EMULATION BASED FULL CHIP LEVEL LOW POWER VALIDATION AT PRE-SILICON STAGE

Kyoungmin Park  
km8228.park@samsung.com

Jaegeun Song  
jaegeun.song@samsung.com

Hyundon Kim  
hyundon.kim@samsung.com

Seonil Brian Choi  
seonilb.choi@samsung.com

Suk Won Kim  
sw4u.kim@samsung.com

1Samsung Electronics Co., Ltd.  
1-1, Samsungjeonja-ro, Hwaseong-si, Gyeonggi-do, Korea

Abstract- IP and block level power estimation and low power feature verification has become a must for mobile AP development. Recently, it has become possible to analyze power consumption based on more realistic scenarios such as running real applications in the OS level so that power consumption is optimized in the full chip level during RTL design stage. This typically requires a full chip environment as opposed to an IP or block level environment to run realistic or benchmark scenarios. Thus if the full chip level environment enables more realistic scenarios at the RTL design stage, power validation for hardware based power optimization features can be achieved before silicon. Also, long scenarios of several milliseconds to minutes can be profiled to optimize power consumption at the full chip level. This paper presents a low power feature validation flow and the result at a pre-silicon stage.

Keywords-Emulation, Power Estimation, Power Validation, Power Verification

I. INTRODUCTION

IP and block level power feature verification using RTL and gate level simulation is a necessity for mobile AP development. However, focusing at the IP and block level doesn’t cover system level power feature verification. At the full chip level, the clock management unit and power management unit does control overall system power management. Power feature verification can be accurate after full-chip integration with those system IP[1]. However, RTL and gate level simulations are slow, so they can not verify various realistic scenarios based on firmware or OS level in full chip[2].

Thus if the full chip level verification environment enables more realistic scenarios at the RTL design stage, power validation for hardware based power optimization features can be achieved before silicon arrives. Furthermore, long scenarios of several milliseconds to minutes can be profiled to optimize power consumption in a full chip level.

This paper describes a validation flow for low power features using realistic scenarios with a full chip level RTL. The power optimization and its validation using the low power features are further discussed. The results
show a good correlation between the power estimation at RTL and the actual measurement at the post-silicon stage using the low power features.

The rest of this paper is organized as follows: This section is in the Introduction. Section II describes traditional power validation flow at the post-silicon stage and the power validation method at the pre-silicon stage which this paper proposes and shows the respective flow chart. In section III, two techniques to reduce Turn Around Time (TAT) in our flow are described. It includes a sampling ratio control scheme and using power critical signals. In section IV we provide the result of applying this flow to our latest mobile AP product and a correlation with silicon measurements. Lastly, section V offers our conclusions of this work.

II. POWER VALIDATION METHOD

In this section, the traditional power validation flow which is used post-silicon is described. Furthermore, the proposed pre-silicon power validation flow is presented with comparisons to the traditional power validation flow.

A. Traditional power validation flow at post-silicon stage

As figure 1 shows validating low power features and optimizing power consumption in a full chip or system level is performed during post-silicon. However, power analysis at the post-silicon stage can’t improve hardware power optimization and can only influence software based power optimization because the hardware design is already fixed. If bugs are found in hardware based low power features, it is difficult to fix them at the post-silicon stage. Fixes can not be applied to the current hardware version and they should be implemented in the next hardware version.

![Figure 1 Power feature validation flow at the post-silicon stage](image)

To mitigate bugs, software workarounds are used or low power features are disabled completely. Thus it is possible that products encounter power issues. Products may not meet important power specification such as mobile AP battery consumption which is defined during product planning.

Silicon power is measured as the current of a power rail or power sources on the board. Next, the power features are enabled and validated and the power savings are studied. Using this method, off-chip factor can affect the validation results, but it is very difficult to debug internal hardware nodes and the off-chip environment when there
is an issue in silicon. The full chip level RTL simulation environment is not secured, so that silicon scenario can not be reproduced easily. This makes that seamless debugging is not supported at post-silicon. Thus this lead to developing design and verification process at a pre-silicon stage during RTL design period.

B. Proposed power validation flow at pre-silicon stage

The proposed flow utilizes an emulation based full chip RTL environment to shift the low power feature validation and a power analysis from the post-silicon stage to the pre-silicon stage. Figure 2 shows the new validation flow based on the pre-silicon stage. This proposed environment is for validating the hardware power features by using a almost the same firmware/OS validation environment in pre-silicon as is used in post-silicon.

![Diagram](Figure 2 Power feature validation flow at the pre-silicon stage)

First, the power features can be verified with the profile switching activity from the emulation result. For example, when enabling low power features at any point during normal operation scenarios of the firmware, the low power feature, i.e., clock gating, can be verified with a switching activity graph. If debugging is required, the emulator can provide the waveform for all signals with full visibility. This enables effective debugging[3]. Furthermore, if the design needs to be modified after debugging, hardware can be modified right away. This is in contrast to hardware debugging during post-silicon. The hardware issue debugging during post-silicon is difficult because getting full visibility in hardware design is limited. Also it is not easy to modify the hardware design after silicon is released and thus pushes fixes to the next design version.

After low power feature verification, a power calculation tool can estimate the power savings by analyzing the FSDB of SAIF for interesting simulation time periods. With this result, the effect from each low power feature can be estimated separately. This enables a software development team to optimize the power at the system level using
already validated low power features. This is in contrast to the traditional flow where system level power can be optimized only after low power features are validated on real silicon. This can reduce the post-silicon software development time and impact the overall TAT of a product by shortening the development cycle.

III. METHODS TO REDUCE TAT

The overall flow for executing a long scenario based on the full chip was developed. First, a power profiling methodology for executing long multi-second scenarios is deployed. When a realistic scenario is profiled with a full waveform dump, it takes several days for running emulation, dumping the waveform and profiling switching activities. To reduce the TAT for these steps with the guarantee of good accuracy, a ‘sampling ratio control’ flow is evaluated and proposed.

Second, a power estimation methodology is developed to handle long scenarios. When the power value is calculated for a long realistic scenario with full signal dump, it also takes several days for dumping all signals and calculating the power value.

To reduce TAT for power estimation, it is proposed that 1) power regions of interest are selected from power profiling, 2) specific signals referred to as power critical signal are used from the power calculation tool as opposed to the full signal dump.

A. Sampling ratio control

Signals are dumped(sampled) for all cycles by default on the emulator. If the sampling ratio is controlled, the same signals can be dumped for only 10 cycles out of 1000 cycles. When switching activity is profiled using the sampling ratio control, significant reduction on TAT was achieved without losing much accuracy.

Figure 3 shows how much correlation can be achieved with different sampling ratios. The first graph depicts switching activity for the full cycle dump by default, the second graph shows a one of 32 for full cycle, and the third graph displays switching activity with only one of 256 for full cycle. The whole activity trend shows high correlation for all three graphs. Thus sampling ratio can be controlled with consideration of accuracy by use case or scenario[6].

![Figure 3 Sampling ratio control - corelation](image)

Table 1 shows the reduction on disk spaces and run time in the full chip level during the period of 100ms. The evaluated design size is around 50 million gates. The dumped file size is reduced up to 150X and emulation run time is faster by up to 20X. Profiling is post processing work and 15 cores of the computing farm are used to generate switching activity information. Profiled file size is reduced by up to 241X and post processing run time is faster by up to 176X.
When the realistic scenarios like an OS boot-up or android application runs on an emulator, this technique will be very useful to profile power with a short TAT.

<table>
<thead>
<tr>
<th>Sampling ratio (Full chip, 100ms)</th>
<th>Dumping</th>
<th>Profiling</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>File size</td>
<td>Runtime</td>
</tr>
<tr>
<td>Full cycle</td>
<td>144G</td>
<td>4h 47min</td>
</tr>
<tr>
<td>1 of 2</td>
<td>77G</td>
<td>2h 1min</td>
</tr>
<tr>
<td>1 of 4</td>
<td>42G</td>
<td>1h 8min</td>
</tr>
<tr>
<td>1 of 16</td>
<td>12G</td>
<td>27min</td>
</tr>
<tr>
<td>1 of 64</td>
<td>3.4G</td>
<td>18min</td>
</tr>
<tr>
<td>1 of 256</td>
<td>966M</td>
<td>15min</td>
</tr>
</tbody>
</table>

Table 1 Sampling ratio control - reduction resources

B. Power critical signal flow

The power estimation tool can select the signals which are most critical to calculating power. The number of these power critical signals is normally around 10 percent of the full signals. So, when only power critical signals are used to make FSDB or SAIF and used to calculate power on the power estimation tool, TAT will be decrease significantly[4][7]. Figure 4 shows how power critical signals are generated and used in emulation.

The flow is almost the same with using full signals, but two steps are added: power critical signals generation and fsdb generation only for power critical signals. The power estimation tool can generate power critical signals at the elaboration step with no additional effort. When power regions of interest are selected from the power profiling step, then an FSDB or SAIF file can be generated only for the regions of interest with the power critical signals from power estimation tool. Then, the power estimation tool uses this FSDB or SAIF file to calculate the power value.

Figure 4 Power critical signal flow
A power critical signal flow was applied to calculate an average power and results were achieved as describe in the Table 2. The number of signals was reduced by 92.5% and this reduction directly impacts the FSDB generation time and power calculation time with proper average power difference. A power critical signal flow does not effect the emulation run and dumping time or the switching activity profiling time. Thus the result of those steps are the same as using full signal flow.

The evaluated design size is around 50 million gates, the emulation time is around 100ms and the region of interest is 10ms. This region is generated to the FSDB and is used for calculating the average power.

<table>
<thead>
<tr>
<th>Full chip SOC</th>
<th>Emulation running and profiling</th>
<th>Power Calc.</th>
<th>Power diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># signal</td>
<td>Run+dump (100ms)</td>
<td>Profiling (100ms)</td>
</tr>
<tr>
<td>Full signal</td>
<td>44M</td>
<td>5.3h</td>
<td>24h</td>
</tr>
<tr>
<td>Critic. signal</td>
<td>3.3M</td>
<td>5.3h</td>
<td>24h</td>
</tr>
</tbody>
</table>

Table 2 Power critical signal flow-result

IV. EXPERIMENTAL RESULT

Using the proposed flow, our latest mobile AP was validated at the full chip level with special care toward low power features such as various clock gating schemes at the RTL design stage before silicon. While validating low power features, several hardware bugs were corrected and firmware versions were fully developed to enable those low power features properly.

Table 3 shows the idle state power impact of system level low power features. The number is normalized to explain the effect of the power features. When ‘Low power feature 1’ is applied, the idle power consumption is reduced by 54.2%. When ‘Low power feature 1’ and ‘Low power feature 2’ are applied together, the idle power consumption is reduced by 77.8%.

<table>
<thead>
<tr>
<th>Power Rail</th>
<th>w/o low power feature</th>
<th>Low power feature 1</th>
<th>Low power feature 2</th>
<th>Low power feature 1+2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory I/F</td>
<td>20</td>
<td>15</td>
<td>6.5</td>
<td>5.7</td>
</tr>
<tr>
<td>Data Bus</td>
<td>100</td>
<td>40</td>
<td>49</td>
<td>21</td>
</tr>
<tr>
<td>Total Power</td>
<td>120</td>
<td>55</td>
<td>55.5</td>
<td>26.7</td>
</tr>
<tr>
<td>Reduction</td>
<td>-</td>
<td>54.2%</td>
<td>53.8%</td>
<td>77.8%</td>
</tr>
</tbody>
</table>

Table 3 Result of power validation – Idle power

Table 4 shows the active power consumption when the same low power features are applied. When ‘Low power feature 1’ and ‘Low power feature 2’ are applied together, the active power consumption is reduced by 27.2%. The region of active power consumption is the most heavily operated, so the power reduction level is less then the reduction of idle power consumption.

<table>
<thead>
<tr>
<th>Power Rail</th>
<th>w/o low power feature</th>
<th>Low power feature 1</th>
<th>Low power feature 2</th>
<th>Low power feature 1+2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory I/F</td>
<td>60.8</td>
<td>43.6</td>
<td>60.8</td>
<td>41.8</td>
</tr>
<tr>
<td>Data Bus</td>
<td>100</td>
<td>74.6</td>
<td>103.3</td>
<td>74.6</td>
</tr>
<tr>
<td>Total Power</td>
<td>160.8</td>
<td>118.2</td>
<td>164.1</td>
<td>116.4</td>
</tr>
<tr>
<td>Reduction</td>
<td>-</td>
<td>26.5%</td>
<td>-2%</td>
<td>27.2%</td>
</tr>
</tbody>
</table>

Table 4 Result of power validation – Active power

The average TAT without using sampling ratio and the power critical signal flow for running emulation with firmware with various low power features is about 15 hours. This includes power calculation only focused on the
memory interface and data bus operation based on the full chip level. The FSDB which is used to calculate power periods is a 10ms subset of the full scenario runtime. To select the interest region from the whole scenario, first, the sampling ratio control scheme was used to profile the whole scenario period of around 5 sec. Then only a 10ms period was selected to calculate average power. So the TAT was almost the same with the first result, but the total scenario period was expanded by a factor of 500.

As Table 4 shows, with real silicon when low power features are applied at the OS level with the same scenario, the estimated power values from the RTL design stage are tightly correlated with the result of OS level measurement. But on some other samples, the correlation is varies. The OS level power measurements is affected by various on-chip and off-chip environment effects like the board and package[5].

However, the proposed flow only focuses on the chip environment without any off-chip environment. Thus not an absolute power value suitable for silicon correlation but suitable for evaluating power saving effects of various power features.

<table>
<thead>
<tr>
<th>IDLE Power</th>
<th>Active Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimat. result</td>
<td>Silicon result</td>
</tr>
<tr>
<td>118.6%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 5 Silicon correlation

V. CONCLUSION AND FUTURE WORK

A full chip level low power feature validation and estimation flow at the pre-silicon stage is proposed. Using the proposed flow we can ‘shift-left’ or perform the full chip power optimization and low power feature validation from the post-silicon stage to the pre-silicon stage. It significantly improves discovering hardware design issues at an early stage and is expected to reduce SOC project schedules using the emulation based power analysis solution. This proposed solution is applied to our mobile AP development based on these experimental results. Furthermore, post-silicon power scenarios are used during pre-silicon to expand power coverage to the whole chip.

This proposed solution is just the beginning stage to cover full chip or system level power analysis at pre-silicon using an emulation environment. Many factors need to be considered to enhance power analysis at pre-silicon stage, for example, verifying between the power source(i.e. PMIC) and SOC is critical. If this PMIC and SOC verification environment can be produced, power issues related to peak power between the PMIC and SOC can be investigated and solved. However, integration of analog components should be supported on the emulation environment. This is an important area for the industry to collaborate on[1].

Dynamic Voltage and Frequency Scaling(DVFS) support is required to operate the firmware or OS level scenario in an identical fashion to silicon. The software development team analyzes and optimizes power while controlling DVFS. Currently, frequency scaling is handled partially in the emulation environment but voltage control and effects to the power analysis are challenging topic in this area.

TAT reduction enhancements for long scenarios in the power analysis flow has been proposed in this work. Software developers typically request a fast environment to analyze power during pre-silicon. New solutions are being developed to meet the software development team’s requirements for speed. For example, calculation logic for switching activity can be inserted into the emulation hardware so switching activity can be generated in real time during an emulation run. In terms of TAT, significant enhancement of power profiling and estimation flow will be achieved using this solution.

The hybrid emulation platform is another promising solution to improve our proposed environment. It depends on the configuration, but hybrid emulation speed is faster than the pure emulation environment. Thus, if hybrid emulation can be used to profile and analyze power, the TAT might be shortened significantly.
ACKNOWLEDGMENT

We would like to thank several others in the SOC power, emulation and firmware teams, and SW development teams that were involved in this effort.

REFERENCES