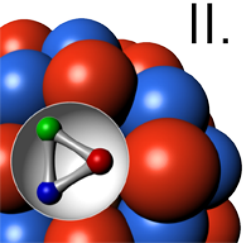
The Compute Node / ATCA System

Björn Spruck for the Giessen Group
II. Physikalische Institut, JLU Gießen
20.10.2011
Online Tracking EVO Meeting

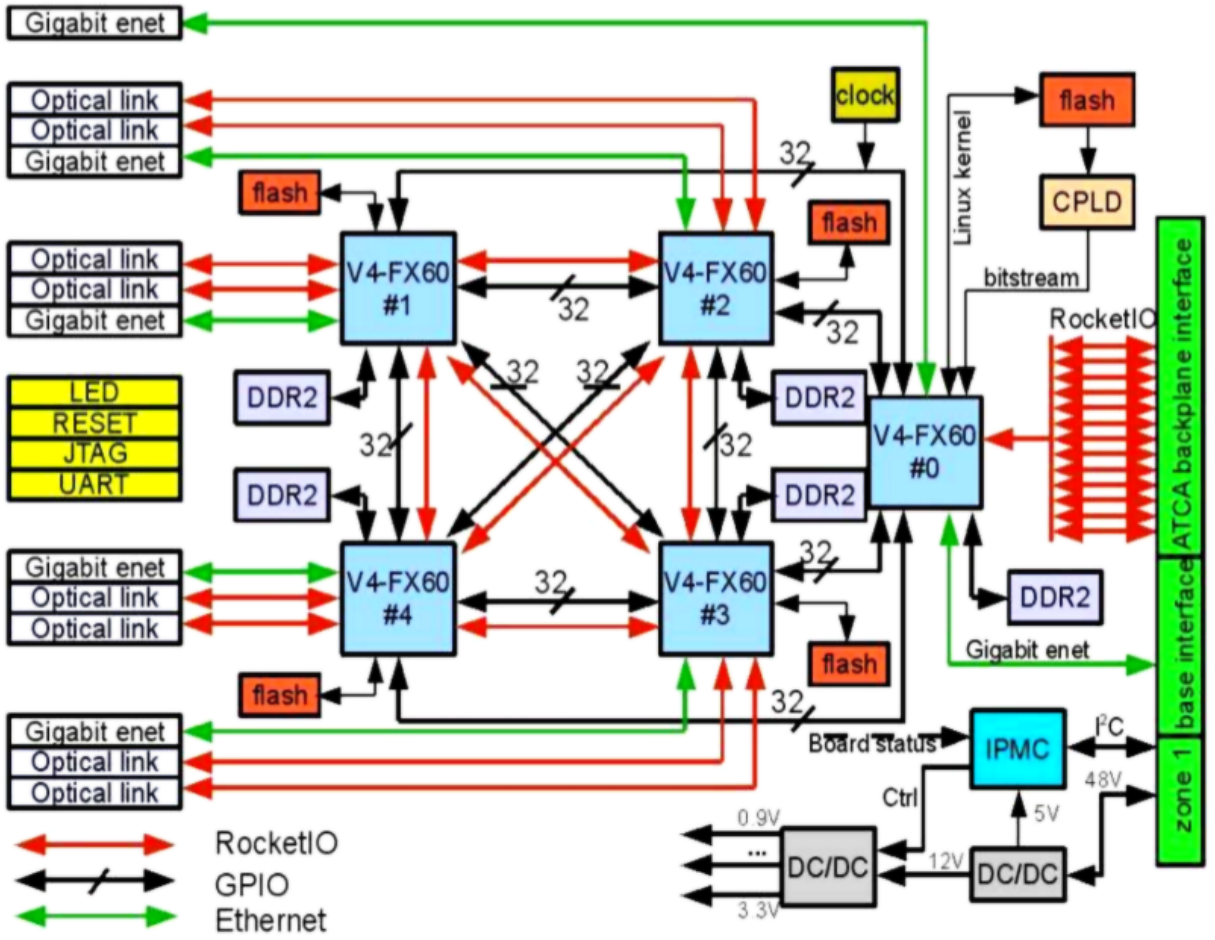


II. Physikalisches Institut
JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

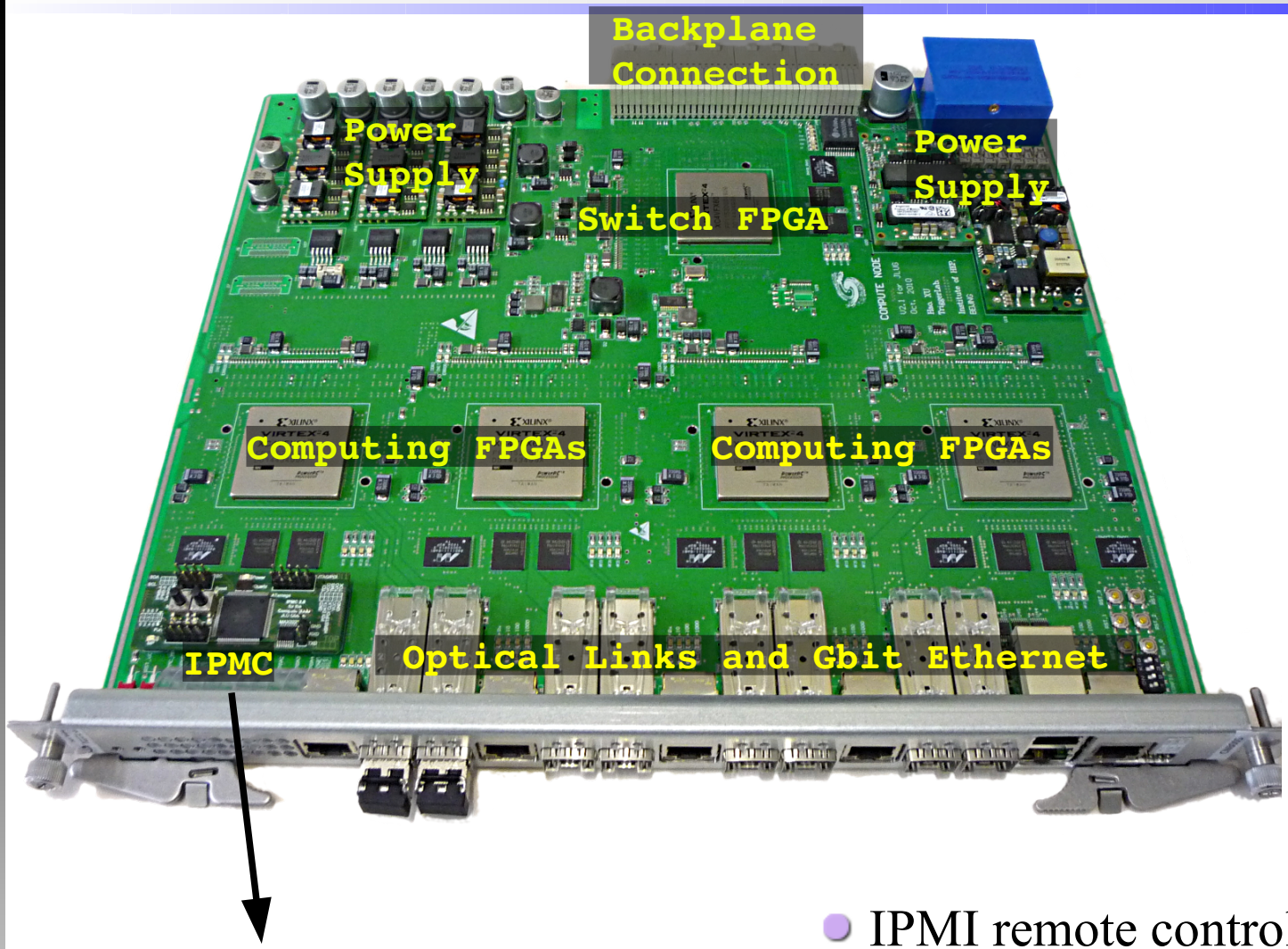
bjoern.spruck@physik.uni-giessen.de

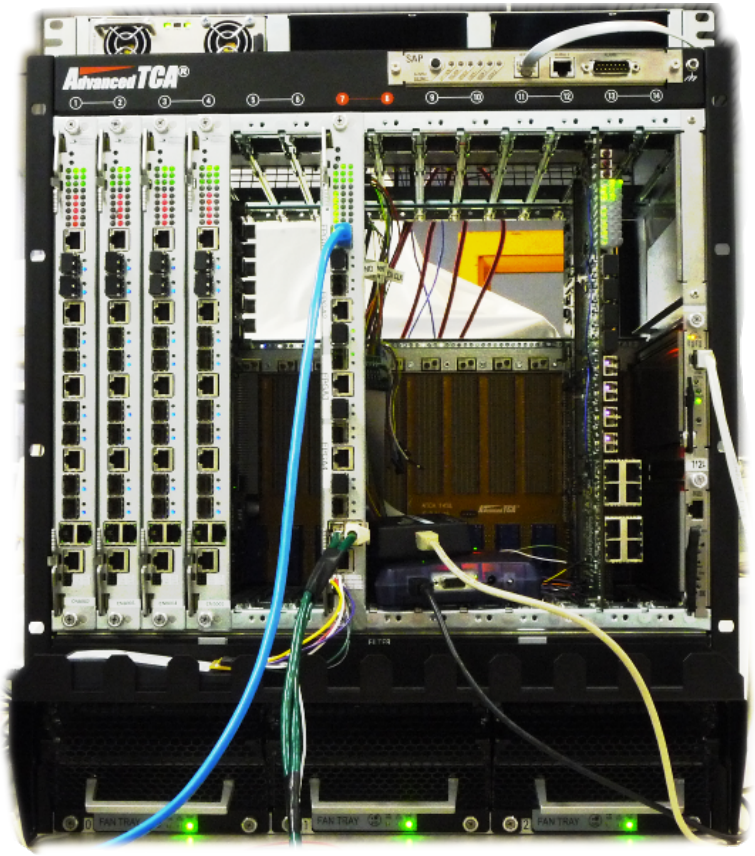
The Compute Node (CN)

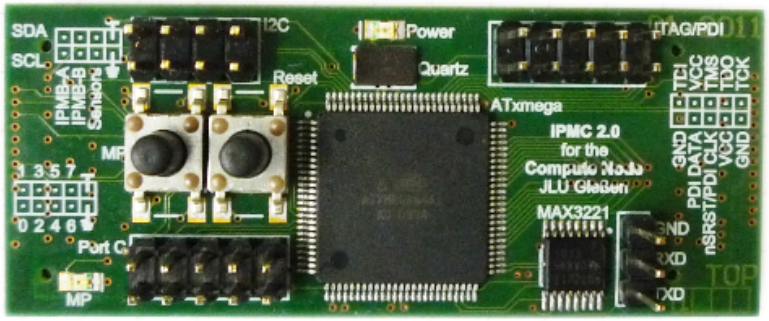
- Developed in close cooperation with Trigger Lab of IHEP Beijing, China
- High Performance Computing
  - 5 Virtex-4 FX60 FPGA
    - (upgrade: Virtex 5)
  - 5*2GB DDR2 RAM
    - (upgrade: 4GB)
  - interconnected by RocketIO
- ~32Gbps Bandwidth
  - 8 Optical Link (3Gbps each)
    - (upgrade: 6.5Gbps)
  - 5x Gigabit Ethernet
- 13x RocketIO to backplane (full mesh)
- 2 embedded PowerPC in each FPGA for slow control
- ATCA compliant



Compute Node, 2nd version

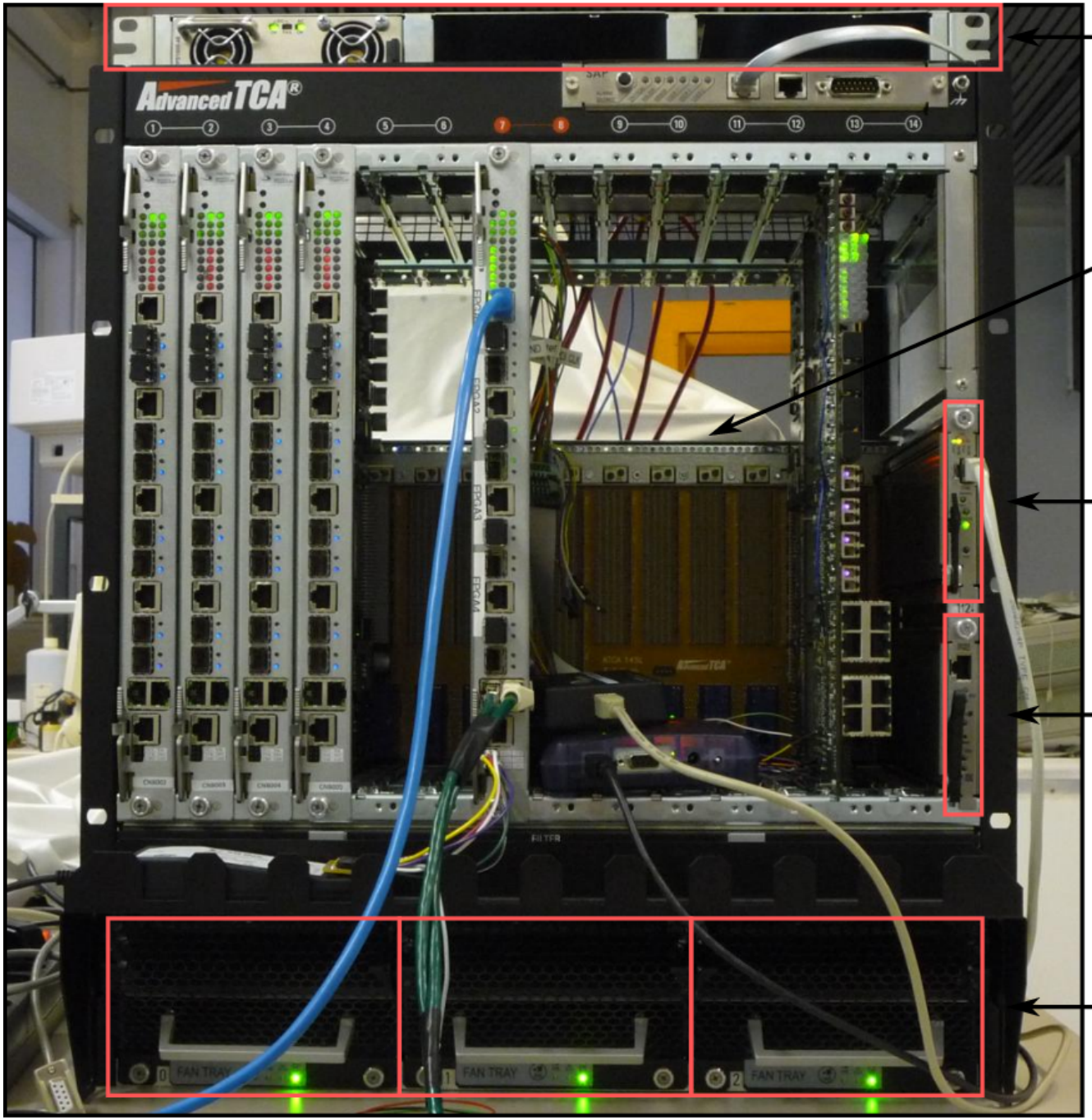






- IPMI remote control
  - Small piggy-back board
  - Functions: power on/off, power negotiation, health monitoring, board reset, bitstream selection

ATCA Shelf



3 kW Power Supply

Full Mesh Backplane

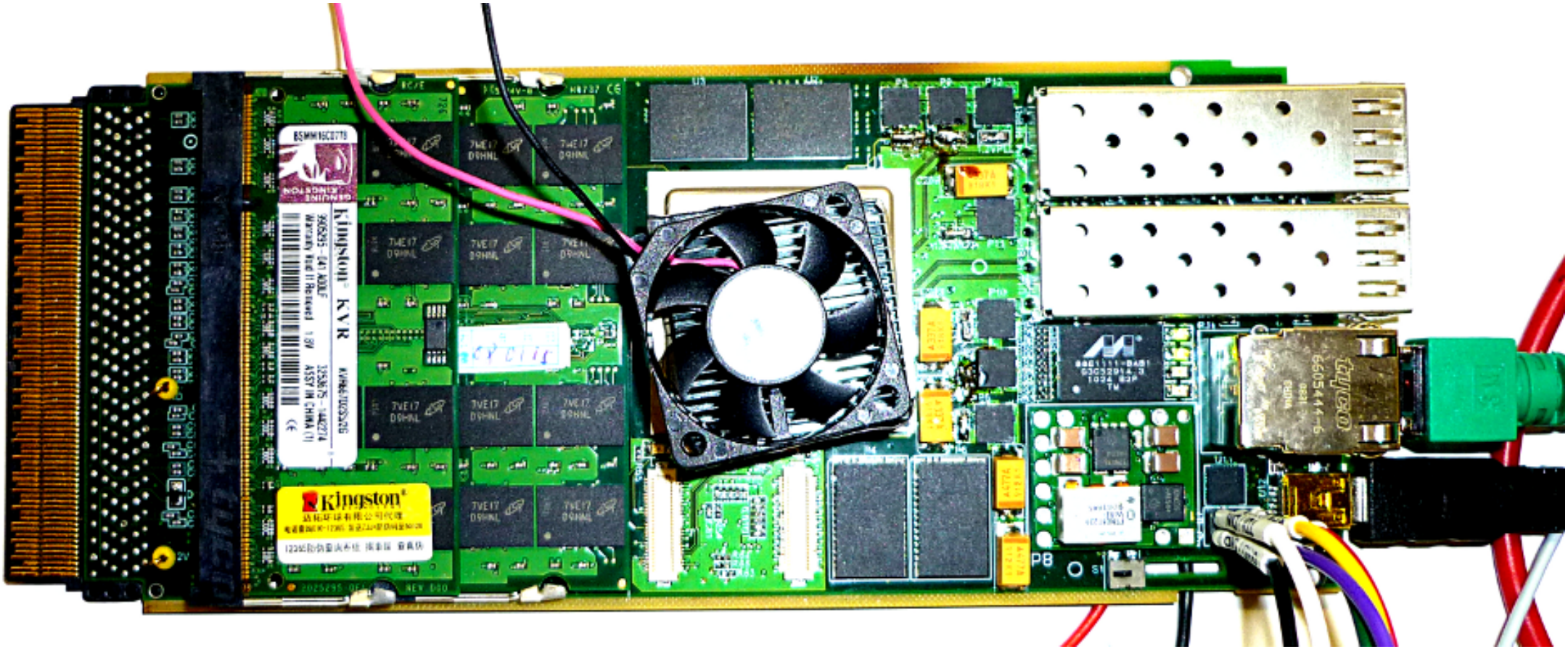
Primary
Shelf Manager

Backup
Shelf Manager

Fan Trays

Update: xTCA Board

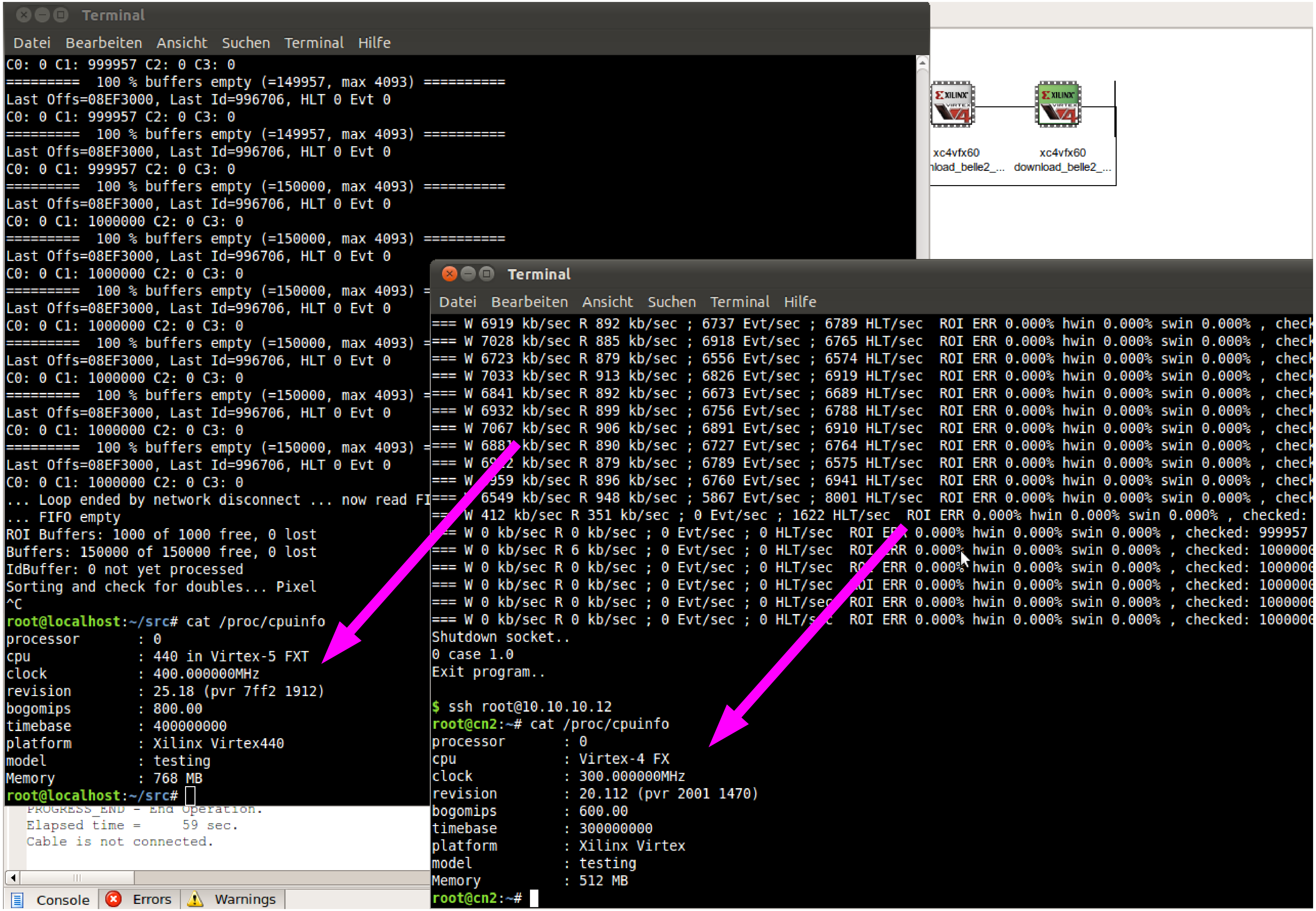
- Two prototype boards, one in Gießen
  - Daughterboard can run standalone!
- Motherboard close to PCB production



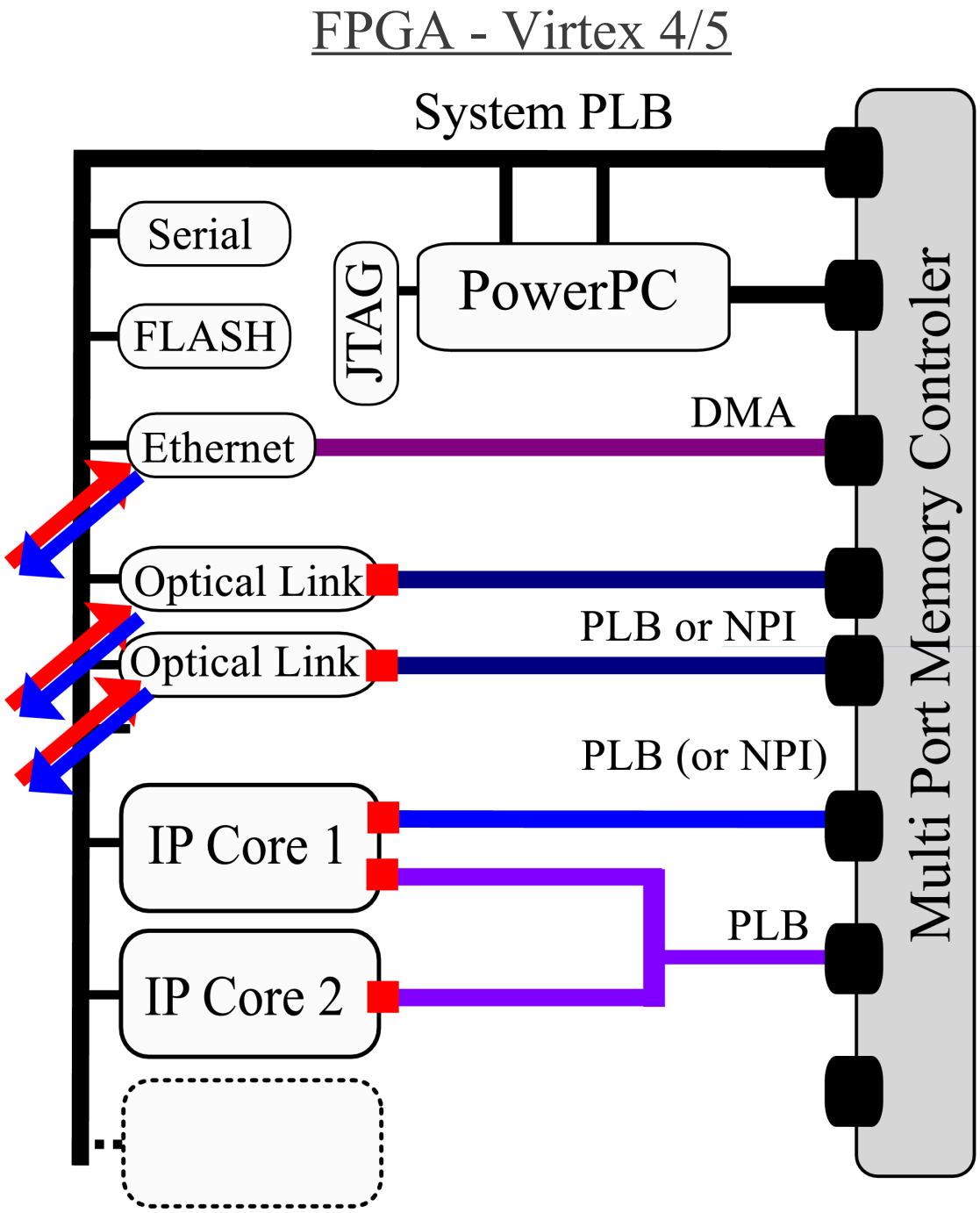
Environment

- Operating System:
  - PowerPC embedded linux, kernel 2.6.37 (from 2011)
  - Two setups
    - Standalone from initramfs
    - Complete system by NFS if available; including gcc, ssh, ...
  - Software TCP/IP (slow)
- Full memory is accessible by PowerPC, but not controlled by Linux
  - Full memory and hardware addresses (IP core) accessable from user program
- Memory access from PowerPC and IP Core → cache problems possible!
  - Use hardware registers for control
- **PowerPC controls the IP cores, but does no computing itself.**
  - **Finally, no data transfered by PowerPC, only "slow" control**

Linux on Virtex4 and 5, Example "Firmware"



System Layout



- **Simple data flow:**
  - Receive by optical link
  - Preprocess / write to memory
  - Read buffered data
  - Process data
  - Write to buffer
  - Send out by optical link

Backups

Real Hardware