Low-Latency Solutions for Storage-Hungry Embedded Applications

“Flash-on-Ethernet?”

Dr. Endric Schubert, MLE
Dr. David Boggs, MLE
Dr. Ulrich Langenbach, Fraunhofer HHI
Dr. Peter Gregorius, Fraunhofer HHI
Abstract

System-of-systems are loosely coupled Embedded Systems which greatly benefit from the high performance of modern SSD technology. Machine visioning, medical imaging, and advanced driver assist systems are among those storage-hungry applications. However, communication latency and bandwidth in between the systems have a significant impact on the overall robustness, cost and performance.

Current techniques based on fieldbuses such as CAN, Flexray, have begun to hit the bandwidth wall and are more and more replaced by multi-Gigabit Ethernet plus techniques for hardware-acceleration of networking protocol stacks.

We present a proof-of-concept implementation specifically targeted for storage-hungry System-of-systems. Integrated into a modern FPGA with multicore ARM CPUs to run Open Source Linux, single-chip solutions become possible which provide full compatibility with all relevant network and storage interface protocols and can reach userspace latencies within few microseconds.
Missing Link Electronics is …

We are a Silicon Valley based technology company with offices in Germany. We are partner of leading electronic device and solution providers and have been enabling key innovators in the automotive, industrial, test & measurement markets to build better Embedded Systems, faster.

Our mission is to develop and market technology solutions for Embedded Systems Realization via pre-validated IP and expert application support, and to combine off-the-shelf devices with Open-Source Software for dependable, configurable Embedded System platforms.

Our expertise is I/O connectivity and acceleration of data communication protocols, additionally opening up FPGA technology for analog applications, and the integration and optimization of Open Source Linux and Android software stacks on modern extensible processing architectures.

MLE is a technology partner of Fraunhofer Heinrich-Hertz-Institute, a Certified Xilinx Alliance Partner, a member of the Altera Design Service Network, and an active contributor to the Open Source software ecosystem.
Founded in 1949, the German Fraunhofer-Gesellschaft undertakes applied research of direct utility to private and public enterprise and of wide benefit to society. With a workforce of over 23,000, the Fraunhofer-Gesellschaft is Europe’s biggest organization for applied research, and currently operates a total of 67 institutes and research units. The organization’s core task is to carry out research of practical utility in close cooperation with its customers from industry and the public sector.

Fraunhofer HHI was founded in 1928 as "Heinrich-Hertz-Institut für Schwingungsforschung“ and joined in 2003 the Fraunhofer-Gesellschaft as the "Fraunhofer Institute for Telecommunications, Heinrich-Hertz-Institut.“ Today it is the leading research institute for networking and telecommunications technology, “Driving the Gigabit Society”.
Hardware Acceleration of Protocols

- PCI Express
  - Gen 2: 3Gbps
  - Gen 3: 6Gbps
- NVM Express
- TCP/IP
- UDP
- 10GE

Timeline:
- 2011
- 2012
- 2013
- 2014
Systems-of-Systems
SSD Architecture Overview

- Others: Flash Controller
- MLE: Interface Controller

(Courtesy: SNIA.org)
Flash-on-Ethernet

- Fast Network i/f
  - 10GbE / 40GbE
- Dependable Protocol
  - iSCSI
- Future-proof SSD i/f
  - NVMeExpress
Benefits of NVMexpress

- Built for PCIe and Flash
- Multi-Queue Facilitates Acceleration
Benefits of Ethernet

Ethernet Speeds – Log

Distance vs Speed

Ethernet operates at different speeds over different distances depending on the media:
- backplanes up to 1m
- Twinax to 15m
- Twisted pair to 100m
- Multimode fiber to 5km
- Single-mode fiber to 40km

(Courtesy: Ethernet Alliance)
Benefits of Ethernet

Add Some History and Map it to Port Volume

Server Class Adapter & LOM Ethernet Ports

Source data: Crehan Research, 2012
IEEE 802.3 Higher Speed Ethernet Consensus Ad Hoc
September 2012
400 Gigabit/s Ethernet

WDM & Next Generation Optical Networking 2014 conference in Nice, France (courtesy Xilinx)
Benefits of iSCSI

L2+L3 10 Gbps Ethernet Switch Bandwidth – Data Center

Switch Bandwidth in Gigabytes

© 2014 Dell’Oro Group
Challenges with iSCSI

Computational complexity, Unpredictable latency

Data Encapsulation into Network Packets

- Ethernet Header
- IP
- TCP
- iSCSI
- Data
- CRC

Delivery of iSCSI Protocol Data Unit (PDU) for SCSI functionality (initiator, target, data read/write, etc.)

Reliable data transport and delivery (TCP Windows, ACKs, ordering, etc.) Also demux within node (port numbers)

Provides IP “routing” capability so that packet can find its way through the network

Provides physical network capability (Cat 5, MAC, etc.)

SCSI to iSCSI Mapping

SCSI Command and Data

- iSCSI PDU
- Header
- Data
- IP Packet

iSCSI PDU alignment with packets varies
Hardware Acceleration Enabled by Modern All-Programmable SoC

Programmable I/O
Programmable Software
Programmable Logic

“Put the processing burden where it belongs!”
Network Protocol Acceleration Technology from FhG HHI

<table>
<thead>
<tr>
<th>Year</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>Mask-less lithography systems &lt;br&gt; Published in 2006 &lt;br&gt; XILINX VIRTEX-II</td>
</tr>
<tr>
<td>2008/09</td>
<td>1GbE TCP/IP stack &lt;br&gt; Demonstrated at 2009 IFA &lt;br&gt; Uncompressed full HD video transfer</td>
</tr>
<tr>
<td>2010</td>
<td>10GbE TCP/IP stack &lt;br&gt; Uncompressed full HD video transfer &lt;br&gt; Mask-less electron beam lithography</td>
</tr>
<tr>
<td>2012</td>
<td>10GbE TCP/IP stack &lt;br&gt; PCIe IP core &lt;br&gt; Uncompressed full HD video transfer &lt;br&gt; High Frequency Trading &lt;br&gt; High Performance Computing &lt;br&gt; Mask-less electron beam lithography</td>
</tr>
</tbody>
</table>

Fraunhofer Heinrich Hertz Institute

contact-HS@hhh.fraunhofer.de
Network Protocol Acceleration Technology from FhG HHI

Hardware Accelerated Internet Protocol

- System Management
- Applications, e.g. Video Processing
- TCP/IP
- 10 GbE

High Speed Hardware Architectures

- Driver
- API
- APP

Fraunhofer Heinrich Hertz Institute

contact-HS@hh.fraunhofer.de
Network Protocol Acceleration
More Than TOE

Fraunhofer HHI
- Entire TCP / UDP protocol processing inside FPGA
- Option to run Application Layer processing in HW, too!

State-of-the-Art
- Software-only
- TCP Offload Engine (TOE)
- requires CPU
Network Protocol Acceleration Processing Devices

- Built on top of leading FPGA technology, MLE offers a complete integrated device solution encapsulating Network Protocol
- Acceleration for 1GbE, 10GbE, and beyond, plus Application Specific Protocol handling in hardware and/or software.
- Example FPGA technology
  - Altera: Stratix-5, Arria-10, Stratix-10
  - Xilinx: Virtex-6, Virtex-7, Kintex-7, Zynq-7, Kintex-Ultrascale, Virtex-Ultrascale
Network Protocol Acceleration
Best in Class Performance

- Stand-alone TCP/IP & UDP/IP stack
- Point-to-point 1GbE or 10GbE
- Full line rate of $\text{TPR}_{\text{max}} = 9.5896$ Gbps
- TCP R/W latency of $\text{TTR(W)} \geq 1.4 \mu s$
- UDP R/W latency of $\text{TUR(W)} \geq 0.75 \mu s$
- Round trip time of $\text{RTT}_{\text{min}} \geq 2.25 \mu s$

(2013 benchmarking data from Fraunhofer HHI)
FPGA Implementation via High-Level Synthesis

- Xilinx XAPP1209 – Designing Protocol Processing Systems with Vivado HLS
Flash-on-Ethernet Architecture

Configurable, elastic system
Balance data rates for Latency and Bandwidth

- SSD
  200k IOPS
  800 MB/s
  PCIe Gen2 x2
  10GbE

- SSD
  800k IOPS
  3 GB/s
  PCIe Gen3 x4
  40 GbE
Flash-on-Ethernet Lab Setup at MLE

- Avnet Mini-ITX
- Xilinx Zynq 7045
- PetaLinux
- AHCI SSD via PCIe
- NPAP
- 10GbE
Preliminary Results

- Good determinism
- Reasonable Latency
Outlook

- Embrace Faster Ethernet: 25GbE, 40GbE

(Courtesy: Brad Booth, Microsoft, 25G Ethernet CFI)
References


Contact Info

Missing Link Electronics
www.missinglinkelectronics.com

Endric Schubert
endric@MLEcorp.com
Phone US: +1 (408) 320-6139
Phone DE: +49 (731) 141149-66