A systems approach to building modern high-speed links

Vladimir Stojanović

Integrated Systems Group
Massachusetts Institute of Technology
Course Map

- **Show modern, systems approach to link design**

- **Day 1**
  - Link environment as communication channel
  - Links as bandlimited communication systems

- **Day 2**
  - Link implementations (signal processing, synchronization)
  - Lab 1

- **Day 3**
  - System modeling (noise sources, ISI)
  - Lab 2
Link System Implementations

- Signal processing – Tx, Rx
  - Equalization
  - Modulation
- Synchronization
  - CDR
  - CDR and Eq interaction
Modern Link Architecture

- Equalization, Multi-level signaling, Partial-response
- Sophisticated timing recovery
  - Multi-level, second-order
Voltage mode reduces supply current by a factor of 4
- Not as clear when pre-driver and regulator power added
Transmitter: Output Drivers

- On-chip clock speed limited to 6-8FO4
- Need to send more bits/clock – multiplex data
Transmitter: Time-Interleaved DACs

- DACs enabled by overlap of two clocks
  - Need precise clocks
  - Fast clocks limit interleaving
  - Capacitance DACs loads output
DAC Output Circuitry

- \( R \cdot C = 25 \Omega \cdot 4.3 \text{pF} \rightarrow 1.5 \text{GHz bandwidth} \)

- Predriver \( V_{ddReg} \) controls output current
Transmit Equalizers: Analog FIR

Simple PAM2/PAM4 Tx

Equalizing 5-Tap 2P/4P Tx
Transmit Equalizers: Analog FIR

Original 5-Tap PAM2/PAM4 equalizing transmitter

Shared equalizing transmitter

Total gate = 3(7/8 + 5/8) = 4.5W/L

[Zerbe 2003]
Linear CT receiver equalizer – analog high-pass

- Tunable RC tail degeneration
- Sensitive to common-mode

[Farjad-Rad 2003]
Rx Continuous Time Linear Equalizer

- Source-connected RC-pole
  - Capacitor becomes a short at high-frequencies, increasing gain
- NOTE: Max gain still limited by Gain*BW product of source-coupled differential pair
Variable Cap Rx CTLE Implementation

- Deal with process variation by tuning C & therefore tuning zero location
Rx CTLE Implementation → Getting More Gain

- Build peaking amplifier by use of inductors
  - Area intensive

[Partovi 2003]
Banwidth peaking, Tx and Rx

- Insert inductors between interleaved stages
  - Approximate lumped LC transmission line
  - Phase adjustment compensates delay
  - Nice-idea – hard to do

[Ellersick 2001]

Integrated Systems Group
Rx CTLE Implementation Issues

- Linearity a challenge
  - Especially when input swings vary greatly in amplitude
- Limited by gain-bw of diff-pair stage
  - Tuned to a data rate with limited range; channel match
  - Sensitive to PVT variations
  - Sensitive to device mismatch, linearity
  - Difficult to offset cancel, calibrate
- Multi-stage issues
  - High gain can lead to clipping in multi-stage design
  - Original design issues become even more difficult
  - Tuning is tricky
- Performance results
  - In 90nm, can generally get ~4-6dB of gain/stage at 10Gb/s if you do things right
  - Not a lot of gain – leads to multi-stage & inductively peaked designs
Receiver with Analog FIR DFE

Loop latency an issue
- Use for tap 2 and more, sometimes just for reflections

[Zerbe 2003]
DFE Feedback Timing:
Splitting Edge & Data Gives Some Relief

- Separate data & edge feedback & summers help
  - Still need to make 1UI data timing – still very tight
  - Allows separate coefficients & ‘edge-DFE’

Courtesy E. Chen, K. Yang UCLA

Integrated Systems Group 18
Complete Di-Bit (Half-rate) DFE

- Combine prDFE and fully split even and odd paths
- Improved tolerance to DCD errors
- Can potentially use separate even/odd coefficients
One phase of double-data rate Rx

- **DFE**
  - First tap unrolled
  - Second – current-mode (still a speed issue)
  - Third – current-mode DAC

[Leibowitz 2007]
Reduction the critical path: Merge sampler & mux

- **StrongArm sampler directly drives domino mux to reduce critical path on first tap**

Integrated Systems Group
Receiver front-end is becoming complicated

- Lots of samplers
  - Data even, odd – multi-level
  - Edge even, odd – multi-level
  - Adaptive sampler

- Lots of DACs
  - Multi-level thresholds
  - Offset calibration
  - Adaptation, on-chip scope
Hardware re-use: Dual-mode receiver

- PAM4
Hardware re-use: Dual-mode receiver

- PAM4

pre-amp with offset

Comparator

Integrated Systems Group
Hardware re-use: Dual-mode receiver

- PAM2
Hardware re-use: Dual-mode receiver

- PAM2 with loop-unrolled DFE tap
Hardware re-use: Dual-mode receiver

- PAM2 with loop-unrolled DFE tap
  - Leverage multi-level properties of signals in loop-unrolling
  - Re-use PAM4 receiver hardware (slicers and CDR)
Adaptation with minimum overhead

- Adaptive sampler
  - Generates the error signal at reference level
- Monitors the link
  - Adjustable voltage and time reference
  - On-chip sampling scope
- Can replace any other sampler - calibration

Dual-loop adaptive algorithm

- Data level reference loop
  \[ d\text{Lev}_{n+1} = d\text{Lev}_n - \text{step}_{\text{dataLev}} \text{sign}(e_n), \quad \hat{x}_n > 0 \]

- Equalizer loop
  \[ \hat{w}_{n+1} = \hat{w}_n + \text{step}_w \text{sign}(e_n) \text{sign}(\hat{x}_n) \]

- Scale the equalizer - output Tx constraint

[Diagram showing initial eye, mid-way equalized, and equalized signals with error and step functions]
Dual loop convergence – 4 tap example

- Hard to estimate analytically
- Experimental results show
  - Both loops are stable within wide range 0.1 – 10x of relative speeds
Partial response adaptation - start

- Extend data filter by one bit \((msb_n, msb_{n-1})\)

(a) Update loops only on \((msb_n, msb_{n-1}) = (1,1)\)
  - Finds \(d\text{Lev}(1,1) - "1+\alpha"\)

(b) Update loops only on \((msb_n, msb_{n-1}) = (0,1)\)
  - Finds \(d\text{Lev}(0,1) - "1-\alpha"\)
Partial response adaptation - end

\[ 2\alpha = d_{\text{Lev}}(1,1) - d_{\text{Lev}}(0,1) \]

Iterate \( \alpha \) finding and equalization loops

- \( \text{msb}_n, \text{msb}_{n-1} \) filter tolerates one tap post-cursor ISI \( \alpha \)
Continuous adaptation with unknown data

- Stop adaptation when data not random enough
  - Use a cross-correlator circuit

Integrated Systems Group
Spectral gating results

DFE adaptation without spectral gating

DFE adaptation with spectral gating

Adaptation Iteration

Integrated Systems Group
Noise detection

- On-chip scope can analyze system noise
- Multi-level samplers especially vulnerable
  - Differential properties degrade with threshold offset
Link System Implementations

- Signal processing – Tx, Rx
  - Equalization
  - Modulation
- Synchronization
  - CDR
  - CDR and Eq interaction
Clocking : Terminology

**Synchronous**
Every participant gets same frequency and phase.

**Mesochronous**
Every participant gets same frequency, but unknown phase. Requires a way to recover the phase from the data. Coding (e.g. 8b/10b) is often used to make sure there are sufficient data edges.

**Plesiochronous**
Every participant gets nearly the same frequency, slowly drifting phase. Requires a way to detect when the Rx clock has drifted 1/2 cycle from Tx clock.

**Asynchronous**
Dispense with clocks altogether, use (e.g.) request/acknowledge 4-phase handshake to ensure correct sequencing of events.

Can do with or without CDR

Needs CDR!
Clock and Data Recovery

- Recovering clock from the data
  - Can recover clock completely, or just phase
  - Just phase: need a reference clock

- Why?
  - Allows separate xtals on different boards
  - Don’t have to match trace lengths, delays
  - Easier system design / clock distribution

- Why Not
  - Expensive: takes area, power
  - Requires coding or transition density or at least a training sequence
    - 8b10b coding uses 10b to xfer 8b of info; 20% BW loss
Example CDR: PLL Technique

- Simple bang-bang PLL
  - Observe data with phase detector
  - Filter Early/Late & drive VCO

- Advantages
  - Good frequency range
  - Low Jitter

- Challenges
  - Phase offset
  - Lock time - startup sequence
  - Loss of lock - coding dependant
  - How to integrate multiple PLLs?
    - Harmonic locking problems
Dual PLL Problem: Harmonic Locking

- Potentially serious in highly integrated plesiochronous systems where residual phase error is close to noise injection in magnitude.
Dual-Loop CDR

- Combination of
  - Core PLL provides multiple phases at frequency
  - Periphery DLL mixes and makes desired phase

- Advantages
  - Avoids harmonic locking
    - Easy to integrate many
  - Rapid CDR lock time
  - CDR very stable
    - Digital = flexible filtering, control
    - Can even 'hold' phase state

- Challenges
  - Limited Freq offset from PLL
  - Jitter not as low as PLL
CDRs require transition density

- PLL based CDRs require to keep lock even in mesochronous
- DLL based require for plesiochronous tracking

- Often coding is used to guarantee
- Can alternately use a scrambler + XOR
CDR Issues: Jitter

- CDR Jitter starts out worse than PLL jitter
- Also can have ‘dither jitter’: phase wander when locked
CDR Dither Jitter

- Caused by need to track plesiochronous differences
- Dither jitter set by
  - Latency of the loop usually 10-20 cycles
  - Step size usually 1% Tsymbol
  - # of averages usually 16+

- The last two along with transition density set the tracking rate – (CDR loop is first order)
  - Conflicting requirements between tracking and jitter
2x Oversampling

- Generate **early/late** from $d_n, d_{n-1}, e_n$
  - Simple 1st order loop, cancels receiver setup time
- **Jitter on data Clk ≠ PLL output**
  - Base is linear PLL jitter
  - Can add non-linear phase selector noise from CDR
Use information from data-level sampler (already there for adaptation)
- Eliminate edge sampler, eclk
- Curvature to waveform here
  - Use a comparator to differentiate between dlev & signal
  - Decide later if 0 or 1 from comparator means early or late
- Even more transition-like level behavior in 4-PAM mid-data levels
Many transition types → PAM2 CDR unusable
4-PAM Eye With Density

- Offset transitions clearly visible
- But good transitions exist…
PAM4 Edges & CDR Approach #1

- Offset edge sampler to data level to get mid-levels
  - Best available edge-rate in FSE system
  - Requires edge samplers placed accurately on data level
PAM4 Edges & CDR Approach #2

- Use all minor transitions and one major transition
  - More transitions to choose from, no voltage offset required
  - Poorer edge-rate from minor transitions
Measured 4-PAM CDR Performance

- 2-PAM CDR on 4-PAM data
  - 60ps p-p @ 8Gb/s

- 4-PAM CDR uses only minor transitions
- Lower dither jitter
  - 35ps p-p @ 8Gb/s
Partial response CDR

- Four signal levels
Partial response CDR

- Four signal levels
- Offset edge samplers for transitions with ISI
  - Otherwise timing error
- Need to filter edges – similar to PAM4
Dual-mode CDR

- PAM4

---

PAM2 with loop-unrolled DFE tap
- Leverage multi-level properties of signals in loop-unrolling
- Re-use PAM4 receiver hardware (slicers and CDR)
Goal is to transfer 8 bits @ f1 on chip 1 to 8 bits @ f2 on chip 2
- First encode and transfer data based on local clock f1
- Then recover data and clock (f1) on chip 2
- Elastic buffer (FIFO) used to transfer data from f1 to f2
- Finally, decode to get 8 bits @ f2
Plesiochronous System Impact

- Packets must have appropriate slack time
  - How else to recover timing difference?
  - Idle characters must be recognized for slack
- Need FIFOs deep enough
  - Set by maximum frequency difference & maximum data length
- CDR track rate
  - Must be able to track maximum difference including dither
- Frequency difference, protocol, maximum data packet size different system components
Second-order loops – PI control

- Solves frequency tracking vs. jitter problem
  - I branch tracks the constant frequency difference
    - Very low tracking rate & bandwidth (thus low jitter)
    - Sub-1ppm tracking capability
  - P branch tracks normal phase shifts
    - Small step size & dither

[Lee04]
CDR : Coding & Transition Density

- CDR requires transition density to keep lock
- AC - coupling requires DC-balance
- Plesiochronous operation requires packets & null characters
- .....Coding as solution
- Typical code : 8b10b from IBM
  - 8 bits into the link => 10 bits on the wires
  - Raw data rate must be 25% faster than effective data rate
    - 6.25Gb raw for 5Gb effective
  - 8b10b code guarantees
    - DC balance
    - Transition density : 2 transitions every 10 bits
    - Reserved codes, control characters
8b10b Code Overview

- DC Balanced within every code word
64/66 Code Overview

Data Codewords have “01” sync preamble

\[
\begin{array}{c|c}
0 & 1 \\
\hline
& 64 \text{ bit data field (scrambled)}
\end{array}
\]

Mixed Data/Control frames are identified with a “10” sync preamble. Both the coded 56-bit payload and TYPE field are scrambled

\[
\begin{array}{c|c}
1 & 0 \\
\hline
8\text{-bit TYPE} & \text{combined 56 bit data/control field (scrambled)}
\end{array}
\]

00,11 preambles are considered code errors and cause the packet to be invalidated by forcing an error (E) symbol on the HARI output

- Much lower overhead
- Poorer DC balance & transition properties
Link System Implementations

- Signal processing – Tx, Rx
  - Equalization
  - Modulation
- Synchronization
  - CDR
  - CDR and Eq interaction
CDR & EQ – General Issues

- Fundamental issue – conditioning signal edges affects CDR edge-position…
- CDR edge-position effects observed ISI
  - Can affect both Tx & Rx coefficients
  - What is best solution for lowest BER???
TxEQ Pulse-Shaping Effects on CDR Lock Position

Effect depends on TxEQ taps
1 pre-cursor taps, 1 main, 2 post-cursor taps

Blue: Raw single bit response

Purple: Final result with adaptation & CDR locking

[Ren 2007]
3 pre-cursor taps, 1 main, 0 post-cursor taps

Significant phase-shift in lock position
Edge/DFE interaction with first post-cursor tap; will move lock point
- Proper prDFE implementation can avoid this entirely

Unlike TxEq, since taps 2-N are not convolved with channel they do not interact with CDR
Conclusions

- Backplane links limited by the channel
- ISI is large in baseband links
  - Can’t completely compensate
    - (At least not with reasonable area/power)
  - Residual ISI also increases CDR jitter
- Generally have low BER requirements
  - Accurate noise statistic important
  - Many of large noise source are bounded
- Power constrained transmitter
  - PAM4 and PAM2 with simple DFE are attractive solutions
- What to do above 10Gb/s?
Next Challenges

- More efficient spectral usage
  - Multi-level, multi-tone modulations
    - Need improved PSR of all circuits in the path
    - Control/calibrate offset and mismatch
- Coding
- Density issues
  - Improve energy-efficiency
- Control of complex architectures
  - Deal with crosstalk
- Lots of opportunity for design!
Bridging the gap: Multi-tone link

- Challenge – balancing the inter-symbol and inter-channel interference
  - Microwave filter techniques
  - Custom signal processing

Integrated Systems Group
Works well on channels with notches

- Frequency Response
  - Chip to Chip
  - Multi-Drop (Memory)
  - Backplane

- Return to bus architecture
  - Can arrange stubs to create controlled notches

Integrated Systems Group
Oversampled Tx filter controls the whole channel band while shaping sub-channels

[Amirkhany 08]
As technology scales
- Digital FIRs become more attractive
Multi-Tone Measured Eye Diagrams

<table>
<thead>
<tr>
<th>Baseband Modes</th>
<th>4-Channel AMT Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>2PAM</td>
<td>2PAM</td>
</tr>
<tr>
<td>Un-Equalized</td>
<td>Equalized</td>
</tr>
<tr>
<td>12Gb/s</td>
<td>12Gb/s</td>
</tr>
<tr>
<td>4PAM</td>
<td>Equalized</td>
</tr>
<tr>
<td>24Gb/s</td>
<td>18Gb/s</td>
</tr>
<tr>
<td>Ch1</td>
<td>Ch2</td>
</tr>
<tr>
<td>Ch3</td>
<td>Ch4</td>
</tr>
</tbody>
</table>
| Equalized – Post Processed |  }

Tx sampled with oscilloscope
With Rx mixing and integration in Matlab

Oscilloscope measurements

Integrated Systems Group
Some cited references

- **Equalization and clock recovery for a 2.5-10-Gb/s 2-PAM/4-PAM backplane transceiver cell**
  IEEE Journal of Solid-State Circuits
  Volume 38, Issue 12, Dec 2003 Page(s):2121 - 2130

- **A second-order semidigital clock recovery circuit based on injection locking**
  Hiok-Tiaq Ng; Farjad-Rad, R.; Lee, M.-J.E.; Dally, W.J.; Greer, T.; Poulton, J.; Edmondson, J.H.; Rathi, R.; Senthinathan, R.;
  IEEE Journal of Solid-State Circuits
  Volume 38, Issue 12, Dec 2003 Page(s):2101 - 2110

- **0.622-8.0 Gbps 150 mW serial IO macrocell with fully flexible preemphasis and equalization**
  Farjad-Rad, R.; Hiok-Tiaq Ng; Edward Lee, M.-J.; Senthinathan, R.; Dally, W.J.; Nguyen, A.; Rathi, R.; Poulton, J.; Edmondson, J.; Tran, J.; Yazdanmehr, H.;
  12-14 June 2003 Page(s):63 - 66

- **A serial-link transceiver based on 8-GSamples/s A/D and D/A converters in 0.25-μm CMOS**
  IEEE Journal of Solid-State Circuits
  Volume 36, Issue 11, Nov. 2001 Page(s):1684 - 1692

- **A 62.5 Gb/s multi-standard SerDes IC**
  21-24 Sept. 2003 Page(s):585 - 588

- **G. Sheets, J. D’Ambrosia “The Impact of Environmental Conditions on Channel Performance,”**
  *DesignCon 2004*
To probe further

- C-C. Yeh and J. R. Barry, “Adaptive Minimum Bit-Error Rate Equalization for Binary Signaling,” *IEEE Transactions on Communications*, vol. 48, no. 7, July 2000

If you have any questions send me mail
vlada@mit.edu