

# PPA (V) : 可变工作电压下的每瓦性能优化

## 作者

庄定铮

新思科技

资深产品经理

## 摘要

每瓦性能已成为保证设计质量的优先考量因素之一，这导致了技术重心和设计功耗优化方法的转变。可变工作电压对于优化每瓦性能结果具有很大的潜力，但这需要探索出准确且高效的方法论。新思科技 Fusion Design Platform 采用基于单一 RTL-to-GDSII 数据模型的独特理念，提供了全流程电压优化和收敛方法，可在要求最严苛的半导体领域实现理想的每瓦性能。

## 高性能计算 (HPC) 持续推进高级节点 PPA 超越极限

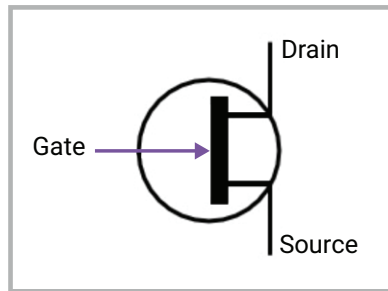
高性能计算 (HPC) 是半导体行业中增长最快的设计领域之一，为云数据中心、人工智能、移动计算、自动驾驶汽车等多种应用提供了驱动力。尽管这个领域持续为实现最高设计性能而努力，但功耗会使多个应用领域的设计性能受到限制。例如，能源和冷却成本会直接影响数据中心的盈利能力，而手机必须谨慎地平衡性能和电池寿命之间的关系。

因此，除了既有的性能、功耗和面积 (PPA) 标准外，每瓦性能已成为 HPC 设计质量的优先考量因素之一，致使技术重心和设计功耗优化方法发生了转变。

另一方面，为了实现最佳的每瓦性能目标，HPC 设计采用了最先进的 FinFET 流程。虽然创新的“鳍片”对电子流具有出色的控制能力，但要实现更快的切换和更低的漏电流，需要比同等的平面结构更大的切换功率。此外，HPC 设计的计算密集型工作负载也会导致几乎不间断的切换，致使功率分布严重受到动态功率主导。

## 动态功率优化的新机会

完成一次切换所需要的功率是由晶体管的栅极电容 (FET) 消耗的。根据图 1 所示的功耗方程，在相同频率下，消耗的功率与栅极电容成线性关系，但与工作电压的平方成正比。



$$\text{动态功率} = \text{电容栅极} * \text{电压}^2 * \text{频率}$$

图 1: FET 晶体管动态功率方程

既有的动态功率优化技术以减小晶体管的尺寸为目标，这会直接导致栅极电容降低。然而，降低工作电压对于降低动态功率具有更大的潜力。如图 2 所示，关于 7nm FinFET 设计的案例表明，仅降低 5% 的工作电压就可以使动态功率降低 9%。

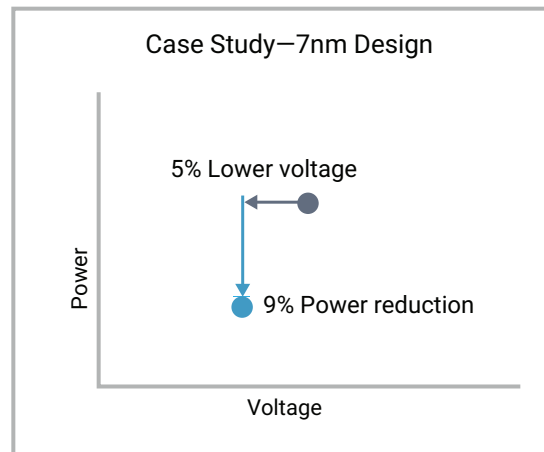


图 2: 工作电压与动态功率案例研究

虽然降低工作电压对动态功率优化非常有效，但在半导体设计行业，这还没有成为一种广泛部署的技术。接下来，我们将探讨导致这一结果的挑战。

## 过去：电压决定与设计优化无关

在以往的半导体设计环境中，工作电压 (Vdd) 是预先定义好的，与生产设计环境无关。这个过程涉及到在晶圆厂对晶体管设备进行分析，同时需要对一小部分单元进行内部模拟，以确定合理的最低工作电压。然后，预先确定的工作电压将推动所有设计的技术库特性、设计优化和签核收敛。

如图 3 所示，对每瓦性能进行优化的设计人员将针对不同性能目标执行多次运行任务来探索预设电压内的合适范围。

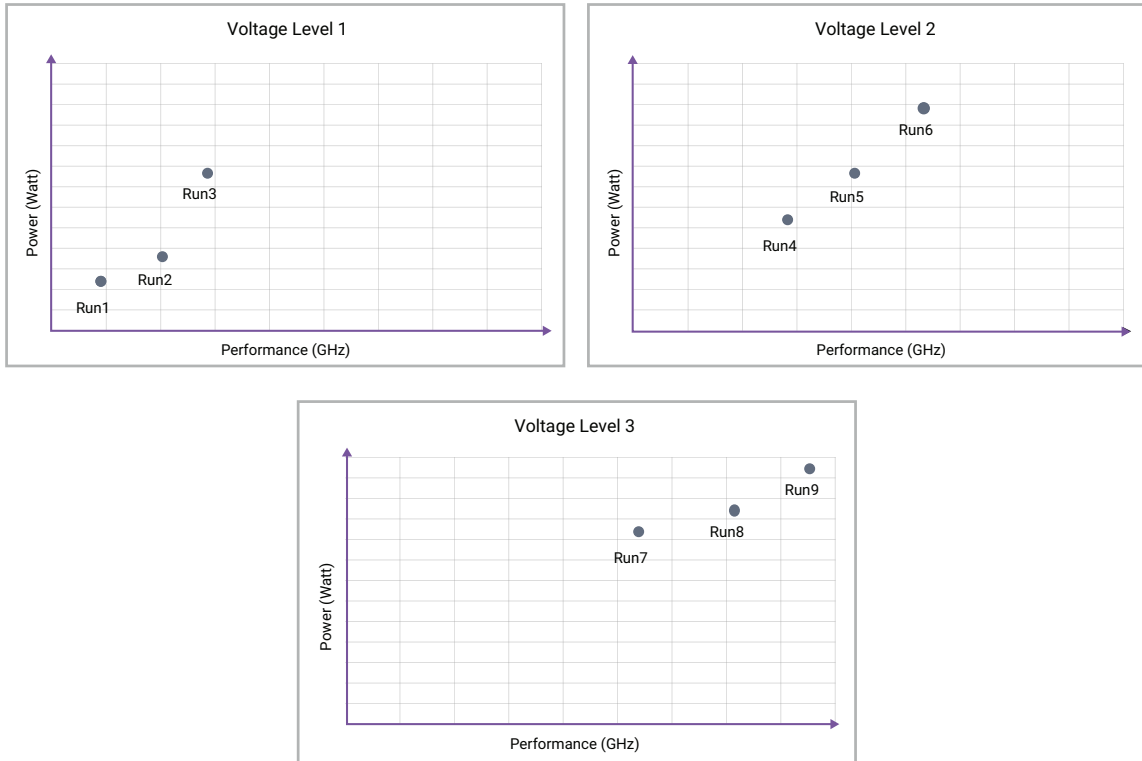


图 3：离散电压水平的每瓦性能优化

假设一个项目在预定电压水平达到了所有 PPA 目标。在这种情况下，由于缺乏技术库，致使没有可行的方法探索在较低工作电压下进一步优化的方式。

因此，在流片后测试过程中，在流片测试台上探索更低的工作电压（即 Vmin 分析），并找到设计能够持续正常运行的最低工作电压，已经成为一种普遍做法。最终结果将用于影响是否对新的库进行特性抽取。这个反馈环路可能需要几个月时间，或者需要多个设计周期才能对提高每瓦性能产生显著影响。

## 在设计流程中进行自由的电压探索

电压水平和时序响应之间的非线性限制了线性电压插值的部署，使其只能在更高的标准电压水平及两个紧密相邻的库之间部署。2017 年，PrimeTime® 解决方案融合了经代工厂认证的先进电压缩放技术，可在较广范围内的任何电压水平下进行精确分析。如图 4 所示，在较宽间隔之间或较低电压水平下可以实现准确签核的电压缩放结果。

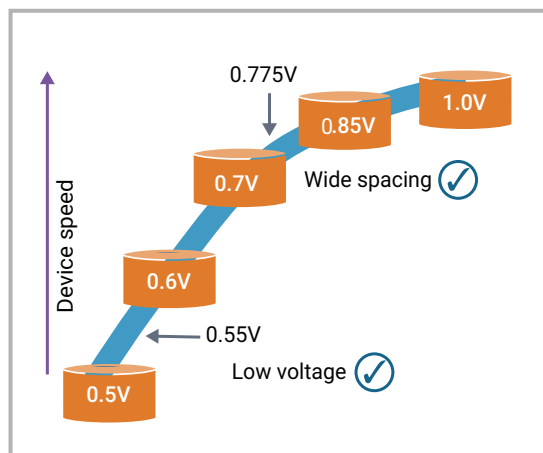


图 4：PrimeTime 的先进电压缩放技术

现在，设计人员能够“扫描”电压范围，在无限的电压水平下试运行相同的设计方案，并最终找到最优的 PPA 或每瓦性能目标。如图 5 所示，尽管事实证明 PrimeTime 解决方案准确且高效，但手动扫描过程耗时较长，且需要消耗大量资源。通过探索，运行次数随着电压水平候选者呈线性增长。

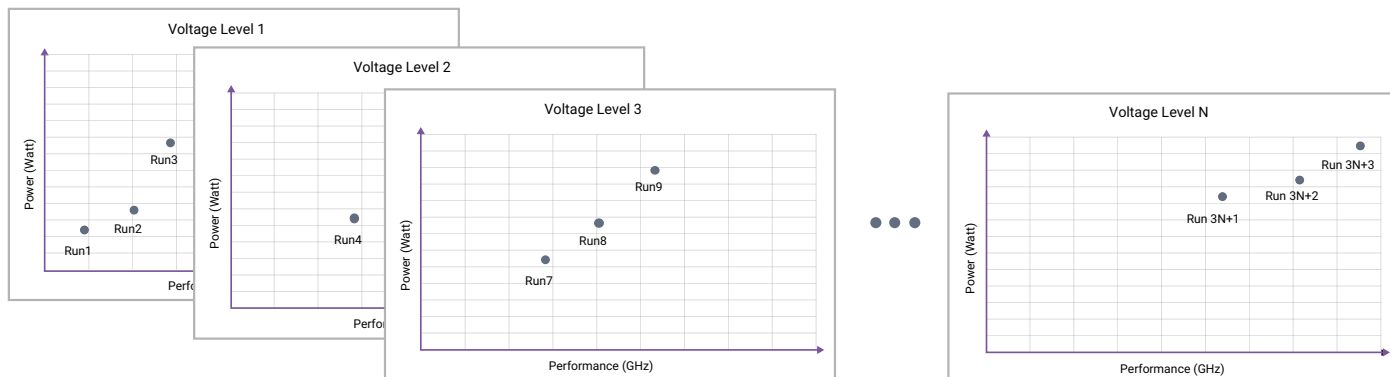


图 5：通过扫描电压范围实现每瓦性能优化

## PPA (V) 优化：在设计优化过程中引入电压作为变量

Fusion Compiler™ RTL-to-GDSII 解决方案和 IC Compiler II™超收敛布局布线解决方案是业界唯一在实现和 PPA 优化过程中部署了新思科技最值得信赖的黄金签核解决方案的数字设计实现解决方案。独特的 Signoff Fusion 技术可在实现环境中无缝启用 PrimeTime、PrimePower 和 StarRC™分析引擎，以实现准确的时序、功率和互连 RC 指导，包括 PrimeTime 先进的电压缩放技术。

通过在设计优化期间引入电压作为变量，Fusion Compiler 和 IC Compiler II 可以扩展每瓦性能的合适范围。通过缩放工作电压并同时推动更高性能，更低功耗和更小面积等指标，优化引擎可以在单次优化运行中自行探索工作电压水平，而无需像以前的电压扫描那样需要昂贵的额外迭代。

在设置了固定频率目标的设计流程中，可变工作电压提供了进一步探索的机会，可以通过更低的工作电压直接降低动态功率，同时最大程度减少泄漏和对面积的影响，从而降低总体功率。如下图 6 所示，通过采用低 14% 的工作电压，在满足相同频率目标的前提下，5nm HPC 设计能够使总体功率降低 26%：

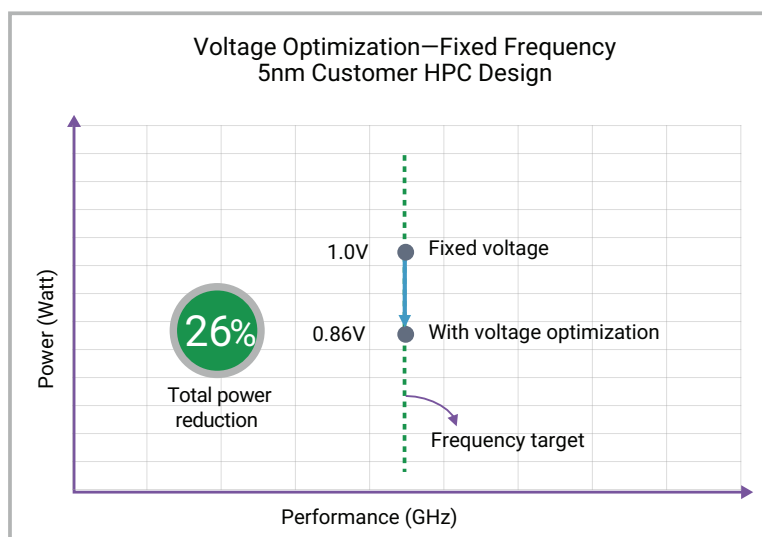


图 6：以更低的工作电压降低总体功率

在设置了固定总体功率目标的设计流程中，可变工作电压可以消除以前为找到可能的最高频率而需要进行的频率和电压扫描过程。优化引擎可以自行探索电压范围，并在单次运行中找到最佳可能的频率，而由于时间和资源的限制，其他方法无法做到这一点，如图 7 所示：

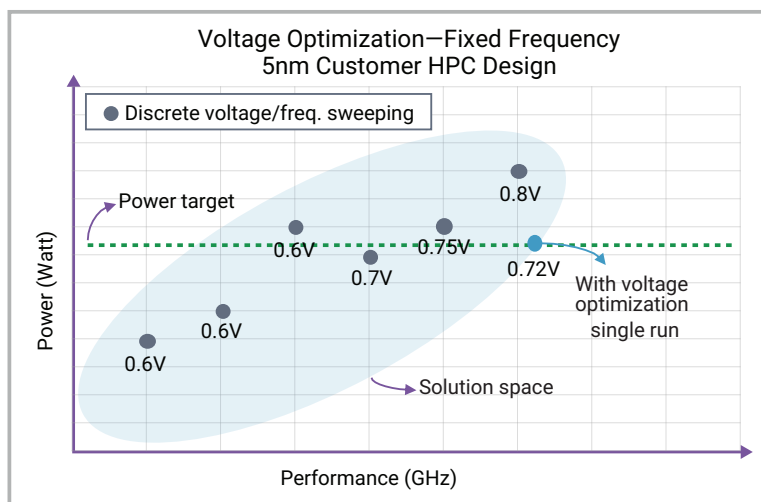


图 7：通过设定固定功率目标而提高频率

## PPA (V) 签核：在设计分析和签核过程中引入电压鲁棒性

电压优化提供了 PPA 优化机会，提升了 PPA 曲线，并且增强了 SoC 设计的每瓦性能。其创新的本机电压扫描功能扩展了优化范围的探索，可在较低的工作电压下实现尽可能高的每瓦性能。

降低工作电压可以显著提高每瓦性能，同时还消除了固定电压设计流程无意产生的过多裕量。除此之外，高级节点中较高的单元和功率密度还需要更先进的工作电压压降分析和裕度确定方法，以防止与压降相关的设计失败。

PrimeShield™ 解决方案扩展了 PrimeTime 的核心技术，并通过引入本机电压分析功能而应对这一挑战。与根据时序裕量报告关键时序路径的静态时序分析不同，这种新的分析方法根据电压裕度这个新指标报告关键路径。如图 9 所示，这一新指标表示一个路径仍然满足时序要求所需的每单元或每路径最小压降。

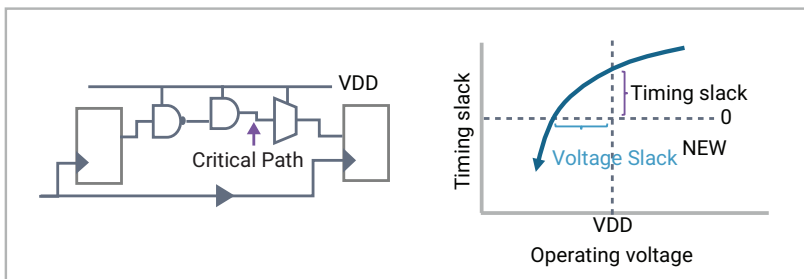


图 8：关键路径的电压裕度报告

时序关键路径通常采用最强驱动单元确保路径延迟满足性能要求。通常情况下，这些强驱动单元对电压变化不太敏感，如图 10 所示。同时，不太关键的时序路径可能采用对电压更敏感的较弱驱动单元，在设计过程中遇到工作电压下降时，这些驱动单元会更早失效。这些诱发风险的路径并不能使用静态时序分析方法直接找到，通常需要大量的电压扫描工作才能发现。

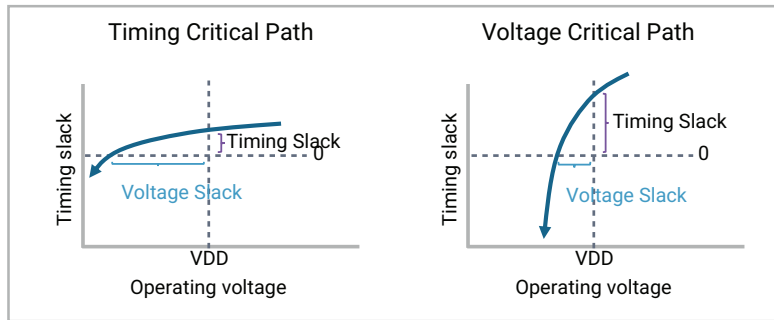
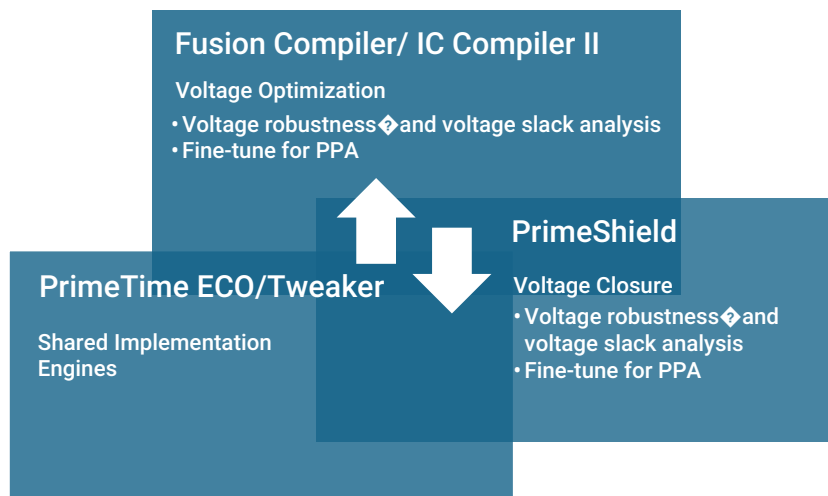


图 9：时序与电压关键路径的电压裕度分析

PrimeShield 的电压裕度分析提供了直接而且高效的途径，用于分析和报告此类信息，以供进一步优化。为了适应高效的优化方法，除了电压裕度报告外，PrimeShield 的电压鲁棒性分析还对电压关键路径进行了瓶颈分析，目的是识别对电压变化最敏感的单元，并设定时序失效的最高风险。

新思科技最受欢迎的 ECO 解决方案— PrimeTime ECO 和 Tweaker，可提供 ECO 指导，以改进上述指标。通过对电压敏感的单元更换为敏感度较低的单元，ECO 的变化可以提高设计的抗压降能力，或者在所有签核时序场景中进一步微调工作电压。

该技术还可以通过方法实现电压裕度的均匀性并改善电压裕量。通过确保在需要的地方保证电压裕量，并消除整个设计中任何未处理的风险点，设计人员无需在全局应用过高裕度，同时获得较低工作电压所带来的功率优势。



## 结论

随着半导体行业（尤其是 HPC 设计领域）持续推崇更高的每瓦性能，Fusion Compiler 和 IC Compiler II 的电压优化能力（基于新思科技的黄金签核解决方案）通过在优化过程中引入工作电压作为变量，为有效增强高级节点设计的每瓦性能提供了一种不同的方法。

PrimeShield 创新的设计鲁棒性和电压裕度分析基于相同的核心基础。新的分析指标使设计人员能够有效地找出电压鲁棒性的瓶颈，提高电压裕度设定效率，并发现直接微调工作电压的机会。

新思科技 Fusion Design Platform 独特地基于单个 RTL-to-GDSII 数据模型而构建，可提供全流程电压优化和收敛方法论，为要求最严苛的半导体领域提供最佳的每瓦性能结果。

