Mbps 28 %TX 31 %RX 0 retrans 0.07 msRTT
  - n3: 8170.6930 Mbps 28 %TX 35 %RX 0 retrans 0.08 msRTT
  - n6: 8167.1622 Mbps 28 %TX 30 %RX 0 retrans 0.09 msRTT
  - n7: 8167.5251 Mbps 28 %TX 31 %RX 0 retrans 0.06 msRTT
  - n5: 8165.5400 Mbps 21 %TX 29 %RX 0 retrans 0.06 msRTT
  - n4: 8160.1735 Mbps 21 %TX 29 %RX 0 retrans 0.05 msRTT
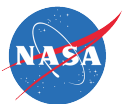
  - Aggregate throughput:
    49003.5589 Mbps

Nuttcp Sample Test Results With One HotLava Systems
6x10GE Tambora 120G6 NIC (5-of-8) [Source: Bill Fink/GSFC]

- 60 Second Receive Test With MaxPayload = 128
    - n3: 9059.9860 Mbps 25 %TX 30 %RX 0 retrans 0.12 msRTT
    - n6: 8391.6758 Mbps 16 %TX 26 %RX 0 retrans 0.12 msRTT
    - n4: 8389.4628 Mbps 16 %TX 23 %RX 0 retrans 0.11 msRTT
    - n2: 9057.1408 Mbps 23 %TX 30 %RX 0 retrans 0.10 msRTT
    - n7: 8391.6331 Mbps 16 %TX 29 %RX 0 retrans 0.11 msRTT
    - n5: 8385.0556 Mbps 16 %TX 23 %RX 0 retrans 0.10 msRTT

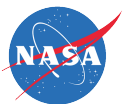   - Aggregate throughput:
        51674.9541 Mbps

Nuttcp Sample Test Results With One HotLava Systems 6x10GE Tambora 120G6 NIC (6-of-8) [Source: Bill Fink/GSFC]

- 60 Second Transmit Test With MaxPayload = 256
    - n6: 9220.9229 Mbps 29 %TX 29 %RX 0 retrans 0.06 msRTT
    - n3: 9224.9003 Mbps 29 %TX 39 %RX 0 retrans 0.07 msRTT
    - n4: 9217.6819 Mbps 23 %TX 30 %RX 0 retrans 0.06 msRTT
    - n7: 9220.6031 Mbps 29 %TX 30 %RX 0 retrans 0.05 msRTT
    - n5: 9217.3856 Mbps 23 %TX 31 %RX 0 retrans 0.06 msRTT
    - n2: 9224.8250 Mbps 29 %TX 34 %RX 0 retrans 0.05 msRTT

    - Aggregate throughput:
      55326.3188 Mbps   (94.55 % of TMT)

Nuttcp Sample Test Results With One HotLava Systems 6x10GE Tambora 120G6 NIC (7-of-8) [Source: Bill Fink/GSFC]

- 60 Second Receive Test With MaxPayload = 256
  - n2: 8673.6754 Mbps 22 %TX 29 %RX 1 retrans 0.12 msRTT
  - n3: 8671.5590 Mbps 24 %TX 30 %RX 0 retrans 0.10 msRTT
  - n6: 8673.8524 Mbps 16 %TX 28 %RX 0 retrans 0.12 msRTT
  - n7: 8671.3342 Mbps 16 %TX 27 %RX 0 retrans 0.10 msRTT
  - n4: 8673.6880 Mbps 17 %TX 24 %RX 0 retrans 0.10 msRTT
  - n5: 8666.1076 Mbps 16 %TX 24 %RX 0 retrans 0.11 msRTT

  - Aggregate throughput:
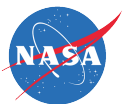   52030.2166 Mbps   (88.91 % of TMT)
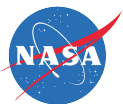
Nuttcp Sample Test Results With One HotLava Systems 6x10GE Tambora 120G6 NIC (8-of-8) [Source: Bill Fink/GSFC]

- 30 Second Bi-Directional Test With MaxPayload = 256
    - n2tx: 6834.9746 Mbps 32 %TX 48 %RX 0 retrans 0.07 msRTT
    - n6tx: 6314.6360 Mbps 33 %TX 21 %RX 0 retrans 0.19 msRTT
    - n3tx: 6195.1905 Mbps 32 %TX 38 %RX 0 retrans 0.06 msRTT
    - n4tx: 8393.6009 Mbps 28 %TX 32 %RX 0 retrans 0.05 msRTT
    - n7tx: 7489.9029 Mbps 32 %TX 27 %RX 0 retrans 0.06 msRTT
    - n7rx: 6627.6585 Mbps 11 %TX 33 %RX 0 retrans 0.23 msRTT
    - n3rx: 3264.0248 Mbps 25 %TX 33 %RX 0 retrans 2.28 msRTT
    - n2rx: 5199.5641 Mbps 37 %TX 32 %RX 0 retrans 0.10 msRTT
    - n6rx: 5117.8068 Mbps 11 %TX 33 %RX 0 retrans 0.33 msRTT
    - n4rx: 6649.6623 Mbps 15 %TX 32 %RX 0 retrans 0.27 msRTT
    - n5rx: 5815.6919 Mbps 12 %TX 34 %RX 0 retrans 6.02 msRTT
    - n5tx: 7214.4784 Mbps 32 %TX 28 %RX 0 retrans 0.07 msRTT

    - Aggregate TX throughput: 42442.7833 Mbps
    - Aggregate RX throughput: 32674.4084 Mbps
    - Total aggregate throughput: 75117.1917 Mbps