300Mn gate Data Centre SoC challenges and PPA insights.

March 2020
Presenter: Maulik Patel & Milan Dalwadi
Agenda

- Introduction
- PnR Implementation Challenges
- Sign-off Challenges
- Full Chip STA challenges
- Conclusion
Introduction

- Increasing demand for internet
- Need for data centre ASIC SoC for ultra faster data processing
- Trillions of Logic gates in SoC.
- Implementation challenges for Timing closure, Congestion and Sign-off
- Improve Power Performance Area using advance features of EDA tool.
Data Centre ASIC Specifications - Overview

**ASIC Specifications**

**Dimensions**: 26mm x 25mm (Part of 2.5D Package) ~650mm²

**Technology**: 16nm FF TSMC, 15+1 Metal layers

**Power**: 450W, 6Track std cell

**Primary clock frequency**: 1.4 GHz and 1.6GHz

**Data throughput**: 2TBPS total channel

**Timing sign off for 50+ Corners**
Data Centre ASIC Challenges

- All the blocks are having very high density ~70%, to minimize chip dimension.
- Very high performance (frequency 1.6 GHz) to meet target throughput
- Very low data latency for high speed data processing
- Severe Congestion issues due to high density & Complex Logic
- 150+ hard macro in 90% blocks
- Physical verification challenges due to dimension
- Turn around time more than 2 weeks (PnR + Signoff) per block
- Aggressive Active & Leakage Power requirement

<table>
<thead>
<tr>
<th>Stage</th>
<th>Run Time (Hours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Floorplan</td>
<td>20</td>
</tr>
<tr>
<td>Placement</td>
<td>52</td>
</tr>
<tr>
<td>CTS</td>
<td>74</td>
</tr>
<tr>
<td>Route</td>
<td>80</td>
</tr>
<tr>
<td>Signoff</td>
<td>40</td>
</tr>
</tbody>
</table>
PNR Challenges - Timing & Congestion

- Place to Route huge timing degradation due to Layer assignment for HFN signals varies between Place & Route.
- Upper metal used for Global routing at placement for HFN nets.
- RC scaling to make net parasitic values more pessimistic during place and clock, which improved timing correlation between placement and route stage.
- In PnR used 10% extra RC pessimism to overcome it.

- One of the Critical block in terms of High ULVT count, Severe Congestion and Huge DRC count.
- High percentage of ULVT causes High leakage power as this blocks was instantiated multiple times at Full Chip.
- Complex logic introduced criss crossing of signals all over the core area.
- Analysis of Logic module connection using ICCOMPILER-II Data Flow Fly Lines utility.

| Stage       | Setup (WNS|TNS|FEP) | Hold (WNS|TNS|FEP) | DRC | ULVT% |
|-------------|----------------|----------------|------|-------|
| Route Before| -0.144|-27.5|3.9k | -0.070|0.5|476 | 4432 | 14.27 |
| Route After | -0.021|-0.748|56  | -0.003|0.05|100 | 130  | 3.89  |

Congestion : 1.27%
PNR Challenges - HFN Issue During Optimization

- Single clock gater to control the entire design at Pre-CTS
- CTS Clones clock gates for HFN nets, may cause DRV failure
- Used ICCOMPILE-II inbuilt utility to build buffer trees for HFN nets.
- Layer Promotion for HFN nets to minimize RC
Clocking Methodology

- Mesh Structure Distribution from PLL to Clock repeater.
  - Manual Tap route from PLL to Clock repeater.
  - Custom routing & buffer insertion in Clock repeater to extend the reach point to all accessible clock terminal throughout chip.
  - Uniform clock pitch distribution & custom routing from spine in horizontal & vertical direction to access chip sections. Equidistant clock terminals (pitch value defined based on simulation) drops to pushdown at block level.
  - Pushdown clock terminals are tapped to L1 header inside block, from L1 header CCOPT or Custom clock tree is built.
Signoff challenges - Signal EM & Dynamic IR

- Numerous EM violation. Due to lib driven max_cap value.
- Based on analysis, we chose 60ff max_cap limit for all the blocks.
- Improvement in SigEm and design QoR by restricting cap value in PnR.
- Slight jump in utilization as tool was added more numbers of buffers to fix cap.
- Significant Improvement from 2500 viols to 200 viols by applying max_cap 60ff limit.

- Dynamic IR drop is a drop in the voltage due to the high switching activity of transistors.
- It happens when there is an increasing demand for current from the power supply due to switching activities of the chip.
- Dynamic IR drop evaluates the IR drop caused when a large number of circuitry switches at the same time.
- Localize IR drop may introduce setup and hold viols due delay variation of transistor.
Signoff challenges - Dynamic IR drop

- Designing the chip at lower technology node, dynamic power consumption is very important when you have highly utilized blocks.
- Mesh clock structure is used to minimize latency & skew. To achieve this we have to use high drive strength buffers.
- Dynamic power loss happens when there is a high switching activity in localized area.
- High cell density in local region.
- Distance between two clock cells can be controlled by using advanced placement feature of ICC. So while building clock tree so that switching activity can be controlled in localize region.
- We have also provided keepout around the clock buffers and in post route we have inserted decap cells.

<table>
<thead>
<tr>
<th>Dynamic IR Clock Results</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamic IR clock value</td>
<td>75mV</td>
</tr>
<tr>
<td>Dynamic IR clock count</td>
<td>200</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dynamic IR Clock Results (with workaround)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamic IR clock value</td>
<td>51mV</td>
</tr>
<tr>
<td>Dynamic IR clock count</td>
<td>3</td>
</tr>
</tbody>
</table>
**Fullchip STA challenges**

- Data Center Soc typical contains Hundred Milions + Instance count, Hundreds of Blocks, Multiple Clusters.
- Lots of Interface to meet Critical Timing & Latency targets.
- Hundred of data/address Buses with width of 500-1000 bps
- STA at single scenario would require roughly 7000G, Multiple CPU, couple of days.
- Tedious job to Analyze timing reports
- Divide - run and Analyze to reduce runtime & TAT
  - Clusterwise Timing (Within cluster interface timing)
  - Clockwise Timing
  - Interface Wise Timing (cluster-cluster interface timing).
  - Blockwise Timing (Interface and internal timing for particular block).
1. Cluster wise timing:
   - Generate histogram, Summary, Native report for each cluster
   - Addressing violations based on most violating cluster.

2. Clockwise timing:
   - 2000+ Clock in the design.
   - Segregating the reports clockwise and address based on clock group.

3. Interface Timing:
   - Inter/Intra cluster timing violations for long paths through feedthrough.
   - Meeting Critical Latency targets, Adjusting the IO constraints and Routing the nets in higher layers
   - Similar way we had identified the major timing violations and closed the full chip timing.

4. Block Wise Timing:
   - Help to analyze block level timing and interface timing.
   - Reduced TAT & Less Iterations.
Fullchip STA challenges

Check Timing:

- Shows potential timing problems
- Finding no clock, unconstrained endpoints, ideal clock and loop violations.
- Check for Clock specs and Constrained Registers

Coverage analysis:

- Coverage report is most important sanity check while doing full chip timing analysis.
- % of total paths being timed.
- Reports endpoints remaining untested with specific reasons like no clock, constant_disabled, false_paths etc.
- Missing Constraints.
- Important to maintain functionality & Performance of the design.
Discrepancy B/w Block & Full Chip Level Timing

- Interface delay modeling uses 60-40% of clock period to allocate I/O delay to block.
- Next Budgeting based on all Routed Blocks to adjust I/O delays.
- Timing discrepancy b/w block and full chip due difference in timing window at the interface.
- Difference in arrival time at block boundary caused due to extra delta delay and failed timing at Block level.
- Developed scripts to generate accurate arrival times for all the blocks using Primetime commands to overcome this issue.
- 200-300 new timing violations per block due to discrepancy of delta delay
- Extra Crosstalk delay at full chip needs to model at block level
- Adjusted I/O delay at block level with the use of scripts
- Fixing new violations at Block Level with new SDC
Conclusion – TTM with right PPA for Networking

Why schedule and TTM?

- Data center market is growing at double digit growth and will continue growing till 2025
- At the same time, it has competitors developing similar products
- Silicon is heart of digital economy which is multi billion in scale and hence launching the product on time is critical.

Summary:

Lot of challenges for Timing, Power, Area in data centre ASIC needs to meet schedule

Achieved TTM

- Advance feature of Tools
- Faster convergence of block
- Focus on Achieving Target PPA
- Experienced team and subject matter experts
Thank You, Questions/Queries?

Thank You