The Dolphin D330 Cluster Adapter Card

Petter Gustad, Svein Erik Johansen, Kåre Løchsen
Dolphin Interconnect Solutions AS
Olaf Helsets vei 6, P.O.Box 70 Bogerud
N-0621 Oslo, Norway
{pegu,sej,kl}@dolphinics.no

Abstract

A description of the Dolphin D330 Cluster Adapter Card and its main components and features will be given. The focus will be on the differences between the previous D320 and the new D330.

1 Introduction

Dolphin Interconnect Solutions has been a provider of chipsets, adapter cards, and software for multiprocessor systems for almost a decade. Dolphin was involved in the IEEE standardization group for the Scalable Coherent Interface standard[3]. The objective for SCI was to make a bus which could keep up with the vast increase in microprocessor performance. The gap between processor and bus performance was widening as microprocessor performance was increasing at a rate faster than bus performance. The SCI committee concluded that wider and wider shared buses did not scale very well and point to point links was the only way to go[6]. The SCI specification was approved by the IEEE in 1992.

Several cache coherent systems based upon SCI have been built, e.g. the Data General AViiON AV 25000[4]. Surprisingly, many of todays systems don’t implement a cache coherent memory controller, but rather utilize the high bandwidth and low latency provided by SCI. SCI is also very flexible in terms of different topologies which can be built, e.g. simple ringlets, switches, torus’ etc.

Dolphin introduced the first SCI based IO bus attachment for the SBus based SUN Sparc computers in the early 1990’s. Later the PCI to SCI Bridge (PSB) chip was introduced which enabled a low cost interface to SCI from the PCI bus.

2 The D330 Cluster Adapter

The D330 card is the next generation cluster adapter card from Dolphin and will replace the current D320 card. The major enhancements are given by the LC3 and PSB66 which are both new generations of their ancestors, namely the LC2 and PSB64. The major differences between the D320 and the D330 cards are outlined in table 1.

<table>
<thead>
<tr>
<th>Feature</th>
<th>D320</th>
<th>D330</th>
</tr>
</thead>
<tbody>
<tr>
<td>Link Controller</td>
<td>LC2</td>
<td>LC3</td>
</tr>
<tr>
<td>Signal levels</td>
<td>IEEE 1596.3</td>
<td>IEEE 1596.3</td>
</tr>
<tr>
<td>Number of queues</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>Raw link bandwidth</td>
<td>500 MB/s</td>
<td>666 MB/s</td>
</tr>
<tr>
<td>Trans. Controller</td>
<td>PSB64</td>
<td>PSB66</td>
</tr>
<tr>
<td>PCI Frequency</td>
<td>33 MHz</td>
<td>66 MHz</td>
</tr>
<tr>
<td>Max packet size</td>
<td>64 byte</td>
<td>128 byte</td>
</tr>
<tr>
<td>Virtual Channels</td>
<td>no</td>
<td>yes</td>
</tr>
</tbody>
</table>

aWorst case process

The D330 is compatible with the D320 card at both the SCI link layer and at the PCI interface. However, the D330 has to be run at the lower D320 speed when connected on the same SCI ringlet.

3 The Link Controller

Figure 1 shows a block diagram of the LC3. The LC3 implements the interface between the SCI physical link and its backend attachment, namely the B-Link. Some key issues with respect to the LC2 are outlined below:

- 8 transmit and receive queues, 16 total (doubled with respect to the LC2)
- 2 wire serial port, e.g. load config data from EEPROM, management port, statistics etc.
5 The D330 Cluster Adapter Card

The D330 is based on a 8 layer FR4 type PCB, which consists of the PSB66, the LC3 and a FPGA which implements the reset and maintenance controller. A SRAM is used to store the ATT, Address Translation Table, for the PSB66 which describes mapping from PCI address space to SCI address space and vice versa. Two serial EEPROM’s contain configuration settings for the LC3 and the PSB66. The D330 uses JTAG to load firmware and for test purposes.

6 Software

The Dolphin Interconnect Resource Manager is a portable driver available under Windows NT, Windows 2000, Linux (X86, Sparc and Alpha), Solaris, Lynx, and VxWorks. Similarly, the SISCI[1] (Software Infrastructure for SCI) interface is an API for user applications to utilize the services provided by SCI. Scalii[5] has developed a MPI software library for their torus connected D320 based systems. There is also an open source MPI library available from RWTH Aachen[7].

The D330 card can replace the D320 card with none or few modifications. Several system vendors are currently integrating the D330 into their systems. Some typical applications areas for the D330 are compute and database clusters, storage area networks and RAID systems, embedded systems, e.g. signal processing, and high availability systems.

The D330 is compatible with the SCI switches available from Dolphin, which includes the LC2 based D515 as well as the LC3 based D535 crossbar switch.

7 Conclusion and future work

The D330 card will also be available as part of the popular Wulfkit[2] which includes the adapter card, the mezzanine card, cables and Scalii’s software CD. There will also be a companion D333 PMC card as well as the D335: a 1 slot 2D, i.e. dual LC3, card available to build torus topologies.

References