

SuperRange: Wide Operational Range Power Delivery Design for both STV and NTV Computing

Xin He^{*†}, Guihai Yan^{*}, Yinhe Han^{*}, and Xiaowei Li^{*}

^{*}State Key Laboratory of Computer Architecture,
Institute of Computing Technology, Chinese Academy of Sciences

[†]University of Chinese Academy of Sciences
{hexin,yan_guihai,yinhes,lxw}@ict.ac.cn

Abstract—The load power range of modern processors is greatly enlarged because many advanced power management techniques like dynamic voltage frequency scaling, Turbo boosting, and Near Threshold Voltage technologies are incorporated. However, the power saving may be offset by power loss in power delivery; moreover, as the efficiency of power delivery varies greatly with different load conditions, conventional power delivery designs cannot maintain high efficiency over the entire voltage range. We propose SuperRange, a wide operational range power delivery scheme. SuperRange complements the power delivery capability of on-chip voltage regulator and off-chip voltage regulator. Experimental results show SuperRange has an average 70% power conversion efficiency over wide operational range which outperforms conventional power delivery schemes. And it also exhibits superior resilience to power-constrained systems.

I. INTRODUCTION

Nowadays processor design has gradually changed from performance-oriented paradigm to power efficiency-oriented paradigm because of the approaching end of Dennard’s scaling [1]. To boost efficiency, the state-of-the-art processors provide multiple performance states (P-states) to enable flexible power management, which has even been marked as a selling point by processor vendors.

Implementing P-states requires dynamic voltage frequency scaling (DVFS) techniques which dictate a power delivery subsystem supporting a wide range of voltage levels. For example, to enable all P-states of an Intel Pentium processor, the power delivery has to provide from 0.9V at P5, the lowest performance state, to 1.4V at the P0, the highest performance state [2]. The operational voltage range will be expanded when incorporating some advanced techniques such as Turbo Boost [3] and Near Threshold Voltage (NTV) [4] technology. Turbo Boost requires higher-than-nominal voltage to support instant over-clocking, while NTV uses lower-than-nominal voltage to achieve aggressive power saving.

However, the power delivery subsystem is never ideal. It draws power when transfers power from power source to the loading processor, which decreases the power conversion efficiency (PCE), i.e. the ratio of processor power to the total power drawn from the board power source. Within optimal range, PCE can reach up to 90% with well-designed on-board voltage regulator (VR), referred as off-chip inductor based VR (Off-VR). However, out of that optimal range, PCE decreases significantly. This can greatly dampen and even totally eat up the power benefit achieved by wide operational range power management. This argument can be confirmed by Fig. 1(a). The result shows a typical Off-VR’s PCE on different output voltage levels ranging from 1.1V to 2.3V [5].

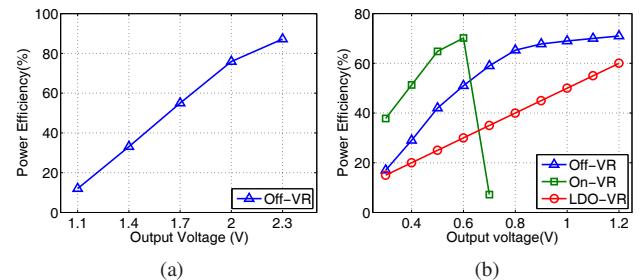


Fig. 1. Power conversion efficiency at different output voltage levels

The PCE reaches the optimal point (89%) at the highest voltage of 2.3V, but plunges to merely 12% at the lowest 1.1V. Clearly, in the low-voltage states, the power efficiency cannot be high because the power lost in the delivery has dominated the total power.

The optimal PCE of conventional Off-VR resides at the high voltage end, but gradually decreases at lower voltage levels. This is not surprising since the power delivery is designed for peak performance. This conventional wisdom becomes less effective when comes to the state-of-the-art processors with wide operational voltage range. The future processors are likely required to work in both traditional super-threshold voltage (STV) mode and aggressive power-saving NTV mode to achieve ultra-high efficiency [6]. This requirement, however, cannot be served well with current power delivery designs. The looming challenge is *how to design a power delivery subsystem that can provide wide operational range and, more importantly, high PCE across the whole range.*

One intuitive solution is to design multiple Off-VRs whose optimal points are evenly located in the operational range. However, we find that designing an Off-VR-based power delivery subsystem targeting low voltage range is impractical because of the degraded regulation quality and voltage scaling speed. Also, it’s neither physically practical nor scalable because today’s processors have been very power intensive, even one Off-VR solution has greatly burdened the board design, never mention multiple Off-VR solutions. Another traditional solution is to use an on-chip LDO-VR to collaborate with Off-VR. Although the LDO-VR can be easily integrated on-chip, the PCE still remains a big problem since LDO-VR’s PCE is limited to output/input voltage ratio, as confirmed in Fig. 1(b).

We resort to on-chip switched capacitor based VR (On-VR) to complement the Off-VR at the lower-end voltage. The rationale is twofold: 1) On-VR doesn’t dictate on-board resources, but only on-chip silicon. Given the dark-silicon, trading the increasingly cheaper silicon for power delivery should be practical; 2) On-VR’s optimal point usually resides in low voltage range compared to the Off-VR, and hence it is just right to complement Off-VR in terms of optimal PCE. As Fig. 1(b) shows the efficiencies of Off-VR and On-VR change with varying load levels under fixed input voltages. The load current is scaled with output voltage. Off-VR can achieve high conversion efficiency at high load levels, but when the load shifts to low levels, Off-VR’s PCE degrades significantly. By contrast, On-VR maintains high PCE at the low load levels,

The work was supported in part by National Basic Research Program of China (973) under grant No. 2011CB302503, in part by National Natural Science Foundation of China (NSFC) under grant No.(61076037, 61376043, 60921002, 61100016) and the Fundamental Research Funds for the Central Universities(KYF-2012-T36).

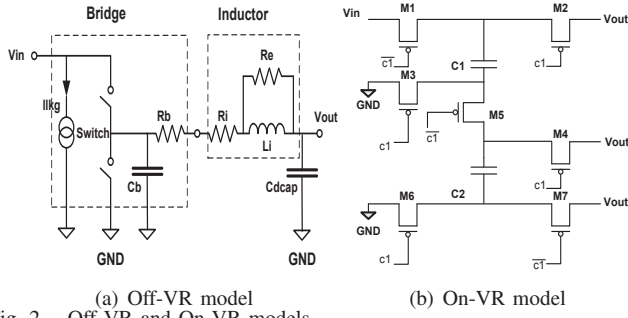


Fig. 2. Off-VR and On-VR models

but cannot feed high load levels. This observation provides an unique opportunity to achieve wide operational range delivery by judiciously building the synergy between Off-VR and On-VR. In particular, we make the following contributions:

- 1) We explore the design space of wide operational range power delivery design. We thoroughly evaluate the feasibility of possible design options, which motivates an optimal design.
- 2) We propose a wide operation range power delivery scheme, called SuperRange, to maximize PCE over the whole range. SuperRange builds the synergy between the Off-VR and On-VR.
- 3) On top of SuperRange, we further propose a VR aware power management algorithm. This algorithm is used to search for optimal processor power states to maximize performance under given power budget.

The rest of the paper is organized as follows. Section II gives the background. Section III analyzes the optional power delivery scheme and Section IV details proposed SuperRange scheme. Section V describes the experiment. We introduce related work in Section VI. Finally Section VII summarizes this paper.

II. BACKGROUND

Voltage regulator is the key component for delivering power to microprocessors. It has always been challenging to achieve high PCE. The widely used VRs are 1) Linear VR and 2) Switch-mode VR.

The most commonly used linear VR is low-dropout regulator (LDO). The voltage regulation is achieved through dropout voltage tuning. LDO VR only provides lower-than-input voltage levels, and its efficiency is limited to the ratio of output to input voltage. Thus LDO VR can achieve high PCE only when the output voltage is close to the input voltage, and linearly degrades with the increasing gap between input and output voltage.

The traditional switch-mode VRs refer to switching inductors VRs. Fig. 2(a) shows a typical Off-VR's model [7]. The voltage conversion is achieved by two parts: the bridge and the inductor. In the bridge part, the two power switches switch on and off asynchronously at a specified switching frequency. This periodical operation generates a square wave of voltage to charge and discharge the inductor. In the inductor part, a low-pass output filter LC is employed to filter the square wave. From Fourier analysis, finally the output voltage value is equal to the average value of the square wave, $V_{out} = D \times V_{in}$, D is duty cycle which is tunable to support different output voltage levels.

This type of VRs is still imperfect in conversion efficiency. It suffers from three kinds of power losses: switching loss, resistive loss in the bridge, and the conductive loss in the inductor. Similar to LDO VRs, it achieves high efficiency at high load level, but degrades fast at low load where the switching loss dominates the total VR power. Moreover, such type of VRs usually is a discrete component and can only reside off-chip due to large form factor.

In recent years, another type of switch-mode VR design, switch capacitor VR, goes increasingly mature. Switch capacitor VR consists of multiple switches and capacitors organized into specific topologies. The regulation is achieved by using capacitors

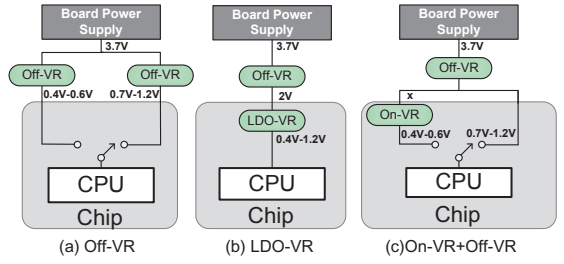


Fig. 3. Top view of optional power delivery designs

as energy storage elements and periodically charging/discharging them. Fig. 2(b) represents a typical switch capacitor VR's model. Basically, it is implemented into a serial-parallel topology, and achieves 3:1 voltage conversion ratio. During charging phase, switch M1, M5, M7 are on, others are off, capacitor C1 and C2 are serially connected and charged, and in discharging phase, switch M1, M5, M7 are off, others are on, capacitor C1 and C2 are parallel connected and discharged. The circuit is periodically switched between two configurations at a specified switch frequency.

Switch capacitor VR can gain high efficiency at a few voltage levels near its target conversion ratio. Compared with inductor mode VRs, switch capacitor VRs can be optimized to deliver power efficiently at low load levels, and they can be built on chip thanks to small physical size and compatible manufacturing process with the host chips. Even though switch capacitor VR lacks the ability to deliver power at all load levels, it's still a promising scheme to complement with other schemes.

III. SUPERRANGE DESIGN SPACE EXPLORATION

In this section, we aim to find an efficient power delivery design towards wide operational range. This design would have the capability to convert the board voltage to a wide output voltage range. In this design, the board voltage is set to 3.7V which follows conventional cases [8]. The output voltage levels range from NTV 0.4V to STV 1.2V [6].

To derive the optimal design is challenging. We therefore first explore the design space and analyze optional designs. These designs are evaluated from two aspects: 1) power conversion efficiency and 2) regulation quality, explained as follows:

The power conversion efficiency (η) of a voltage regulator is defined as the ratio of the loading processor power to the power drawn from power source, that is

$$\eta = \frac{P_{load}}{P_{load} + P_{loss}} = \frac{1}{1 + \frac{P_{loss}}{P_{load}}}. \quad (1)$$

Clearly, we cannot achieve high efficiency if the power loss is comparable to and even overwhelming the load power. The PCE is obtained based on existing Off-VR and On-VR models[9][10].

The regulation quality refers to output voltage ripple. To maintain high voltage integrity, the upper bound of voltage ripple is usually no more than 10%.

A. Off-VR scheme

One intuitive solution is to design multiple Off-VRs whose optimal output points are evenly locate in the operational range, as shown in Fig. 3(a). There are three sources of power loss: 1) switching loss P_{cap} , 2) resistive loss P_{res} in the bridge, and 3) conductive loss P_{ind} in the inductors:

$$P_{loss} = P_{cap} + P_{res} + P_{ind}, \quad (2)$$

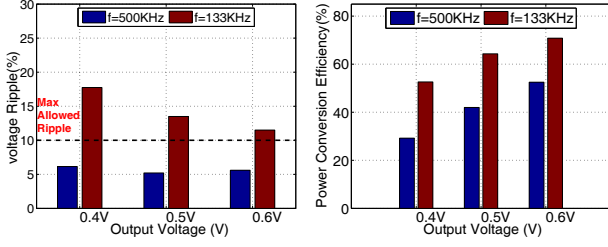
where

$$P_{cap} = C_0 V_{in}^2 f, \quad (3)$$

$$P_{res} = R_c I_{rms}^2 = R_c (I_L^2 + \frac{I_R^2}{3}), \quad (4)$$

$$P_{ind} = R_i I_{rms}^2 = R_i (I_L^2 + \frac{I_R^2}{3}). \quad (5)$$

In this design, we find the main culprit of low efficiency at low load levels is high switching loss P_{cap} , while P_{res} and P_{ind}



(a) Output voltage ripple comparison (b) PCE comparison
Fig. 4. Voltage ripple and efficiency comparison while reducing frequency from 500KHz to 133KHz

are relative small. The most effective approach to reduce P_{cap} is to reduce switching frequency (f). However, decreasing switching frequency will incur new problems: 1) Reduced quality of regulation, i.e. larger output voltage ripple V_R , because $V_R \propto 1/f$; 2) Longer response time. Processor power state transition would take longer, as the switching frequency is reduced; 3) Increased P_{res} and P_{ind} . This is because the effective current I_{rms} and resistance R_i will increase with the reducing frequency.

These tradeoffs can be clearly observed in Fig. 4. The spice simulation and power modeling results show how the voltage ripple and efficiency change with reducing frequency. The initial switch frequency is 500KHz [11] given suggested f is several hundreds of KHz. To shift the range to low end, the f has to be reduced to 133KHz. Unfortunately, the output ripple goes over 10% guard line, but PCE merely increases 10~15%. The power saving is doomed to be offset by the increased load power due to higher voltage to tolerate the large ripple. Moreover, using one more Off-VR further burdens PCB board design and incurs high board area overhead. Hence, Off-VR scheme is not a recommended design.

B. LDO VR scheme

Another solution is to use a large range LDO VR. LDO has the ability to deliver voltage levels below its input voltage. However, the achievable PCE of LDO regulator is limited to the output/input voltages ratio (V_o/V_i) and quiescent current (I_q) as follows:

$$\eta = \frac{I_o V_o}{(I_o + I_q) V_i}. \quad (6)$$

In this analysis, we ignore the quiescent current's impact on PCE, because it is two orders of magnitude smaller than output current I_o . Fig. 3(b) shows this power delivery scheme. A high efficiency Off-VR serves as the frontend and feeds the step-down voltage to a LDO regulator. The LDO regulator can be built on-chip or off-chip, we follow the most conventional design and build it on-chip. In this baseline, LDO delivers power from 2V to voltage from 0.4V to 1.2V. The PCE is limited because of the low ratio of output voltage to input voltage. Especially at near threshold region, the PCE is less than 30%. So using LDO regulator is not an efficient solution to support low load levels.

C. On-VR scheme

We propose to use On-VR to deliver low output voltage levels as Fig. 3(c) shows. To serve NTV efficiently, one should reduce the gap between input and output voltage. An Off-VR first steps 3.7V power source down to an intermediate value, then the On-VR further converts the voltage to 0.4V~0.6V. Because the PCE is the product of PCEs of On-VR and Off-VR, both of them should be taken into careful consideration. The On-VR model follows a high efficiency design [10]. The power loss is shown as follow:

$$P_{loss} = P_{C_{fly}} + P_{R_{sw}} + P_{bott-cap} + P_{gate-cap}, \quad (7)$$

where $P_{C_{fly}}$ is switch capacitor loss, $P_{R_{sw}}$ is switch conductance loss, $P_{bott-cap}$ is parasitic capacitor switching loss and $P_{gate-cap}$ is gate parasitic capacitance switching loss. Fig. 1(b) shows PCE of On-VR when converts voltage from the intermediate level 2V to 0.4V~0.6V. Clearly, it achieves high conversion efficiency at low output voltage levels because it uses lossless low power MOSFETs.

However, On-VR alone lacks the ability to support high voltage range. We therefore need to use the frontend Off-VR as the complement in high voltage range. Specifically, the power delivery steers to Off-VR and shut down On-VR when the loading processor needs STV, or steers to On-VR when NTV is engaged. The on-chip power switches have been well studied [12][13]. In this scheme, STV is readily supported, but NTV is worthy to be further clarified.

The voltage transfer function of On-VR can be formulated as follows [10]:

$$V_o = \alpha V_{in} - I_o R_i, \text{ where } R_i \propto 1/f. \quad (8)$$

In this equation, V_o is output voltage, V_{in} is input voltage, R_i is inner resistance, f is switching frequency and α is a topology specific parameter which defines the conversion ratio which is constant given an On-VR design. According to the equation, the output voltage can be modulated by 1) tuning resistance of On-VR by changing f and holding the V_{in} , or 2) tuning input voltage produced by Off-VR and keeping the resistance constant. However, these two approaches are not equivalent to each other, and we find the second one is preferred. The reason is explained as follows:

The first way is to adjust operation state of On-VR fed with fixed input voltage from frontend Off-VR. The voltage is firstly converted from 3.7V to a lower value 2V by Off-VR. Its PCE can be optimized to 80%~87%. Then On-VR converts the voltage from fixed voltage 2V to 0.6V. Further voltage scaling, i.e. 0.4V~0.6V, is realized by tuning the switching frequency of On-VR. This design can regulate the voltage to near threshold region with high regulation quality, but the PCE of On-VR suffers from the decreasing switching frequency. The optimal operation point of On-VR is limited in a narrow range near 0.6V. Simulation result shows this design has an average efficiency 58% to support NTV. Although it's 20% higher than conventional design, the efficiency is still very low at the lowest 0.4V output voltage.

Another alternative is to tune the Off-VR state to generate a variable voltage V_{in} and use fixed On-VR to step this intermediate voltage to NTV. For example, the Off-VR first changes the duty cycle and steps the source voltage to a value between 1.3V and 2V, then the voltage value is converted to 0.4V~0.6V range by On-VR. Because On-VR doesn't need to decrease its switching frequency, the PCE can stay around 90%. However, given $V_o I_o \approx 90\% \times I_{in} V_{in}$ in On-VR, the output current of Off-VR, I_{in} , is very low, which renders the the Off-VR less efficient. In our case, the efficiency drops to 44% on average. Although such solution is still not optimal, it sheds light on the way to further optimization: *improve the efficiency of frontend Off-VR under low current mode*, which is the key design consideration in SuperRange scheme.

IV. THE PROPOSED SUPER RANGE SCHEME

In this section, we use a multi-phase Off-VR (4 phases in this paper) as the frontend to build SuperRange. This design solves the low efficiency problem of Off-VR at low current condition when feeding the On-VR.

A. The Proposed Power delivery design

Fig. 5 shows the top view of the SuperRange scheme and loading processor. The processor working voltage covers NTV and STV. The STV levels are directly provided by 4-phase Off-VR, while the NTV levels are produced by On-VR which uses the Off-VR as its frontend. The On-VR is a serial-parallel single topology VR with 3:1 conversion ratio. The Off-VR is a conventional Off-VR, similar to commercial regulator LTC3733 [11].

The multi-phase Off-VR provides an opportunity to improve the efficiency of frontend Off-VR under low current mode. Multi-phase technology is commonly used to increase the load current. Each phase delivers one equal portion of total load current. Modern Off-VR has the capability to dynamically change the number of working phases [14]. Therefore some phases can be disabled to adapt to the low current mode, without degrading PCE. In our design, when delivering to NTV, we find that a

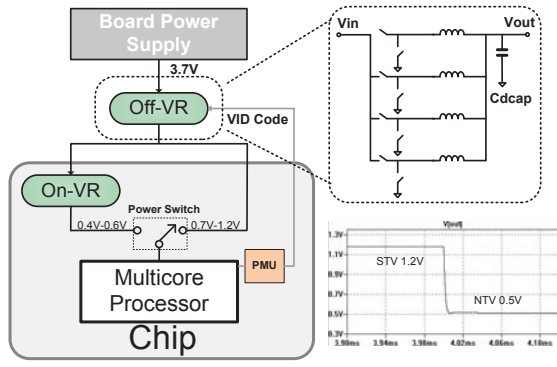


Fig. 5. Implementation of proposed power delivery

single working phase can afford the load, so other phases are disabled. The PCE increases over 70%. Note that reducing number of working phase inevitably incurs larger ripple, but the increased current ripple can be easily removed by using a relative large inductance inductor. The increased inductor size will not burden board design since we use only one Off-VR in this design. In our case, we find a 1.5uH inductor is competent to ease the extra ripple which is confirmed in Fig. 6. An inductor with inductances larger than 1.5uH can reduce the ripple to a level lower than 7%. Thus we can dynamically reduce the number of working phase when more load current is required.

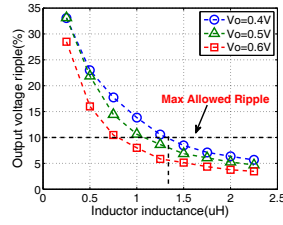


Fig. 6. Output ripple with varying inductor size

Basically, SuperRange has two working modes:

1) Supporting STV: Voltage conversion to STV is performed by Off-VR. The Off-VR receives and decodes the VID, and power delivery is directly steered to Off-VR. It changes output voltage based on VID.

2) Supporting NTV: The support to NTV region is achieved by a two step conversion. Off-VR first decodes VID and power delivery is steered to On-VR. Then the Off-VR sets to single working phase, and the output voltage is initially regulated to a variable intermediate level based on VID. In our case, the intermediate level is 1.3V, 1.6V or 2.0V with carefully consideration of On-VR's resistance. After this, the On-VR steps the voltage to corresponding near threshold value.

By doing so, the proposed scheme can provide high conversion efficiency over the entire load spectrum.

B. A VR-Aware Power Management Algorithm

Supply voltage scaling decision is made by on-chip power management unit (PMU). PMU is a micro-controller and runs power management routine which makes decisions on core power states (voltage and frequency) and active core count selections. PMU collects power data from on-chip current sensors and performance statistics from performance counters. Power management routine then uses data collected as input and selects supply voltage level to achieve certain power management goals. Once PMU makes a decision on the selection of supply voltage level, the PMU sends the corresponding VID (Voltage Identification) code to Off-VR. The detailed PMU design is beyond the scope of this paper.

To demonstrate the potential of SuperRange, we present a VR-aware power management algorithm employed by PMU. The goal of this algorithm is to maximize system performance under given power budget. It's worth to note that SuperRange can also be applied to other management algorithms, not limited to this one.

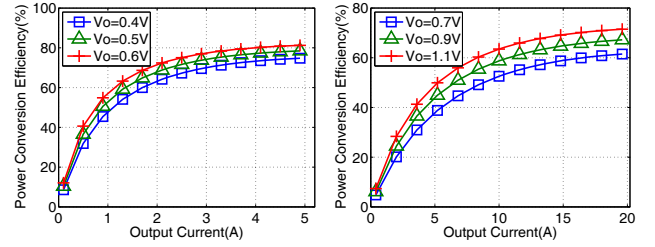


Fig. 7. Per-Phase efficiency at different load conditions

We first exploit the efficiency of SuperRange at different load conditions as Fig. 7 shows. Fig. 7(a) and (b) shows the PCE when supporting NTV and STV levels, respectively. Clearly, the efficiency of the SuperRange strongly correlates with output voltage and current. Usually excessively low current cannot yield high efficiency, hence should be avoided as much as possible.

Algorithm 1 Optimal Configuration Search under SuperRange

Input: Core power $P_{v_{all}}$, Performance B_v at voltage level v , Processor core count N_{all} , Power budget P_B ; SuperRange PCE η_v
Output: Output voltage level V_o ; Active core number N_o

- 1: for each v do
- 2: $P_v = \frac{P_{v_{all}}}{\eta_v}$ //total power when all core active
- 3: end for
- 4: Select highest voltage v_H at which P_v is smaller than P_B , and get B_{v_H} ;
- 5: Select lowest voltage v_L at which P_v is higher than P_B ;
- 6: for $i = N_{all}; i > 0; i --$ do
- 7: $P = \frac{P_{v_L}}{i}$; // P_{v_L} is obtained by current sensor
- 8: Collect B_i through performance counter;
- 9: if $P \leq P_B$ then
- 10: $B_{v_L} = B_i$;
- 11: break;
- 12: end if
- 13: end for
- 14: $(V_o, N_o) = \max(B_{v_H}, B_{v_L})$; //final configuration gives higher Bips

Then we consider system power efficiency. The power efficiency is defined as the ratio of performance (billion instruction per second) to total power, a.k.a. BIPS/Watt. Without considering the imperfect PCE, the power efficiency increases quadratically with lower voltage. This is because core power consumption reduces cubically while the performance degrades linearly with lower voltage. However, there is a tradeoff when taking the imperfect PCE into account. Although high voltage benefits high PCE, it degrades the application power efficiency more. The following algorithm is used to tackle this tradeoff and the goal is to maximize the performance under power budget.

The load condition, the efficiency of power delivery system and the application performance serves as inputs to the algorithm. The load condition and performance can be measured through on-chip current sensor and performance counter. We assume a homogeneous multi-core processor, but the basic principle is also applicable to heterogenous processors which is supposed to be future work.

The problem can be formulated as follows:

- Given: 1) Application power and performance in different voltage levels; 2) SuperRange PCE under different load conditions; 3) Processor core count; 4) Power budget.

- Determine: the number of active cores and the supply voltage level to maximize performance under power budget.

The problem is solved with Algorithm 1. Basically, it's a two-step process:

Step 1: The algorithm first computes the total power P_v when all cores are active at each voltage level. The power is the sum of measured core power and VR power loss. Then it compares the power values with given power budget, then selects the lowest level

TABLE I. BASELINE ARCHITECTURE CONFIGURATIONS

Parameter	Value
Core number	16
Power delivery	Hybrid
LLC capacity	32MB
LLC feature	32B block, 8-way
Cache coherence	distribute directory-based MESI
On chip interconnection	mesh + router
Memory controller	1
Memory bandwidth	10Gb/s
Area (mm^2)	253.03

TABLE II. VOLTAGE REGULATOR PARAMETERS

Parameter	Off-VR	On-VR
Topology	Buck	Switch Cap
V_{in}	3.7	1.2-2.0
V_{out}	0.7-2.0	0.4-0.7
Freq (Mhz)	0.5	300
No. of phases	4	20
L per ph (uH)	1.5	N/A
intrinsic resistance (mohm)	32	56
C_{fly} (nF)	N/A	20
Area (mm^2)	3.8	0.084

V_L under which P_v is higher than P_B and highest level V_H under which P_v is lower than P_B . The optimal voltage setting will be chosen between the two voltage levels.

Step 2: The algorithm then calculates the maximum number of active core at V_L which satisfies power budget, and get the corresponding performance. Then it compares the performance with the performance B_{V_H} at V_H , and picks the configuration which gives the higher performance.

V. EXPERIMENT

A. Experimental Setup

Baseline Architecture: The baseline is a multi-core processor consisting of 16 OoO cores, though SuperRange essentially can support any type of architectures. The baseline is deployed on a 32MB last level cache (LLC). The LLC is organized in eight banks; each bank is 8-way associative. Distributed directory based MESI cache coherence is enabled to maintain the shared data consistency. The processor connects to main memory through a DDR3-1333 memory controller with 10.6Gb/s bandwidth. Detailed information is provided in Table I.

Voltage Regulators: The efficiency model of off-chip inductor based VR is based on converter model [9], and inductor and switches parameters are derived from LTC3733's datasheet [11]. The switch capacitor VR follows an existing high efficiency model [10], and use a 3:1 serial-parallel topology voltage regulator. We also build their LTspice models to simulate their behavior. Detailed configurations including VR parameters and area cost are listed in Table II. The area overhead of On-VR is $0.084mm^2$, which is less than 0.5% area of the baseline processor.

Simulation Infrastructure: We use a full-system simulator gem5 [15] as our simulation infrastructure. In addition, McPAT [16] is plugged in gem5 to evaluate power consumption and area overhead. The technology is configured to 32nm technology node. We extend McPAT to calculate power in different performance states. We build the Out-of-Order (OoO) core resembling Alpha-EV6. The core configuration is detailed in Table III. The cores have nine performance states, each state corresponds to a specified voltage/frequency setting: Pstate1 (1.2v, 1.9GHz), Pstate2 (1.1v, 1.7GHz), Pstate3 (1.0v, 1.5GHz), Pstate4 (0.9v, 1.3GHz), Pstate5 (0.8v, 1.1GHz), Pstate6 (0.7v, 0.9GHz), Pstate7 (0.6v, 0.7GHz), Pstate8 (0.5v, 0.5GHz), Pstate9 (0.4v, 0.3GHz).

Benchmarks: We use Parsec benchmark suite [17] to evaluate our design, because it targets the general purpose processors and aims to represent emerging workloads in the near future.

Power Domains: The cores are powered by SuperRange scheme. For other power consuming components, we follow Intel Nehalem family processor design [18], i.e. the LLC and memory

TABLE III. MICROARCHITECTURAL PARAMETERS

Parameter	Value
Processor	Out-of-Order core
Fetch/Decode Width	4
Branch-Predictor Type	Hybrid 2-Level
Reorder Buffer Size	128
Unified Load/Store	64
FP/INT Register File Size	64/64
Supply Voltage (V)	0.4:0.1:1.2
Clock rate (GHz)	0.3:0.2:1.9
Inst/Data cache	32kB, 2-way
Access Time	2ns
Miss Penalty	12ns
Area (mm^2)	13.65

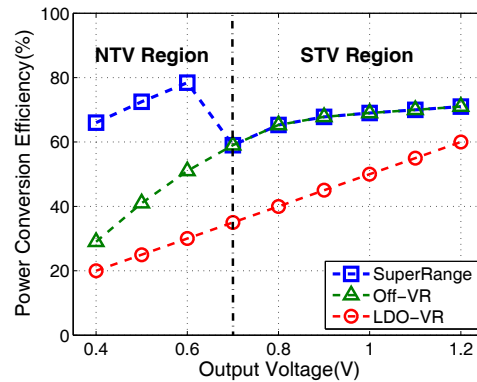


Fig. 8. Power conversion efficiency over the entire operational range

controllers are powered by their own off-chip VR. Because LLC and memory controller contribute to fixed and relative small portion of system power, our work focuses on core power delivery.

B. Experimental Results

First, we show the overall power conversion efficiency of SuperRange in Fig. 8. From the results, we can see that PCE at NTV increases by 40% compared with conventional Off-VR design and the average PCE over entire operational range can be nearly 70%. This is because SuperRange not only implements the synergy between Off-VR and On-VR to improve PCE, but also propose to optimize the PCE of frontend Off-VR while maintaining On-VR a high PCE (around 90%) at NTV levels. Since our scheme fully explores the efficiency benefit of both Off-VR and On-VR, we can conclude that the wide operational range now can be well supported.

Second, we study the performance potential under a constant power budget. To simulate a power-constrained system, the highest power budget is able to power up eight cores at Pstate1, the highest performance state, and half of the cores have to stay in dark. If more active cores turn out to yield higher performance, we need to reduce the voltage to enable lower performance states, at the risk of large power loss in the delivery. Fig. 9 shows the maximum performance delivered by the target processor under three power delivery schemes (as shown in Fig. 3) with 25% of highest power budget. The results show that the proposed SuperRange obviously outperforms the other two schemes. This is because with the highest conversion efficiency over the large voltage range, the processor has more flexibility to tune between the single core performance and multiple cores parallelism. This flexibility provides the applications more opportunities to successfully enable the optimal multi-threading configurations. By contrast, the LDO-VR and Off-VR lead to less flexibility because the efficiency loss during the low voltage range may exclude some multi-threading configurations which turn out to be performance optimal if without such efficiency loss.

The proposed SuperRange scheme not only outperforms LDO-VR and Off-VR at the constant power budget, but also shows superior resilience on even tighter power constraints. Fig. 10 shows the normalized performance of the target processor under shrinking

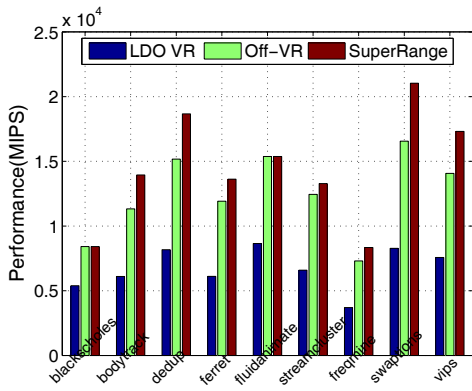


Fig. 9. Performance comparison in power-constrained system

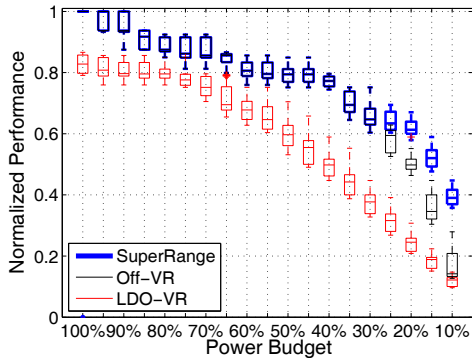


Fig. 10. Performance under increasingly tighter power constraint

power budget. X axis sets the system power budget, Y axis shows the corresponding maximum achievable performance under three delivery schemes. Each box shows distribution of the performance running Parsec Benchmarks. The results show that at loose power budget, SuperRange behaves as well as Off-VR scheme. When the budget gets tighter, the system prefers NTV levels. The result shows the achievable performance under SuperRange is far higher than ones in LDO-VR and Off-VR by 171% and 52% on average, correspondingly. The result demonstrates that SuperRange is a more promising solution in the dark silicon era.

VI. RELATED WORK

Design for high efficiency power delivery has been a hot topic for years. Yan et al. presented a hybrid power delivery scheme to explore different power phases, but they didn't take the varying PCE into consideration [19]. Ng et al. and Le et al. proposed a high efficiency switched-capacitor DC-DC converter. Ng's design aims at large input voltage range, but the output voltage level is limited [20]. Le's design implements a multiple topology converter with increased design complexity and on-chip area overhead. It's unsuitable for powering microprocessors [10]. Kim et al. designed a integrated 3-level DC-DC converter, their design targets at fast DVFS and doesn't have a high efficiency at low load level [21]. All the designs stated above lack the ability to support a wide operational range efficiently.

Another line of prior work considers system level strategies. Cho et al. proposed a system level power management method based on dynamic voltage regulator scheduling to achieve high conversion efficiency over a large power range [13]. They choose the most efficient voltage regulator with changing load levels. Sinkar et al. proposed a workload aware voltage regulator configuration tuning method to optimize system power [22]. Amelifard et al. introduced a reconfigurable power delivery network design by dynamically changing the topology of voltage regulators to support different load levels [23]. Ghasemi et al. implemented per-core voltage domains by using on-chip low dropout converters, their

method imposes low area overhead since the converters share components with power gating circuit [24]. However, the LDO converter has a low conversion efficiency at low load levels. Differing from these designs, our work addresses the efficiency problem through a hybrid power delivery scheme, more importantly, highlight wide operation range which is missed in prior work.

VII. CONCLUSION

In this paper we propose a large operational range power delivery scheme, SuperRange, by exploring the advantages of both on-chip and off-chip VRs. We thoroughly analysis the efficiency behavior of existing VRs, and ensure that SuperRange can provide high power conversion efficiency over the entire voltage range. Moreover, We propose a VR aware power management algorithm. This algorithm heuristically finds optimal processor configurations (active core numbers and VF setting) to maximize performance under given power budget. Experimental results show that SuperRange is well suitable to support wide power range for future power-constrained systems.

REFERENCES

- [1] H. Esmailzadeh, E. Blem, R.S. Amant, K. Sankaralingam, D. Burger, "Dark silicon and the end of multicore scaling," *JSCA*, 2011.
- [2] Intel, "Enhanced Intel SpeedStep Technology for the Intel Pentium Processor," *White Paper*, 2004.
- [3] E. Rotem, A. Naveh, D. Rajwan et.al, "Power management architecture of the 2nd generation Intel Core microarchitecture, formerly codenamed Sandy Bridge," *Hot Chips*, 2011.
- [4] R. G. Dreslinski, M. Wiecekowsk, D. Blaauw et.al, "Near-Threshold Computing: Reclaiming Moore's Law Through Energy Efficient Integrated Circuits," *Proceedings of the IEEE*, vol. 98, no. 2, pp. 253-266, 2010.
- [5] J. Kim, M. A. Horowitz, "An efficient digital sliding controller for adaptive power-supply regulation," *JSSC*, vol. 37, no. 5, pp. 639-649, 2002.
- [6] S.Jain et.al, "A 280mV-to-1.2V Wide-Operating-Range IA-32 processor in 32nm CMOS," *ISSCC*, pp. 66-68, 2012.
- [7] G. Schrom, P. Hazucha, F. Paillet, D. S. Gardner, S. T. Moon, T. Karnik, "Optimal Design of Monolithic Integrated DC-DC Converters," *ICICDT*, 2006.
- [8] W.Kim, M.S. Gupta, G.Y.Wei, D.Brooks, "System level analysis of fast, per-core DVFS using on-chip switching regulators," *HPCA*, 2008.
- [9] Y.Choi, N.Chang, T.Kim, "DC-DC Converter-Aware Power Management for Low-Power Embedded Systems," *TCAD*, pp. 1367-1381, 2007.
- [10] H. P. Le, S. R. Sanders, E. Alon, "Design Techniques for Fully Integrated Switched-Capacitor DC-DC Converters," *JSSC*, vol. 46, no. 9, pp. 2120-2131, 2011.
- [11] Linear Technology, "LTC3733: 3-Phase, Buck Controllers for AMD CPUs," *Datasheet*.
- [12] B.Amelifard, M. Pedram, "Design of an Efficient Power Delivery Network in an SoC," *ISLPED*, 2007.
- [13] Y. Cho, Y. Kim, Y. Joo, K. Lee, N. Chang, "Simultaneous optimization of battery-aware voltage regulator scheduling with dynamic voltage and frequency scaling," *ISLPED*, pp. 309-314, 2008.
- [14] P. Zumel, C. Ferañdez, A. de Castro, O. Garcia, "Efficiency improvement in multiphase converter by changing dynamically the number of phases," *PESC*, 2006.
- [15] gem5, "The gem5 Simulator System," <http://www.m5sim.org>, 2012.
- [16] S. Li, J.H. Ahn, R.D. Strong, J.B. Brockman, D.M. Tullsen and N.P. Jouppi, "McPAT: an integrated power, area, and timing modeling framework for multicore and manycore architectures," *Micro*, pp. 469-480, 2009.
- [17] C. Bienia, S. Kumar, J.P. Singh and K. Li, "The PARSEC Benchmark Suite: Characterization and Architectural Implications," *PACT*, pp. 72-81, 2008.
- [18] Intel, "Energy-Efficient Computing: Power Management System On The Nehalem Family Of Processors," *Intel Technology Journal*, vol. 14, no. 3, 2010.
- [19] G. Yan, Y. Li, Y. Han, X. Li, M. Guo, X. Liang, "AgileRegulator: A Hybrid Voltage Regulator Scheme Redeeming Dark Silicon for Power Efficiency in a Multicore Architecture," *HPCA*, 2012.
- [20] V. Ng, S. Sanders, "A 92%-Efficiency Wide-Input-Voltage-Range Switched-Capacitor DC-DC Converter," *ISSCC*, pp. 282-284, 2012.
- [21] W. Kim, D. Brooks, G. Y. Wei, "A Fully-Integrated 3-Level DC-DC Converter for Nanosecond-Scale DVFS," *JSSC*, vol. 47, no. 1, pp. 206-219, 2012.
- [22] A. A. Sinkar, H. Wang, N. S. Kim, "Workload-aware voltage regulator optimization for power efficient multi-core processors," *DATÉ*, pp. 1134-1137, 2009.
- [23] B. Amelifard, M. Pedram, "Optimal selection of voltage regulator modules in a power delivery network," *DAC*, pp. 168-173, 2007.
- [24] H. R. Ghasemi, A. A. Sinkar, M. J. Schulte, N. S. Kim, "Cost-effective power delivery to support per-core voltage domains for power-constrained processors," *DAC*, pp. 56-61, 2012.