Agenda

• Introduction
• Electrical I/O
• Optical I/O
• Q&A
Moore's Law - Evolution

- **Gordon Moore published observations**
  - April 1965 - Electronics: transistors per chip 2X every 12 months.
- **David House while at Intel in 1980's**: performance doubles every 18 months.
- **Feb. 2003** - ISSCC Gordon Moore keynote: 2X every 24 to 36 months.

"Moore's Law has been the name given to everything that changes exponentially in the industry.

Carver Mead introduced the term *Moore's Law* after seeing Gordon Moore's original paper.
Chip performance will improve by combining cores, memory controllers, very large L2/L3 caches, I/O hubs, and special purpose processors using high-performance on-chip busses.

Increasing cache capacity will improve memory subsystem performance, reducing the required off-chip memory bandwidth and improving latency.
System level considerations

Fundamental ‘game changes’:
- Bandwidth: I/O bandwidth is becoming the limiting factor to leverage CPU performance!
- Size/Distance: The critical I/O link distance is between 2mm for on-MCM links and 30m between racks.
- Power/area/C4’s: Processor I/O power consumption, chip area and available numbers of C4 terminals are most limiting factors.

For the future, several required I/O classes can be identified:
- On-MCM
- On-Board (zero-one connector)
- Backplane (1m, two connectors)
- Shelf-to-shelf, rack-to-rack

---

**Pentium 1 Vitals Summary Table**

- **Introduction date**: March 22, 1993
- **Process**: 0.8 micron
- **Transistor Count**: 3.1 million
- **Clock speed at introduction**: 60 and 66 MHz
- **Cache sizes**: L1: 8K instruction, 8K data

---

*Source: The Earth Simulator*
Electrical I/O
12.5 GByte/s RX I/O Macro:
10 lanes @ 10 Gbps with scrambling and FEC

- 10 GByte/s on 5 mm² area with < 1.5 W @ 1V power supply
- Layout with stackable footprint

Diagram:
- RX frontend
- Phase rotator
- Clock buffers
- Data logic (ECC, de-skew, scrambling, training...)
- PR Control
- BIST interface
- Data output @ ¼ rate = 2.5 - 3.2 GHz bus clock
- Quarter rate data from individual channels
- 32+1 (40 raw)
12.5 GByte/s RX I/O Macro C4 “4 on 8” Footprint

- RX frontend
- Phase rotator
- CDR logic
- PR Control
- Clock buffers
- Data logic (ECC, de-skew, scrambling, training,...)
- Quarter rate data from individual channels
- RF PLL
- BIST interface
- Data output @ ¼ rate = 2.5 - 3.2 GHz bus clock
- 32+1 (40 raw)

Phase rotators
Shared RF PLL
One channel analog front-end

Chip photo for 90nm CMOS SOI

- Signal
- Analog GND
- Analog VDD
- Digital GND
- Digital VDD
- Other (Ref..)
- Macro level logic
- CDR logic loop

Dimensions:
- 2.8448mm
- 0.9144mm
Game changers: I/O terminal to I/O device/macro wiring

**C4 to I/O wiring:**
- Up to 3 Gbps: Short on-chip data wiring is allowed but total I/O BW is limited due to slow speed of the lanes.
- More BW is obtained by optimizing the speed (= data-rate) for each C4.
- **BUT:** The on-chip wiring distance between C4 terminal and I/O circuits has to be optimized for a jitter budget closure.

**Consequences:**
- The I/O macro and the C4 I/O terminals have to be placed in close proximity.
- If the I/O has to own the area around the C4 terminals, the optimum data-rate is such that the area is just filled.
Game changers: I/O area & speed vs C4 pitch

"What if" question: Transition from “4 on 8” C4 pitch (200um) to “1 on 2” (50um)

**I/O area:**
- Up to 3 Gbps: Minimum area ‘somewhere’ on chip (see previous chart)
- Above 3 Gbps: I/O macro owns area beneath a C4 square

![Diagram of I/O area comparison between Conventional C4 and “1 on 2” C4]

**Consequences:**
- C4 pitch and CMOS $f_t$ number plus target link distance determine the optimum I/O data-rate.
- Narrow C4 pitch plus medium-speed link design enhance aggregate chip throughput
Area/performance example: CMOS SOI 90nm Phase Rotator Layout Comparison

**CML Type**
- Active area: 76 µm X 61 µm
- Speed 13 GHz
- Configuration: 6 phases input to 1 phase output

**CMOS Type**
- Active area: 80 µm X 30 µm
- Speed 8 GHz
- Configuration: 16 phases to 1 phase output
- 6-to-1 configuration area: 30 µm x 30 µm

0.2 x area @ 0.6 x speed ⇒ 3x bandwidth per area improvement
Area shrinkage demonstrator: Measured 10 Gbps operation

Measurement setup:

![Measurement setup diagram]

Measured outputs:

![Measured outputs graph]
Optical I/O
Equalization

- 8-inch FR4 Board (measured S-parameters)
- HyperBGA Package (TX & RX)
- 700fF for ESD/C4 (TX & RX)

Channel S21 response:

- No Equalization
  \[ H(z) = 1 \]
- 1tap Equalization
  \[ H(z) = 0.79 - 0.21z^{-1} \]

Remark: Only minor improvement in jitter for higher order FIR filters (for this channel)
I/O link distance enhancement

- Short electrical CMOS link design for optimum power & area
- Optical extension for optimum distance
Our Technology-Approach

Electronics: Cable $\rightarrow$ Printed Circuit

Optics: Fiber $\rightarrow$ Integrated Waveguides
Waveguide Manufacturing

FR4 substrate

Lower cladding

Epoxy resin

Cu layer

Woven glass fiber bundles

FR4

Core layer

WG patternning

Upper cladding

WG core

upper cladding

lower cladding

FR4

50 μm

50 μm

35 μm

35 μm
Propagation losses

Experimental results
- Consistent losses for WG width of 30, 50, and 70 µm
- Clearly increased losses in the 2nd and 3rd telecom window

Increased losses come from 2nd and 3rd overtones of hydrocarbon (C-H) bond vibrations (absorption peak @ 3.39 µm)

Possible solution: Fluorination, i.e. replacing C-H groups by C-F groups

0.04 dB/cm loss @ 850nm
Bending Losses

Mask layout

Photograph @ \( \lambda = 640 \text{ nm} \)

Micrograph of 50-\( \mu \text{m} \)-WG bends

Measurement results

- 0.1 dB loss per 180°-bending of radius R = 20 mm
Crossing Losses

Mask layout

Micrograph of 30-µm-crossing

Top view

Purpose

- Waveguide channel routing

Measurement results

- Loss per 90°-crossing: 0.02 dB (@ 850 nm)
  (Example: 100 crossings add up to only 2 dB)
Y-Splitters

Micrograph of 50-µm-splitters

Purpose
- Required for non point-to-point links

Measurement scheme

Experimental results
- 0.10 dB excess loss for 50%:50% splitting (@ λ = 850 nm)
- 0.17 dB excess loss for use as combiner (@ λ = 850 nm)
Waveguide Density

("pseudo-standard")

At 10 Gbps channel modulation, this delivers an aggregate data density of 1 TByte/s per inch.

Increase in waveguide channel density
Experiments: 12.5 Gbps Signal over WG Spiral

- Open eye diagrams at 12.5 Gbps through 100 cm waveguide spiral
- Modal dispersion & loss not critical
Classification of coupling approaches

Assumption: Electrical tracks and optical waveguides are both parallel to board surface. OE-components emit/accept light perpendicular to the chip surface.

Consequence: A 90°-bend is required in this E-O path.

Question: Is this task moved to the optical domain or to the electrical domain?

- more "electronics-friendly" (standard package and orientation), but more complex optical part (especially for 2D)
- separation of active OE-component and passive board (repairability); board (w/o OE) is sealed; interface to optics is on surface (servicability)
- etc.

- more "optics-friendly" (effort in optical domain minimized), but more complex electrical part (flex)
- closer interlock between passive and active parts; board (w/o OE) has open slot; interface to optics is within board
- etc.

For an honest and realistic evaluation, a system-level view is crucial.

At this point in time, both approaches have to be considered in more detail.
Waveguide based approach for I/O link extension

• **Approach**
  - Plug-in self-aligned coupler modules
  - Collimated beam coupling concept

• **Features**
  - Simple waveguide structures
  - Potentially cheap and mass-producible

VCSEL

without lens at 65 mm

with lens at 970 mm
I/O Link Technology: Recent progress

Optical wave-guides on electrical FR4 test board
4 x 12 channels = 600 Gb/s aggregate data-rate @ 12.5 Gb/s channels

- Loss <0.05 dB/cm
- 30 - 70 um square wave-guide cross-section
- 85um buried below surface for protection

Wave-guides based on Acrylates (Dupont)


4 polymer stripes with 12 wave-guides in each stripe
Experiments: On-Board Link at 10 Gbps
Video of Passive Positioning

- **Live 10 Gbps link**
  - Setup runs at 10 Gbps, using OE-module #4
- **Passive placement of OE-module**
  - OE-module #4 is removed
  - OE-module #3 is inserted
  - As soon as #3 has electrical contact, the 10 Gbps eyes can be seen again
- **Zoomed views of setup**
  - Scope screen with two 10 Gbps eyes; time axis detail
  - Pattern generator speed
  - Zoom to IR-camera view, showing waveguide between packages, illuminated by some scattered light
Summary & Conclusions

- CMOS technology at 90 nm node (and follow-on nodes) has the inherent potential for >> 10 Gbps signaling rates.
  ➔ Technology folks have done their job

- First level of packaging (chip I/O terminals) is the most limiting factor for leveraging of the CMOS technology potential.
  ➔ Packaging folks are on: Denser C4 pitch would help

- Equalization concepts help increase signaling rates, but as speed goes up, complex equalization increases area significantly.
  ➔ Distance enhancements via optical extension is one potential solution

- Polymer waveguides have the potential to serve as the next generation wires.
  ➔ Physicists & chemicists have done their job

- Optical packaging and optical backplane connector are not finished yet
  ➔ Packaging folks are on: Cheap optical packages are required
Acknowledgements

- This presentation would not have been possible without the excellent inputs from…
  - various development groups in IBM’s STG division
  - the IBM Research division, in particular the Zurich…
    - I/O Link Technology team
    - Photonics team
    - “Movie star” Chistoph Berger, ZRL who did most of the videos