Compact Packaging and Liquid Cooling Technology for Exascale

Including Achievements from ExaNeSt and Concepts and Achievements from EuroEXA

by Peter Hopton, Fabien Chaix, e.a.; Iceotope, UK; and FORTH, Heraklion, Crete.
The QFDB architecture

- **FPGA**
  - XCZU9EG-ffvc900
  - Network FPGA
    - XCZU9EG-ffvc900
    - Storage FPGA
      - XCZU9EG-ffvc900

- **Memory**
  - 16 GB _DDR4 SODIMM_
  - 16 GB _M.2 (NVMe) 512GBYTE_
  - 64 MB _DDR4 SODIMM_
  - 64 MB _QSPI_
  - 64 MB _QSPI_

- **Connections**
  - 12+12 LVDS
  - 2 GTH
  - 10 GTH
  - RGMI
  - 10x UART
  - 4PS- GTR
  - 2x UART
The QFDB implementation

Rev. A version
→ DDR issue (erroneous swap)
→ 48V → 12V regulator issue

Rev. B version
→ 48V → 12V regulator issue
→ SSD (configuration) issue

<table>
<thead>
<tr>
<th>Key facts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimension</td>
</tr>
<tr>
<td>Height</td>
</tr>
<tr>
<td>PCB</td>
</tr>
<tr>
<td>Estimated TDP</td>
</tr>
</tbody>
</table>
## Current QFDB Achievements

<table>
<thead>
<tr>
<th>Feature</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zynq functions</td>
<td>Validated individually,</td>
</tr>
<tr>
<td>DDR-RAM</td>
<td>Validated</td>
</tr>
<tr>
<td></td>
<td>Linux runs with DDR4-2133</td>
</tr>
<tr>
<td></td>
<td>Bare-metal test pass with DDR4-2400</td>
</tr>
<tr>
<td>QSPI memories</td>
<td>Validated on Linux</td>
</tr>
<tr>
<td></td>
<td>Minor issues with bare-metal</td>
</tr>
<tr>
<td>Gigabit Ethernet</td>
<td>Validated</td>
</tr>
<tr>
<td>LVDS links</td>
<td>Validated, ~5% of traces are problematic</td>
</tr>
<tr>
<td>I2C busses</td>
<td>Validated all but the M.2 device</td>
</tr>
<tr>
<td>UART links</td>
<td>Validated</td>
</tr>
<tr>
<td>Clock generators</td>
<td>Validated</td>
</tr>
<tr>
<td>M.2 Solid State Disk</td>
<td>Validated on Rev. A,</td>
</tr>
</tbody>
</table>
Remote RAM Linux setup

- We need Linux to validate board sub-systems:
  - PCIe, QSPI, Gigabit Ethernet
- Successful test of external RAM to boot.
  ➔ Accessed memory from another board
  ➔ Forward memory access through Network FPGA
ExaNeSt Tracks

**Track 1**
- 4 QFDBs
- 800W Max.
- 1.5 per u (b2b)
- 45C Inlet
- Passive I/O

**Track 2**
- 16 QFDBs
- 3.2kW Max
- 2 per u (b2b)
- 45C Inlet
- Onboard Switch
QFDB x Track-1 Mezzanine
EuroEXA Concepts

- Shipping Container Optimised Cabinets (2 rows)
- 2MW per shipping container
- Enables 3D stacking to minimise distances

Batteries and Rectifiers for 250kW of load

In 40ft ISO or 10mx2.4m
Advancing to Track 2

- >4x Thermal Improvement Required
- 7x Thermal Improvement Targeted
- Addition of an integrated high speed switch
- Large thermal and mechanical challenge
What This Means For Track 2

<table>
<thead>
<tr>
<th>Key facts</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimension</td>
<td>21” Rack x 762mm Deep</td>
</tr>
<tr>
<td>Nodes</td>
<td>16 x QFDB (or Next Gen)</td>
</tr>
<tr>
<td>Height</td>
<td>1 O-u</td>
</tr>
<tr>
<td>PCB</td>
<td>Megtron-6</td>
</tr>
<tr>
<td>Estimated TDP</td>
<td>3.2kW</td>
</tr>
<tr>
<td>I/O (Uplink)</td>
<td>800Gb/s</td>
</tr>
<tr>
<td>I/O (Total)</td>
<td>2Tb/s</td>
</tr>
</tbody>
</table>
Tackling the 7x Cooling Challenge
Proving it will work!

Scope

• Track 2 horizontal chassis
• Thermal proxy QFDBs x16, nominal
  TDP 200W (3.2kW overall)

• Critical conditions
  • Junction temperature < 100°C
  • Maximum PC temperature < 70°C

• Three tests defined
  • Critical SC inlet temperature
  • Critical TDP duty
  • Critical SC flow-rate
Proving it will work!

Results at 45C Inlet
Proving it will work!

Results at 45°C Inlet
Thermal Validation Summary

Summary

- Based on stated nominal conditions, the following critical conditions were determined:
  - SC feed temperature is 47°C
  - TDP load factor is 1.1 (1,084 W)
  - SC flow-rate is 1.6

- In each case, the Primary Coolant max temperature was found to be the limiting factor.

- These findings still need further verification experimentally.
Many Thanks
Please Connect at thecoolingguy.me

Presenter – Peter Hopton
Dissemination Lead, EuroEXA, ExaNeSt
Founder, Iceotope
http://thecoolingguy.me