Equalized On-chip Interconnect: Modeling, Analysis, and Design

Byungsub Kim

Integrated Systems Group
Massachusetts Institute of Technology

Advisor: Vladimir Stojanović
Interconnects dominate performance and power

IBM blue gene/L
Interconnects dominate performance and power

IBM blue gene/L

racks

Cables
Interconnects dominate performance and power

IBM blue gene/L

Cables

and more hidden
Interconnects dominate performance and power

IBM blue gene/L

Cables

racks

node cards
Interconnects dominate performance and power

IBM blue gene/L

racks

Backplane interconnect

Cables

node cards
Interconnects dominate performance and power.

IBM blue gene/L

- Racks
- Cables
- Node cards
- Compute cards
Interconnects dominate performance and power

IBM blue gene/L

racks

PCB wiring pattern

Cables

node cards

compute cards
Interconnects dominate performance and power

IBM blue gene/L

racks

Cables

node cards

compute cards

power 440 soc
Interconnects dominate performance and power

IBM blue gene/L

racks

on-chip wires

power 440 soc

Cables

node cards

compute cards
What are these interconnects?:
RC-dominant wires

Many core processor networks
(Tilera 64core)

- Short distance (1-2cm, 2-4cm, <10cm), high wire density, large channel loss
- Power and metal area budget

Packages: silicon carrier
(Patal DAC06)

PCB wires
Increasing demand for global interconnects

- Mesh – cheap local interconnect but low performance
- Higher radix networks – expensive global interconnect but high performance
Interconnects in NoC design hierarchy

NoC topologies

Kim ICCAD07

Terascale

Power

Offered BW

Offered BW

NoC trade-offs

Rep.

Eq.

Trade-off curve of interconnect

Trade-off curve of interconnect

Throughput Density

(Gbps / um)

Width

Space

Rep., Width

Rep., Space

NoC

Interconnect

Link parameter

Link trade-off

- wire, circuit parameter ↔ Interconnect metrics ↔ NoC trade-offs

[Diagram showing NoC topologies, Terascale, power, offered BW, and link parameter and trade-off with equations and graphs]
Thesis contributions

- Fast CAD tool for link design space exploration (Kim ICCAD07) – joint wire + circuit optimization
- Charge Injection FFE + Pre-distortion (Kim ISSCC09)
- Trans-impedance receive amplifier
- Infrastructure for \textit{in-situ} signal and energy characterization of the on-chip links (KimJSSC10)
Modeling and analysis

Power, latency, Offered bandwidth

Data rate density (Gb/s/um)  
Energy/bit (pJ/b)  
Latency (ps)

Interconnect topologies,  
Wire width & pitch,  
Transistor size

NoC topologies,  
VC buffer size,  
Channel width...

many core processor

Throughput Density (Gbps/um)  
Circuit+wire
Review: Equalization

No equalization

Feed forward equalization (FFE)

FFE + decision feedback equalization (DFE)
Equalization versus repeater

- Equalization
  - Faster data rate
  - Lower latency
  - Even lower power consumption by voltage swing reduction

![Diagram of equalization and repeater systems with voltage swing and channel attenuation graphs.](image-url)
Joint wire + circuit optimization

- Link = Driver (impedance) + Wire + Receiver (impedance)
- Optimize for power and performance

![Diagram of wire and circuit optimization](image-url)
Channel model – channel transfer function

Set telegrapher’s equation with Thevenin model
Channel model – comparison to SPICE simulation

- Closed form solutions for $T(f)$ and $T_c(f)$

$$T(\omega) \approx \frac{2e^{-\omega R_c C_o}}{(Z_c(\omega) + Z_L(\omega))\left(\frac{1}{R_s} + \frac{1}{Z_c(\omega)}\right)}$$

- Exponential form: small tap count FFE&DFE
- Impact of the impedance

- Good match with SPICE simulation
CAD flow

Technology information

Transistor: spice model
Wire: metal conductance, dielectric constant, etc.
CAD flow

R, C model for LCM & Inverter
Linearized RC switch extraction
Transistor: spice model
Wire: metal conductance, dielectric constant, etc.
Technology information

Circuit Model
Normalized R(Ohm-um), C(fF/um)
switch model database

Circuit type: LCM|Inverter, W_LCM, V_s, V_p

Circuit type: LCM | Inverter
W_LCM, V_s, V_p
CAD flow

\[ R = \begin{bmatrix} r_o & r_c \\ r_c & r_o \end{bmatrix}, \quad L = \begin{bmatrix} l_o & l_c \\ l_c & l_o \end{bmatrix}, \]

\[ G = \begin{bmatrix} g_o & g_c \\ g_c & g_o \end{bmatrix}, \quad C = \begin{bmatrix} c_o & c_c \\ c_c & c_o \end{bmatrix} \]
CAD flow

Channel model

Transfer function: $T(f)$, $T_c(f)$

R, C model for LCM & Inverter

Wth, Sp

target wire length: l

2D RLGC matrices database

Wth, Sp

2D field solver

Technology information

Transistor: spice model
Wire: metal conductance, dielectric constant, etc.

Linearized RC switch extraction

Circuit type: LCM | Inverter

Wire Model

Circuit type: LCM, Vs, Vp

Circuit Model

Normalized

$R$(Ohm-um), $C$(fF/um)

switch model database
CAD flow

Equalization coefficient: \( w, y_1 \)

Link architecture: FFE, DFE tap numbers

Transfer function:
- \( T(f) \)
- \( T_c(f) \)

Channel model

Link performance model

2D RLGC matrices database

Wire Model

R, C model for LCM & Inverter

Target wire length

2D field solver

Wire: metal conductance, dielectric constant, etc.

Technology information

Normalized
- \( R(\text{Ohm-um}) \)
- \( C(\text{fF/um}) \)

Circuit type:
- \( \text{LCM} | \text{Inverter} \), \( W_{\text{LCM}}, V_s, V_p \)

Linearized RC switch extraction

Circuit type:
- \( \text{LCM} | \text{Inverter} \), \( W_{\text{LCM}}, V_s, V_p \)

Data rate density, latency, eye opening, sampling delay \( T_d \)
CAD flow

- Energy-per-bit ($E_b$)
- Equalization coefficient: $w$, $y_1$
- Link power model
- Link performance model
- Transfer function: $T(f)$, $T_c(f)$
- Channel model
- R, C model for LCM & Inverter
- Target wire length: $l$
- Circuit type: LCM | Inverter, $W_{LCM}$, $V_s$, $V_p$
- Circuit type: LCM | Inverter, $W_{LCM}$, $V_s$, $V_p$
- Technology information
  - Transistor: spice model
  - Wire: metal conductance, dielectric constant, etc.
- Data rate density, latency, eye opening, sampling delay ($T_d$)
- Link architecture: FFE, DFE tap numbers
- Link strategy
- 2D RLGC matrices database
- 2D field solver
- Linearized RC switch extraction
- 2D RLGC parameters
- Circuit Model
  - Normalized $R$(Ohm-um), $C$(fF/um)
  - switch model database
- Wire Model
  - 2D RLGC matrices database
  - Wire Model
  - Normalized $R$(Ohm-um), $C$(fF/um)
  - switch model database
CAD flow

Connection to network architecture optimizer
energy-per-bit, data rate density, latency

Design selection based on metrics

Equalization coefficient: \( w, \gamma \)

Link power model

Energy-per-bit (\( E_b \))

target data rate density

RLGC parameters

Wire Model

2D RLGC matrices database

\( W_{th}, S_p \)

2D field solver

Linearized RC switch extraction

Technology information

Transistor: spice model
Wire: metal conductance, dielectric constant, etc.

Target eye constraint:
\( w_c \_\text{eye} \geq \text{eye}_{\text{target}} \)

Data rate density, latency, eye opening, sampling delay (\( T_d \))

\( R, C \) for LCM & Ci
\( N_C \)
\( R(\text{Ohm}-\text{um}) \)

R, C matrix database for LCM & Inverter

Circuit type: LCM | Inverter,
\( W_{LCM}, V_s, V_p \)

Circuit type: 

Data Rate Density (Gbps/um)

Energy/Bit (pJ/Bit)

Equalized, 30mV Eye
Equalized, 50mV Eye
Equalized, 90mV Eye
Repeated

\( 0 1 2 3 0.5 1 1.5 2 2.5 \)

\( 0 1 2 3 \)
Tool verification

- Tool finds optimal parameters
- Verified against auto-generated spice-netlist
Run time comparison: x30,000

<table>
<thead>
<tr>
<th>Design space</th>
<th>LMSE optimal Td</th>
<th>Brute Force Equation</th>
<th>SPICE</th>
</tr>
</thead>
<tbody>
<tr>
<td>423K points</td>
<td>0.74</td>
<td>180</td>
<td>21,460</td>
</tr>
<tr>
<td>Run time (h)</td>
<td>0.74</td>
<td>180</td>
<td>21,460</td>
</tr>
<tr>
<td>Normalized</td>
<td>x1</td>
<td>x244</td>
<td>x 29,00</td>
</tr>
</tbody>
</table>

- Design space trade-off curves are computed over 423,000 designs within 44.4 min
- Run time is improved by ~x30,000

Network design

- Energy, latency, data rate density
Link design space exploration: 90nm

- Energy vs. data rate density trade-off
- 2-10x energy improvement
Design space exploration: 32nm

- 2x less latency improvement with 10x less power
Provides proper metrics for the NoC simulators
Connection to network simulation

- Power versus offered bandwidth trade-offs of Clos NoCs topologies with repeated and equalized interconnects.
Circuit techniques for improved equalized interconnects

- 3-tap Nonlinear Charge-Injecting Feed Forward Equalizer (FFE) improves driver power efficiency
- Trans-impedance Amplifier (TIA) before 1-tap DFE improves the bandwidth-power-amplitude trade-off
Review: 2-tap FFE driver

\[ Y_{FFE} = w_0 d_0 - w_1 d_{-1} \]

- Data dependent summation/subtraction

\[ d_i \]
\[ d_0 \]
\[ -w_1 \]
\[ + \]
\[ Y_{FFE} \]

\[ D \]
\[ w_0 \]
Conventional: voltage dividing (VD) driver

- Current wasted in subtraction
Conventional: voltage dividing (VD) driver

- Current wasted in subtraction
Conventional: current mode logic (CML) driver

- Current wasted in subtraction
Conventional: current switching (CS) driver

Current wasted in subtraction
Proposed: Charge injection (CI) driver

- No current subtraction: no current waste
Power comparison

Average supply current (mA)

Typical off-chip  On-chip

R of VD

Icml
Ivd
Ics
Ici
The percentage of the eye reduction divided by the percentage of the coefficient perturbation.

\[
S_{x_i} = \frac{\frac{\Delta E_{ye}}{E_{ye}}}{\frac{\Delta x_i}{x_i}}
\]

\(\Delta x_n = 0, n \neq i\)
A large error due to subtraction of large currents

Coefficient errors do not attenuate
Eye sensitivity: CI FFE ($I_0$)

- A small error by small $I_0$ current
Eye sensitivity: CI FFE \( (I_1) \)

- Large error but attenuated by the channel
Resolution requirement comparison

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Resolution Requirement (bits)</th>
</tr>
</thead>
<tbody>
<tr>
<td>W0</td>
<td>7-8b</td>
</tr>
<tr>
<td>W1</td>
<td>8-9b</td>
</tr>
<tr>
<td>W2</td>
<td>6-7b</td>
</tr>
<tr>
<td>W0</td>
<td>7-8b</td>
</tr>
<tr>
<td>W1</td>
<td>8-9b</td>
</tr>
<tr>
<td>W2</td>
<td>6-7b</td>
</tr>
<tr>
<td>W0</td>
<td>7-8b</td>
</tr>
<tr>
<td>W1</td>
<td>8-9b</td>
</tr>
<tr>
<td>W2</td>
<td>6-7b</td>
</tr>
</tbody>
</table>

- **CI FFE**
  - >10x difference on eye sensitivity
  - 10x less accurate hardware requirement
- 3-tap Nonlinear Charge-Injecting Feed Forward Equalizer (FFE) improves driver power efficiency
- Trans-impedance Amplifier (TIA) before 1-tap DFE improves the bandwidth-power-amplitude trade-off
- Strong driver: Large current, $I_1$, $I_2$, $I_1 + I_2$
- Weak driver: Small current, $I_0$

3-tap CI FFE transmitter

<table>
<thead>
<tr>
<th>Current (mA)</th>
<th>Time (ns)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>0.5</td>
</tr>
<tr>
<td>2</td>
<td>0.6</td>
</tr>
<tr>
<td>2.5</td>
<td>0.7</td>
</tr>
<tr>
<td>3</td>
<td>0.8</td>
</tr>
<tr>
<td>3.5</td>
<td>0.9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Voltage (V)</th>
<th>Time (ns)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Middle</td>
<td>1.5</td>
</tr>
<tr>
<td>High</td>
<td>2</td>
</tr>
<tr>
<td>Low</td>
<td>2.5</td>
</tr>
</tbody>
</table>

Static pre-distorted FFE coefficients

Transition Signals:
$P_i^+, P_i^-, N_i^+, N_i^-$,
$P_i^+, P_i^-, N_i^+, N_i^-$, $P_i^+, P_i^-, N_i^+, N_i^-$
3-tap CI FFE transmitter

- Strong driver: Large current, $I_1, I_2, I_1 + I_2$
- Weak driver: Small current, $I_0$
3-tap CI FFE transmitter

- An array of binary weighted transistor
- A skewed enable NAND gate
3-tap CI FFE transmitter

- The decoding only require small overhead
Nonlinearity and predistortion

- Inver-like strong drivers are power-area efficient but non-linear
- Static pre-distortion compensates non-linearity
- Static pre-distortion is only possible in CI FFE
CI FFE operation

- 0001000

- 0.5 2 2.5 3 3.5

- 0.6

- 0.7

- 0.8

- 0.9

- 0.5

- 0.6

- 0.7

- 0.8

- 0.9

- 1.5 2 2.5 3 3.5

- 1.5

- 2.0

- 2.5

- 3.0

- 3.5

- 1.5

- 2.0

- 2.5

- 3.0

- 3.5

- (I_0 + I_1 + I_2)

- VREF

- Middle

- High

- Low

- Static pre-distorted FFE coefficients
CI FFE operation

0001000

Weak Driver

Strong Driver

Latch & Decoding Block

Static pre-distorted FFE coefficients

-\( I_0 \)
-\( I_0 \)
-\((I_0+I_1+I_2)\)

Middle

High

Low

1.5 2 2.5 3 3.5

Time (ns)

1.5 2 2.5 3 3.5

Current (mA)

0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0

Voltage (V)

Middle

Low

Mux

+ + + +

1 2 1 2

- - - -

1 2 1 2

P, P, N, N

P, P, N, N

P, P, N, N

P, P, N, N

8 Transition Signals:

\( P_1, P_2, N_1, N_2, P_1', P_2', N_1', N_2' \)

\( A_{<19:0>} \)
CI FFE operation

- 0001000

- Strong Driver
  - Static pre-distorted FFE coefficients

- Weak Driver

- Latch & Decoding Block

- CLK
  - 8 Transition Signals:
    - $P_0^*, P_1^*, N_0^*, N_1^*$,
    - $P_0^*, P_1^*, N_0^*, N_1^*$

- Current (mA) vs. Time (ns)
  - High
  - Middle
  - Low
  - $I_0$, $I_1$, $I_2$, $-(I_0 + I_1 + I_2)$
Trans-impedance amplifier receiver

- TIA achieves high bandwidth, small static power, large amplitude simultaneously
  - 50% static current, 150% voltage amplitude for the same bandwidth
Chip fabrication: 90nm CMOS

- Tx : 16x70um, Rx : 16x40um
- 10mm wire on M8 layer over M7 under M9
Chip fabrication: 90nm CMOS

- Tx: 16x70um, Rx: 16x40um
- 10mm wire on M8 layer over M7 under M9
in-situ measurement support
Channel measurement

- High-loss channel
  - 25dB @ 690MHz, 40dB @ 2GHz, 46dB @ 3GHz
  - 50% delay and 90% settling time of step response: 1.4ns and 5.5ns i.e. 8.6 and 33 UI at 6Gb/s
- Losses in most other literatures on Eq. are up to 30dB
Eye measurement

(a) DFE 3-tap FFE 6Gb/s (b) No DFE, 3-tap FFE 4Gb/s (c) No DFE 2-tap FFE 2Gb/s

- Vertical eye height is set about 100mV for trade-off curve extraction
- First on-chip measurement
- Demonstrated good eye quality (>50%, max 80%)
Sensitivity measurement

- **Sensitivity < 3 (10x better)**: good match with the analysis
- Low resolution and cheap (power & area) hardware (inverter-like) can equalize a high-loss channel (>40dB)

(a) 6Gb/s  (b) 4Gb/s  (c) 2Gb/s
As data rate increases

- DC bias (TIA) energy decreases
- Channel compensation energy (TxStr) increases

Optimal power efficiency: balancing between static (TIA) and active (channel, digital) energy
Comparison to relevant works

- 2x and 3x performance improvement for ~ additional 30% and 120% energy
- The only measurement >2Gb/s/uum-3Gb/s/uum
Conclusion of thesis

- **Modeling and CAD tool**
  - The first analytical link model for an RC-dominant link
  - Fast CAD tool improves x30,000 runtime
  - Equalized interconnects provides 2x-10x improved power efficiency than repeaters
  - First demonstration of Eq. CAD + NoC CAD simulation

- **Circuit – CI FFE Tx and TIA Rx**
  - CI FFE is introduced for the first time
  - CI FFE saves power by 2x than VD and CS FFE and 4x than CML
  - CI FFE allows simple static pre-distortion to utilize a cheap coarse inverter-like drivers. Demonstrated for the first time
  - CI FFE is 10x more robust to coefficient error than CS FFE
  - TIA-termination saves 50% static current and achieves 150% amplitude for the same bandwidth than a resistive termination
  - Area-energy efficient link operates over 40dB loss!
Publication & talk list

- **Conference**

- **Journal**

- **Workshop**
Acknowledgement

Adviser: Vladimir Stojanović
Committee: Anathan Chandrakasan, Jowel Dawson
Funding: Intel, IBM, NEC
Fabrication: TAPO, IBM Kansas City Plant
Industry: Ian Young, Alexandra Kern, Sam Palermo,
John Bulzacchelli, Dan Friedman, Yong Liu,
Timothy Dickson, Troy Beukema
Technical help: Fred Chen, Ranko Sredojević, Sanquan Song, Ajay Joshi, Jose Bohorquez, Matt Park, Min Park,
Sungwon Chung
Entertaining help –
ISG-coffee mates: Ranko Sredojević, Ajay Joshi, Michael Georgas,
Maxine Lee
Old CSG-CAG buddies: Alfred Ng (Coolman), Jaewook Lee, Vinson Lee (Real Crazyman), Jason Kim (Coolguy)
Beer-night gangsters: Min Park, Hyun ho Boo, Sunghyuk