

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

An Ageing-Aware and Temperature Mapping Algorithm For Multi-Level Cache Nodes

EMMANUEL OFORI-ATTAH, MICHAEL OPOKU AGYEMAN (Senior Member, IEEE)

Faculty Of Science and Technology Technology, University of Northampton, UK (e-mails: Emmanuel.Ofori-Attah@northampton.ac.uk; Michael.OpokuAgyeman@northampton.ac.uk)

Emmanuel Ofori-Attah (e-mail: Emmanuel.Ofori-Attah@northampton.ac.uk).

ABSTRACT Increase in chip inactivity in the future threatens the performance of many-core systems and therefore, efficient techniques are required for continuous scaling of transistors. As a result of this challenge, future proposed many-core system designs must consider the possibility of a 50% functioning chip per time as well maintaining performance. Fortunately, this 50% inactivity can be increased by managing the temperature of active nodes and the placement of the dark nodes to leverage a balance working chip whilst considering the lifetime of nodes. However, allocating dark nodes inefficiently can increase the temperature of the chip and increase the waiting time of applications. Consequently, due to stochastic application characteristics, a dynamic rescheduling technique is more desirable compared to fixed design mapping. In this paper, we propose an Ageing Before Temperature Electromigration-Aware, Negative Bias Temperature Instability (NBTI) & Time-dependent Dielectric Breakdown (TDDB) Neighbour Allocation (ABENA 2.0), a dynamic rescheduling management system which considers the ageing and temperature before mapping applications. ABENA also considers the location of active and dark nodes and migrate task based on the characteristics of the nodes. Our proposed algorithm employ Dynamic Voltage Frequency Scaling (DVFS) to reduce the Voltage and Frequency (VF) of the nodes. Results show that, our proposed methods improve the ageing of nodes compared to a conventional round-robin management system by 10% in temperature, and 10% ageing.

INDEX TERMS Dark-Silicon, Power Management, Run-time Mapping, Many-Core Systems, Dynamic Voltage Frequency Scaling (DVFS), Task Migration, Power States, Dark Core

I. INTRODUCTION

EXCESSIVE thermal rise in many-core systems caused by uncontrolled leakage power as transistor size decreases leads to the Dark-Silicon challenge, where a section of many-core chips have to powered-off or under-clocked to function within a power budget; whilst considering the temperature threshold as well as meeting application demands [1]–[4]. Thermal Design Power (TDP) and Thermal Safe Power (TSP) [5] are examples of power budgeting techniques used to prevent chip damage with the latter offering more power for resource activity. Consequently, during run-time, DVFS, Power Gating (PG), Task Migration (TM) and Near Threshold Computing (NTC) are used to control the power.

However, inefficient exploitation of Dark-Silicon can lead to a poverty-stricken performance system. This is because, different resource activation patterns result in different ther-

mal profiles and thus, presents an opportunity for performance optimization. An optimized activation of resources can determine the maximum allowable power budget.

Contiguous resource activation offers better application performance however, the heat produced from active working components leads to thermal hotspots and limits the number of activated resources. This leads to more power being consumed. Additionally, this can also threaten the reliability and lifetime of the chip.

On the other hand, non-contiguous resource activation can degrade the performance of the system without respect for applications which require inter communication. As a matter of fact, the communication overhead caused by the distance between nodes which require inter communication would require processors running faster which eventually leads to more power being consumed [6]. Therefore, to leverage the

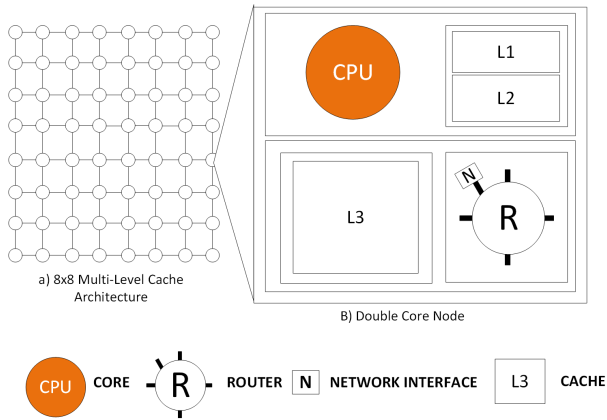


FIGURE 1: Multi-Level-Cache Architecture

disparity between performance and temperature, the non-active nodes, which we refer to as "dark" or "dumso" interchangeably are placed in-between active nodes. Over the years, various techniques have been proposed which exploit the use of dark nodes to improve the performance of active cores. An active node can function at a higher frequency when dark nodes are placed near it. However, over use of dark nodes can lead to the starvation of newly arrive applications as a result of insufficient nodes in a large management systems [7].

Moreover, this not only causes a performance challenge, but it also introduces expensive packaging and cooling costs coupled with reliability and ageing of resources with the latter being vastly ignored by many proposed techniques. As a matter fact, ITRS predict that as technology scales further down, future chips will age faster [8].

To overcome these challenges, several design aspects have been proposed in literature [9]–[13], [13], [14]. One design aspect which can be exploited in different ways is architectural heterogeneity: whether this being the implementation of different sizes of resources as depicted in [15] or incorporating CPU/GPU [16], [17]. Yang et al. [15] proposed the Quad-core cluster. The Quad-core cluster consists of four different types of cores (High Performance, General Purpose, Power Saving, Low Energy). Based on the application demands, the appropriate core is selected for computation. Additionally, DVFS and TM are employed to scale down VF and migrate tasks to cores at run-time [18]–[20]

Another technique widely used is the efficient use of management systems [21], [22]. Others on the other hand, reduce power consumption in Network-on-Chip (NoC) Components [23]–[27] for an optimized performance. Fig. 1 depicts a Multi-Level Cache Node comprised of core, router and a Multi-Level Cache Architecture (MCA). Although MCA is mentioned in this work, we only address the power consumption of the nodes. We do not extensively address power related issues as this will be done in future work. In this work, we seek to improve the temperature and lifetime of nodes by using the lifetime of nodes as the main factor to allocate ap-

plications and tasks. The highlights of the proposed scheme are as follows:

This paper addresses the challenges by proposing a Dark-Silicon patterning approach that handles repeated workloads where the temperature and lifetime of nodes are considered. The applications and tasks are assumed as fixed. The main contributions of this paper can be summarised as:

- We present two Dark-Silicon algorithms to improve the lifetime reliability of many-core systems.
- The first algorithm improves the lifetime of nodes by migrating tasks from hot nodes to dark nodes at run-time.
- The second algorithm monitors and improve the lifetime of nodes by migrate tasks after an epoch. Additionally, the algorithm monitors the Lifetime of two neighbouring nodes and then selects the node with the highest MTTF to improve the Lifetime.
- Additionally, the proposed method reduces temperature by ensuring an active node is surrounded by dark nodes to allow it to function at higher frequency.
- The proposed approach is compared with the conventional round-robin against the following parameters: Temperature, Hotspot, Lifetime and Utilization.
- Results from the experiment conducted shows how optimised the proposed method is.

The paper is organized as follows: Section II briefly discusses related work about heterogeneous nodes, dynamic application mapping which considers computation and communication intensive application. Section III presents an observation into the round-robin mapping algorithm and in Section IV, both proposed approaches are presented. Section V presents the experimental results. Finally, Section VI concludes the paper and discusses future work.

II. RELATED WORK

The allocation of resources in many-core systems have been the focus of technology since the emergence of the Dark-Silicon phenomenon. However, [16] states that, previous task-resource allocation only considers mapping applications contiguously without regard for generated heat amongst resources. Active components generate heat and over time spread among neighbouring resources causing thermal hotspot. Consequently, over time, this affects the lifetime reliability of the resources. This causes accelerated EM, NBTI & TDDB and leads to function slower than the other [28].

Work that considers heat dissipation does not emphasise the lifetime reliability of systems and count node utilization as a crucial metric when optimising performance.

Xiaohang Wang et al. [7] proposed a virtual mapping algorithm to estimate the number of dark cores required for an application to efficiently execute. The aim of the virtual mapping algorithm is to prevent inefficient use of dark cores to ensure there are enough free cores for incoming applications. This algorithm considers communication and

computation applications. However, node utilization is not considered as a metric when choosing the first node.

Kanduri et al. [29] presented adBoost, a thermal aware performance boosting system which boost the performance of active cores by the efficient use of dark cores. The algorithm employed in this system maps applications spatially to avoid hot spot. First Node selection is selected based on a finding a node that is far away using MapPro from an active node with sufficient nodes around it for mapping of application task. Unfortunately, node utilization is not used as a property when selecting the first node.

The following work, however, addresses this challenge by considering computation and communication demands when activating resources.

Reza et al. [16] proposed a resource management system for heterogeneous NoCs which examines the performance requirement of an application (Communication or Computation Intensive) before distributing its tasks amongst the appropriate resources. In this management system, the chip is partitioned into clusters with a set of CPUs and GPUs. The architecture present employs a CPU from each cluster and assigns them as Cluster managers to monitor and configure the resources within that cluster. This information is then feed-backed to a global manager which uses the BalancedMap mapping application algorithm to activate the feasible resources based on the application. The BalancedMap algorithm maps communication intensive applications to CPUs and computation intensive application to GPUs.

The global manager is assigned by calculating the CPU with the shortest distance from all the assigned Cluster Managers. The BalancedMap Mapping Algorithm works by classifying the application into two groups (Communication and Communication Intensive). During CPU and GPU node selection, the node with the lowest peak energy is selected. If more than one node appears to meet the demand, the node with the lowest utilization is selected to improve the thermal hotspot. Power consumption is minimized by configuring the link, node voltage and the shutting down of idle routers when links are not carrying any communication traffic. TM is employed when a node being utilized exceeds it warning threshold.

Rahmani et al. [30] considers reliability in Dark-Silicon systems by employing reliability-aware power allocator to monitor the ageing of nodes and to reduce the amount of workload on stressed out regions that are experiencing fast ageing. Consequently, the voltage/Frequency of the nodes are scaled down to prevent the nodes from functioning at full throttle.

Rathore et al. [31] on the other hand, proposed a HiMap, a hierarchical mapping approach which reduces the lifetime of nodes by mapping application to healthier nodes whilst also placing dark nodes amongst the selected region. This mapping approach considers Process Variation when assessing the ageing and reliability of the node. The assessment is done by checking nodes for process variation, temperature, and ageing. Weaker nodes are then used as dark nodes.

Mohammed et al. [32] proposed a technique which uses both TS and DVFS for an optimise system performance. TM is used to move cores from active to dark cores. In this architecture, the dark and active cores are all running concurrently and therefore makes it easier for task to be swap amongst them. Unfortunately, this increases the temperature and thus, DVFS is applied.

Rathore et al. [33] proposed a technique to optimise the lifetime reliability for many-core systems. Similar to our proposed technique, they employ clusters. Although their proposed system maps tasks to the node, hotspot can still occur because of mapping tasks to high MTF nodes in the same area. Alternatively, Ansari et al. [34] proposed a Thermal-Aware Standby-Sparing Technique (TASS) in a heterogeneous multi-core system. The multi-core system consists of two pairs of cores: High Performance Core (HP) and Low Performance Core (LP). The main tasks of applications are scheduled on the HP whilst the backup tasks are scheduled on the back up cores. Unfortunately, there is no mention of high temperature between active components which are next to each other. This can still hotspot. In our purposed method, we compare between neighbours to reduce hotspot.

In our previous work [35], we presented architectural saving techniques to improve the power efficiency for many-core systems. Based on the study conducted, it can be deduced the fraction of powered-off nodes in Dark-Silicon chips can be improved by power budgeting, architectural heterogeneity, NoC interconnect, Cache Memory, and run-time management. Particularly, resource allocation using application mapping through run-time management methods.

Similar to the above-mentioned techniques, we approach the Dark-Silicon challenge by permitting only 50% of the nodes available to function at full through per time. However, unlike previous proposed approaches that statically and randomly generate dark and active nodes, our proposed algorithm gracefully selects dark and active nodes based on their current lifetime and temperature.

III. PROBLEM STATEMENT

At any given time, A number of applications dynamically arrive in the system. Each application spawns several T threads. $\{A_0, A_1, A_2, A_3, \dots, A_{A-1}\}$. The objective is to allocate each application to several N nodes $\{N_0, N_1, N_2, N_3, \dots, N_{N-1}\}$ to meet the application deadline without aggravating the temperature of the chip. Fortunately, the overall temperature depends on several factors: The number of actively working nodes, the position and the voltage frequency of nodes also have an impact. Therefore, the heat generated can be model as:

$$H = \sum \mathcal{P}^i \times \tau^{amb} \times \tau^i$$

Where \mathcal{P}^i is the power dissipated by the node, τ^{amb} is the temperature of neighbouring nodes and τ^i is the ambient temperature.

Contiguous mapping techniques spawn application threads next to each other without considering the temperature and heat generated by neighbouring nodes. The heat dissipated by active nodes haphazardly affects neighbouring nodes and pushes nodes to reach their critical temperatures even when they are operating under low V/F. Pattern 1 in Fig.2 is an example of such an approach. Non-contiguous mapping avoids this challenge by dispersing application and threads across the chip. Unfortunately, dispersing tasks across the chip increases the communication latency between threads which require inter communication and thus, increases the power consumption and temperature.

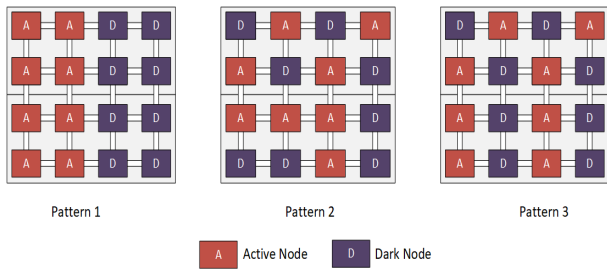


FIGURE 2: Dark-Silicon Patterns

A naive mapping of two applications can cause the temperature of the chip to be aggravated. Fig. 2 depicts three different patterns. Pattern 1 is likely to result in a high temperature chip because all the nodes that have been selected for mapping are in the same region. Additionally, incoming applications will be continuously mapped to the same active nodes. Pattern 2 and 3 on the other hand have dispersed applications across the chip. This offers low temperature; however, consideration of node performance must be considered when making such a decision to prevent bottlenecks. This can either be, the temperature of the node, the lifetime of the node, the location of the node or the effect that a mapping of a thread to a node can have on an incoming application.

Therefore, different mapping approaches results in various thermal profile of the chip. An optimal approach can allow up to 10 applications to be mapped without aggravating the temperature of the many-core system. Hence, our proposed approach.

IV. PROPOSED APPROACH

This section presents the proposed technique. The proposed technique aims to optimise the performance of the many-core systems under several thermal constraints. Due to modern processors supporting several low-power states termed C-states (C0, C1, C2....) [36], we employ DVFS to manage the frequencies and to shut down parts of the nodes based on their thermal profile.

Fig. 3 shows the overall architecture of the considered system. The platform contains a set of homogeneous processing tiles organised in a 2-D mesh-based topology and connected by an NoC.

A. PROPOSED FRAMEWORK

The proposed many-core platform is partitioned into several clusters of equal sizes in an $N_X \times N_X$ grid of MCA tiles.

The clusters are formed by specifying the number of nodes in the system, the type of mesh required, the application with the highest number of tasks and the number of tasks. 3 shows a 64-node system with 4 X 4 sized clusters. To determine the number of nodes per cluster in a 64-node system, the application with the highest number of tasks is considered. If an application has 8 tasks, then four clusters will be created with 16 nodes (8 active and 8 dark). This is because ABENA employs the 50% dumso rule dark-silicon.

Each cluster has been allocated a module to monitor the temperature and lifetime of nodes and to utilise DTM techniques to the nodes accordingly. Each module then returns the average temperature and lifetime of its cluster to a centralised resource manager. Based on the information collected, applications are assigned to the right cluster. Additionally each node has per-core DVFS capability [37].

1) Proposed Algorithm

Majority of dynamic mapping applications do not consider ageing even though some nodes age faster than others. To improve performance, temperature and power budget are usually used as measurements or parameters for enabling and disabling nodes. Over time, permanent faults caused by Electromigration (EM), Negative Dependent Temperature Instability (NBTI), and Time Dependent Dielectric Breakdown (TDDB) reduces the lifetime and increases the systems Mean Time To Failure (MTTF).

Therefore, efficient techniques are required to improve the lifetime of nodes as well maintaining the performance. Hence why we proposed a new method which uses the MTTF as main option for choosing dumso and active nodes. The proposed approach is not limited to just one phenomenon, it can be applied to other ageing mechanisms. However, for the purpose of this study, we consider TDDB. The MTTF of a node or system is an immediate value generated under existing circumstances. This can change depending on the circumstance, parameters, and application characteristics. In contrast, Aging is accumulated over a period of time.

The proposed algorithm (ABENA) is detailed in Algorithm 1. ABENA Monitors the following parameters at runtime: the temperature of all the clusters T_c , the average MTTF of all clusters $MTTF_c$, the temperature of all active nodes $A = \{a_0, \dots, a_0\}$, the temperature of all dark nodes $D = \{d_0, \dots, d_0\}$, the MTTF of all nodes $MTTF_n$, and the frequencies of all nodes F . Given $n \times n$ number of nodes, ABENA divides the many-core system into several C_{SET} . We employ this Dark-Silicon cluster mapping approach because it allows flexibility between restricting application threads to a specific area whilst also allowing the tasks to be dispersed in that region. For each C , we calculate and compare the average $MTTF$ by monitoring all $MTTF$ of the N in each cluster. ABENA selects the first C for mapping by calculating the average $MTTF$ of nodes from

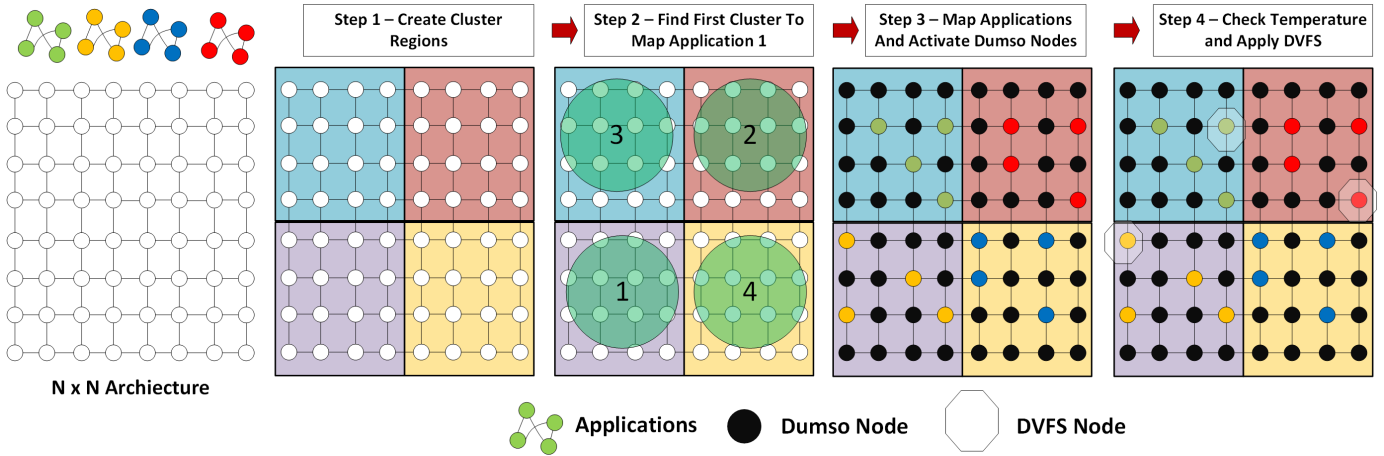


FIGURE 3: Proposed Architecture

Algorithm 1: ABENA Algorithm

Data: $C_1 \dots C_N$
Result: Low to High Clusters

- 1 N : Nodes C : Clusters ;
- 2 C_{SET} : Cluster Set ;
- 3 *ReturnAverage* : Return Average of the Cluster ;
- 4 $MTTF^o$: Mean Time To Failure ;
- 5 T_{TRS} : Temperature Threshold ;
- 6 $Temp$: Temperature ;
- 7 $N_{Frequency}$: Temperature ;
- 8 **for** $C \in C_{SET}$ **do**
- 9 $ReturnAverage \leftarrow N_K \div (\sum N_1^{MTTF} \times N_K^{MTTF})$;
- 10 **end**
- 11 $FirstCluster \leftarrow C$ with the Highest $MTTF$ Average
- 12 * Initial Mapping * Select 50% of the nodes with the lowest temperature nodes from each cluster
- 13 **if** $N_{Temp} > T_{TRS}$ **then**
- 14 $N_k \leftarrow ApplyDVFS N_{Frequency} - 10\%$
- 15 **else**
- 16 **end**

Algorithm 2: NANA Algorithm

Data: $NC_1 \dots C_N$ Node Cluster

- 1 **foreach** $NC_1 \in C_N$ **do**
- 2 **if** $N_1 > N_2$ **then**
- 3 $N_k \leftarrow AssignThread$ **if** $N_1^{Temp} > T_{TRS}$: **then**
- 4 $N_k \leftarrow ApplyDVFS N_{Frequency} - 10\%$
- 5 **else**
- 6 **else**
- 7 **end**

each C . The highest $MTTF$ cluster is selected as the prime candidate for mapping. Additionally, if two clusters return the same $MTTF$, the C with the lowest T is chosen (line 2-5). In line 5 - 6, active nodes from each region are chosen based on their T . Step 4 in Fig. 3 depicts this approach. Unfortunately, this mapping approach is unpredictable. In a large many-core system, hotspot can easily occur when four neighbouring nodes with the lowest T is selected. To reduce this ABENA implements a Neighbour-Temperature-Aware Node-Allocation (NANA) (Algorithm 2) which forms clusters between two neighbouring nodes.

This ensures that, A nodes are always surrounded by the D to hotspot. The active node between the two neighbouring nodes is selected based on the node with the Highest $MTTF$. (Line 1). If the active node's temperature is more than the T_{TRS} , the F is decreased. The proposed design is implemented with different voltage levels. In the experiment conducted, the frequency is reduced by 10%. The 10% is not a special value that guarantees the temperature violations. From a design perspective, reducing the frequency by 1% will not have any significant effect and thus we use 10%. Additionally, if the temperature of the nodes does not change, the task is migrated unto the neighbouring dumso node. Although, ABENA aims to keep 32 nodes active, if all applications have been mapped, the idle active nodes are shut down to allow the working nodes to function at a higher frequency. ABENA assumes that a source node goes to C1 State at the initial phase of migration maintaining the contents of its L1/L2 cache.

V. PERIODIC WORKLOADS

Algorithm 3 depicts the algorithm ABENA-P, the periodic version of ABENA for mapping applications. ABENA-P checks for the $MTTF$ of neighbouring nodes and assign tasks to nodes with the highest $MTTF$ (Line 3). After assigning the tasks, ABENA-P then checks the temperature of the node and reduce the frequency of the node until it meets the temperature threshold. When ABENA-P is applied, intervention

TABLE 1: The used symbols description

Symbol	Description
N	Nodes
A_{SET}	Set of all active nodes
A	Active node
D_{SET}	Set of all dark nodes
D	Dark node
C_{SET}	Set of all Clusters
C	Cluster
F	Frequency
T	Temperature
T_c	Set of transient temperature of all clusters
T_n	Set of transient temperature of all nodes
T_{TRS}	Temperature Threshold
$MTTF_c$	Average MTTF of a cluster
$MTTF_n$	Average MTTF of nodes

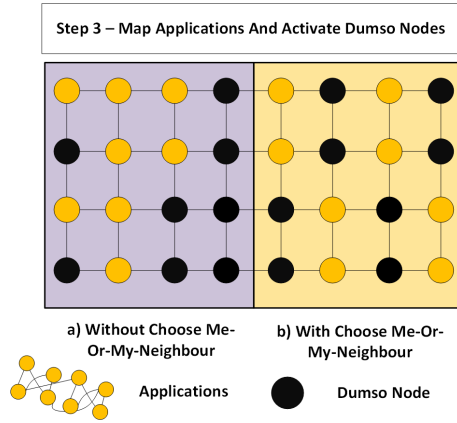


FIGURE 4: 4 x 4 Architecture With the Proposed Ageing.

happens once every epoch, which is in months or hours resulting in minimal or no performance penalty.

Algorithm 3: ABENA Periodic

```

Data:  $Epoch_1, \dots, Epoch_N$  : Epoch
Data:  $N$  : Nodes
Data:  $NC_1, \dots, NC_1$  : Neighbouring Clusters
Data:  $D$  : Dark node
1 foreach  $Epoch_1 \in \mathcal{E}poch_N$  do
2   if  $N_{1MTTF} < N_{2MTTF}$  then
3      $N_1 \leftarrow D$  if  $N_{1Temp} > T_{TRS}$  : then
4        $N_k \leftarrow ApplyDVFS N_{Frequency} - 10\%$ 
5     else
6   else
7      $N_2 \leftarrow D$  if  $N_{1Temp} > T_{TRS}$  : then
8        $N_k \leftarrow ApplyDVFS N_{Frequency} - 10\%$ 
9     else
10    end
11 end

```

Consequently, to accommodate more applications, ABENA utilises the Flow Node Algorithm. This enables an application which is not heavily dependent on inter-communication to be mapped to nodes in different clusters.

Similar to [31], at the end of every each epoch, applications are mapped to healthy nodes.

VI. ENVIRONMENTAL EVALUATION

This section presents the experimental setup, results, and discussions of the proposed work.

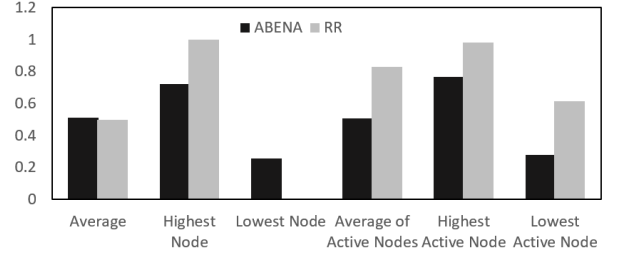


FIGURE 5: Normalised comparison of the average peak temperatures of the chip by running the two algorithms on a 4.0 GHZ frequency

A. EXPERIMENTAL SETUP

Experiments and evaluations are conducted using an extended version of Sniper [38] called LifeSim [39], which integrates McPat [40] for power simulation, HotSpot [41] to model the thermal profile and generate thermal values. Additionally, LifeSim also integrates RAMP [42] to model the lifetime of the nodes. There are other power models such as Watch [43] and PowerTrain [44] which can also be used to estimate the cores' power profiles. However, to the best of our knowledge, McPat and HotSpot are the most widely tools used in literature to measure and assess temperature and power.

RAMP [42] is a dynamic reliability management tool which calculates the expected lifetime of cores based on the current temperature and its utilization. RAMP divides the processor into a few structures – ALUs, FPUs, register files, branch predictor, caches, load-store queue, instruction window – and applies the analytic models to each structure as an aggregate. Our proposed architecture uses the lifetime of nodes generated by RAMP as the parameter for assigning applications.

LifeSim simulator consists of a set of applications which spawn several threads which run simultaneously and can be picked from any of the available benchmarks. When the simulation starts, application threads are spawned on nodes specified by a mapping file. The experimental setup of the proposed work is shown in 8. During the simulation, LifeSim uses the previously mentioned integrated tools to generate the following values for analysis: McPAT to generate power values, HotSpot to generate thermal values using the MCPAT generated power values and RAMP to estimate the lifetime of the nodes.

LifeSim provides two modes of run-time. These are preemptive and run-time non-preemptive scheduling. For this this particularly case study, we use non-preemptive

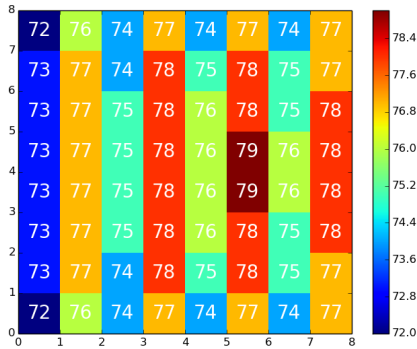


FIGURE 6: Temperature of ABENA

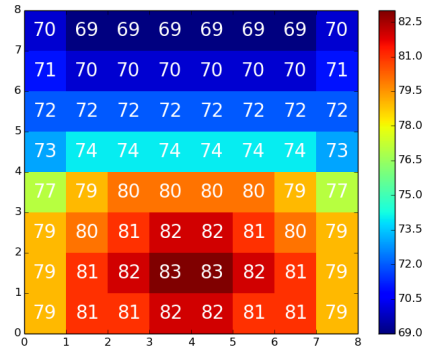


FIGURE 7: Temperature of RR

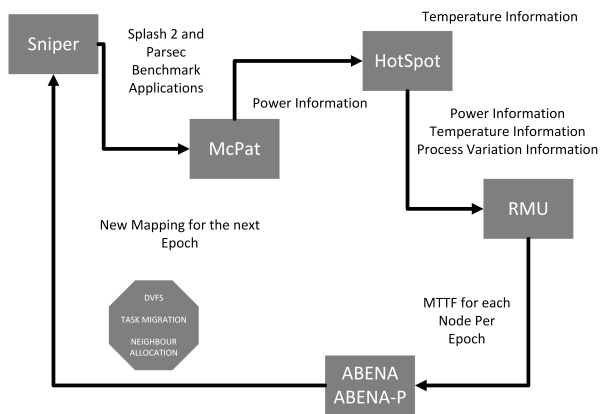


FIGURE 8: The environmental setup of the proposed method

scheduling. Non-preemptive scheduling decisions are made ahead of time and therefore does not delay mapping. Non-preemptive also employs epoch schedule. Unlike run-preemptive scheduling where intervention occurs every specified second, non-preemptive scheduling occurs after every epoch which is usually in months and hours and thus time overhead and performance is very minimal. Similar to [45] we could not measure the switching overhead. Snipersim does not simulate the operating system and the context switch managed by it but only the user-space. In theory, context switching usually takes $5-7n$ but in this case study, intervention occurs after every epoch [33]. The profiling results (power, temperature and MTTF) of are logged in parallel to the simulation, each sampling interval. We invoke our proposed method during this process to compare and analyse the results being generated to assign applications, reduce the frequency of nodes, and migrate tasks among nodes. We compared the proposed method (ABENA-P) with a Dark-Silicon round-robin mapping algorithm with realistic computation applications (radix, oceans, fit and barnes from Splash-2 benchmark [46]) to verify and validate ABENA. By default, ABENA uses the default length epoch (1 Month).

Table 2 depicts the parameters for the simulated configura-

tion. In our previous paper [47], we used a 45nm processor to evaluate ABENA in a 4×4 architecture. To prove that it can be applied to large scale architectures, we evaluated ABENA on many-core system that consists of 64 nodes, where 32 are active and 32 are dumso. The nodes are connected using an 8×8 mesh NoC with 22nm processor. Each node consists of a private 32 KB L1 data cache, 32 KB L1 instruction and a 512KB L2 Cache.

To get accurate readings, we modified sniper to support Dark-Silicon and conducted the experiment 10 times. We compared our proposed method with the conventional Round-Robin (RR). This comparison was conducted into two stages. At the first stage, we conducted a study on both algorithms with 4.0 GHZ frequency to examine the worst-case scenarios; Primarily, this was done to test if the activation of nodes in a specific area affects the overall temperature. The threshold temperature was set to 80°C . The second stage experiment was conducted at a frequency of 2.0 GHZ. In this phase of the experiment, the DVFS in ABENA was turned-off. The parameters that the results are conducted on are temperature, MTTF, Hotspot and Utilization. In this research, the failure mechanism that is used is TDDDB. Wu et al. concluded in [48] that temperature and voltage affect the lifetime of ultra-thin gate oxide due to TDDDB. As a result of this, the MTTF due to TDDDB, $MTTF_{TDDDB}$, at a temperature, T , and a voltage, V , can be model as:

$$MTTF_{EM} \propto \left(\frac{1}{V}\right)^{(a-bT)} e^{\frac{(X + Y/T + ZT)}{kT}}$$

where the values $a = 78$, $b = -0.081$, $X = 0.759\text{ev}$, $Y = -66.8\text{evK}$, and $Z = -8.081 - 4\text{ev/K}$ are fitting parameters [48].

TABLE 3: Workload Characteristics

Parameter	Value
Temperature threshold	80°
Parameters	Temperature, Average MTTF, Utilisation

TABLE 2: Summary of Simulation Parameters

Parameter	Value
Nodes	64
Core Frequency	2.0 - 4.0 GHZ
Core Size	5.71 x 5.71
L1 Data Cache	32KB, 4-way set-associative, 64B Block Size
L1 Instruction Cache	32KB, 4-way set-associative, 64B Block Size
L2 Cache	512KB, 4-way set-associative, 64B Block Size
L3 Cache	8129, 23 Block Size
Applications	Radix, oceans, fft, barnes
Technology	22 nm
Network size	8 X 8 NoC
Temperature threshold	80 °
Failure Mechanism	TDDB
Algorithm °	
ABENA-P	

B. COMPARATIVE RESULTS AND ANALYSIS

Table 3 shows the workload characteristics. For a wide range of results, we compared both approaches with two frequencies (2.0 & 4.0 GHZ)

1) Evaluation of Temperature between ABENA and RR using 4.0 GHZ Frequency

Fig. 11 shows a normalised temperature comparison between ABENA and RR. The RR architecture achieves a 3% advantage over ABENA because 50% of its nodes are not utilized. In contrast, the highest node from both algorithms shows that RR has the highest temperature caused by having active nodes functioning next to each other. The active nodes in ABENA achieves a 30% advantage of RR. In addition, ABENA reduces hotspot in its architectures. Fig. 7 is an image of the average temperature of ABENA and RR. By placing dark nodes next to each active node and using DVFS, we minimise hotspot and keep the temperature of the nodes below the temperature cap. This is reflected in the highest temperature of ABENA being 79°C compared to 82°C with RR. ABENA achieves a better temperature performance for all the active nodes.

2) Evaluation of MTTF and Utilisation between ABENA and RR using a 4.0 GHZ Frequency

Fig. 9 shows a normalised comparison of the MTTF from simulated results between ABENA and RR. ABENA outperforms RR by 20% when comparing the average of active nodes. As previously mentioned, due to half of nodes in RR being inactive, the average of RR is higher than ABENA.

However, ABENA outperforms RR when comparing the node with the lowest and highest active MTTF. By using the lifetime as a parameter, the lifetime of a node is evaluated before a task is assigned to it.

Fig. 10 shows the utilization comparison between both approaches. It can be concluded from Fig. 10 that the gap

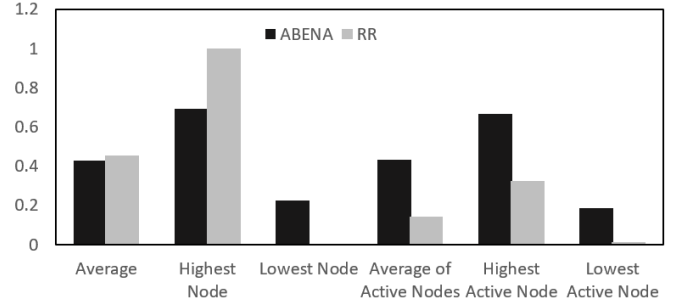


FIGURE 9: Normalised comparison of the average MTTF between ABENA and RR using Frequency 4.0 GHZ

between the highest node and lowest from ABENA is smaller compared to RR. This proves how architectures which implement RR suffers perform it is implemented. There is 30% reduction in this gap when ABENA is applied.

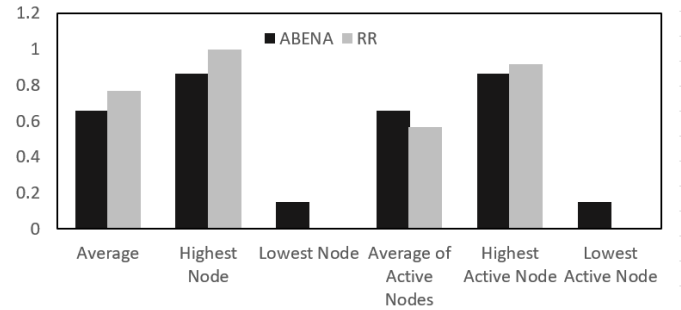


FIGURE 10: Normalised comparison of the average Utilization of the chip by running the two algorithms

Moreover, based on the results presented, it is obvious that ABENA outperforms the conventional round-robin. ABENA ensures that the nodes are functioning below the temperature cap to prevent thermal issues whilst improving the lifetime of nodes. However, we also notice that the introduction of DVFS in ABENA provides an advantage for ABENA. Therefore, we conducted another set of experiments where the frequency was reduced to 2.0 without the use of DVFS.

3) Evaluation of Temperature between ABENA and RR using 2.0 GHZ Frequency

Fig. ?? shows the average temperature of the two methods when they are evaluated with a frequency of 2.0. One can see that the average temperature of all the algorithms is below the temperature threshold but once again RR achieves the lowest average temperature. For the rest of the results, it can be concluded that ABENA outperforms RR in every area. The highest node and lowest node from the study revealed that ABENA minimises the temperature by 10% and 15% respectively. Consequently, ABENA outperforms RR by 20% when comparing the average of active nodes. Additionally, the highest node and lowest node from ABENA also outperforms RR by 10% respectively. This proves that even without the

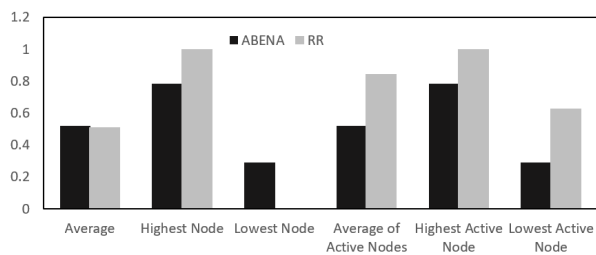


FIGURE 11: Normalised comparison of the average peak temperatures of the chip by running the two algorithms

TABLE 4: Comparison between ABENA and Round-Robin

Parameter	Round-Robin	ABENA
Hotspot	High	Low
Lifetime	Medium	Low
Temperature	Medium	Low

DVFS, ABENA achieves a better temperature performance because of the pattern at which active and dark nodes are selected. Fig. ?? shows the temperature of ABENA. Similar to Fig. 11, the temperature is evenly spread across the chip. There is evenly spread of temperature across the chip.

4) Evaluation of MTTF and Utilisation between ABENA and RR using a 2.0 GHZ Frequency

Fig. 14 compares the MTTF of the two methods. Similar to the results from the frequency 4.0 GHZ, the highest node from both algorithms is RR. However, ABENA outperforms RR when it comes to the highest and lowest MTTF node by 10%. This further highlights the disparity between the highest node and the lowest. Additionally, ABENA outperforms RR by 10% in average of active nodes, 15% and 10% in highest and lowest active node respectively.

Fig. 15 shows the idle percentage of nodes used in both algorithms. From the image, it is obvious that the idle percentage of nodes in the round-robin algorithm is high and presents a major issue for the lifetime of nodes. As a matter of the fact, there is 40% gap between the highest and lowest node from RR. In contrast, the highest node and lowest node from ABENA is a 20% gap. Additionally, ABENA outperforms RR in every area when comparing active node utilization.

VII. SUMMARY OF CONTRIBUTION

Moreover, it can be concluded that ABENA outperforms RR in hotspot minimisation, temperature, MTTF and utilization. Fig. 4 summaries the results. By activating a dark node next to dark node, ABENA improves the temperature performance of many-core systems by 10%; Primarily, the active nodes. In addition, ABENA also improves the lifetime of nodes by more 5% by efficiently utilization nodes. Throughout the study, ABENA bridged the disparity between the lowest and highest utilized. ABENA achieved this by using the lifetime of nodes as an additional and main parameter before assigning tasks.

VIII. CONCLUSION AND FUTURE WORK

In this paper, we presented ABENA, a run-time lifetime algorithm to optimise the lifetime of nodes whilst improving the temperature and reducing hotspot. The temperature, MTTF and utilization under two different frequencies were evaluated and compared to the traditional round-robin algorithm. The results obtained show that, ABENA outperforms the conventional algorithm with or without DVFS. Although results show that the average temperature of both approaches for MTTF are similar, the average of active nodes shows that ABENA improves the average lifetime of all active nodes by 20%. This is evident by comparing the highest MTTF Node of both approaches. With more stress being placed on more nodes, the lifetime of the chip will diminish faster. ABENA delivers a balanced chip. Furthermore, this effectively improves the hotspot performance of the chip. This is because, by ensuring that there is always at least one dark node next to an active node, the temperature reduces. In contrast, the conventional round-robin algorithm creates hotspot which exceeds the temperature cap even though half the nodes of the chip are dark. Additionally, this shows that with a larger scale network and an increased frequency, the round-robin algorithm will cause thermal issues. In our future work, we will focus on task migration in every epoch to efficiently migrate tasks among active and dark nodes in a single workload. We will also evaluate the impact of consistent task migration.

IX. ACKNOWLEDGMENT

The supporting dataset for this manuscript is available from the University of Northampton Research Explorer at <http://doi.org/10.24339/b7abfe3f-312f-4c0f-a94b-5ae4067fb3a0>

REFERENCES

- [1] D. Juan, S. Garg, J. Park, and D. Marculescu. Learning the optimal operating point for many-core systems with extended range voltage/frequency scaling. In International Conference on Hardware/Software Codesign and System Synthesis (CODES+ISSS), 2013.
- [2] A. Kanduri, M. Haghbayan, A. Rahmani, P. Liljeberg, A. Jantsch, and H. Tenhunen. Dark silicon aware runtime mapping for many-core systems: A patterning approach. In 33rd IEEE International Conference on Computer Design (ICCD), 2015.
- [3] T. S. Muthukaruppan, M. Pricopi, V. Venkataramani, T. Mitra, and S. Vishin. Hierarchical power management for asymmetric multi-core in dark silicon era. In 50th ACM/EDAC/IEEE Design Automation Conference (DAC), 2013.
- [4] Muhammad Shafique, Siddharth Garg, Tulika Mitra, Sri Parameswaran, and Jörg Henkel. Dark silicon as a challenge for hardware/software co-design: Invited special session paper. In Proceedings of the International Conference on Hardware/Software Codesign and System Synthesis, CODES '14, 2014.
- [5] S. Pagani, H. Khdr, J. J. Chen, M. Shafique, M. Li, and J. Henkel. Thermal safe power (tsp): Efficient power budgeting for heterogeneous manycore systems in dark silicon. IEEE Transactions on Computers, 2017.
- [6] W. Liu, L. Yang, W. Jiang, L. Feng, N. Guan, W. Zhang, and N. Dutt. Thermal-aware task mapping on dynamically reconfigurable network-on-chip based multiprocessor system-on-chip. IEEE Transactions on Computers, 2018.
- [7] X. Wang, A. K. Singh, B. Li, Y. Yang, H. Li, and T. Mak. Bubble budgeting: Throughput optimization for dynamic workloads by exploiting dark cores in many core systems. IEEE Transactions on Computers, 2018.

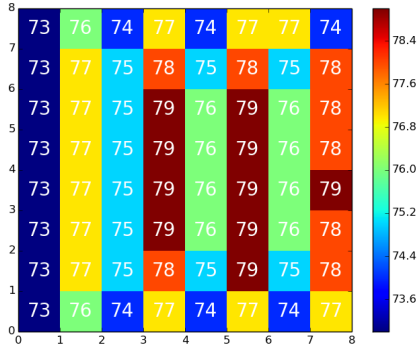


FIGURE 12: Temperature of ABENA

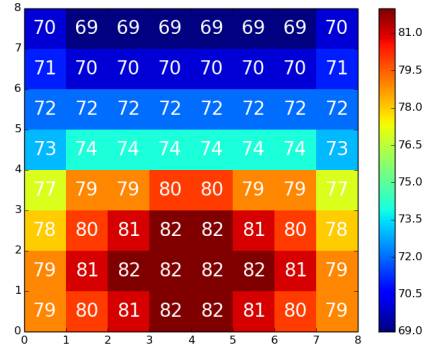


FIGURE 13: Temperature of RR

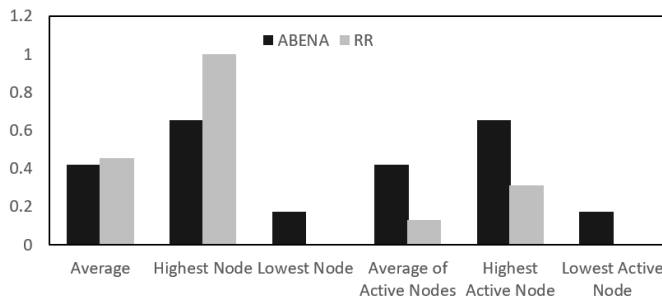


FIGURE 14: Normalised comparison of the average MTF temperatures between ABENA and RR using Frequency 2.0

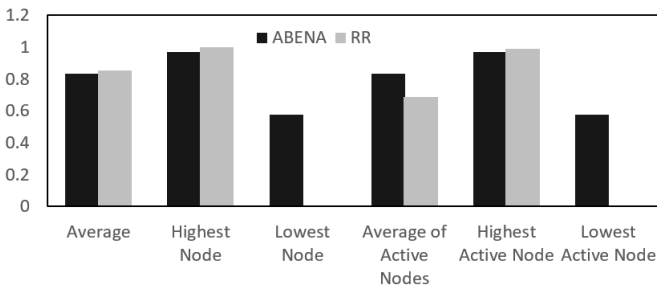


FIGURE 15: Normalised comparison of the average Utilization of the chip by running the two algorithms

[8] ITRS. Online. available: <http://public.itrs.net>. In International Technology Roadmap for Semiconductors (ITRS), 2014.

[9] E. Ofori-Attah and M. O. Agyeman. A survey of recent contributions on low power noc architectures. In Computing Conference, pages 1086–1090, 2017.

[10] Emmanuel Ofori-Attah and Michael Opoku Agyeman. A survey of low power noc design techniques. New York, NY, USA, 2017. Association for Computing Machinery.

[11] Emmanuel Ofori-Attah, Xiaohang Wang, and Michael Opoku Agyeman. A survey of low power design techniques for last level caches. In ARC, 2018.

[12] Emmanuel Ofori-Attah, Washington Bhebhe, and Michael Opoku Agyeman. Architectural techniques for improving the power consumption of noc-based cmps: A case study of cache and network layer. Journal of Low Power Electronics and Applications, 7:14, 2017.

[13] Iglá Hoxha and Michael Opoku Agyeman. High performance hetero-

geneous multicore architectures: A study. In ISCSIC: 3rd International Symposium on Computer Science and Intelligent Control, Amsterdam, The Netherlands, September 25-27, 2019.

[14] Emmanuel Ofori-Attah and Michael Opoku Agyeman. A survey of low power noc design techniques. In Sören Sonntag, José Manuel García Carrasco, José Luis Abellán Miguel, and Daniel Müller-Gritschneider, editors, Proceedings of the 2nd International Workshop on Advanced Interconnect Solutions and Technologies for Emerging Computing Systems, AISTECS@HiPEAC 2017, Stockholm, Sweden, pages 22–27. ACM, 2017.

[15] L. Yang, W. Liu, N. Guan, M. Li, P. Chen, and E. H. M. Sha. Dark silicon-aware hardware-software collaborated design for heterogeneous many-core systems. In 22nd Asia and South Pacific Design Automation Conference (ASP-DAC), 2017.

[16] M. F. Reza, D. Zhao, and M. Bayoumi. Power- thermal aware balanced task-resource co-allocation in heterogeneous many cpu-gpu cores noc in dark silicon era. In 31st IEEE International System-on-Chip Conference (SOCC), 2018.

[17] Ying Zhang, Lu Peng, Xin Fu, and Yue Hu. Lighting the dark silicon by exploiting heterogeneity on future processors. In ACM/EDAC/IEEE Design Automation Conference (DAC), pages 1–7, 2013.

[18] I. Anagnostopoulos, V. Tsoutsouras, A. Bartzas, and D. Soudris. Distributed run-time resource management for malleable applications on many-core platforms. In 50th ACM/EDAC/IEEE Design Automation Conference (DAC), 2013.

[19] E. L. d. S. Carvalho, N. L. V. Calazans, and F. G. Moraes. Dynamic task mapping for mpsoes. IEEE Design Test of Computers, 2010.

[20] J. Castrillon, A. Tretter, R. Leupers, and G. Ascheid. Communication-aware mapping of kpn applications onto heterogeneous mpsoes. In DAC Design Automation Conference, 2012.

[21] T. S. Muthukaruppan, M. Pricopi, V. Venkataramani, T. Mitra, and S. Vishin. Hierarchical power management for asymmetric multi-core in dark silicon era. In 50th ACM/EDAC/IEEE Design Automation Conference (DAC), pages 1–9, 2013.

[22] M. H. Haghbayan, A. Miele, A. M. Rahmani, P. Liljeberg, and H. Tenhunen. A lifetime-aware runtime mapping approach for many-core systems in the dark silicon era. In Design, Automation Test in Europe Conference Exhibition (DATE), pages 854–857, 2016.

[23] C. Feng, Z. Liao, Z. Lu, A. Jantsch, and Z. Zhao. Performance analysis of on-chip bufferless router with multi-ejection ports. In In IEEE 11th International Conference on ASIC (ASICON), pages 1–4, 2015.

[24] C. Fallin, C. Craik, and O. Mutlu. Chipper: A low-complexity bufferless deflection router. In In IEEE 17th International Symposium on High Performance Computer Architecture, pages 144–155, 2011.

[25] B. K. Daya, L. S. Peh, and A. P. Chandrakasan. Towards high-performance bufferless noes with sceptor. In IEEE Computer Architecture Letters, pages 62–65, 2016.

[26] Cheng Li and P. Ampadu. A compact low-power edram-based noc buffer. In In Low Power Electronics and Design (ISLPED), IEEE/ACM International Symposium on, pages 116–121, 2015.

[27] N. Nasirian and M. Bayoumi. Low-latency power-efficient adaptive router design for network-on-chip. In In 28th IEEE International System-on-Chip Conference (SOCC), pages 287–291, 2015.

- [28] K. A. Bowman, S. G. Duvall, and J. D. Meindl. Impact of die-to-die and within-die parameter fluctuations on the maximum clock frequency distribution for giga-scale integration. *IEEE Journal of Solid-State Circuits*, 2002.
- [29] A. kanduri, M. haghbayan, A. M. Rahmani, M. Shafique, A. Jantsch, and P. Liljeberg. adboost: Thermal aware performance boosting through dark silicon patterning. *IEEE Transactions on Computers*, 2018.
- [30] A. M. Rahmani, M. Haghbayan, A. Miele, P. Liljeberg, A. Jantsch, and H. Tenhunen. Reliability-aware runtime power management for many-core systems in the dark silicon era. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 2017.
- [31] V. Rathore, V. Chaturvedi, A. K. Singh, T. Srikanthan, R. Rohith, S. Lam, and M. Shafique. Himap: A hierarchical mapping approach for enhancing lifetime reliability of dark silicon manycore systems. In *Design, Automation Test in Europe Conference Exhibition (DATE)*, 2018.
- [32] M. S. Mohammed, A. K. Al-Dhamari, A. A. ab Rahman, N. Paraman, A. A. M. Al-Kubati, and M. N. Marsono. Temperature-aware task scheduling for dark silicon many-core system-on-chip. In *8th International Conference on Modeling Simulation and Applied Optimization (ICM-SAO)*, 2019.
- [33] Trevor E. Carlson, Wim Heirman, and Lieven Eeckhout. Sniper: Exploring the level of abstraction for scalable and accurate parallel multi-core simulation. In *SC '11: Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis*, pages 1–12, 2011.
- [34] Mohsen Ansari, Sepideh Safari, Sina Yari-Karin, Pourya Gohari-Nazari, Heba Khdr, Muhammad Shafique, Jorg Henkel, and Alireza Ejlali. Thermal-aware standby-sparing technique on heterogeneous real-time embedded systems. *IEEE Transactions on Emerging Topics in Computing*, pages 1–1, 2021.
- [35] Emmanuel Ofori-Attah, Xiaohang Wang, and Michael Opoku Agyeman. A survey of system level power management schemes in the dark-silicon era for many-core architectures. *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*, 5(15), 9 2018.
- [36] Forum U.E.F.I. Advanced Configuration and Power Interface (ACPI) Specification., 2020.
- [37] Wonyoung Kim, Meeta S. Gupta, Gu-Yeon Wei, and David Brooks. System level analysis of fast, per-core dvfs using on-chip switching regulators. In *2008 IEEE 14th International Symposium on High Performance Computer Architecture*, pages 123–134, 2008.
- [38] Trevor E. Carlson, Wim Heirman, Stijn Eyerman, Ibrahim Hur, and Lieven Eeckhout. An evaluation of high-level mechanistic core models. *ACM Transactions on Architecture and Code Optimization (TACO)*, 2014.
- [39] R. Rohith, V. Rathore, V. Chaturvedi, A. K. Singh, S. Thambipillai, and S. Lam. Lifesim: A lifetime reliability simulator for manycore systems. In *IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*, 2018.
- [40] 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. In *42nd Annual IEEE/ACM International Symposium on Microarchitecture (MICRO)*, 2009.
- [41] R. Zhang, M. R. Stan, and K. Skadron. Hotspot 6.0: Validation, acceleration and extension. In *University of Virginia, Tech. Rep*, 2015.
- [42] Jayanth Srinivasan, Sarita V. Adve, Pradip Bose, and Jude A. Rivers. The case for lifetime reliability-aware microprocessors. page 276, March 2004.
- [43] David Brooks, Vivek Tiwari, and Margaret Martonosi. Wattch: A framework for architectural-level power analysis and optimizations. *SIGARCH Comput. Archit. News*, 28(2):83–94, May 2000.
- [44] W. Lee, Y. Kim, J. H. Ryoo, D. Sunwoo, A. Gerstlauer, and L. K. John. Powertrain: A learning-based calibration of mcpat power models. In *2015 IEEE/ACM International Symposium on Low Power Electronics and Design (ISLPED)*, pages 189–194, 2015.
- [45] Vijeta Rathore, Vivek Chaturvedi, Amit K. Singh, Thambipillai Srikanthan, and Muhammad Shafique. Longevity framework: Leveraging online integrated aging-aware hierarchical mapping and vf-selection for lifetime reliability optimization in manycore processors. *IEEE Transactions on Computers*, 70(7):1106–1119, 2021.
- [46] S. C. Woo, M. Ohara, E. Torrie, J. P. Singh, and A. Gupta. The splash-2 programs: characterization and methodological considerations. In *Proceedings 22nd Annual International Symposium on Computer Architecture*, pages 24–36, 1995.
- [47] Emmanuel Ofori-Attah and Michael Opoku Agyeman. ABENA: An Ageing before Temperature Electromigration-Aware Neighbour Allocation for

Many-Core Architectures, page 146–152. Association for Computing Machinery, New York, NY, USA, 2021.

- [48] J. Srinivasan, S.V. Adve, P. Bose, and J.A. Rivers. The case for lifetime reliability-aware microprocessors. In *Proceedings. 31st Annual International Symposium on Computer Architecture*, 2004., pages 276–287, 2004.



EMMANUEL OFORI-ATTAH is a Lecturer and a PhD student at the University of Northampton. He has extensive knowledge in programming and has industrial experience as a Full Stack Developer. Previously, he worked as a Graduate Teaching Assistant for the Faculty of Science and Technology between 2013 - 2018. He graduated at the University of Northampton with a BEng in Computer Systems Engineering and has won several awards.

These include the University of Northampton's 3 Minute Thesis (3MT) Competition (People's Award and Judges Awards) and an award from the Worshipful Company of Information Technologists (WCIT). He is currently the author of 3 Journals with over 6 conference publications since his inception to field. Emmanuel's research interest is in High Performance Computing and Embedded Systems (VLSI SoC design, computer architecture, re-configurable computing, embedded systems, and Internet-of-Things (IoT)).



MICHEAL OPOKU AGYEMAN is a Professor and Program Leader of Computer Systems Engineering at the University of Northampton (UoN), UK. He is a Chartered Engineer (CEng) of the IET, Chartered Manager (CMgr) of CMI and a Fellow of Higher Education Academy (UK). Michael represents the Research Community of UoN at the University Senate. He is the Postgraduate (PGR) Lead for the Faculty of Arts Science and Technology and Co-Chairs the University's PGR Supervisory Forum. Michael has over 10 years' experience in Embedded Systems Engineering. Previously, he was with Intel Embedded System Research Group of the Chinese University of Hong Kong (CUHK) as a Research Fellow. He received the PhD from Glasgow Caledonian University, UK.

He is the author of 5 books, 2 book chapters and over 70 publications in major journals and conference proceedings. Michael's main research interest spans across 3 main strands and disciplines: 1. Embedded Systems High Performance Computing (VLSI SoC design, computer architecture, reconfigurable computing, wired and wireless NoCs, smart rehabilitation solutions, embedded systems and Internet-of-Things (IoT)); 2. Business Administration (neuromarketing, advertising and market research) and 3. Pedagogy. He has been a Guest Editor of the *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*. He is a technical committee member of several conferences such as IEEE ICCSN, IEEE ICBD and IEEE ICCT and serves as a reviewer of several conferences and journals including IEEE Access. His work on wireless NoC has attracted 2 best paper awards in IEEE/IFIP EUC 2016 and Euromicro DSD 2016, respectively. He is also the recipient of the 2018 International Changemaker of the year award in the first UK Ashoka U changemaker Campus.

...