Chelsio S3 iSCSI Driver for Linux

Introduction

The Chelsio T3 ASIC based Adapters (S310, S320, S302, S304, Mezz cards, etc. series of products) support iSCSI acceleration and iSCSI Direct Data Placement (DDP) where the hardware handles the expensive byte touching operations, such as CRC computation and verification, and direct DMA to the final host memory destination:

- iSCSI PDU digest generation and verification

  On transmitting, Chelsio S3 h/w computes and inserts the Header and Data digest into the PDUs. On receiving, Chelsio S3 h/w computes and verifies the Header and Data digest of the PDUs.

- Direct Data Placement (DDP)

  S3 h/w can directly place the iSCSI Data-In or Data-Out PDU's payload into pre-posted final destination host-memory buffers based on the Initiator Task Tag (ITT) in Data-In or Target Task Tag (TTT) in Data-Out PDUs.

- PDU Transmit and Recovery

  On transmitting, S3 h/w accepts the complete PDU (header + data) from the host driver, computes and inserts the digests, decomposes the PDU into multiple TCP segments if necessary, and transmit all the TCP segments onto the wire. It handles TCP retransmission if needed.

  On receiving, S3 h/w recovers the iSCSI PDU by reassembling TCP segments, separating the header and data, calculating and verifying the digests, then forwarding the header to the host. The payload data, if possible, will be directly placed into the pre-posted host DDP buffer. Otherwise, the payload data will be sent to the host too.

The cxgb3i driver interfaces with open-iscsi initiator and provides the iSCSI acceleration through Chelsio hardware wherever applicable.

Using the cxgb3i Driver

The following steps need to be taken to accelerates the open-iscsi initiator:

1. Load the cxgb3i driver: "modprobe cxgb3i"

   The cxgb3i module registers a new transport class "cxgb3i" with open-iscsi.

   - in the case of recompiling the kernel, the cxgb3i selection is located at:

     ```
     Device Drivers
             SCSI device support --->
                     [*] SCSI low-level drivers  --->
                             <M>   Chelsio S3xx iSCSI support
     ```

2. Create an interface file located under /etc/iscsi/ifaces/ for the new transport class "cxgb3i".

   The content of the file should be in the following format:

   ```
   iface.transport_name = cxgb3i
   iface.net_ifacename = <ethX>
   iface.ipaddress = <iscsi ip address>
   ```

   - if iface.ipaddress is specified, <iscsi ip address> needs to be either the same as the ethX's ip address or an address on the same subnet. Make sure the ip address is unique in the network.

3. edit /etc/iscsi/iscsid.conf The default setting for MaxRecvDataSegmentLength (131072) is too big; replace with a value no bigger than 15360 (for example 8192):

   ```
   node.conn[0].iscsi.MaxRecvDataSegmentLength = 8192
   ```

   - The login would fail for a normal session if MaxRecvDataSegmentLength is too big. A error message in the format of "cxgb3i: ERR! MaxRecvSegmentLength <X> too big. Need to be <= <Y>." would be logged to dmesg.

4. To direct open-iscsi traffic to go through cxgb3i's accelerated path, "-I <iface file name>" option needs to be specified with most of the iscsiadm command. <iface file name> is the transport interface file created in step 2.