CMOS High-Speed I/O

- Background, Circuits, and Future Trends -

Hirotaka Tamura

Nov. 26th, 2004
Outline

Background
- Trends in IC bandwidth
- Pin-bandwidth bottleneck

Circuit solutions for bandwidth bottleneck
- Multi-bit, multi-port I/O applications
- Clock recovery scheme
- Receiver and transmitter front ends

High-speed-I/O Future trends
On-chip/off-chip data rate trends

- Clock Frequency (MHz) vs. Data Throughput (Mbytes/s)

- Data rate trends with percentages per year:
  - 33%/Year
  - 44%/Year

- External I/O’s (Off-Chip Communications):
  - 10G Base-T
  - Infiniband
  - RapidIO
  - Hyper Transport
  - Fibre Channel

- Processors:
  - ISA
  - PCI
  - PCI-EX
  - Serial-ATA

- Buses:
  - ATA (HDD)
  - Ethernet

- On-chip/off-chip

- Clock Frequency (MHz)
  - 75
  - 80
  - 85
  - 90
  - 95
  - 2000

- Data Throughput (Mbytes/s)
  - 1
  - 10
  - 100
  - 1000
  - 10000

- Year
  - 05
  - 10
Growth Rates of Performance and BW

C: Performance, N: Number of grids, f: frequency

\[ C \equiv N \times f \]

1.68/year 1.26/year 1.33/year

B: bandwidth needed for the performance

\[ B = kC^\alpha \quad \alpha \approx 0.7 \quad \text{(typical)} \] (Rent’s law)

1.44/year

B_{chip}: Maximum obtainable bandwidth

\[ B_{chip} = B_{pin} \times N_{pin} \]

1.25/year 1.20/year 1.04/year
Trend in BW Demands

Relative Performance

Year

(1) B (x 1.44/year)

(2) C (x 1.68/year)

(3)

page 5

FUJITSU
BW Supply and Demand Gap

Relative Performance

Year

10^3

10^2

10^1

10^0

95 00 05 10

(1) BW demand (x 1.44/year)

(2) BW supply (x 1.25/year)

(3)
Implication of Pin-Bandwidth Bottleneck

Before bottleneck
- Pin BW is not the primary concern
- Data rate per pin does not have to be maximized

After bottleneck
- Pin BW can be the most significant bottleneck
  - Both the pin data rate and the number of pin should be maximized
  - Need continuous improvement in circuit topology
Outline

Background
- Trends in IC bandwidth
- Pin-bandwidth bottleneck

Circuit solutions for bandwidth bottleneck
- Multi-Gigabit, Multi-bit I/O applications
- Clock recovery scheme
- Receiver and transmitter front ends

High-speed-I/O Future trends
HSIO Application (@late 90’s)

- High-speed link for server and storage
- 2.5GByte/s/direction (1.25Gb/s x 2 Byte x2)

Cable (5-20m twisted pair)

Multi-processor server

Mesh Link

R : Router
N : Node

High-speed link

Switch chip

I/F chip

DRAM

MPU
Issues (@late 90’s)

Multi-Gb/s, multi-bit, multi-port link

High data rate
- 1.25 Gbps/signal/direction

Use both PCB and cable
- 1.25 Gbps@5m
- 625 Mbps@20m
- Needs equalization
- Amplitude control

Parallel link
- Need per-bit de-skew
- Many transceivers
  (126TRX/chip)

Clocking
- Should allow 100-ppm
  frequency deviation
- 7 clock domains
Clock Recovery

- Means to track the input phase needed
- Use feedback from phase detector to phase generator

Issues:
- How to build
  - Phase detector
  - Phase generator
  - Feedback path for multi-bit link
Bang-bang Phase Detection

Clock-phase extraction from 0101 pattern

Phase extraction from data

### Data-decision clock

<table>
<thead>
<tr>
<th>Data</th>
<th>Data-decision clock</th>
<th>Phase-decision clock</th>
</tr>
</thead>
<tbody>
<tr>
<td>0101</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Edge

<table>
<thead>
<tr>
<th>Edge</th>
<th>Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>early</td>
</tr>
<tr>
<td>1</td>
<td>late</td>
</tr>
</tbody>
</table>

### D(n-1) B D(n)

<table>
<thead>
<tr>
<th>D(n-1)</th>
<th>B</th>
<th>D(n)</th>
<th>Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>early</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>late</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>early</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>late</td>
</tr>
</tbody>
</table>
Linear or Bang-bang?

- **Linear detector**
  - Well-defined Gain and loop behavior
  - May have systematic phase error

- **Bang-bang detector**
  - Can use the same decision circuit for both the data decision and phase detection
  - No systematic phase error
  - Output compatible with logic circuit

- We chose bang-bang detector for our link
Clock Recovery for Multi-bit, Multi-port Link

- **Clock forwarding**
- **Per-bit clock recovery**

![Diagram showing clock recovery mechanisms](image-url)
Clock Recovery for Multiple Clock Domains

- Can we place many PLLs on a single chip?
- VCOs run with different frequencies ($\Delta f \sim 100$ ppm)
- Concern: interaction through injection locking

![Diagram showing PLLs on a chip]
Phase Generation by interpolation

- Generate phase by weighted sum of reference clocks
- No oscillator, thus no injection locking
- Weights are controlled by DACs
- Works as a digital-to-phase converter

\[ (1-x)sn(t) + x cs(t) \]
Phase Interpolator in 0.11µm CMOS

- Ideal
- Calculated
- Output phase delay [ps]

- $100\text{ps} = \pi/4$
- Current DACs

- Deviation $<1\text{ps}$
Phase-control Feedback Loop

- Feedback path is composed of logic circuits in parallel-data clock domain

- Diagram showing:
  - Serial input
  - Front-end
  - DEMUX
  - Data channel
  - Boundary channel
  - PDC
  - Digital filter
  - Phase-to-Digital Converter
  - Parallel data
  - Reference clock
  - Parallel-data clock domain
  - Phase interpolator
Second-order Feedback Loop

- Two registers that store frequency and phase errors
- No tracking error for static frequency difference
- Equivalent to second-order PLL
Front-end Comparator

Amplifier or regenerative latch

- Regenerative latch is faster for a given technology

\[ A = g_m R_L, \quad \tau \equiv R_L C_L, \quad \tau_0 = C_L / g_m \]

\[ \tau \approx n R_L C_L = n A \tau_0 \]

Regenerative latch

\[ \Delta V = \Delta V_0 \exp \left[ (A - 1) t / \tau \right] \]

\[ \tau_{latch} = \frac{\tau}{A - 1} \approx \frac{\tau}{A} = \tau_0 \]
Strong-Arm-Type Decision F/F Circuit

Schematic

Waveform

Evaluation: Precharge

Time [nsec]
Example of CML-based Comparator

- Aperture time compatible with 40Gb/s signal
- Operate with 10-GHz reduced-swing clock

Output eye diagram

40Gb/s PRBS $2^{23}-1$ (BER < $10^{-12}$)
10Gb/s Rx Front end and DEMUX

Integrating Sampler

Decision Circuit

16-to-32 DEMUX

10 Gb/s

16 bit, 622 MHz

DATA(BDRY) 32 bit, 312 MHz

16-to-32 DEMUX

DCODE (BCODE)

PI

Dynamic NORs

Φ0 Φ1 Φ2 Φ3

Decision 0

Φ0 Φ1 Φ2 Φ3

Decision 1

Φ0 Φ1 Φ2 Φ3

Decision 2

Φ0 Φ1 Φ2 Φ3

Decision 3

Φ0 Φ1 Φ2 Φ3

622 MHz 312 MHz
10Gb/s MUX

32-to-4 MUX
- 32 bit, 312 MHz
- upper 16
- lower 16
- clock 312 MHz
- 4-phase 2.5 GHz \( \phi_0-\phi_3 \)
- selectors

4-way 2.5 GHz
- selectors

4-to-1 MUX
- \( \phi_0 \), \( \phi_1 \), \( \phi_2 \), \( \phi_3 \)
- INV-NOR
- 1-UI-delayed data
- 4-to-1 MUX

Output stage
- Main
- Out 10 Gb/s
- Pre

32-to-4 MUX
- Control
10Gb/s Transmitter Output Waveforms

- High-speed monitor signal pins
- Control pins
- Socket
- Evaluation board
- 10Gb/s I/O pins
- XFI(ASIC) eye template
- 10Gb/s Transmitter Output Waveforms
Inter-symbol Interference (ISI) Compensation

Need to compensate for ISI due to signaling media

- Tx: Pre-emphasis
- Rx: Equalization

No opening in eye patterns
Second-order equalizer

“1” Flat gain over the signal BW

“S” Differentiator

Transfer function: 
\[ H(s) = as^2 + bs + c \]
Transmitter Equalizer (FIR filter)

1-bit shifters

Input

Output

MUX
Effects of Equalization

- 6.4Gb/s Transceiver with Tx FIR filter and Rx equalizer

![Diagram of signal transmission through a 30m cable with and without Tx equalization, and the output of the Rx equalizer.](image-url)
Example of Macro Layout

- XAUI Tx + Rx
- 3.125Gbps x 4ch
- PLL
- 4-ch. Rx array
- 4-ch. Tx array
- 2.5Gb/s 16-ch Rx
- Parallel out
- Parallel In
- 8-ch. Rx array

Dimension:
- 1360µm
- 2280µm
- 3688µm
ASICS using High-Speed I/O Macros

- **0.18um**
  - 2.5Gbps x 72bit
  - 780Mbps x 44bit
  - Area: 14.962mm-sq
  - Core logic: 7.5Mgates
  - Package: FCBGA1225

- **0.18um**
  - 2.5Gbps x 128bit
  - Area: 16.283mm-sq
  - Core logic: 9Mgates
  - Package: FCBGA1089

- **0.11um**
  - 3.125Gbps x 96bit
  - Area: 16mm-sq
  - XAUI(4x3.125Gbps) x12ports
  - Package: FCBGA728

System-level evaluation done
## Power Consumption (in 0.11µm CMOS)

### 6.4 Gb/s transceiver (12-ch Tx, 12-ch Rx)  
(with 5-tap Tx FIR filter + Rx equalizer)

<table>
<thead>
<tr>
<th>Component</th>
<th>Power Consumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tx channel</td>
<td>150 mW / channel</td>
</tr>
<tr>
<td>Rx channel</td>
<td>90 mW / channel</td>
</tr>
<tr>
<td>PLL</td>
<td>90 mW</td>
</tr>
<tr>
<td><strong>Power per transceiver</strong></td>
<td>255 mW/ch</td>
</tr>
</tbody>
</table>

### 10Gb/s transceiver (4-ch Tx, 4-ch Rx)

<table>
<thead>
<tr>
<th>Component</th>
<th>Power Consumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tx channel</td>
<td>137 mW / channel</td>
</tr>
<tr>
<td>Rx channel</td>
<td>129 mW / channel</td>
</tr>
<tr>
<td>PLL</td>
<td>15 mW</td>
</tr>
<tr>
<td>5-GHz buffer / etc.</td>
<td>75 mW</td>
</tr>
<tr>
<td><strong>Power per transceiver</strong></td>
<td>311 mW/ch</td>
</tr>
</tbody>
</table>
Outline

- Background
  - Trends in IC bandwidth
  - Pin-bandwidth bottleneck

- Circuit solutions for bandwidth bottleneck
  - Clock recovery scheme
  - Receiver and transmitter front ends

- High-speed-I/O Future trends
CMOS High-speed I/O Speed Prediction

1. Clock period
   - Limited by clock tree bandwidth
   - Minimum period $T_{\text{min}} : 8 \times (\text{FO-4 delay})$ for FO4 inverter chain

2. Transmitter unit interval
   - Minimum unit interval: $1 \times (\text{FO-4 delay})$

3. Receiver unit interval
   - Minimum unit interleaved: $1 \times (\text{FO-4 delay})$

4. Overall performance
   - Should use multi-phase clock
   - Minimum unit time $1 \times (\text{FO-4 delay})$ feasible

## Speed Estimation for 0.11µm CMOS

![Signal Diagram]

<table>
<thead>
<tr>
<th>Item</th>
<th>Minimum 1UI value</th>
<th>@0.11µm</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CMOS Logic</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clock period</td>
<td>chip, 16 x FO4 delay</td>
<td>610 ps</td>
</tr>
<tr>
<td></td>
<td>local, 8 x FO4 delay</td>
<td>300 ps</td>
</tr>
<tr>
<td><strong>Tx bit time (UI)</strong></td>
<td>1 x FO4 delay</td>
<td>40 ps</td>
</tr>
<tr>
<td><strong>Rx bit time (UI)</strong></td>
<td>latch, 4 x FO4 delay</td>
<td>150 ps</td>
</tr>
<tr>
<td></td>
<td>total, 1 x FO4 delay@4-way</td>
<td>40 ps</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>1 x FO4 delay</td>
<td>40 ps</td>
</tr>
</tbody>
</table>

*Note: 1UI is the time it takes for a signal to propagate through one FO4 delay.*
End of Moore’s Law (for High BW CMOS)

Data rate per pin [Gbps]

Speed limit: 17%/year

High-BW chip

High-end μP

End of Moore’s Law (for High BW CMOS)
Optical or Electric?

Back plane (FR4 model)

- 10dB
- 30dB
- 50dB

<table>
<thead>
<tr>
<th>Length [m]</th>
<th>Speed [Gb/s]</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>0.1</td>
<td>100</td>
</tr>
<tr>
<td>0.1</td>
<td>10</td>
</tr>
</tbody>
</table>

References:
Conclusions

- Exponential growth in CMOS IC performance
  - High-speed I/O emergence was inevitable
- Limitation in exponential trend
  - Cannot sustain the exponential growth forever
    - Need new wiring topology for high-end extreme
    - Low power, ease of use, etc required for low end
- A big picture needed to decide what to do now
Backup Slides
CMOS, 2-PAM. Rx Equalizers


Injection locking

シューベルト Phenomena in oscillators

・Oscillator is synchronized with an injected clock signal if the injection frequency is close to the free running frequency of the oscillator

Analysis methods

・Adler equation (R. Adler, Proc IRE, vol34, p.351, 1946)

\[
\frac{d\theta}{dt} = \omega_0 + \frac{\omega_0}{2Q} \frac{A_{\text{inj}}}{A} \sin(\theta_{\text{inj}} - \theta)
\]


\[
\frac{d\theta}{dt} = \varepsilon \gamma(t + \theta(t)) n(t)
\]
First-order feedback loop

- Phase error is integrated into the phase register
- No phase error for static phase error (skew)
- Tracking error for dynamic phase error (i.e., frequency difference)
Open-loop Gain of Second-Order Loop

\[ H = \frac{G_1}{s^2} + \frac{G_2}{s} \]

- 40dB/decade $G_1/s^2$
- 20dB/decade $G_2/s$

$\omega_z^2 = 4\xi^2$

$\omega_1$ and $\omega_2$:
- $\omega_1 = \omega_2 = \frac{G_1}{G_2}$
- $\omega_2 = G_2$
**Sampling Function Calculation**

\[ S(\tau) = \frac{V_{\text{offset}}(\tau)}{\varepsilon} \]

\[ \varepsilon \cdot \delta(t - \tau) \]

\[ V_{\text{offset}}(\tau) \]

\[ D_{\text{out}}(\tau) \rightarrow 0 \]

\[ t = 0 \]

\[ \delta(t - T_0) \]

\[ \text{Ideal sampler} \]
Clock signal rises at $t=0$ and output observed at $t=T_0$

$$Out(t) = \int_{-\infty}^{T_0} h(t, \tau) In(\tau) d\tau$$

$$Out(T_0) = \int_{-\infty}^{T_0} h(T_0, \tau) In(\tau) d\tau$$

$$= \int_{-\infty}^{T_0} S(\tau) In(\tau) d\tau$$

$$S(\tau) \equiv h(T_0, \tau)$$

**Sampling function**
Sampling Function Example

- **PMOS switch**
- **Strong-ARM latch**
- **CML latch**

![Diagrams of PMOS switch, Strong-ARM latch, and CML latch](image)

Below is a graph showing the time in nsec and the voltage over time for each component.