proteanTecs

White Paper

Application-Specific Power Performance Optimizer
Based on Chip Telemetry

Table of contents

1 | Abstract

As datacenter power consumption continues to pose cooling and cost challenges, and battery driven devices are expected to last longer between charges, the search for advanced power management mechanisms continues.

A modern design must balance between maximizing performance, consuming the least amount of power, and guaranteeing no failures in field. The latter requires safety margins which translate into guard-bands that could otherwise be used to boost performance and/or reduce power. Either the design is "squeezed" to the max and therefore it risks failure since there were insufficient guard-bands to overcome in-field events or power/performance is "left on the table" to provide those guard-bands "just in case".

If we consider the power consumed by the digital logic (not including PHYs, PLLs and different analog elements connected to other power supplies, besides the core VDD), then CMOS power consumption is a combination of dynamic and static power consumption, according to:

$$P = P_{dynamic} + P_{leak} \sim C * V_{dd2} * f + V_{dd} * I_{leak}$$

A variety of dynamic power management methods focus on lowering the clock rate and/or the operating voltage, while still meeting the system performance requirements.

Adaptive Voltage Scaling (AVS) uses in-chip structures to automatically adjust the supply voltage to adapt for effects such as process variation, aging, and temperature.

However, some effects cannot be effectively tracked in a conventional manner. Varying workloads may cause IR drops in different areas of a chip and in different magnitudes. Guard bands are taken based on the characterization of worst-case workload scenarios (what they are assumed to be).

More advanced solutions include critical path emulators that closely mirror the behavior of a chip. Since they are not subject to the same workloads as the real logic circuits they aim to emulate, a safety guard-band must still be included.

Due to these limitations of monitoring structures, the best known AVS methods are limited in how much they can reduce the voltage and hence the power consumption.

proteanTecs' AVS Pro™ monitors the margin to timing failure of millions of real paths, in real time, under real workloads, to reduce voltage to the lowest point that still allows error-free functionality. It also provides an inherent safety-net, enabling fast frequency and voltage scaling to maintain error free functionality when events like voltage drops occur. These features provide the capability to be "workload aware" and maximize the power reduction, while avoiding failure. In this paper we will demonstrate how this solution can save between 9%-14% dynamic power when compared to conventional methods.

2 | Power and Performance Guard Bands

The power, performance, reliability equilibrium

A semiconductor's performance depends on its clock frequency of operation. For the digital logic to be able to operate at a given frequency, it must be supplied with a minimum voltage. This critical voltage should be sufficiently high so that all circuit paths can be sampled correctly in a single clock cycle. Under ideal circumstances, this could be derived from static circuit-level timing.

Real life variations from the ideal models may cause this voltage to be insufficient. Fabrication process variation, environmental conditions like temperature, unexpected operational workloads and on-chip noise sources may require the supply of higher than critical voltage. This will ensure the reliable operation of all circuits even in the worst-case operating scenarios, but at the same time compromise either power (higher dynamic voltage) or performance (lower frequency). This relationship is illustrated in **Figure 1.**
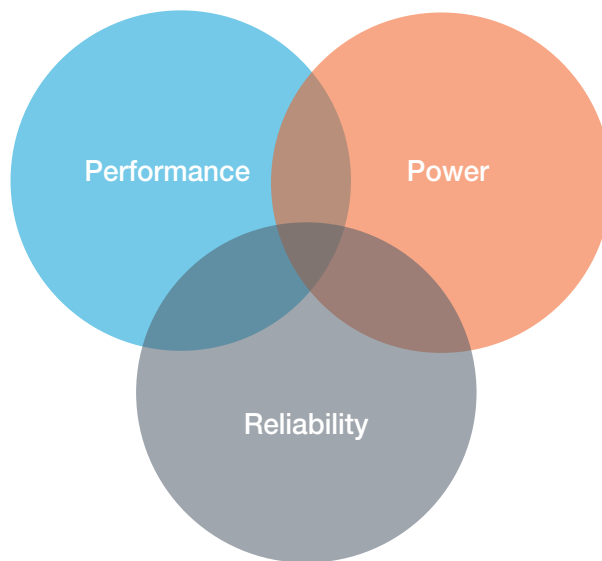

Figure 1: Relationship between Performance, Power, and Reliability

Built-in Guard-band: Process Variation

Fabricated chips have process dependent variation, which is reflected in this minimum voltage required to support a target frequency. Optimally, this voltage would be found by testing the chip at iteratively reduced voltages, at the ATE. The last voltage at which the functional test passes (aka "Last Pass") is the Vddmin for this device.

In practice, not all companies have the time, funds, inclination, or expertise to search for Vddmin per chip. The procedure itself can be time consuming, and operational costs high. Therefore, such companies may focus their characterization efforts on finding the Vddmin for the slowest chips, under the most challenging conditions, thus guaranteeing the full population will be functional. Other companies may compromise and search for the Vddmin per "bins" of chips (but not of individual chips); e.g., one Vddmin for "low power" (slower) chips, another for "high performance" (faster) chips, etc.

Each approach is a trade-off between characterization time, complexity, and power savings.

The result of such compromises is that some chips will have a supplied voltage that is an "overkill" (higher than the minimum required). Figure 2 shows that Vddmin based on "slow" material is an overkill for faster chips, creating a positive timing slack. For these chips, Vddmin could have been lower (conserving power) or conversely the clock frequency could have been increased (higher performance).

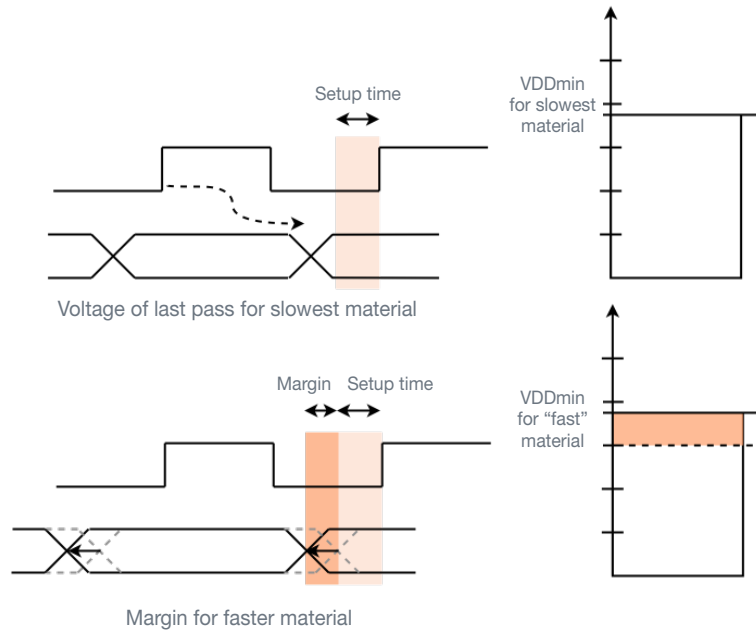Figure 2: Vddmin for "slow" chips means extra timing margin for "faster" chips

Note: proteanTecs provides an application for high accuracy Vddmin prediction per chip during the High-Volume Manufacturing testing stage, which takes process variation into account, providing an accurate Vddmin per chip while minimizing the overhead described above. This application is based on a trained model created in the proteanTecs SW platform, and edge SW with this trained model running at the ATE, to infer the Vddmin per chip in real time, while the chip is under test. This is out of scope for this paper.

Note: when Vddmin static setting is fused per chip at the ATE, the Vddmin takes into account the process variation and thus no or partial process guard-band is left (built-in) during lifetime operation, since it was already all or partially taken into account.

Lifetime Operation: Guard Bands Noise, Environment and Degradation

Guard-bands are also needed to compensate for field issues that were either unmodelled, untested in production testing environments, or unexpected but probable in Lifetime Operation (or some combination of these).

> Guard-band for aging effects: Over time, chips develop (for example) bias temperature instability and hot carrier injection, which cause age-related performance degradation and increase the voltage required to operate correctly.

> Guard-band for environmental conditions ensures correct operation at the worst-case temperature, e.g., to accommodate the slower circuit operations at high temperatures.

> Guard-band for operating characteristics: the actual functional workloads in the field may exceed the one(s) used in production testing environment (ATPG, functional). In this case, it must accommodate:

  • The extra path activity that is not exercised in production testing environment

  • The voltage droops due to the extra workload

  • The temperature difference caused by the extra workload

> Guard-band for DC voltage IR drop: to compensate for inaccuracy of voltage regulator/sense loop.

All these factors may lead to increased path delay. The applied (fused) operational voltage of a robust design will include enough extra voltage beyond the Vddmin, to overcome these potential factors that may arise during lifetime operation.

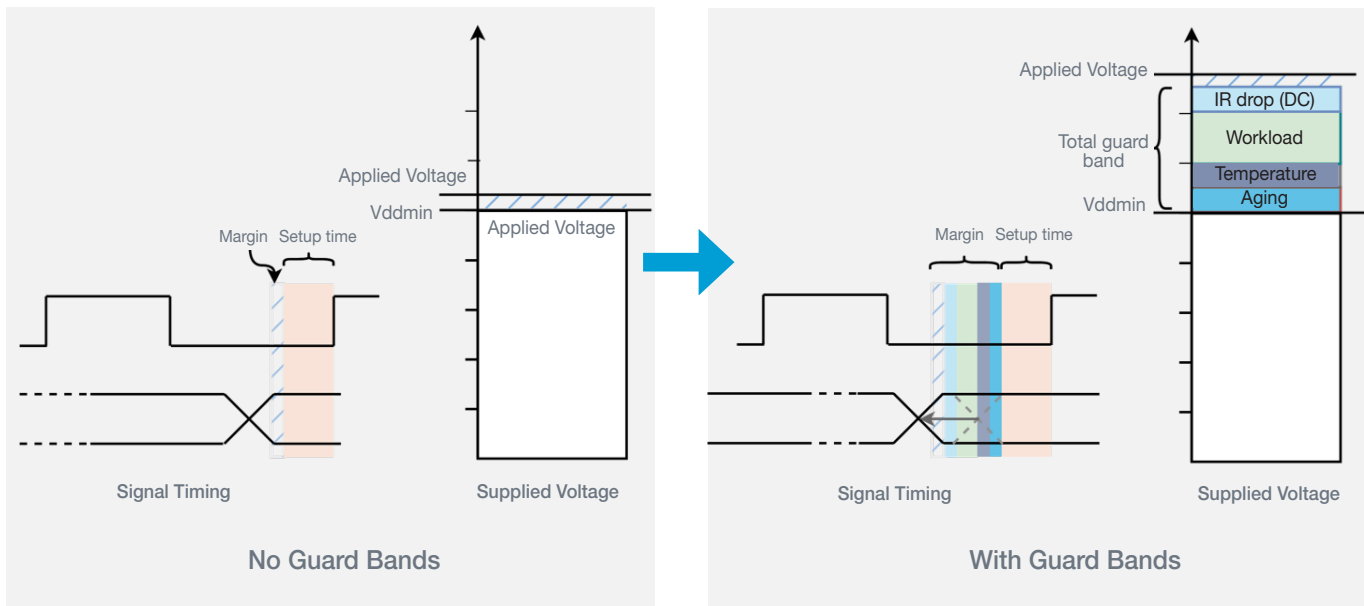Figure 3: Voltage added to compensate for various path delay degradation factors

When are the Guard-bands Needed?

Figure 3 illustrates the voltage overhead needed for robust in-mission performance. Guard-bands are added on top of the Vddmin to account for the different possibilities during the lifetime of the device. Since these can potentially happen simultaneously, their sum needs to be considered.

Note: If the Vddmin was not set per individual chip, taking into account its specific process variation, then the set Vddmin will exceed the minimum requirement of part of the population. By how much depends on how generalized the Vddmin search was. In the specific case where one Vddmin is provisioned to all the chips, this means that it is high enough for "slow" material and therefore an overkill for most of the population that will be faster material. This is indicated in the illustration as the process built-in "guard band" that can also be reclaimed by the AVS Pro.

Note: proteanTecs provides a Vddmin prediction application during High-Volume Manufacturing testing.

Since the consequence of timing failure can be catastrophic, the total guard bands described in the previous two chapters is provisioned. But in fact, in most real-life cases, all these factors will not come to play simultaneously.

Figure 4 shows use cases that illustrate the actual vs required voltage.

Note: For illustration purposes, these examples assume that the Vddmin is set exactly according to the particular process variation and thus voltage cannot be reduced on account of process variation.

Case A represents a worst case, where most of the provisioned guard band is required. For example, this may be a device that has aged enough to use up most of the aging guard band, the DC voltage guard-band and temperature guard-bands are required and in addition, it is running at the worst functional workload creating local voltage droops and exercising performance limiter paths. Still, since the applied guard-band is the sum of all the effects in the worst case, a small guard-band is left.

However, case B represnets a more realistic case. Here, much less guard-band is required, for example if the functional workload running at this time is less stringent, temperature guard-band is not required, not all the DC voltage guard-band is required, and aging has no effect yet.
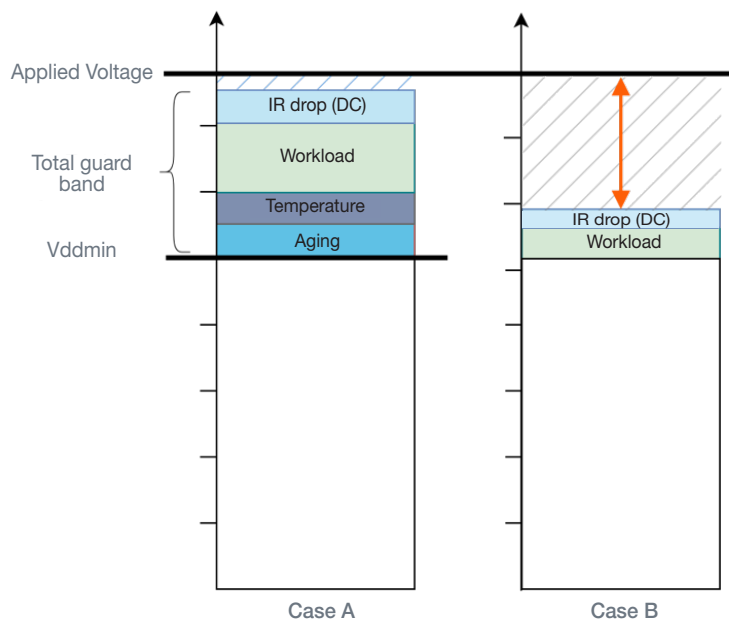

Figure 4: Actual vs required voltage examples

3 | Dynamic Power Management Methods

Dynamic Voltage and Frequency Scaling (DVFS)

DVFS is a software-controlled method in which clock rates and voltages are adjusted according to the performance required by the application or due to the need to reduce the power consumed by the device, to reduce the temperature from a dangerous level. It is based on the fact that different levels of computational power are required at different times, together with the fact that a dangerous temperature threshold is crossed. Software selects between pre-defined operating frequency and voltage pairs. This way, at any given time, and per application, the device runs at the minimum frequency to meet the system's processing requirements, and at the minimum voltage to allow error free operation at this frequency. When needed both the frequency and voltage are reduced to cool the device.

What DVFS does not address is the power optimization at a certain {F, V} pair. While each pair is optimized for the expected average software workload, the same guard-bands are needed to accommodate the performance-degrading environmental, operational, aging and noise effects described previously.

Adaptive Voltage Scaling (AVS)

AVS dynamically adjusts voltage levels according to the current operating conditions, thereby minimizing the voltage guard-band to the minimum required, per the existing conditions. These conditions include the operating frequency, as well as all the performance degrading factors mentioned previously. It is implemented as a closed-loop system comprised of on-chip structures, an AVS controller and a voltage regulator and clock throttler to act. It can be similarly utilized in designs with a single voltage operating point or in conjunction with DVFS, where the latter determines the F, V pairs to operate at per SW application or when the device temperature is higher than a given threshold, and the former optimizes the voltage per each pair, minimizing the voltage for each operational frequency.

The accuracy of AVS implementation depends on the structures used to calculate the minimum required voltage. Some common examples of these are:

❯ **Process, Temperature, Voltage** sensors allow AVS to customize the supply voltage per inherent silicon differences such as process variation, to "flatten" the power consumption curve, as well as dynamic differences such as temperature and voltage.

❯ **Critical Path Emulation** circuits closely model the behavior of the chip and adapt to environmental and process variations. These can be based on Ring Oscillator (ROSCs) that operate at the same voltage as that of the rest of the design. The ROSC's measured frequency indicates the voltage-frequency relationship for the chip under the conditions in which it is operating. Alternatively, the path emulator may be based on "delay chains" of standard cells, where the time taken for a process to traverse the chain simulates the performance of the chip.

4 | proteanTecs AVS Pro™

Gaps in Current Solutions

The different AVS methods described in the chapter above, "Adaptive Voltage Scaling (AVS)", attempt to "trim" the guard-bands on the chips when they are not needed. However, such methods fall short of their objective and thus cannot eliminate the need for provisioning voltage guard-band.

For starters, the measurements provided by on-chip sensors are local to their position and connectivity in a chip. Since the degree of freedom in their positioning and connectivity is limited, so is their accuracy in reflecting the different effects on the design. Moreover, different circuits in a design may be affected differently by these changes. For instance, aging will be different on a ring oscillator or critical path emulator, than on the real paths, since toggle rate has a big impact on aging.

But most importantly, these sensors do not monitor the more dynamic effects that can occur in a chip during its lifetime, and therefore cannot completely trim the overhead. For instance, while the applied voltage would take the worst/most stressful workload into account, in actual use, the workload may be significantly lower than that most of the time. In those cases, conventional methods will not be able to dynamically customize the voltage to the workload. Ring oscillators and critical path emulators do not experience the exact same effects as the circuits they supposedly emulate; it is not the same workload that exercises them. Without high coverage monitoring of the real paths and the real minimum voltage or maximum frequency limiters (performance limiters), deciding to reclaim guard-bands, let alone protect against causes for failure, may expose the device to risk.

The bottom line is that despite these methods, there will still be significant guard-bands provisioned to overcome their shortcomings.

Figure 5 shows the minimum required voltage over time. Some factors are more dominant than others. Some factors change slowly over time, and yet others can change at the rate of change of the software workload that runs on the device. The best-case optimization would be to track these changes precisely, to tailor the required voltage to fit the actual need at any point in time.
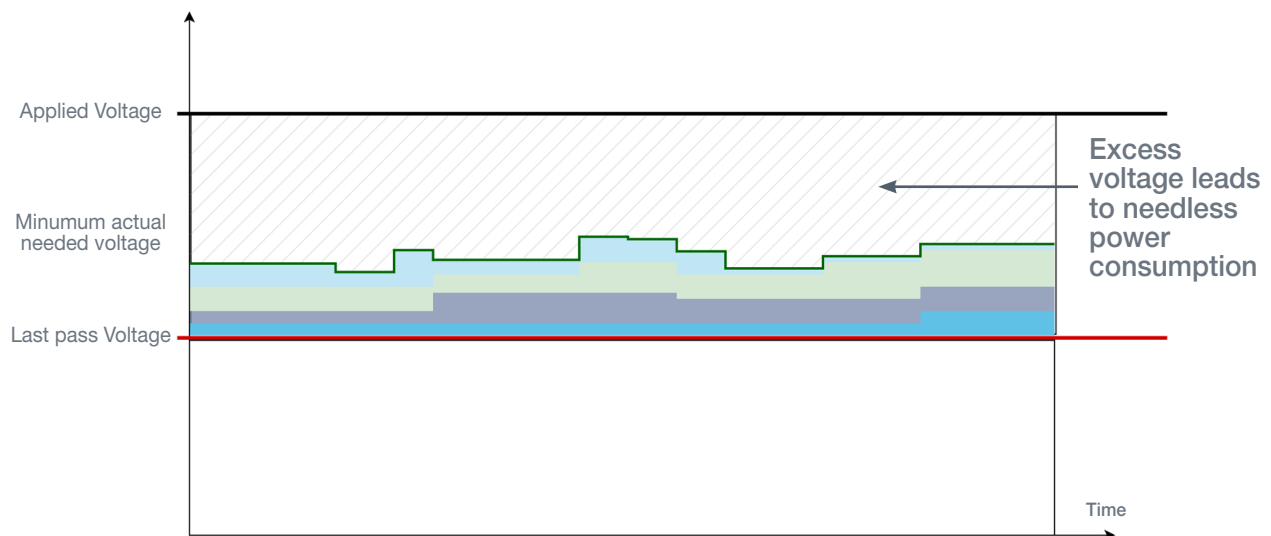

Figure 5: Minimum required voltage over time

New Approach: Monitoring Margin to Timing Failure

Why Monitor Margin to Timing Failure?

In a healthy, functional chip, signals propagate from source to target flip flops fast enough for the latter to properly sample them. VLSI designers will sign off a project to production only after ensuring this will be so even under the least accommodating conditions expected for the device, since a failure to do so means a failure of the chip.

The voltage guard-bands described in the chapter "Power and Performance Guard-Bands" above, are added in order to counter the degradation in path propagation time (i.e., the performance) due to the factors described, that would risk timing not being met.

Since the parameter that directly indicates digital failure is the available timing slack of a logic path, and since at any given time this slack is a result of all degrading factors (e.g., temperature, Voltage droops, DC voltage etc.), it would make sense that monitoring timing slack itself is superior to monitoring the characteristics that affect it. By measuring the phenomena directly, instead of indirectly, the exact measure of required voltage could be applied, regardless of the collection of factors that caused it. The alternative is based on characterizing the indirect measurement when a timing failure occurs, which could be an extensive and inconclusive endeavor. The direct measurement gives precise feedback to the voltage scaling.
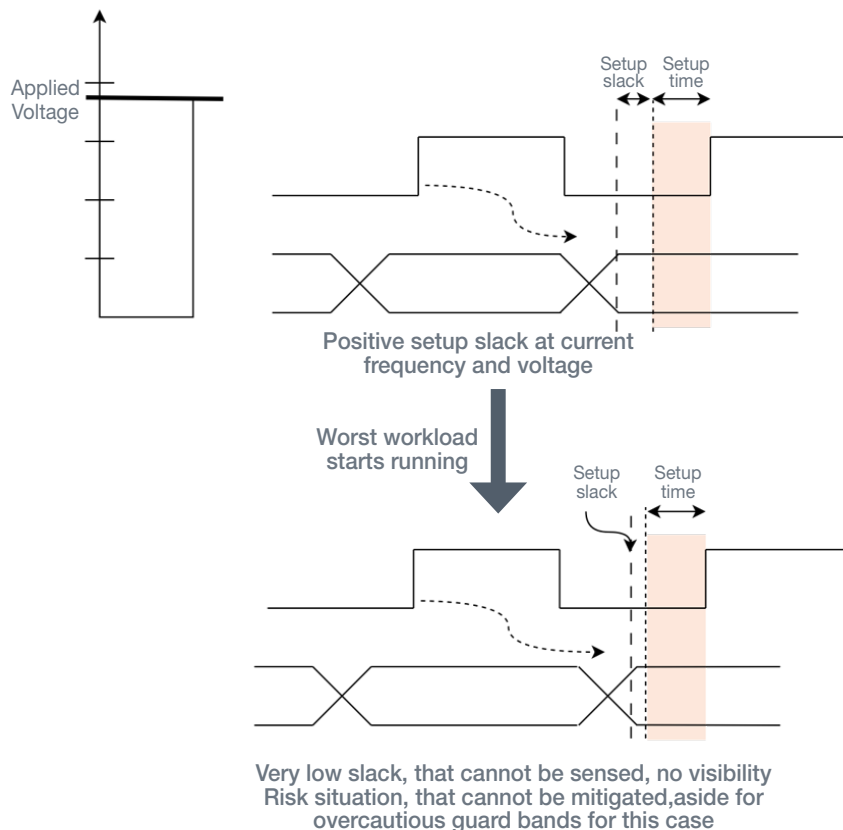
Figure 6: Signal Timing change due to increased workload, limited visibility

By providing a very high coverage of a design's real minimum Vdd limiting paths' timing margin, at a real time resolution, a device could adjust the voltage appropriately. When one or more of the effects mentioned above manifest themselves, the path delay will increase, thus reducing the timing margin-to-failure. If one or more of the paths approach failure (still there is guard-band left to avoid failure) as a result, the PMIC may increase the applied voltage to mitigate this.

Figure 7: Scaling voltage based on margin

Introducing proteanTecs Margin Agents

proteanTecs patented Margin Agents continuously monitor millions of logic paths to identify the lowest margin to timing failure. The timing margin is affected by which paths are being exercised and at what rate (referred here as "workload"), voltage droops, temperature hot spots, noise (e.g., cross talk), DC voltage, latent defects that manifest in operation, and aging.

Margin Agents connect to the inputs of flip flops that terminate logic cones or memory outputs (see figure 8). One Margin Agent can connect to the inputs of up to 32 such FFs. Since 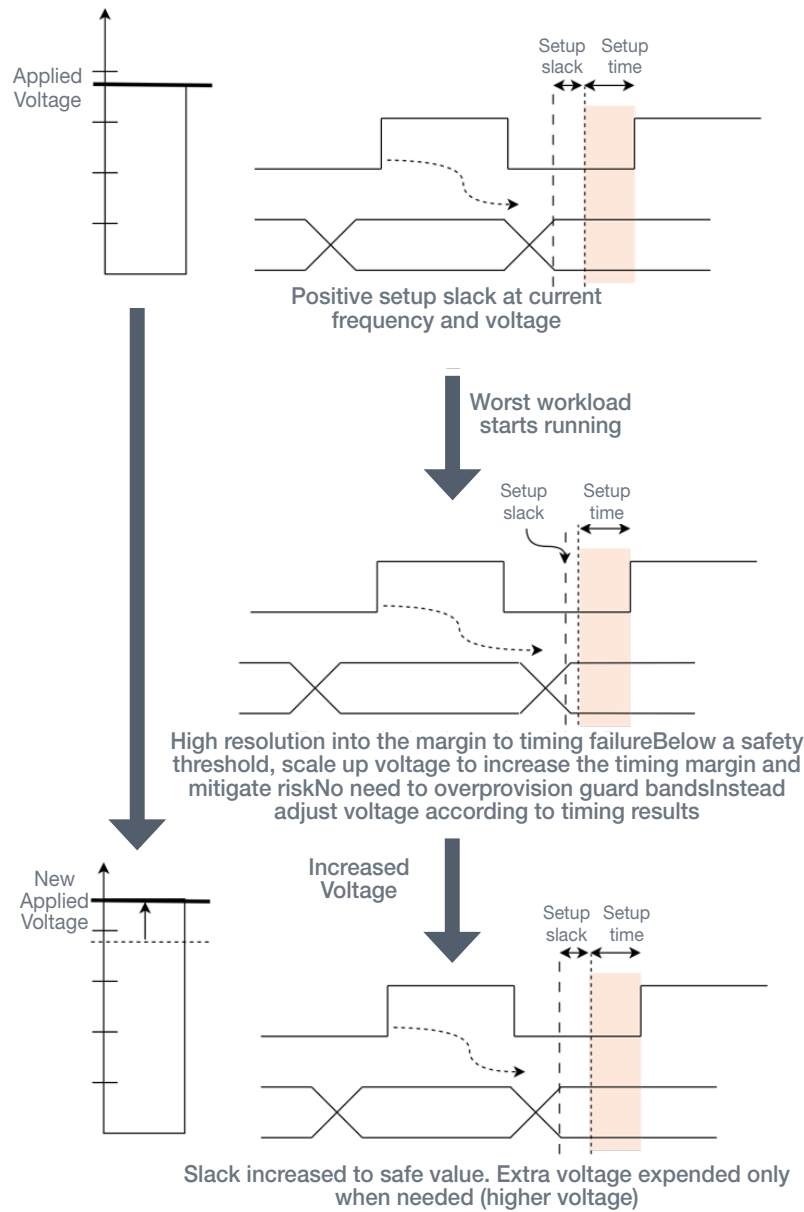they have a small footprint and are flexible to place in a design, Margin Agents can provide very high coverage of the design's logic and especially on the real performance limiter paths. They also measure at a high frequency (@ clock) so that their outputs can be acted on in real time.

Margin Agents do not require a dedicated test or mode to run in. They measure while a device is being tested or functions in the field. They can also measure timing margin during capture at speed, in scan tests.

Figure 8: proteanTecs Margin Agent schematics

The Margin Agent budget is decided using the advanced proteanTecs IP compiler (Agent Advisor). The result of the Agent Advisor is RTL that the designer integrates into their design. However, connectivity of the Margin Agents to the monitored FFs is done automatically in the PnR phase (not in RTL) by provided proteanTecs tools that makes the task seamless and guarantee coverage of the performance limiters paths.

This is the fundamental capability that allows AVS Pro to provide very high coverage of performance limiters paths, independent of the process variation of the chip, the operating conditions, the operating workload, IR drop and so on. Power Management systems can use the timing margin measurements to understand how much lower the voltage can be reduced without risking timing failure.

The AVS Pro solution

The Application

proteanTecs AVS Pro capitalizes on the ability of Margin Agents to provide high coverage monitoring of a chip's performance limiting paths' margin-to-failure to enable two main features.

1. Protection against timing failures

2. Optimization of power and performance

Protecting against timing margin failure

When the margin to timing setup for any of millions of paths falls below "Minimum Timing Margin" at any point in time, due to any reason, an alert is asserted. Falling below this threshold means there is critically low timing margin and thus a higher is a risk of failure. This constitutes a "safety net" against failure and is characterized during product characterization. The high coverage, low latency alert allows fast reaction by the host system, by throttling the clock to a safe frequency (if applicable) while scaling up the voltage.

Power and Performance optimization by utilizing excess margins

When the minimum margin to timing failure is met by all circuits in a design, the excess margin may be trimmed to increase performance or to decrease power consumption.

AVS Pro Firmware (FW) measures the worst margin in the blocks for the respective power domain, targeted for voltage scaling, and incrementally reduces the voltage (by updating the voltage regulator) until a "Minimum Reduction Margin" is reached. This limit would be the highest between the least amount of voltage needed at that time and a user-configurable minimum allowed voltage. The latter limit may be determined during product characterization or may be dictated by the specifications of design components (e.g., memory minimum voltage), or how much of the guard-band is allowed to be reclaimed. It may be set more pessimistically or optimistically at the user's discretion.

Real World Scenarios

Figure 9 shows the results of a real system, where the timing margin data and the internal DC voltage are extracted from the device (by proteanTecs Agents) to show the AVS Pro operation on a dashboard.

A. Steady state - under the given workload, applied voltage and other conditions, the worst margin is constant and steady.

B. AVS Pro enabled - the FW identified the margin gap from the current worst margin and the Minimum Reduction Margin and commences voltage reduction (Voltage Adjustment) by indicating the new target voltage to the voltage regulator.

C. Minimum Reduction Margin reached - AVS Pro continues the Protection Mode, guarding against margin falling below the Minimum Timing Margin.
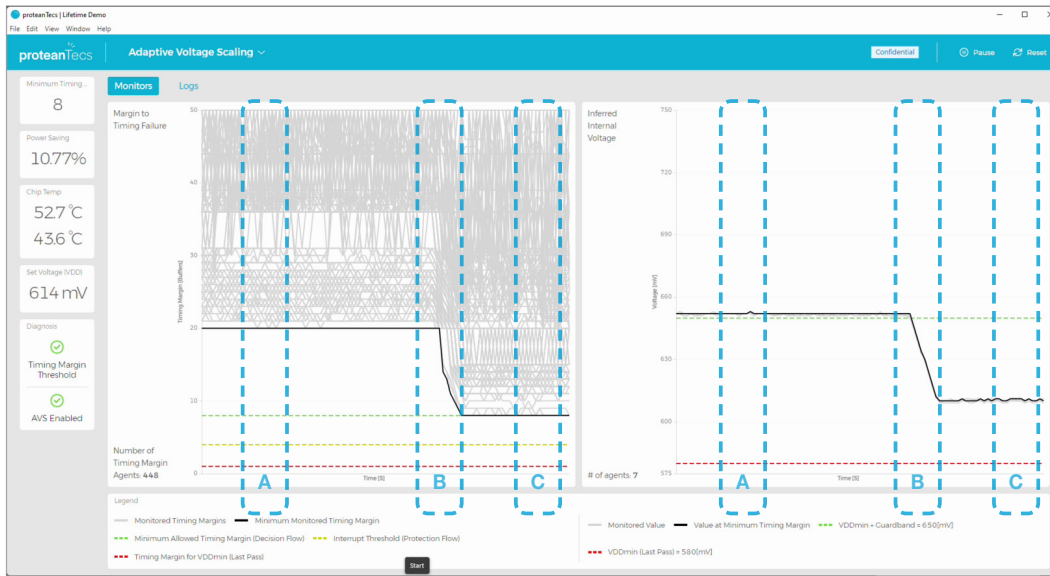
Figure 9: AVS Pro voltage optimization in a real system

Legend

Left window: margin measurements

• Grey lines: readouts of 448 Margin Agent over time

• Black line: the worst margin measured (note: may be a different Margin Agent at different times).

• Green line: the "Minimum Reduction Margin"

• Yellow line: the "Minimum Timing Margin"

• Red line: "Last pass" margin (minimum measurable margin, correlative to "Last pass voltage"), under which the device will fail.

Right window: in-chip actual voltage (by other proteanTecs Agents, out of scope here).

By running AVS Pro in the system, 12.5% of dynamic power was reduced. The total power reduction is even bigger due to the fact that leakage Power will also be reduced but was not counted here.

Figure 10 shows what happens when a protection event takes place.

D. Protection event taking place - due to any number of circumstances, e.g., increased workload that has caused an IR drop in some of the logic, the margin dropped to (below) the Minimum Timing Margin threshold.

E. Recovery - the FW instructs the voltage regulator FW to increase the voltage to a safe value (in this case Vop). Once Steady State is achieved again, the sequence starts anew.
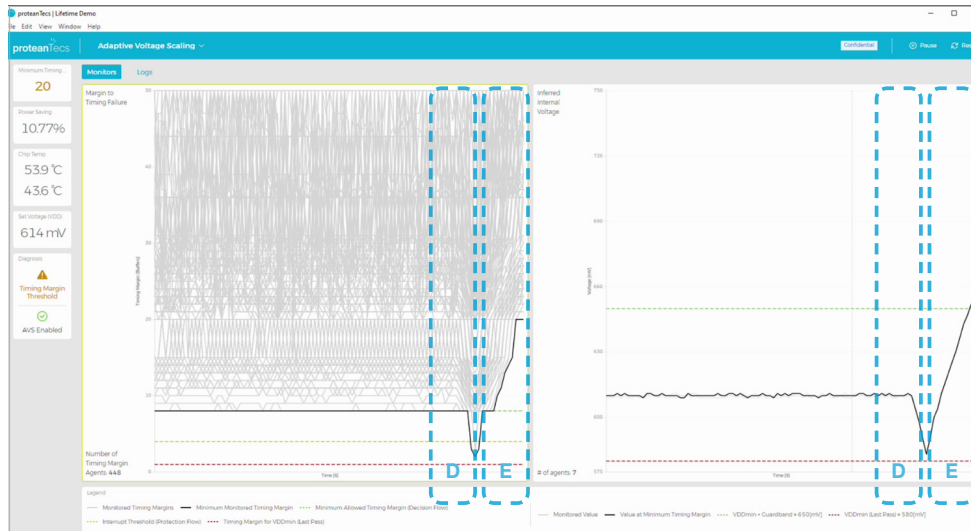
Figure 10: AVS Pro recovery from performance degrading event

5 | Summary

Due to the nature of the lifetime operation of semiconductors, performance and power guard-bands must be taken to avoid failure. The holy grail would be to tailor the power to the actual device requirements over time, so that the device gets (only) what it needs, when it needs it. But the various practices that exist today cannot achieve this, each eventually reducing some guard-bands, but always leaving too much on the table, thus leaving power on the table.

proteanTecs' AVS Pro tracks the precise margin-to-failure of millions of logic paths, including the performance limiting paths. By monitoring functional workloads, AVS Pro not only allows reduction of the power until the point of minimal guard-band, but also protects from failure. Field use has shown that the AVS Pro has saved customers between 8% and 14% power.

In the race to battle surging power demands, rising costs and reliability risks while enhancing product performance, every mW saved counts.

proteanTecs

proteanTecs Ltd.

www.proteanTecs.com