

DarkGates

A Hybrid Power-Gating Architecture to Mitigate the Performance Impact of Dark-Silicon in High Performance Processors

Jawad Haj-Yahya

Jeremie S. Kim
Efraim Rotem

A. Giray Yağlıkçı
Yanos Sazeides

Jisung Park
Onur Mutlu

SAFARI

ETH zürich

intel®

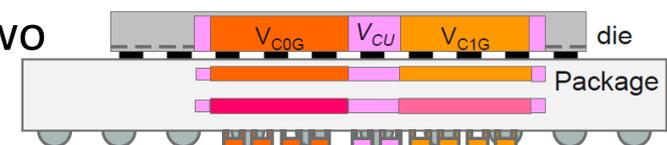
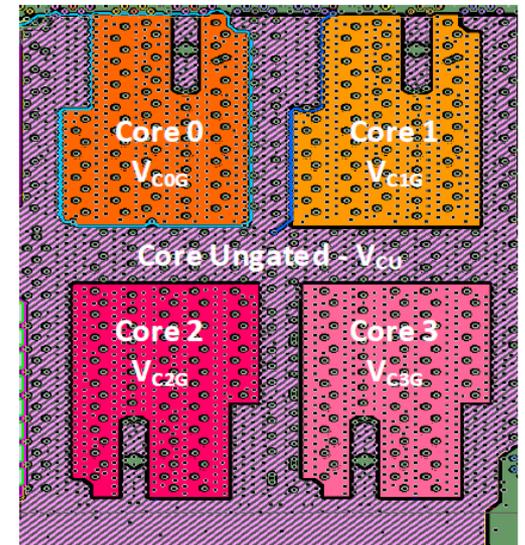
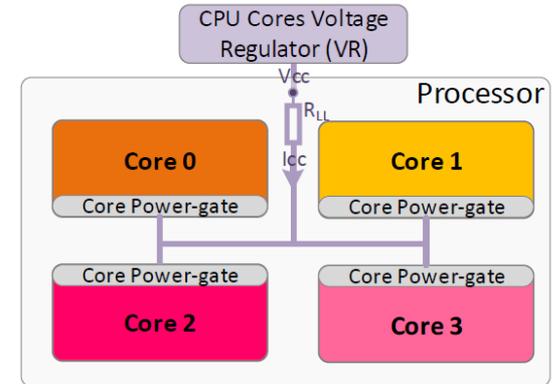


University
of Cyprus

 **HUAWEI**

Client Processor Power Management Architecture

- A high-end client processor is a system-on-chip that typically **integrates three main domains** into a single chip:
 - Compute (e.g., CPU cores and graphics engines)
 - IO
 - Memory system
- We show the architecture used in recent Intel processors with a **focus on CPU cores**
 - Each CPU core has **a power-gate** for the entire core
- Package layout showing
 - An **ungated** main voltage domain (V_{CU})
 - **Four power-gated** voltage domains (one for each CPU core, V_{C0G} , V_{C1G} , V_{C2G} , and V_{C3G})
- Side view of die and package showing
 - The **ungated** main voltage domain (V_{CU}) and two cores' voltage domains (V_{C0G} and V_{C1G})
 - The package's **decoupling capacitors**



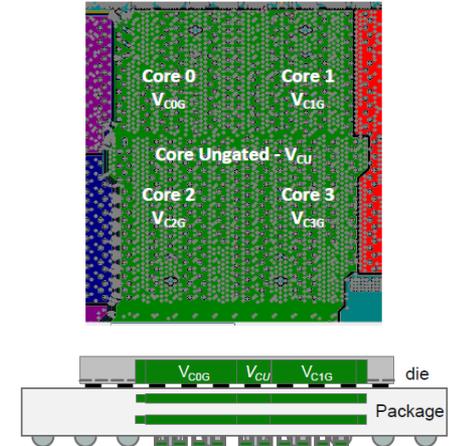
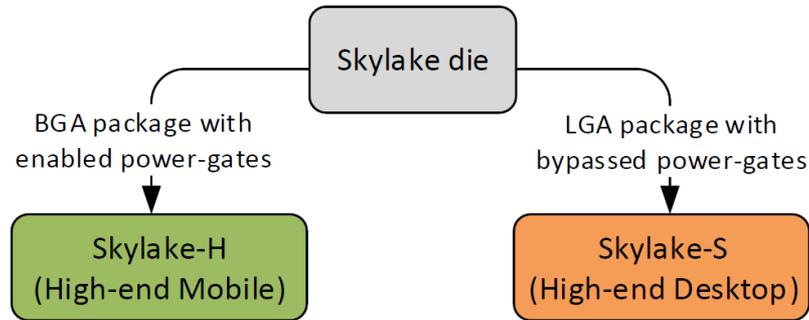
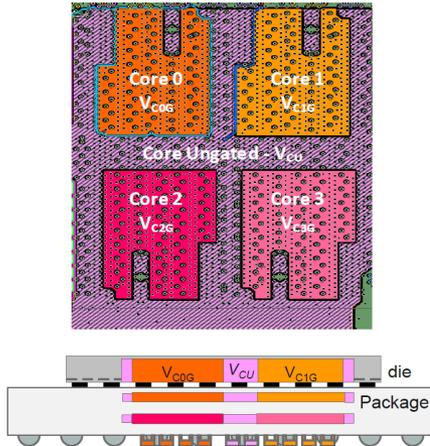
Our Goal

- Based on our experimental analyses, we propose **DarkGates**, a hybrid system architecture that increases the performance of F_{\max} -constrained systems while fulfilling their **energy efficiency requirements**
- We design **DarkGates** with two design **goals** in mind:
 1. Reduce CPU cores' **power-delivery impedance** to reduce voltage guardband, thereby improving the performance of high-end desktops
 2. Meet the **energy efficiency requirements** of desktop devices by enabling deeper package C-states

DarkGates

- We achieve **DarkGates** goals with three key components:
 1. A Power-gate Bypassing technique
 2. Improved power management firmware algorithms
 3. A new deep package C-state

1. Power-gate Bypassing



- Power-gate Bypassing technique is responsible for **reducing CPU cores' voltage drop** in F_{\max} -constrained systems by **reducing system impedance**
- The technique uses the same Intel Skylake die to build:
 - A dedicated package for **Skylake-H** (high-end mobile) with the **power-gates enabled** and
 - A dedicated package for **Skylake-S** (high-end desktop) that **bypasses the power-gates** at the package level by **shorting gated and un-gated CPU core power domains**
- This architecture is feasible since client processors typically **share the same die** between **Skylake-H** and **Skylake-S**

2. Improved Power Management Algorithms

- DarkGates architecture requires the adjustment of three main components of the **power management unit**:
 1. Adjustment of **DVFS firmware** power management algorithms (e.g., P-state, Turbo) to take into account the **new V/F curves for the desktop system**
 2. Adjustments of the **power budget management** algorithm (PBM) to take into account the **additional power consumption** due to the leakage of inactive cores
 3. Adjustment of the **reliability voltage guardband** since DarkGates can change the processor's lifetime reliability due to **powering-on cores more time** compared to baseline

3. New Package C-state for Desktops

- Intel desktop processors that are prior to Skylake (e.g., Haswell, Broadwell) support **up to the package C7 state**
- Desktop's energy efficiency benchmarks (**ENERGY STAR & Intel Ready Mode Technology**)
 - Have large phases in which **the processor is fully idle**
 - Their average power consumption needs **can be met with package C7 state**
- The power consumption of package C7 is **~3x higher** in DarkGates than in the baseline due to the additional cores leakage power
 - Since the **voltage regulator is turned on** in package C7 state and the power-gates are bypassed
- To mitigate this issue, we extend the desktop systems with the **package C8 state**
 - A deeper package C-state in which the voltage regulator of the **CPU cores is off**

Package C-state	Major conditions to enter the package C-state
C0	One or more cores or graphics engine executing instructions
C2	All cores in CC3 (clocks off) or deeper and graphics engine in RC6 (power-gated). DRAM is active .
C3	All cores in CC3 or deeper and graphics engine in RC6 . Last-Level-Cache (LLC) may be flushed and turned off, DRAM in self-refresh , most IO and memory domain clocks are gated, some IPs and IOs can be active (e.g., DC and Display IO).
C6	All cores in CC6 (power-gated) or deeper and graphics engine in RC6 . LLC may be flushed and turned off, DRAM in self-refresh , IO and memory domain clocks generators are turned off. Some IPs and IOs can be active (e.g., video decoder (VD) and display controller (DC)).
C7	Same as Package C6 while some of the IO and memory domain voltages are power-gated . CPU core VR is ON .
C8	Same as Package C7 with additional power-gating in the IO and memory domains. CPU core VR is OFF .
C9	Same as Package C8 while all IPs must be off. Most voltage regulators' voltages are reduced. The display panel can be in panel self-refresh (PSR)
C10	Same as Package C9 while all SoC VRs (except state always-on VR) are off. The display panel is off.

3. New Package C-state for Desktops

- Intel desktop processors that are prior to Skylake (e.g., Haswell, Broadwell) support up to the package C7 state
- Desktop's energy efficiency benchmarks (ENERGY STAR & Intel Ready Mode Technology)
 - Have large phases in which the processor is fully idle

Package C-state	Major conditions to enter the package C-state
C0	One or more cores or graphics engine executing instructions
C2	All cores in CC3 (clocks off) or deeper and graphics engine in RC6 (power-gated). DRAM is active .
C3	All cores in CC3 or deeper and graphics engine in RC6 . Last-Level-Cache (LLC) may be flushed and turned off, DRAM in self-refresh , most IO and memory domain clocks are gated, some IPs and IOs can be active

The new **package C8 state** reduces the CPU cores' **leakage power** and saves even more power in the uncore compared to **the package C7 state**

- The power consumption of package C7 is **~3x higher** in DarkGates than in the baseline due to the additional cores leakage power
 - Since the **voltage regulator is turned on** in package C7 state and the power-gates are bypassed
- To mitigate this issue, we extend the desktop systems with the **package C8 state**
 - A deeper package C-state in which the voltage regulator of the **CPU cores is off**

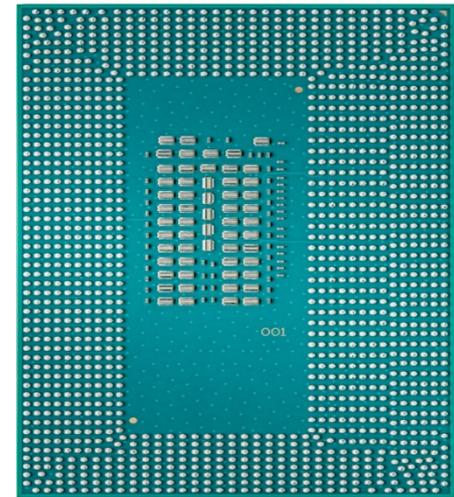
C7	Same as Package C6 while some of the IO and memory domain voltages are power-gated . <u>CPU core VR is ON.</u>
C8	Same as Package C7 with additional power-gating in the IO and memory domains. <u>CPU core VR is OFF.</u>
C9	Same as Package C8 while all IPs must be off. Most voltage regulators' voltages are reduced. The display panel can be in panel self-refresh (PSR)
C10	Same as Package C9 while all SoC VRs (except state always-on VR) are off. The display panel is off.

Methodology

- **Framework:** We implement DarkGates on the Intel Skylake die that targets high-end desktop (Skylake-S) and high-end mobile (Skylake-H) processors
 - For our **baseline** and **DarkGates** measurements, we use the **Skylake-H (mobile)** and **Skylake-S (desktop)**, respectively
- **Configuring the Processor:** We use Intel's In-Target Probe (ITP) silicon debugger tool that connects to an Intel processor through the JTAG port
- **Power Measurements:** We measure power consumption when running energy-efficiency benchmarks by using a National Instruments Data Acquisition (NI-DAQ) card
- **Workloads:** We evaluate DarkGates with three classes of workloads
 - **SPEC CPU2006** benchmarks to evaluate CPU core performance
 - **3DMARK** benchmarks to evaluate computer graphics performance
 - **ENERGY STAR** and **Ready Mode Technology (RMT)** workloads to evaluate the effect of DarkGates on energy efficiency

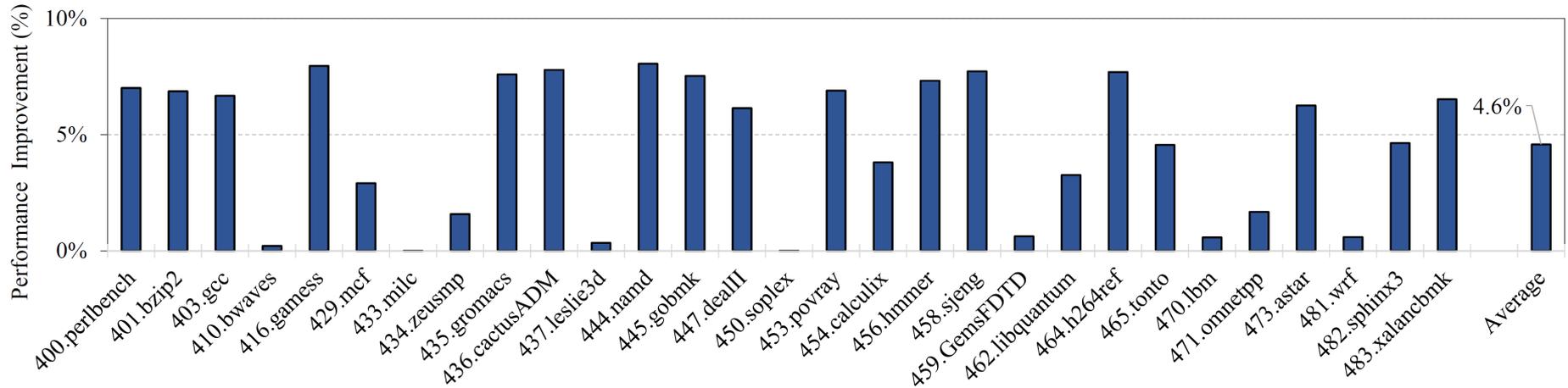


Skylake-S



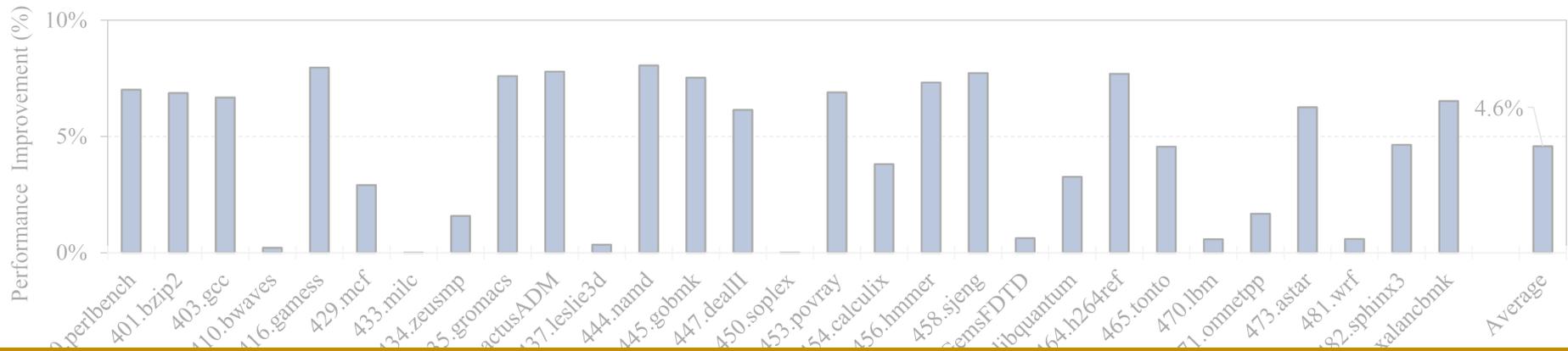
Skylake-H

Evaluation of CPU Workloads



- DarkGates improves real **system performance by up to 8.1%** (4.6% on average)
- The performance benefit of DarkGates is **positively correlated with the performance scalability** of the running workload with CPU frequency
 - **Highly-scalable workloads** (i.e., those bottlenecked by CPU core frequency, such as 416.games and 444.namd) experience the highest performance gains
 - Workloads that are **heavily bottlenecked by main memory**, such as 410.bwaves and 433.milc, have almost no performance gain

Evaluation of CPU Workloads



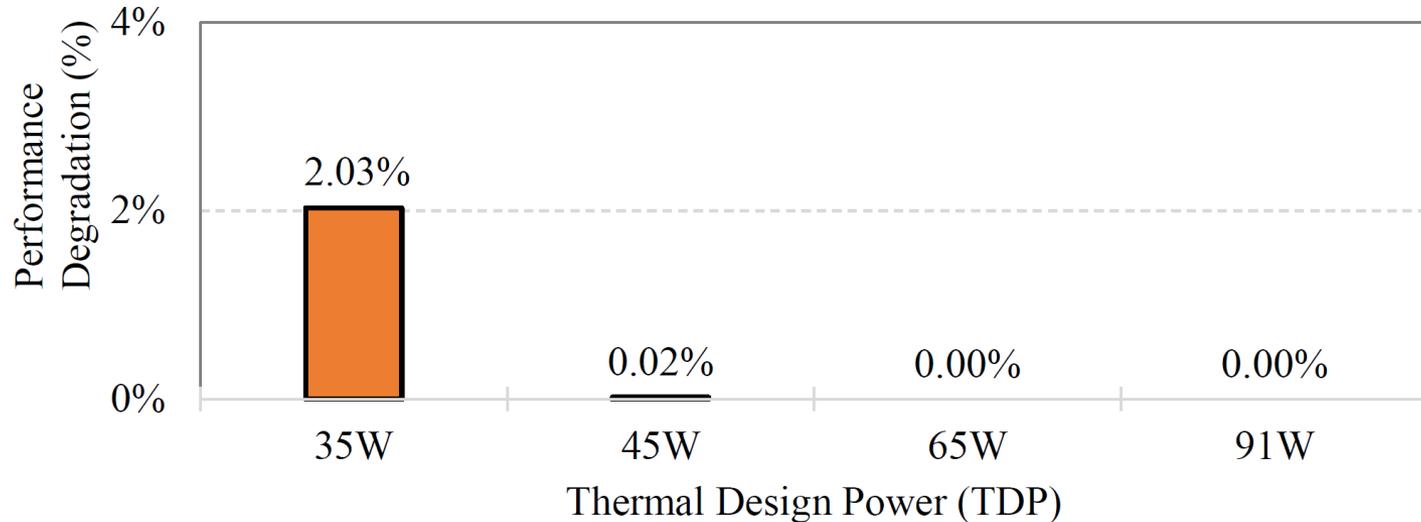
DarkGates significantly improves CPU core performance by reducing the voltage guardband with Power-gate bypassing

Doing so improves the V/F curves and leads to higher CPU core frequency for both thermally-constrained and V_{\max} -constrained systems

performance scalability of the running workload with CPU frequency

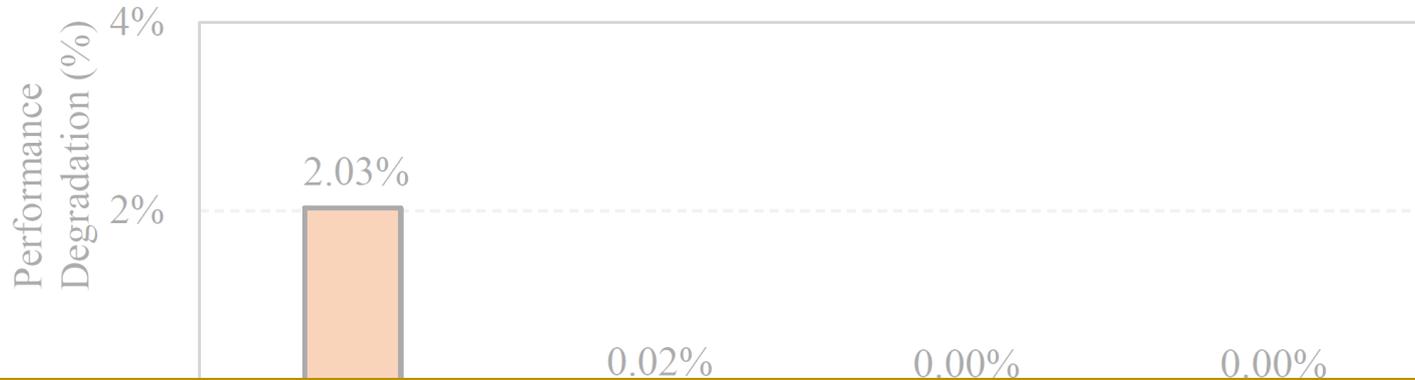
- Highly-scalable workloads (i.e., those bottlenecked by CPU core frequency, such as 416.games and 444.namd) experience the highest performance gains
- Workloads that are heavily bottlenecked by main memory, such as 410.bwaves and 433.milc, have almost no performance gain

Evaluation of Graphics Workloads



- DarkGates provides the **same system performance** for 3DMark for TDP levels $\geq 45W$
 - Since graphics workloads in these systems are **not limited by thermal constraints**
- DarkGates leads to only **2% performance degradation** for a TDP level of 35W
 - The additional **leakage power of the idle CPU cores** forces the power budget management algorithm to **reduce the frequency of the graphics** engine to keep the system within the TDP limit

Evaluation of Graphics Workloads

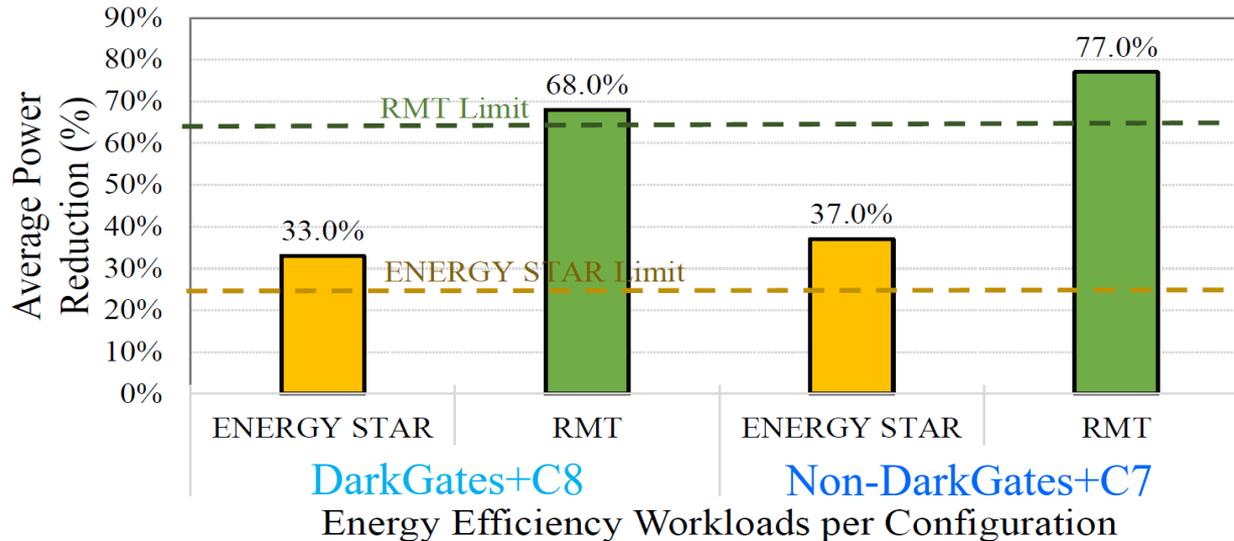


The reduced graphics engine power budget due to the additional leakage power of idle CPU cores **can slightly degrade the performance** of graphics workloads in **thermally-limited systems**

Not a main concern in many real systems that are **not thermally-limited**

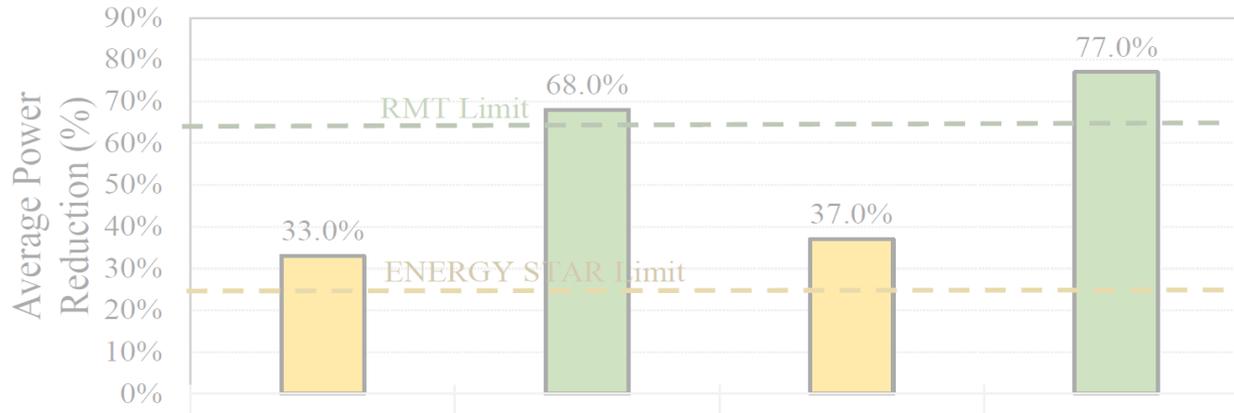
- DarkGates leads to only **2% performance degradation** for a TDP level of 35W
 - The additional **leakage power of the idle CPU cores** forces the power budget management algorithm to **reduce the frequency of the graphics** engine to keep the system within the TDP limit

Evaluation of Energy Efficiency Workloads



- ENERGY STAR and Intel Ready Mode Technology (RMT) efficiency workloads have fixed performance requirements and include long idle phases
- DarkGates system (*DarkGates+C8*) reduces the average power consumption of ENERGY STAR and RMT by 33% and 68%, respectively, on the real Intel Skylake-S system
 - Compared to the baseline where we limit the deepest package C-state to C7 (*DarkGates+C7*)
- The baseline system (*DarkGates+C7*) does not meet the target power limit for both workloads
- The system without DarkGates at the deepest package C-state of C7 (*Non-DarkGates+C7*) shows higher power reduction versus the system with DarkGates at the deepest package C-state of C8 (*DarkGates+C8*)
 - When some of the cores are idle they consume leakage power in *DarkGates+C8*, but they are power-gated in *Non-DarkGates+C7*

Evaluation of Energy Efficiency Workloads



Applying **DarkGates with package C8 state** significantly reduces the average processor power consumption,

thereby **meeting the target average power requirements** of the energy efficiency standards

- ENERGY STAR requirements
- DarkGates 33% and
- Compared to the baseline where we limit the deepest package C-state to C7 (*DarkGates+C7*)
- The baseline system (*DarkGates+C7*) does not meet the target power limit for both workloads
- The system without DarkGates at the deepest package C-state of C7 (*Non-DarkGates+C7*) shows higher power reduction versus the system with DarkGates at the deepest package C-state of C8 (*DarkGates+C8*)
- When some of the cores are idle they consume leakage power *DarkGates+C8*, but they are power-gated in *Non-DarkGates+C7*

DarkGates

A Hybrid Power-Gating Architecture to Mitigate the Performance Impact of Dark-Silicon in High Performance Processors

Jawad Haj-Yahya

Jeremie S. Kim
Efraim Rotem

A. Giray Yağlıkçı
Yanos Sazeides

Jisung Park
Onur Mutlu

SAFARI

ETH zürich

intel®



University
of Cyprus



HUAWEI