

논문 2014-51-7-12

3D NoC 구조에서 성능을 고려한 어댑티브 수직 스로틀링 기반 동적 열관리 기법

(Performance-aware Dynamic Thermal Management by Adaptive
Vertical Throttling in 3D Network-on-Chip)

황 준 선*, 한 태 희**

(Junsun Hwang and Tae Hee Han[©])

요 약

최근 등장한 TSV(Through Silicon Via)기반의 3D 적층 기술은 보다 강력한 발열관리 기법을 필요로 하며 냉각 비용과 폼팩터(form factor)의 제한을 고려했을 때 소프트웨어적인 열관리 기법의 중요성이 더욱 강조되고 있다. 이러한 접근 방식의 유력한 후보 중 하나로 제시되었던 스로틀링을 통한 열관리 기법의 경우, 증가하는 버스 점유율로 인해 전체적인 성능저하를 야기하는 문제점이 있다. 본 논문에서는 향후 TSV 기반 3D SoC의 커뮤니케이션 병목 현상을 해결하기 위한 3D 네트워크-온-칩(Network-on-Chip, NoC) 구조에서 어댑티브 스로틀링 기법을 제안하여, 열관리와 더불어 온-칩 네트워크상의 트래픽 감소를 통해 전체적인 성능향상을 목표로 한다. 본 논문에서는 실험을 통하여 기존의 방식에 비하여 스로틀링으로 인해 저하된 처리량이 최소경로 라우팅 시 최대 72% 향상됨을 알 수 있었다.

Abstract

Recent TSV based 3D Integrated Circuit (IC) technology needs more powerful thermal management techniques. However, because cooling cost and form factor are restricted, thermal management are emphasis on software based techniques. But in case of throttling thermal management which one of the most candidate technique, increasing bus occupation induce total performance decrease. To solve communication bottleneck issue in TSV based 3D SoC, we proposed adaptive throttling technique Experimental results show that the proposed method can improve throughput by about 72% compare with minimal path routing.

Keywords : 3D NoC, Dynamic Thermal Management, Vertical Throttling, TCU

I. 서 론

무어의 법칙에 의한 소자 미세화의 한계를 극복하는

방안의 하나로 'More than Moore'로 지칭되는 3D 적층 반도체 기술에 대한 연구가 활발히 진행되고 있다. 또한 SoC에서 폭증하는 데이터 대역폭 요구량을 해결하기 위한 방안으로 NoC(Network-on-Chip)기술이 최근 각광받음에 따라, 두 가지 기술이 결합된 TSV기반 3D NoC가 차세대 SoC 구조 발전의 중요한 흐름으로 자리 잡을 예정이다.

기존 CMOS 소자에서의 전압 스케일링이 한계에 도달함에 따라 전력밀도의 급격한 증가와 더불어 특히 핫스팟(hot spot)에서의 혁신적인 발열 관리 기술이 시급하다. 열관리 기법은 히트싱크(heat sink)나 쿨링팬

* 학생회원, ** 평생회원, 성균관대학교 정보통신대학 (College of Information & Communication Engineering, Sungkyunkwan University)

© Corresponding Author(E-mail: than@skku.edu)

※ 본 논문은 (주)SK하이닉스의 지원으로 연구되었음. (CMOS 이미지센서의 동작 모드를 고려한 저전력 설계 기술 개발)

접수일자: 2014년03월12일, 수정일자: 2014년05월20일
수정완료: 2014년06월23일

(cooling fan) 같은 인위적 냉각 장치를 이용하는 하드웨어 방식과 시스템 특성을 이용해 알고리즘적인 해결을 도모하는 소프트웨어 방식으로 구분할 수 있다. 시스템의 TDP(Thermal Design Power) 사양에 따라 하드웨어적 냉각 장치가 필요할 수 있으나, 모바일을 포함한 많은 어플리케이션 영역에서는 소프트웨어적인 열관리 기법만이 유효하다. 소프트웨어적 열관리 기법은 기본적으로 DTM(Dynamic Thermal Management)에 기반을 둔다^[1]. DTM 기법은 주어진 TDP에 맞게 파워 게이팅(power-gating) 이나 동적 전압 주파수 스케일링(Dynamic Voltage Frequency Scaling, DVFS)과 같은 다양한 저전력 기술을 적용하여 칩의 온도를 줄이거나 워크로드를 칩 내 자원에 효과적으로 배분하여 핫스팟 발생을 방지한다. 최근 제기된 선실행 방식의 수직 스로틀링(vertical throttling) 기법의 경우 3D 구조에서 수직축의 방열속도가 빠름을 활용하여 수직축 상단부터 스로틀링을 실행해 열관리를 시도한다^[2]. 이 기법은 3D NoC 구조에서 매우 효과적임이 증명 되었으나 스로틀링 된 노드로 인하여 발생하는 성능저하는 여전히 존재한다. 본 논문에서는 TCU(Throttling Control Unit)를 이용한 어댑티브 스로틀링 방식을 통해 스로틀링 된 노드의 네트워크를 이용하여 NoC 전체 트래픽 감소를 유도하고 동시에 빠른 워크로드 처리를 통한 성능향상을 목표로 한다.

II. 관련 연구 및 배경

1. Vertical Throttling

최대 성능 동작 시 높은 전력 밀도로 인하여 시간이 경과함에 따라 TDP를 초과하는 상황이 발생한다. 이를 방지하기 위해 칩의 성능을 낮추거나 멈춰야 하며, 온도는 즉시 저하되는 것이 아니기에 임계온도 보다 낮은 발동온도(Trigger Temperature)를 설정하여 발동온도를 초과할 경우 칩 내부 블록의 성능을 감소시키거나 잠시 멈추는 스로틀링 기법이 제시되었다. 이를 통해 온도 관리를 미리 실행하여 임계 온도를 초과하는 것을 미연에 방지하는 것이 가능하다.

기존 2D 구조에서는 실행시간기반 열관리(Run-time Thermal Management) 기법인 전역적 스로틀링 방식이나 분산적 스로틀링 방식을 통해 열 관리를 하였다. 그러나 칩 내 폭증하는 데이터 트래픽을 고려했을 때 전

역적 또는 분산적 스로틀링 방식은 긴 스로틀링 시간과 칩 내부 네트워크 장악에 따른 성능 저하로 인해 효과적이지 않다^[2]. 또한 3D IC로 들어서면서 TSV를 고려한 방열 방식이 필요하게 되었고, 이에 최근 제시된 수직 스로틀링은 빠른 냉각 시간과 간단한 구동 알고리즘을 통해 성능 저하를 최소화 하였다. 동시에 예측 기반의 열 관리 기법이 수동적 열 관리 기법에 비해 높은 효율을 보임에 따라, 예측 기법을 적용한 선실행 방식의 수직 스로틀링 기법은 단계적 스로틀링을 통하여 긴급 스로틀링 노드의 발생 수를 줄이고 이를 통하여 기존보다 적은 성능감소와 전체적인 발열량 감소 효과를 보이게 된다^[3~4].

2. Thermal-Aware Routing Algorithm

앞서 제시된 선실행 방식의 수직 스로틀링 기법을 통하여 전체적인 발열량 감소와 네트워크 처리량 개선에는 성공하였으나, 스로틀링 노드의 증가로 인해 발생하는 성능저하를 막기 위해 다양한 라우팅 기법이 제시되었다. 스로틀링 노드의 증가는 성능 뿐 아니라 칩 내부 온도에도 영향을 미치는 핫스팟의 발생으로 이어지게 되었고, 따라서 핫스팟 발생을 막기 위해 예상되는 구역에 대해 트래픽을 분산하여 처리해야 한다. 동시에 칩 내부 온도를 낮추기 위하여 열을 분산시킬 필요성이 요구 되는데, 이를 위해 히트 싱크의 위치를 고려한 라우팅 기법이 제시되었다.

하향 라우팅 알고리즘의 경우 기존의 라우팅 알고리즘과 다르게 비 최소 경로 라우팅(Non-minimal Path Routing)을 사용하여 트래픽 분산과 열 분산을 통해 핫스팟의 발생을 미연에 방지하고자 한다^[5]. 3D 구조를 활용함으로써 비 최소 방식의 라우팅이 갖는 부하에 비해 트래픽 분산 및 핫스팟 해소가 가능하여 성능 및 열 문제를 쉽게 해결할 수 있다. 또한 하향 라우팅 알고리즘은 선실행 방식의 수직 스로틀링이 적용된 3D NoC 환경에서 상층부에 몰리는 트래픽을 분산시킬 수 있고 스로틀링 코어에 의해 최소 경로 라우팅이 불가능한 상황에서 빠른 네트워크가 가능하다. 또한 이를 응용하여 열적 부하가 적은 외각 경로를 사용하는 Beltway 방식의 라우팅 알고리즘을 통하여 스로틀링 노드 발생량을 감소시키고 이를 통해 성능을 증가시키는 연구도 진행되었다^[6].

최근 제시된 TAAR(Topology-Aware Adaptive

Routing)기법의 경우에는 스로틀링 된 노드들을 불규칙한 토폴로지와 같은 구조로 인지하여 효율적인 라우팅 경로를 구하고자 하였다^[7]. 이웃 라우터들의 열 상태에 대한 정보와 스로틀링 노드 정보를 토폴로지 테이블 (topology table)을 통해 가장 발열량이 적은 노드를 우선적으로 라우팅하여 온도와 트래픽을 고려함과 동시에, 불규칙적으로 변화하는 스로틀링 노드들에 대해 효과적인 라우팅을 가능하게 하였다.

그러나 이와 같은 연구는 지속적인 비 최소 경로 라우팅을 사용하기에 결과적으로 네트워크 속도가 감소하게 되고 이는 전체적인 성능 저하로 이어지게 된다. 또한 스로틀링 노드의 수가 증가함에 따라 라우팅경로가 제한되어 효율이 급격히 감소한다.

III. TSV 기반 3D NoC 구조에서의 어댑티브 스로틀링 기법

기존의 3D구조에서 스로틀링에 의한 성능저하 개선에 관련된 논문들의 경우 스로틀링 노드발생으로 인한 처리량 저하를 라우팅 알고리즘 개선 등을 이용하여 극복하려 하였으나, 이러한 방식은 제한된 토폴로지 환경에 의해 성능 향상의 한계를 가지게 된다. 본 장에서는 기존의 스로틀링 기법을 개선한 어댑티브 스로틀링 기법을 통해, NoC상의 트래픽 감소를 이용하여 전체적인 성능 향상을 목표로 한다.

1. 3D NoC에서 효율적인 어댑티브 스로틀링을 위한 구조

3D NoC 구조에서 빠른 응답속도를 갖는 효율적인 DTM 구현을 위하여 본 논문에서는 NoC상의 각 노드에 TCU(Throttling Control Unit)을 배치하도록 한다. Thermal-aware 구조에서는 각 IP들에 온도 센서가 내장됨으로, 그림 1과 같이 TCU를 각 노드에 두어 온도 센서와 TCU간의 빠른 통신을 통해, 열 위험 상태에 대한 즉각적인 열관리의 실행이 가능하다^[4]. 또한 빠른 응답속도와 높은 신뢰성, 그리고 온도와의 선형성을 갖는 동적 주파수 스케일링(Dynamic Frequency Scaling) 기반의 스로틀링 기법을 TCU에 적용하여 적은 부하로 효과적인 열관리를 수행하도록 한다^[8].

각각의 라우터들은 이웃 IP들의 상태에 대한 정보를 공유하며 스로틀링의 발동조건을 만족하면 열관리를 실행

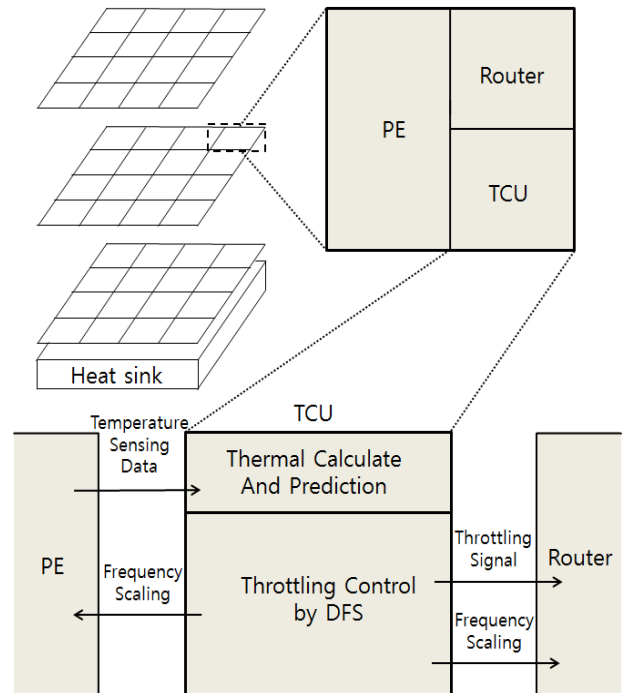


그림 1. 3D NoC에서 DFS를 이용한 TCU 구조.
Fig. 1. Proposed DFS based TCU structure.

한다. 이러한 열관리는 노드 단위로 발생하게 되며 본 논문에서 제시하고자 하는 네트워크 성능의 향상을 위하여, TCU는 IP와 라우터의 스로틀링을 개별적으로 관리할 수 있어야 한다. 이를 통해 IP내에서 증가하는 온도를 감지하여 빠른 열관리가 필요한 긴급 스로틀링을 제외한 경우에 라우터의 기능을 사용할 수 있다.

2. 3D NoC 구조에서의 어댑티브 스로틀링 구현

칩 내부의 코어들이 과열되면서 발생하는 스로틀링 노드의 증가는 네트워크의 성능 감소로 이어진다. 이에 따라 손실된 성능을 보상하기 위해 II장에서 언급한 선실행 방식의 수직 스로틀링 기법과 TAAR기법을 통해 성능을 향상시키고자 하였다. 그러나 스로틀링에 의해 점유된 노드의 증가는 라우팅 경로를 제한하여 전체 트래픽 양을 증가시키게 되고, 실행중인 라우팅에 대해서도 재 라우팅을 요구하게 됨으로 성능저하로 귀결된다. 결국 스로틀링의 실행 순서와 라우팅 알고리즘의 개선으로는 성능 향상에 한계를 가지며 이러한 네트워크 점유를 방지하기 위해서는 각 노드에 IP와 라우터의 개별적 열관리가 요구된다. 본 논문에서는 기존에 제시되었던 선실행 방식의 수직 스로틀링을 기반으로 하는 어댑티브 스로틀링을 통하여, 스로틀링 된 노드 안에서의

개별적 열관리를 통해 네트워크 전체의 트래픽을 줄이고자 한다.

단 이러한 기법을 적용할 시 스로틀링 된 각 노드에서 라우터의 기능을 사용하게 됨으로, 라우터에서 소모되는 전력량으로 인해 방열 속도가 저하 된다. 그러나 NoC환경에서 각 노드의 라우터가 차지하는 온도 및 전력 소모 비율이 낮고, 코어와의 네트워킹이 이루어지지 않아 더 적은 소비전력을 갖게 된다. 물론 이러한 동작 수준 역시 임계 온도를 초과하는 노드에 대해서 치명적으로 동작할 수 있다. II장에서 언급한 선실행 방식의 수직 스로틀링 기법을 통해 긴급 스로틀링 노드의 발생량은 트래픽량에 따라 10~20% 감소함을 보이니^[4] 여전히 비정상적으로 발생하는 긴급 스로틀링 노드는 순간적으로 한계 온도를 초과하게 됨으로 이에 대해 빠른 방열이 가능한 완전한 스로틀링의 수행이 요구된다. 따라서 TCU는 노드의 온도수준에 대한 감지와 예측을 통하여 정확한 정보를 가져야 하며, 이를 통해 긴급노드와 일반노드에 대해 구분하여 스로틀링을 실행하여야 한다.

위의 동작을 위하여 전체 NoC 환경의 노드 상태에 대한 정보는 그림 2와 같이 스로틀링 테이블을 통하여 모든 라우터들이 공유한다. 예를 들어, 1GHz로 동작하는 네트워크 환경에서 스로틀링 테이블은 10ms의 부하로 재구성된다. 즉 전체 네트워킹 시간 중 0.1%의 적은 소요시간으로 관리가 가능함으로^[7] 모니터링 유닛을 통한 관리에 비해 면적과 전력 소모 면에서 적은 부하를 갖으며 이와 같은 관리를 통해 빠른 네트워킹 구현이 가능하다.

스로틀링 테이블은 네트워크상의 각 노드들의 스로틀링 정도를 파악하고 이를 토대로 긴급 스로틀링 여부를 판단하여야 한다. 그림 3과 같이 노드의 성능을 100% 제한하는 스로틀링이 단계적 열관리를 거치지 않

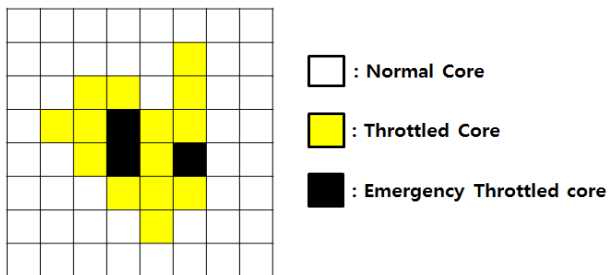


그림 2. 스로틀링 테이블의 예. Fig. 2. Example of throttling table.

고 실행되었다면 그 노드는 긴급 스로틀링 노드로 인식하도록 한다. 그러나 TCU를 통해 단계적으로 열 관리를 거친 경우는 어댑티브 스로틀링이 가능한 노드이므로 라우팅 경로로 이용이 가능하다. 따라서 이러한 구분을 스로틀링 테이블을 통하여 공유하여 라우팅을 수행하여야 한다. 위와 같은 관리를 통하여 라우터는 목적지 까지 최소한의 경로로 라우팅을 실행 할 수 있으며 동시에 열적인 문제에 대한 고려가 가능하여 부담이 가장 적은 노드들을 선택할 수 있다.

각 노드의 라우터들은 패킷을 전송하기 이전에 스로

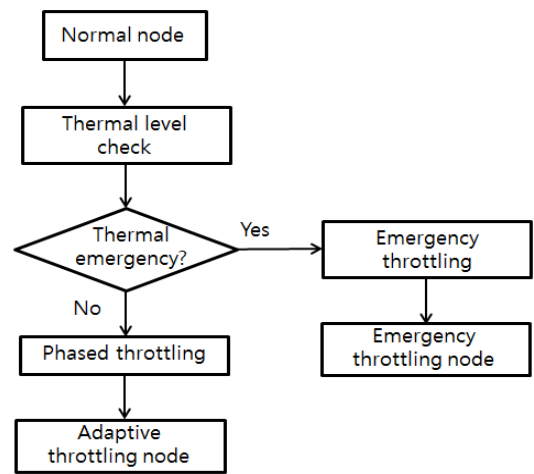


그림 3. 스로틀링 레벨에 따른 노드 구별 동작 순서. Fig. 3. Operation flow of determine node according to throttling level.

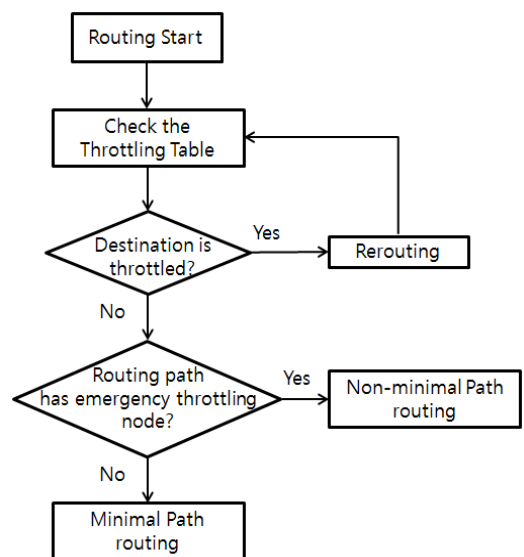


그림 4. 어댑티브 스로틀링이 적용된 라우팅 순서도. Fig. 4. Flow chart of the adaptive throttling-aware routing.

틀링 테이블을 참조하여 목적지와 경로의 스로틀링 상태를 우선 파악하여야 한다. 라우팅 동작순서는 다음과 같다.(그림 4)

- 1) 라우팅 시작을 위하여 스로틀링 테이블을 확인한다.
- 2) 목적지가 스로틀링된 노드인지를 확인하고 스로틀링이 되었다면 라우팅을 재실행한다.
- 3) 설정되는 라우팅 경로에 긴급 스로틀링 노드가 포함되어있는지 확인하여 만약 포함되어있다면 이를 우회하는 비 최소경로 라우팅을 실행한다.
- 4) 앞의 1),2),3) 단계가 충족되었다면 어댑티브 스로틀링을 통한 최소경로 라우팅을 실행한다.

이와 같이 최소 경로의 라우팅을 수행할 조건이 충족되었을 시 기존에 제시되었던 스로틀링 기법에 비해 최소한의 홉(hop)수로 통신이 가능하게 됨으로 전체 네트워크 성능의 향상이 가능하다. 이러한 성능 증가는 스로틀링 노드가 증가함에 따라 점층적으로 증가하게 되며 이를 통해 동일한 workload에 대하여 기존의 방식에 비해 빠른 처리가 가능하다.

어댑티브 스로틀링 기법의 적용을 통해 NoC상의 모

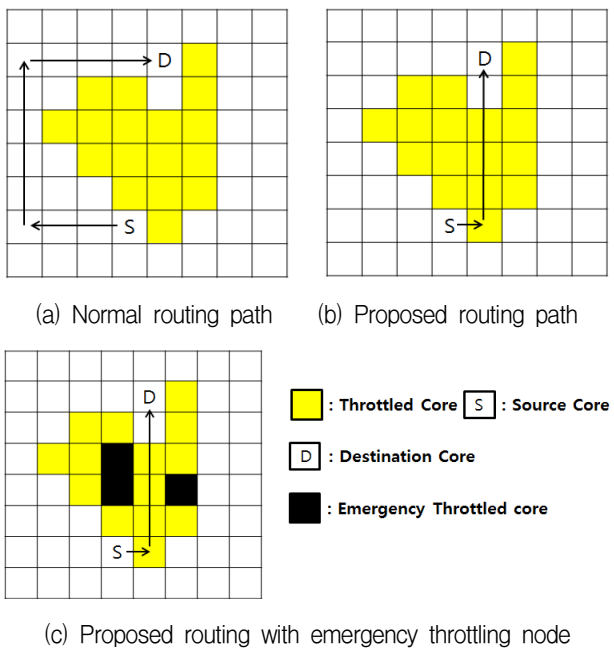


그림 5. 기존 스로틀링 환경과 제안하는 환경에서의 라우팅 경로 비교.

Fig. 5. Comparison of routing path between exist environment and proposed throttling method.

든 노드들은 긴급 스로틀링을 제외한 경우 라우팅 경로로서의 역할이 가능해 진다. 그림 5는 기존 스로틀링 환경과 제시하는 스로틀링 환경에서의 라우팅 경로 차이를 나타낸다. 그림 5(a)와 같이 기존의 스로틀링 환경의 경우 어댑티브 스로틀링이 적용되지 않아 출발지에서 목적지까지 도달하는데 우회된 라우팅이 필요하나 그림 5(b)와 그림 5(c)는 제안하는 환경이 적용되어 스로틀링 노드와 긴급 스로틀링 노드가 존재함에도 최소 경로의 라우팅으로 패킷 전달이 가능해 진다.

IV. 실험 및 분석

이번 장에서는 3장에서 제시한 어댑티브 스로틀링 기법의 성능향상 정도를 정확한 환경에서 실험하며 유사 논문과의 비교 분석을 위해 65nm 공정 환경에서 5.1 GHz의 동작속도와 1.2V의 전압을 사용하며 최대 230W의 전력을 소모하는 인텔사의 80-core processor^[9] 모델을 사용하였다. 이를 통해 소모되는 전력량 및 발열량을 분석하며 어댑티브 스로틀링을 통하여 얻게 되는 성능개선 정도를 비교 및 분석한다.

1. 라우터 전력 소모량 분석

성능 개선을 위하여 실행되는 어댑티브 스로틀링은 스로틀링된 노드의 라우터 동작을 이용하기 때문에 해당 유닛의 전력을 소모하게 된다. 그러나 앞서 언급한 바와 같이 IP부분은 스로틀링을 실행 중임으로 코어와의 패킷 입출력은 불가능하다. 따라서 라우터가 갖는 기능의 일부를 사용하지 않게 됨으로 정상적으로 작동할 때와 비교하여 적은 전력소모량을 갖게 된다. 우리는 Orion 2.0 시뮬레이터^[10]를 통해 전력 소모량을 이전과 비교하여 본 결과 기존 비례 표 1과 같이 약 17%정도 감소한 전력소모량을 보임을 알 수 있었다.

이러한 전력 소모는 추가적인 부하이나, 제시하는 어댑티브 스로틀링을 통하여 재 라우팅으로 인한 부하가 감소하게 되며 동일 워크로드에 대한 처리가 기존의 방

표 1. 어댑티브 스로틀링 시 라우터 전력소비량 비교.
Table 1. Router power consumption comparison of proposed model.

	Normal mode	Proposed mode
Total energy (pJ)	113	94
Total power (mW)	577	479

식에 비해 빠르게 처리됨으로 전체적인 성능 향상 및 신뢰성 면에서 우수하다.

2. 처리량 및 발열량 분석

제안하는 어댑티브 스토틀링을 통한 성능향상을 측정하기 위하여 Hotspot^[11]과 Noxim^[12] 시뮬레이터를 연동시킨 Traffic-thermal co-simulator^[13]를 통하여 실험을 진행하였다. 스토틀링 노드가 랜덤으로 발생하는 경우 스토틀링 노드의 위치에 따라 처리량 차이가 크게 발생할 수 있으므로 정확한 성능 비교를 위하여 반복된 실험을 통한 평균치를 통해 제시하는 모델과 기존의 선 실행 방식의 버티컬 스토틀링 모델과의 성능 차이를 비교하였다. 처리량은 식 (1) 과 같이 단위 시간당 처리한 플릿(Flit)수를 총 노드 수(N_{node})와 전송에 걸린 총 사이클(T)로 나눈 값으로 정의한다.

$$Throughput = \frac{Flits}{N_{node} \times T} \quad (1)$$

그림 6과 같이 패킷 입력 수준(injection rate)이 높아짐에 따라 처리량이 증가한다. 이를 기존의 방식과 비교하였을 경우, 입력 수준 0.1 이상부터 4x4x4에서 8x8x4까지의 토폴로지가 모두 포화되는 처리량 값을 보임에 따라 0.1의 랜덤 트래픽 입력 수준과 1000000 cycle의 시뮬레이션 환경에서 XYZ 라우팅 알고리즘을 통해 저하되는 성능을 우선 측정 후, 제안하는 방식과의 차이를 비교하여 보았다.

그림 7과 같이 4x4x4에서 8x8x4까지 확장하여 비교하였을 경우 스토틀링 노드의 수가 증가함에 따라

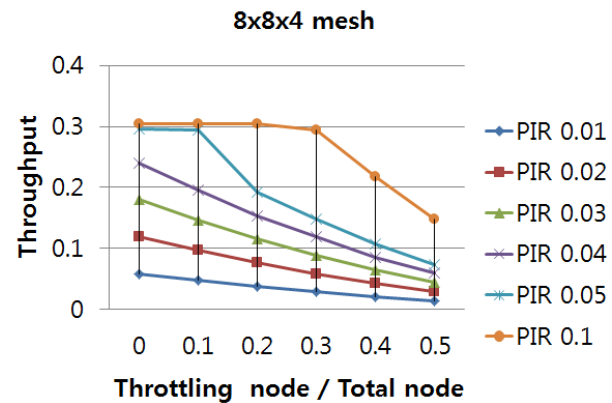
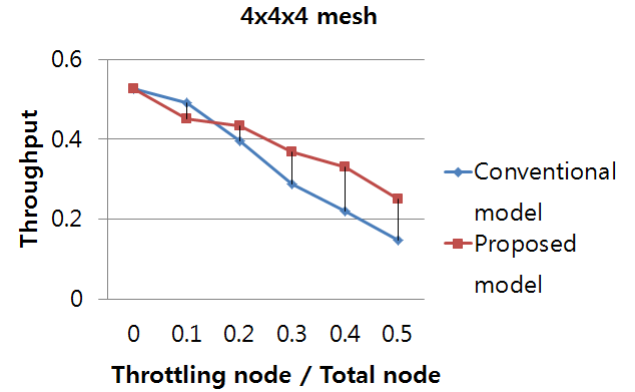
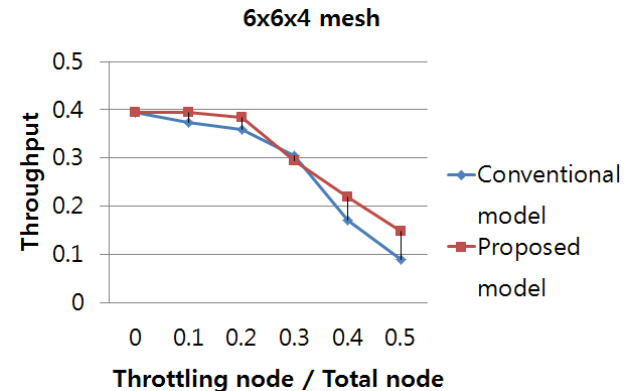


그림 6. 패킷 입력 수준에 따른 처리량 비교.
Fig. 6. Throughput comparison according to packet injection rate.

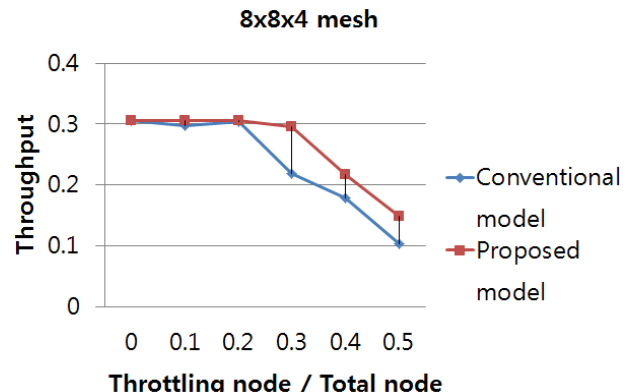
50%에서 최대 75%까지 성능이 저하한다. 토폴로지의 크기가 작을수록 우회경로가 줄어들어 따라 스토틀링 노드에 의한 성능저하가 빠르게 작용하며 토폴로지의 크기가 커지면 상대적으로 성능저하 속도가 줄어드나, 스토틀링 노드의 수가 증가함에 따라 동일한 수준의 성능 감소율을 보인다. 그러나 본 논문이 제시하는 어



(a) 4x4x4 mesh topology



(b) 6x6x4 mesh topology



(c) 8x8x4 mesh topology

그림 7. 토폴로지 크기 별 처리량 비교.
Fig. 7. Throughput comparison of topology by size.

맵티브 스토틀링 기법을 적용함에 따라 이러한 성능저하는 감소하게 된다. 동일한 환경에서 본 논문이 제안하는 방식과의 차이를 실험해 보았을 때, 스토틀링 노드 증가에 따라 5%에서 최대 72%까지의 성능 개선율을 보였다.

비 최소 경로를 사용하는 라우팅 알고리즘에서의 성능 개선을 비교를 위하여 앞서 II장에서 언급한 Downward 라우팅 알고리즘에서 실험을 진행하였다. 위와 같은 경우 수직 스토틀링 기법이 적용됨을 가정하기에 상대적으로 낮은 처리량을 갖는다. 따라서 그림 8과 같이 동일한 비 최소 경로 라우팅 알고리즘을 적용할 경우 스토틀링으로 인한 처리량 감소가 무시할 수 있는 수준이므로 III장에서 언급한 바와 같이 긴급 스토틀링 노드로 인해 비 최소 경로 라우팅을 사용하더라도 기존 모델에 비해 큰 성능향상을 가짐을 보였다.

최종적인 발열량 부하 측정을 위하여 스토틀링을 적용하지 않았을 시 발생하는 발열량과 어댑티브 스토틀링 기법을 적용하였을 시 평균 발열량 차이를 비교하였다. 80-core processor 모델의 경우 평균 성능으로 동작 시 80°C의 발열량을 갖으며 이를 1000000 cycle 동안

시뮬레이션 하였을 시 나타나는 발열량 차이를 비교하였다.

표 2에서 보이는 바와 같이 스토틀링 비적용 모델에 비해 10%~11% 정도의 발열량 감소율을 보였다. 이는 최대 20%의 평균 발열량 감소를 갖는 선실행 방식의 수직 스토틀링 기법에 비해 높은 수준이나 본 논문이 제안하는 방식이 갖는 처리량 개선율이 월등함이 앞의 실험으로 나타났다.

V. 결 론

본 논문에서는 3D NoC 구조에서 열 문제 해결을 위해 기존에 제시되었던 선실행 방식의 수직 스토틀링 기법들에 의해 발생되었던 성능 저하를 극복하기 위하여, 각 노드에서 라우터와 코어간의 개별적인 스토틀링이 가능한 어댑티브 수직 스토틀링의 구조 및 기법을 소개하였다. 제안하는 방식을 통하여 제한되었던 라우팅 경로를 개선하여 최소 경로의 라우팅이 가능함에 따라 기존의 방식에 비해 높은 처리량을 보였다.

제안하는 어댑티브 수직 스토틀링 기법은 기존의 선실행 방식의 수직 스토틀링 기법에 비해 발열량이 12% 정도 증가하나 이는 TDP를 초과하지 않으며 스토틀링 노드 증가에 따라 최대 72%만큼의 처리량이 증가하였다. 또한 긴급 스토틀링으로 인한 비 최소 경로 라우팅 시에는 스토틀링 비적용 시와 비교하여 손실 없는 처리량을 보였다.

그러나 보다 유연한 열 관리를 위해서는 동작 정도를 세분화한 연구가 필요하며, 동시에 보다 효율적인 저전력 기술이 적용된 스토틀링 기법을 통하여 전력 효율 향상과 더불어 발열량의 감소가 가능할 것이다. 또한 본 논문이 제안하는 스토틀링 기법을 보다 효율적으로 사용할 수 있는 라우팅 알고리즘의 제안이 요구된다.

REFERENCES

- [1] J. Kong, S.W. Chung, and K. Skadron, "Recent Thermal Management Techniques for Microprocessors," ACM Computing Surveys, 2011.
- [2] C.H. Chao et al., "Transport Layer Assisted Routing for Run-Time Thermal Management of 3D NoC Systems," to be appeared in ACM

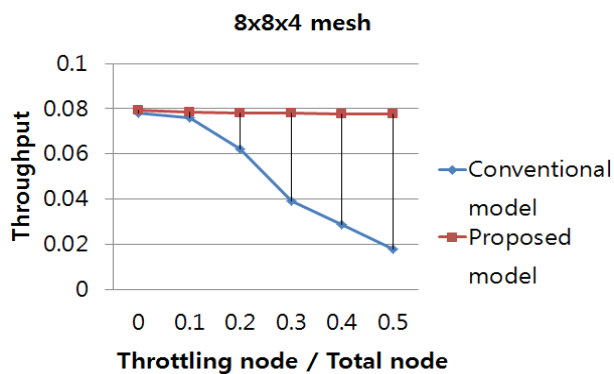


그림 8. Downward 라우팅 알고리즘 적용시의 기존 모델과 제안하는 모델과의 성능비교.

Fig. 8. Throughput comparison between proposed and conventional model in Downward routing algorithm.

표 2. 어댑티브 스토틀링 시 발열량 비교.

Table 2. Temperature value consumption comparison of proposed model.

Model \ Layer	4	3	2	1
Non-throttling	99.9°C	99.5°C	98.1°C	94.4°C
Proposed	90.4°C	89.6°C	88°C	86°C

- Transactions on Embedded Computing Systems, 2011.
- [3] T. Wegner et al., "Impact of Proactive Temperature Management on Performance of Networks-on-Chip," Int'l Symp. System on Chip (ISSOC), pp.116-121, Oct. 2011.
- [4] K.C. Chen et al., "Design of thermal management unit with vertical throttling scheme for proactive thermal-aware 3D NoC systems," IEEE Int. Symp. VLSI Design, Automation and Test (VLSI-DAT), Apr. 2013.
- [5] C.-H. Chao et al., "Traffic- and Thermal-Aware Run-Time Thermal Management Scheme for 3D NoC Systems," in Proc of NOCS 2010, pp. 223-230.
- [6] K.C. Chen, C. Kuo, H.-S. Hung, and A.-Y. (Andy) Wu, "Traffic- and Thermal-aware Adaptive Beltway Routing for three dimensional Network-on-Chip Systems," IEEE international Symposium on Circuits and Systems (ISCAS), May 2013.
- [7] K.-C. Chen, S.-Y. Lin, H.-S. Hung, A.-Y. Wu, "Topology-Aware Adaptive Routing for Non-Stationary Irregular Mesh in Throttled 3D NoC Systems," IEEE Trans. Parallel and Distributed Systems, vol.24, no.10, pp. 2109-2120, Oct. 2012.
- [8] S.W. Chung and K. Skadron, "A Novel Software Solution for Localized Thermal Problems," Proc. Fourth Int'l Symp. Parallel and Distributed Processing and Applications (ISPA), pp. 63-74, Dec. 2006.
- [9] Y. Hoskote, S. Vangal, A. Singh, N. Borkar, and S. Borkar, "A 5-GHz Mesh Interconnect for A Teraflops Processor", IEEE MICRO, vol. 27, pp. 51-61, 2007.
- [10] A. Kahng, B. Li, L.-S. Peh, and K. Samadi, "ORION 2.0: A Fast and Accurate NoC Power and Area Model for Early-Stage Design Space Exploration," in DATE, 2009.
- [11] W. Huang, M.R. Stan, K. Skadron, K. Sankaranarayanan, and S. Ghosh, "HotSpot: A Compact Thermal Modeling Method for CMOS VLSI Systems," IEEE Trans. Very Large Scale Integration Systems, vol. 14, no. 5, pp. 501-513, May 2006.
- [12] Noxim: network-on-chip simulator [Online]. Available: <http://sourceforge.net/projects/noxim/>
- [13] K.-Y. Jheng, C.-H. Chao, H.-Y. Wang, and A.-Y. Wu, "Traffic-Thermal Mutual-Coupling

Co-Simulation Platform for Three-Dimensional Network-on-Chip," in Proc. IEEE Intl. Symp. on VLSI Design, Automation, and Test (VLSI-DAT'10), Apr. 2010.

— 저 자 소 개 —



황 준 선(학생회원)
2012년 가천대학교 전자공학과
학사 졸업.
2012년 9월~현재 성균관대학교
정보통신대학 석사과정.

<주관심분야 : SoC 설계, NoC>



한 태 희(평생회원)
1992년 KAIST 전기 및 전자공학과
학사 졸업.
1994년 KAIST 전기 및 전자공학과
석사 졸업.
1999년 KAIST 전기 및 전자공학과
박사 졸업.

1999년 3월~2006년 8월 삼성전자 통신연구소
책임 연구원.

2006년 9월~2008년 2월 한국산업기술대학교 전
자공학과 조교수.

2008년 3월~현재 성균관대학교 정보통신대학
반도체시스템공학과 부교수.

2011년 5월~2013년 4월 지식경제부 시스템반도
체 PD

<주관심분야 : SoC 아키텍처 및 설계 방법론, 3D
IC, 메모리/스토리지 시스템 구조, 임베디드 SW,
IT 융합 기술>