PIANO™
Physical Interconnect Aware Network Optimizer

INTERCONNECT PHYSICAL OPTIMIZATION

K. CHARLES JANAC
President and CEO
Why is timing closure a problem with latest SoCs?

**Smaller Process Nodes**
- Transistors: Smaller & faster
- Wires: More & relatively slower

**Larger Chips Mean Longer Wires**
- Interconnect the only IP that traverses the chip
- More IP to be connected
- Contains longest wires of any IP

**Performance Requirements Increase**
- Frequencies increasing to 1.6Ghz+
- Latency optimization of critical paths
- Greater bandwidth requirements

**Voltage Drive Strengths Lower**

**NEED**

**AUTOMATED TIMING CLOSURE SOLUTION**
Wire Delays – Can’t Cross a Chip in 1 Clock Cycle

PHYSICAL DISTANCE DICTATES THE NUMBER OF PIPELINE STAGES

- Interconnect Frequency: 1.2GHz = 833ps
- Distance to travel = ~6mm
- Propagation delay = ~400ps/mm in 16nm FinFET; Needs 2400ps to span the distance
- Requires at least 3 pipeline stages and 4 clock cycles to meet timing

Large 14nm FinFET SoC may have >6,000 pipelines with 6K factorial pipeline combinations and 60 timing parameters – Too much for human comprehension!
The Interconnect IP is the Key to SoC Delivery!

THE SOC INTERCONNECT CONNECTS IPS TOGETHER

- Interconnect is only IP that traverses the chip
- Contains the longest wires in the SoC
  - Span long distances
  - Congestion points
- Carries most of the interesting data
- Changes in response to Architecture and Marketing ECOs
- Interconnect helps define the SoC architecture
  - Must support SoC performance requirements
- Often the last IP to be frozen, has to fit into available SoC channel space
  - Time to timing closure becomes schedule-critical!
PIANO™ – Interconnect Timing Closure Assistance

NoC Architecture

SoC Floorplan

Ncore / FlexNoC Logical IP

NoC topology

Fast Iterations

Pipeline configuration & placement

Ncore / FlexNoC Physical IP

Physical Synthesis, P&R

*Patents granted & pending
Three Loops of Interconnect Timing Closure

- **Architecture - Topology Development**
  - Inputs: data traffic, IP list, memory address map, performance goals, technology targets
  - ArterisIP capabilities: MiniFloorplan generation, NIU auto pipeline insertion, NoC placement
  - Verification: SystemC model, simulation
  - Output: Architectural timing convergence estimation

- **RTL Development**
  - Inputs: Production floorplan, functional block boundaries, IP socket locations
  - Arteris capabilities: Production floorplan input, floorplan editor, auto pipeline insertion, layout synthesis interface
  - Verification: Physical Synthesis
  - Output: Timing verified NoC instance

- **Layout – Place & Route, Domain of Customers using Cadence & Synopsys P&R**
  - Inputs: Optimized pipeline scheme reduces P&R iterations, improves interconnect PPA
  - Output: Timing closed layout database
Example PIANO flow
Interconnect Physical Optimization

- What are my latencies?
- What are my bandwidths?
- What are my frequencies?

- Start with a proposed SoC Architecture and Floorplan
- Organize input data, design objectives and constraints
- Generate the required connectivity and viable interconnect topology

- Implement the candidate topology through physical design
- Extract performance data
Logical Interconnect Topology Development

FLEXNOC & NCORE INTERCONNECT IPS DEFINE ARCHITECTURES

• ArChip16 Example: Large SoCs have multiple classes of interconnect
  – Non-coherent, Coherent, Control/Status, Observability, etc.
• Ncore & FlexNoC interconnects are managed separately from IP blocks, increasing design flexibility
Memory Access Traffic Classes

- **Cache Coherent (CC)** within Compute Cluster
- **Low Latency (LL)** to SRAM
- **High Bandwidth (HB)** to DRAM & Cache Fill
- **Best Effort (BE)** for Peripherals & DMA
- QoS for Video
- Multiple functional NoCs interacting
- Physically Constrained
Connectivity Map → Interconnect Connections → Layout

- Connectivity table defines interconnect connections within the floorplan
- Routes must pass through available channels in the floorplan
- Connectivity passes from initiator NIU to switch, to link, to RC buffers and finally to target NIU
Memory NoC: Interconnect Topology – Traffic Classes

Classify your IP connections per class of traffic:

<table>
<thead>
<tr>
<th>Class</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best Effort (BE)</td>
<td>Image system</td>
</tr>
<tr>
<td>Low Latency (LL)</td>
<td>SRAM</td>
</tr>
<tr>
<td>High Bandwidth (HB)</td>
<td>Main/Coherency</td>
</tr>
</tbody>
</table>
Memory NoC: Physical Constraints

• Memory NoC has significant location constraints across die
  - 1 to 4 DDR DRAM channels on Top/Bottom chip sides
  - Re-order Buffer (ROB) serves as gateway to Coherent & Main NoCs
• Some IP have direct memory access, not via ROB
Floorplan enables refinement of NoC topology

- Refine from a crossbar topology to a physically friendly topology
- Create switches that split and converge data flows based on layout
- Minimizes length of wires, latency, switching power, and leakage
Memory NoC:
Traffic classes are mapped onto logical interconnect topology
Mapping Interconnect to Floorplan

- DDR Channels (4) Schedulers and Reorder buffer are fixed in the layout
  - Regions reflect clock and power domains
  - Central Region is used to control congestion
- Data flow to Schedulers and ROB, via regions
Memory NoC: Floorplan Region Generation

- Topology intent passed to optimizer via regions → NoC element assignment to regions
  - Either explicitly or by placement during optimization
- Regions are created for Schedulers, ROB, main channels, some corners
- Regions also used for congestion control
Memory NoC: Physical NoC Topology Defines Latency

- Distances in the end define latencies, along with clock crossings, NoC services
- Memory NoC shows lower latency to SRAM, distance is significantly shorter (1/5 of DRAM3)
- Memory NoC shows long latency to DRAM Channel 3 since on other side of SoC
Summary & Customer Results
Interconnect Physical Optimization

- PIANO isolates individual interconnects so they can be managed quickly and easily
- PIANO places NoC logic into individual channel regions and inserts pipelines
- All PIANO timing estimates correlated to Physical Synthesis to accelerate layout process
PIANO™ Results Proven in a Production 7nm SoC

• Making floorplan friendly NoC took 2 days of on site support
  – capture data
  – decide paths
  – configure NoC
  – build Piano scripts
  – optimize for physical synthesis

• PIANO optimized thousands of pipeline stages automatically in several hours

• Physical Synthesis export used to Guide Synthesis tool to apply same placement as PIANO

• Timing Closure Estimate convergence used as starting point for placement and routing process

• Customer **saved 1.5 Months** of iterations

• This SoC taped out in February 2018
BECOME THE SOC DELIVERY MACHINE YOU WERE MEANT TO BE!
Thank you

CHARLIE.JANAC@ARTERIS.COM