A Bit of Analysis on Self-Timed Single-Bit On-Chip Links

Jonathan Tse, Benjamin Hill, and Rajit Manohar
Computer Systems Laboratory, Cornell University

May 21, 2013
Single-Bit Interconnect

- Transmit data across die(s)
- How best to do that?
- Scope
  - Single-bit links
  - Asynchronous context
  - Delay-Insensitive Encodings
  - Handshaked links
Interconnect Design Challenges

- **Pressure on Wiring Resources**
  - Planar wiring (mostly) plentiful
  - Interconnect heavy-designs (FPGAs, etc)
  - Thru-Silicon Vias (TSVs) comparatively scarce
  - Delay-insensitive encodings expensive

- **Electrical Characteristics**
  - RC characteristics not scaling well
  - Lumped capacitance model invalid
  - Long wires — charge relaxation problem
Efficient Wire Usage

- Synchronous Most Wire-Count Efficient
  - Bundled data, etc. are close
  - Delay insensitive encodings worse
- Asynchronous Protocols Contextually Appropriate
  - 2-phase computation difficult
  - 4-phase dual-rail long distance signaling expensive
Choosing a Protocol

- What does “optimal” mean?
  - Area
  - Energy
  - Throughput
  - Latency
  - Ease of design
  - Robustness

- Approaching Optimality
  - Sizing
  - Circuit family
  - Buffer insertions
  - Metallization choices
Pareto Front

- Three Metrics
  - Throughput
  - Energy
  - Area
- Best Tradeoff
Pareto Front

- Three Metrics
  - Throughput
  - Energy
  - Area
- Best Tradeoff
Pareto Front

- Three Metrics
  - Throughput
  - Energy
  - Area
- Best Tradeoff

Energy vs Throughput

Throughput vs Energy plot with points and line indicating the Pareto front.

Background
- Context 6/21
## Single-Bit Links

<table>
<thead>
<tr>
<th>Link</th>
<th>Handshake</th>
<th>Timing</th>
<th>Voltage</th>
<th>Wires</th>
</tr>
</thead>
<tbody>
<tr>
<td>WCHB</td>
<td>4-Phase</td>
<td>QDI</td>
<td>Full-Swing</td>
<td>3</td>
</tr>
<tr>
<td>RQDI</td>
<td>2-Phase NRTN</td>
<td>RQDI</td>
<td>Full-Swing</td>
<td>3</td>
</tr>
<tr>
<td>STFB</td>
<td>2-Phase RTN</td>
<td>Single-Track</td>
<td>Full-Swing</td>
<td>2</td>
</tr>
<tr>
<td>ATLS</td>
<td>4-Phase</td>
<td>QDI</td>
<td>Ternary</td>
<td>2</td>
</tr>
<tr>
<td>STATS</td>
<td>2-Phase RTN</td>
<td>Single-Track</td>
<td>Ternary</td>
<td>1</td>
</tr>
</tbody>
</table>
## Single-Bit Links

<table>
<thead>
<tr>
<th>Link</th>
<th>Handshake</th>
<th>Timing</th>
<th>Voltage</th>
<th>Wires</th>
</tr>
</thead>
<tbody>
<tr>
<td>→ WCHB</td>
<td>4-Phase</td>
<td>QDI</td>
<td>Full-Swing</td>
<td>3</td>
</tr>
<tr>
<td>RQDI</td>
<td>2-Phase</td>
<td>RQDI</td>
<td>Full-Swing</td>
<td>3</td>
</tr>
<tr>
<td>STFB</td>
<td>2-Phase</td>
<td>Single-Track</td>
<td>Full-Swing</td>
<td>2</td>
</tr>
<tr>
<td>ATLS</td>
<td>4-Phase</td>
<td>QDI</td>
<td>Ternary</td>
<td>2</td>
</tr>
<tr>
<td>STATS</td>
<td>2-Phase</td>
<td>Single-Track</td>
<td>Ternary</td>
<td>1</td>
</tr>
</tbody>
</table>
## Single-Bit Links

<table>
<thead>
<tr>
<th>Link</th>
<th>Handshake</th>
<th>Timing</th>
<th>Voltage</th>
<th>Wires</th>
</tr>
</thead>
<tbody>
<tr>
<td>WCHB</td>
<td>4-Phase</td>
<td>QDI</td>
<td>Full-Swing</td>
<td>3</td>
</tr>
<tr>
<td>RQDI</td>
<td>2-Phase NRTN</td>
<td>RQDI</td>
<td>Full-Swing</td>
<td>3</td>
</tr>
<tr>
<td>STFB</td>
<td>2-Phase RTN</td>
<td>Single-Track</td>
<td>Full-Swing</td>
<td>2</td>
</tr>
<tr>
<td>ATLS</td>
<td>4-Phase</td>
<td>QDI</td>
<td>Ternary</td>
<td>2</td>
</tr>
<tr>
<td>STATS</td>
<td>2-Phase RTN</td>
<td>Single-Track</td>
<td>Ternary</td>
<td>1</td>
</tr>
</tbody>
</table>

### Data

<table>
<thead>
<tr>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Data Waveform" /></td>
</tr>
</tbody>
</table>

### Enable

<table>
<thead>
<tr>
<th>Enable</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Enable Waveform" /></td>
</tr>
</tbody>
</table>

### Parity

<table>
<thead>
<tr>
<th>Parity</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Parity Waveform" /></td>
</tr>
</tbody>
</table>

### Circuit Diagram

![Circuit Diagram](image)
## Single-Bit Links

<table>
<thead>
<tr>
<th>Link</th>
<th>Handshake</th>
<th>Timing</th>
<th>Voltage</th>
<th>Wires</th>
</tr>
</thead>
<tbody>
<tr>
<td>WCHB</td>
<td>4-Phase</td>
<td>QDI</td>
<td>Full-Swing</td>
<td>3</td>
</tr>
<tr>
<td>RQDI</td>
<td>2-Phase</td>
<td>NRTN</td>
<td>QDI</td>
<td>3</td>
</tr>
<tr>
<td>STFB</td>
<td>2-Phase</td>
<td>RTN</td>
<td>Single-Track</td>
<td>2</td>
</tr>
<tr>
<td>ATLS</td>
<td>4-Phase</td>
<td>QDI</td>
<td>Ternary</td>
<td>2</td>
</tr>
<tr>
<td>STATS</td>
<td>2-Phase</td>
<td>RTN</td>
<td>Single-Track</td>
<td>1</td>
</tr>
</tbody>
</table>

### Diagram

- **True**: 
  ![True Diagram]

- **False**: 
  ![False Diagram]
Single-Bit Links

<table>
<thead>
<tr>
<th>Link</th>
<th>Handshake</th>
<th>Timing</th>
<th>Voltage</th>
<th>Wires</th>
</tr>
</thead>
<tbody>
<tr>
<td>WCHB</td>
<td>4-Phase</td>
<td>QDI</td>
<td>Full-Swing</td>
<td>3</td>
</tr>
<tr>
<td>RQDI</td>
<td>2-Phase NRTN</td>
<td>RQDI</td>
<td>Full-Swing</td>
<td>3</td>
</tr>
<tr>
<td>STFB</td>
<td>2-Phase RTN</td>
<td>Single-Track</td>
<td>Full-Swing</td>
<td>2</td>
</tr>
<tr>
<td>ATLS</td>
<td>4-Phase</td>
<td>QDI</td>
<td>Ternary</td>
<td>2</td>
</tr>
<tr>
<td>STATS</td>
<td>2-Phase RTN</td>
<td>Single-Track</td>
<td>Ternary</td>
<td>1</td>
</tr>
</tbody>
</table>

Data

Enable

Enc  Dec

Single-Bit Links — Overview 7/21
# Single-Bit Links

<table>
<thead>
<tr>
<th>Link</th>
<th>Handshake</th>
<th>Timing</th>
<th>Voltage</th>
<th>Wires</th>
</tr>
</thead>
<tbody>
<tr>
<td>WCHB</td>
<td>4-Phase</td>
<td>QDI</td>
<td>Full-Swing</td>
<td>3</td>
</tr>
<tr>
<td>RQDI</td>
<td>2-Phase</td>
<td>RQDI</td>
<td>Full-Swing</td>
<td>3</td>
</tr>
<tr>
<td>STFB</td>
<td>2-Phase</td>
<td>Single-Track</td>
<td>Full-Swing</td>
<td>2</td>
</tr>
<tr>
<td>ATLS</td>
<td>4-Phase</td>
<td>QDI</td>
<td>Ternary</td>
<td>2</td>
</tr>
<tr>
<td>STATS</td>
<td>2-Phase</td>
<td>Single-Track</td>
<td>Ternary</td>
<td>1</td>
</tr>
</tbody>
</table>

![Wire Diagram](image)

![TX RX Diagram](image)
Single Track Asynchronous Ternary Signaling (STATS)
Single Track Asynchronous Ternary Signaling (STATS)

Wire

TX ▶ RX

$\frac{1}{2} V_{DD}$ Supply
Single Track Asynchronous Ternary Signaling (STATS)

- $\frac{1}{2} V_{DD}$ Supply
- Ternary Decode
Single Track Asynchronous Ternary Signaling (STATS)

- $\frac{1}{2} V_{DD}$ Supply
- Ternary Decode
- Sending Tokens
Single Track Asynchronous Ternary Signaling (STATS)

- $\frac{1}{2}V_{DD}$ Supply
- Ternary Decode
- Sending Tokens
- Return to Null

Diagram:

- a) Passgate
- b) Self-Invalidating Driver
- c) Shorted Inverter
Heuristic Optimization

- Global Optimum?
  - Sizing problem is convex
  - Other non-sizing factors to consider

- Heuristic Optimization Techniques
  - General-purpose
  - Non-convex problems
  - Handles local optima
  - Flexible
  - Easy implementation
Genetic Algorithms

- **Selection**
- **Configuration**
- **SPICE**
- **Evaluation**
- **Area Estimation**

- **Energy**
- **Throughput**
- **Correctness**
- **Area**
Evaluation

- Planar Wiring and TSV Cases
- 4-phase Dual-Rail Environment
- Configurations
  - Sizing
  - Circuit Topology
  - $V_{DD}$ Scaling (Non-Ternary)
- Metrics
  - Throughput
  - Energy
  - Area
Planar Evaluation

- Distributed RC Wiring Model
- Dual-Rail Source/Sink
- Insert Buffers
Planar Evaluation

- Distributed RC Wiring Model
- Dual-Rail Source/Sink
- Insert Buffers
Planar Results in 90nm

Energy vs Throughput in 90nm

Area vs Throughput in 90nm
Cross-Technology Planar Results

Energy vs Throughput

- RQDI
- STFB
- WCHB

Area vs Throughput

- STFB
- WCHB

Throughput [MHz]

Area [$\lambda^2$]

Energy [pJ]
Cross-Technology Planar Results

Energy vs Throughput

Area vs Throughput in 90nm

- RQDI
- STFB
- WCHB

Throughput [MHz]

Energy [pJ]

Area [$\lambda^2$]

90nm

65nm

0 1000 2000 3000 4000 5000 6000

0 1000 2000 3000 4000 5000 6000

0 0.2 0.4 0.6 0.8 1 1.2

$10^5$

0 0.2 0.4 0.6 0.8

90nm

65nm
Cross-Technology Planar Results

Energy vs Throughput

Area vs Throughput

Evaluation — Planar
Planar Takeaway Points

- Single-Track Timing Assumption
  - STFB offers benefits in Energy, Area
  - WCHB, RQDI more conservative
- Ternary buffers are expensive
  - Perform poorly in high-resistance environments
  - Ternary conversion cost high
TSV Evaluation

- Pair of Buffers
- No Intermediary Buffers
- TSVs
  - Doesn’t scale with technology
  - Less dense than planar
  - Wire-efficiency important
  - Scale throughput by TSV usage
TSV Results in 90nm

Energy vs Scaled Throughput in 90nm

Area vs Scaled Throughput in 90nm
TSV Results in 90nm

Energy vs Scaled Throughput in 90nm

Area vs Scaled Throughput in 90nm
Cross-Technology TSV Results

Energy vs Scaled Throughput

Area vs Scaled Throughput

Throughput per TSV [MHz]

Energy [pJ]

Area [µm²]
Cross-Technology TSV Results

**Energy vs Scaled Throughput**

- **RQDI**
- **STATS**
- **STFB**
- **WCHB**

**Area vs Scaled Throughput**

- **STATS**
- **STFB**
- **WCHB**

**Throughput per TSV [MHz]**

- 90nm
- 65nm

**Energy [pJ]**

- 0
- 0.5
- 1
- 1.5
- 2

**Area [µm²]**

- 0
- 10
- 20
- 30
- 40
- 50

**Evaluation — TSV 18/21**
Cross-Technology TSV Results

Energy vs Scaled Throughput

Area vs Scaled Throughput
TSV Takeaway Points

- TSVs are highly capacitive
  - STATS good fit
  - STFB unhappy
- STATS efficiently uses TSVs
- Interesting optimization opportunities
Conclusion

- Single-Track Timing
  - Aggressive designs offer clear benefits
  - Difficult to design
  - Not as robust

- Full-QDI
  - WCHB is most robust
  - Small penalty for robustness

- Heuristic Optimization
  - Quick design-space exploration
  - Augment/confirm designer intuition
  - Flexible, easy to implement
  - Pareto front tradeoff
A Bit of Analysis on Self-Timed Single-Bit On-Chip Links

Jonathan Tse, Benjamin Hill, and Rajit Manohar
Computer Systems Laboratory, Cornell University

May 21, 2013
## Link Failure Rates

<table>
<thead>
<tr>
<th>Link</th>
<th>% Planar Failure</th>
<th>% TSV Failure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>90 nm</td>
<td>65 nm</td>
</tr>
<tr>
<td>ATLS</td>
<td>23.94</td>
<td>16.34</td>
</tr>
<tr>
<td>RQDI</td>
<td>25.60</td>
<td>23.93</td>
</tr>
<tr>
<td>STATS</td>
<td>42.40</td>
<td>36.26</td>
</tr>
<tr>
<td>STFB</td>
<td>28.18</td>
<td>21.99</td>
</tr>
<tr>
<td>WCHB</td>
<td>10.67</td>
<td>8.49</td>
</tr>
</tbody>
</table>

Note: \(2856 \leq n \leq 11158\)
## Average Sparse Wiring Energy Percentage Improvements

<table>
<thead>
<tr>
<th>Link</th>
<th>90 nm</th>
<th>65 nm</th>
<th>45 nm</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATLS</td>
<td>47.36</td>
<td>16.93</td>
<td>-24.67</td>
</tr>
<tr>
<td>RQDI</td>
<td>33.71</td>
<td>7.22</td>
<td>13.98</td>
</tr>
<tr>
<td>STATS</td>
<td>27.42</td>
<td>-92.28</td>
<td>-112.87</td>
</tr>
<tr>
<td>STFB</td>
<td>39.04</td>
<td>18.11</td>
<td>12.26</td>
</tr>
<tr>
<td>WCHB</td>
<td>49.66</td>
<td>28.43</td>
<td>20.99</td>
</tr>
</tbody>
</table>
Single-Track Trace

STFB TSV Trace

STATS TSV Trace
Appendix — Analog Behavior

Noise Margin

Level-Shifter Voltage Transfer Curves

Wire [V]

A [V]

0.0 0.6 1.2
0.2 0.3 0.4 0.5

Wire [V]

B [V]

0.0 0.6 1.2
0.2 0.3 0.4 0.5

Wire [V]

C [V]

0.0 0.6 1.2
0.7 0.8 0.9 1.0

Wire [V]

D [V]

0.0 0.6 1.2
0.7 0.8 0.9 1.0
WCHB Level Shifters

Full Swing to Low Swing Pipelined Converter

Low Swing to Full Swing Pipelined Converter