

논문 2010-05-01

데이터 압축을 이용한 고성능 NoC 구조

(A High Performance NoC Architecture Using Data Compression)

김 홍 식*, 김 현 진, 홍 원 기, 강 성 호

(Hong-Sik Kim, Hyunjin Kim, Won-Gi Hong, Sungho Kang)

Abstract : 본 논문에서는 네트워크 온 칩(NoC: network on chip) 구조에서의 내부 데이터 통신의 성능을 최적화 할 수 있는 새로운 온 칩 네트워크 인터페이스 구조를 제안하였다. 제안하는 NoC 구조는 기본적으로 하드웨어 면적을 줄이기 위하여 XY 라우팅 알고리즘을 기반으로 구현되었으며, 전달되는 패킷의 크기 또는 플릿의 개수를 최소화하기 위하여 Golomb-Rice 인코딩/디코딩 알고리즘에 기반을 둔 하드웨어 압축기/해제기를 이용하여 통신되는 데이터의 양을 크게 줄임으로써 네트워크 지연시간을 최소화 할 수 있는 새로운 구조를 제안하였다. 즉 전송될 데이터는 전송자(sender)의 네트워크 인터페이스에서 내장된 하드웨어 인코더를 통해 압축된 형태로 패킷의 개수를 최소화하여 온 칩 네트워크상의 데이터를 업로드 하게 된다. 이러한 압축된 데이터가 리시버(receiver)에 도착하면, 하드웨어 디코더를 통해서 원래의 데이터로 복원된다. 사이클 수준의 시뮬레이터를 통하여 제안된 라우터 구조가 온 칩 시스템의 네트워크 지연 시간을 크게 줄일 수 있음을 증명하였다.

Keywords : NoC, 네트워크 레이턴시, 시스템 온 칩, 무손실 압축

1. 서론

시스템 온 칩 (system on chip) 기반의 응용분야는 과거에 비해서 더 많은 수의 내장 모듈을 필요로 하고 있고, 이러한 경향은 향후 보다 더 심화될 것으로 예상 된다. 또한 공정 정밀도가 나노 스케일로 미세화 되고 트랜지스터 집적도가 증가됨에 따라서 온 칩 인터커넥션은 SoC 설계의 전력소모 및 성능을 결정하는 매우 중요한 요소로 대두될 것으로 예상 되고 있다. 실제로 나노 공정에서의 신호 지연의 약 70-80%정도가 글로벌 인터커넥트 지연에 기인하고 있는 것으로 측정되고 있다[1]. 이에 따라서 버스 기반의 기존 온 칩 커넥션 구조는 확장성 및 성능제약의 한계 때문에 미래의 SoC 설

계에는 적합하지 않을 것으로 예상되며, 패킷 기반의 온 칩 커뮤니케이션 네트워크 구조를 이용하는 새로운 설계 기술인 네트워크 온 칩 (NoC: network on chip) 설계 기술이 온 칩 인터커넥트의 복잡도를 줄일 수 있는 대안기술로서 대두되고 있다[2-5].

NoC 설계는 기존의 오프 칩 네트워크와는 다르게 온 칩 상에 구현되어야 하기 때문에, 면적 및 전력소모에 대한 제약이 매우 크다. 따라서 XY 라우팅 [5]과 같은 가능한 간단한 형태의 라우팅 알고리즘을 사용하게 되는데, 이로 인하여 NoC 구조에서는 네트워크 상태에 따라서 효율적으로 라우팅 알고리즘이 대처하는 것이 어렵기 때문에 네트워크 성능 저하의 문제점이 발생하게 된다.

본 논문에서는 이러한 NoC 라우팅 구조의 한계를 극복하기 위해서 무손실 데이터 압축 기술을 적용하여 전송되는 데이터 패킷의 양을 줄임으로써 온 칩 네트워크의 성능을 향상시킬 수 있는 새로운 NoC 구조를 제안하였다. 이와 같은 온 칩 네트워크 설계 기술에 적용되는 데이터 압축 기술은 기본적으로 무손실 압축기술에 사용되어야 하는데, 본 논문에서는 Golomb-Rice 알고리즘[10]에 기반한 압축 기술을 하드웨어로 구현하였다. 제안하는 데이

* 교신저자(Corresponding Author)

논문접수 : 2009. 12. 22., 수정일 : 2010. 01. 08., 채택확정 : 2010. 2. 2.

김홍식, 김현진, 강성호 : 연세대학교전기전자공학과
홍원기 : 대구대학교 정보통신공학과

※ 이 논문은 2008년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No.2008-0057698)

터 압축 기술 기반의 NoC 구조는 기존의 라우터 구조에 대한 변경 없이 데이터 압축과 해제를 위한 하드웨어를 네트워크 인터페이스 블록 내에 내장함으로써 간단히 구현이 가능하기 때문에, 기존의 다양한 NoC 구조에 쉽게 적용될 수 있을 것이다. 사이클 수준의 시뮬레이터를 통한 실험에 의하면 기존의 방법에 비해서 적은 하드웨어 면적의 추가를 통해서 19.7%의 네트워크 성능을 향상시킬 수 있음을 증명하였다.

II. 기존의 NoC 구조에 관한 연구 정리

1. NoC의 개념

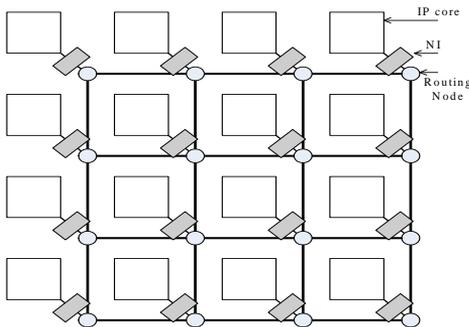


그림 1. 4x4 메쉬구조 기반의 NoC 구조의 예
Fig. 1. An NoC example based on 4x4 mesh architecture

그림 1은 4x4 메쉬 구조를 기반으로 하는 일반적인 NoC 구조를 설명하는 그림이다. 기존의 버스 및 개별적인 포인트 투 포인트 방식의 연결 대신에, SoC 칩 내부에 라우팅 노드를 배치하고 통신 링크를 통해서 서로 연결하는 방식으로 글로벌 커뮤니케이션을 제공하는 방법이 NoC 구조라고 할 수 있다. 이와 같은 NoC 구조는 일반적으로 다음과 같은 세 가지 내부 블록으로 구성된다.

-네트워크 인터페이스(NI: network interface): NI는 내장되는 개별 IP 코어와 NoC의 라우팅 노드에 대한 연결을 제공한다. IP 코어에서 전달받은 데이터를 NoC 프로토콜에 적합한 패키지로 변환하는 역할과, NoC 패킷 데이터를 원래의 데이터 형식으로 변환하는 역할을 포함한다.

-라우팅 노드(routing node): NoC의 프로토콜

에 따라서 NoC 상에 올려진 데이터를 라우팅한다. NoC에서의 라우팅 노드는 온 칩 상에 구현되기 때문에 가능한 단순한 구조를 갖는 것이 일반적이다.

-링크(link): 라우팅 노드를 연결하는 메탈라인으로 정해진 대역폭을 가지고 구현되며, 경우에 따라서 다수의 논리적/물리적 채널로 구현될 수 있다.

2. NoC 분야의 연구동향

NoC 기반의 SoC는 전력소모 및 하드웨어 면적에 대한 제약 조건이 매우 크기 때문에, 이를 고려하여 최적의 성능을 발휘할 수 있는 NoC 구조에 관한 연구가 이루어져 왔다. 초기에는 온 칩 라우터 구조에 관한 연구 및 네트워크 토폴로지에 관한 연구[2-4]가 주를 이루었으며, 최근에는 저전력 구조[6], 신뢰성을 고려한 설계 기술 [7]등 다양한 분야로 연구가 확산되고 있다.

NoC 성능 향상의 관점에서는 먼저 결정 라우팅 알고리즘 기반의 라우터구조에 관한 연구에서 출발하여 가상 채널을 이용한 적응형 라우팅 기술 [8], 파이프라인 채널 구조 [9] 등의 다양한 고성능 온 칩 라우터 구조들이 제안되어 왔다. 하지만 이와 같은 구조들에 대한 연구들은 기본적으로 온 칩 자원 제약의 한계를 벗어날 수 없기 때문에, 성능향상의 폭이 제한적인 것이 사실이다. 따라서 본 논문에서는 데이터 압축 기술을 응용하여 기존 NoC 구조의 성능을 더욱 향상시킬 수 있는 새로운 방법론을 제안하고자 한다.

III. 제안하는 NoC 구조

1. 제안하는 전체 NoC 구조

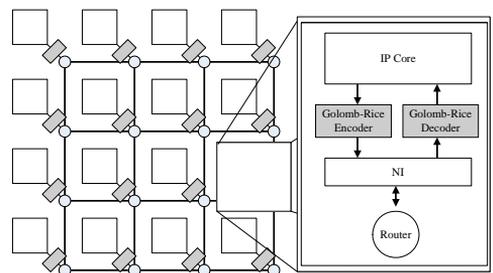


그림 2. 제안하는 NoC 구조
Fig. 2. The proposed NoC architecture

그림 2는 제안하는 NoC 구조를 설명하는 그림이다. 그림에서는 4x4 메쉬 구조를 기반으로 하여 구조를 설명하였으나, 제안하는 방법은 이 밖의 다양한 NoC 구조에도 적용이 가능하다. 기본적으로 NoC 구조에서는 개별 IP 코어에서 데이터 전송이

함으로써 압축을 수행하게 된다. 이 때 매개변수 즉 계수가 2의 거듭제곱이 되는 특수한 경우에는 인코딩 절차가 매우 간단해 지게 되는데, 나눗셈 동작 대신에 쉬프트 동작으로 간단하게 몫과 나머지를 구할 수 있게 된다. 이와 같이 매개변수가 2의 거듭제곱인

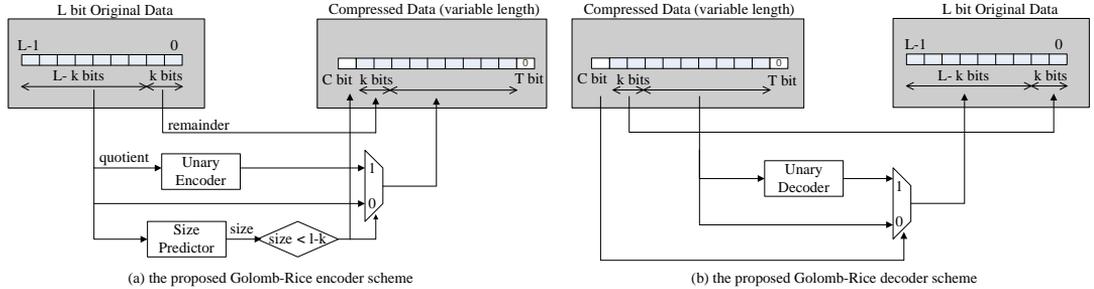


그림 3. 제안하는 Golomb-Rice 인코더 및 디코더의 하드웨어 구조

Fig. 3. The proposed golomb-rice encoder and decoder hardware architecture

이슈되면, 이 데이터는 NI(network interface)를 통해서 패킷 형태로 변환되어 온 칩 라우터에 전달되고, 온 칩 네트워크 자원에 의해서 원하는 목적지의 IP 코어에 데이터가 전달되게 된다.

제안하는 NoC 구조에서는 전송자(sender) IP 코어로부터 데이터 전송이 이슈 되면, NI 입력 부분의 Golomb-Rice 인코더를 이용하여 입력 데이터를 압축하여 전체 패킷의 개수를 줄이게 된다. 압축된 패킷이 온 칩 네트워크를 통해서 리시버에 전달되면, 리시버 (receiver) NI의 Golomb-Rice 디코더를 이용하여 원본 데이터를 복원하게 된다. 이와 같은 압축 기술을 이용한 NoC 구조는 기존의 다양한 라우터 구조와 NoC 토폴로지에 상관없이 적용이 될 수 있는 성능향상에 매우 효율적인 구조라고 할 수 있다.

2. Golomb-Rice 인코더 및 디코더

제안하는 데이터 압축 기술을 NoC 성능향상을 위하여 적용하기 위해서는 기본적으로 무손실 압축 알고리즘을 이용하여야 한다. 이를 위하여 본 논문에서는 Golomb-Rice 알고리즘 [10]을 이용한 무손실 압축 기술을 적용하였으며, 이를 위한 효율적인 하드웨어 구조를 제안하였다.

2.1 Golomb-Rice 알고리즘

먼저 Golomb 코딩에서는 입력 데이터 값을 n 이라고 하면, n 을 매개 변수 m 으로 나눈 몫은 일진(unary) 코딩을 하고 나머지는 이진(binary) 코딩을

경우의 Golomb 코딩을 Golomb-Rice 코딩이라고 한다. 예를 들어서 8비트 입력 데이터 14 (이진수 00001110)를 Golomb-Rice 코딩 (매개변수 22)으로 인코딩한다고 가정하자. 이때 매개변수가 2의 거듭제곱이므로 입력값의 상위 6비트(000011)는 몫이 되고 하위 2비트(10)은 나머지게 된다. 따라서 Golomb-Rice 코딩에 의하면 몫은 이진코딩, 즉 (1110)으로 코딩된다. 몫이 십진수로는 3이기 때문에 연속된 세 개의 1로 일진코딩을 하고 구분자를 위해서 0을 붙여준다. 그리고 나머지는 이진코딩이 되므로 2 비트 (10)을 그대로 사용된다. 따라서 총 비트 수는 6비트로 2비트가 줄게 된다.

2.2 하드웨어 인코더 및 디코더 구조

Golomb-Rice 알고리즘은 무손실 압축알고리즘으로서 알고리즘 복잡도가 매우 낮기 때문에 하드웨어 설계에 적합한 압축 기술이라고 할 수 있다. 따라서 제안하는 NoC 구조에서는 이와 같은 Golomb-Rice 인코딩 및 디코딩 알고리즘을 하드웨어로 구현하여 패킷 압축을 위하여 사용했다. 그림 3은 제안하는 Golomb-Rice 인코더와 디코더의 구조를 설명하는 그림이다.

기본적으로 Golomb-Rice 인코딩 알고리즘의 경우 입력 값을 매개변수로 나눈 몫과 나머지를 각각 가변길이 및 고정 길이로 인코딩하게 된다. 따라서 나머지의 경우는 압축 데이터의 길이가 항상 매개변수의 거듭제곱 값으로 일정하지만, 몫의 경우는 그 길이가 가변적이게 된다. 최악의 경우 압축된 데이터

의 길이가 원본 데이터의 길이보다 길어지는 경우도 낮은 확률로 발생하게 되는데, 이와 같은 문제를 방지하기 위해서 본 논문에서는 길이 예측 로직(size predictor)를 사용하여, 압축되어질 몫의 길이를 계산하여 이 값이 원본의 길이 (L-k 비트) 보다 길거나 같게 되면, 압축데이터를 출력하는 대신, 원본 데이터를 압축없이 출력하도록 하였다. 따라서 압축 및 미압축의 경우를 구분하기 위해서 C 비트가 추가되게 된다. 또한 압축 데이터의 길이가 가변적이기 때문에, 현재 데이터의 종료여부를 알려주기 위하여 T 비트를 이용하였다.

Golomb-Rice 디코딩의 경우는 입력 데이터가 압축 데이터인지 미압축 데이터인지를 구분하기 위한 C 비트를 먼저 확인하여, 압축 데이터인 경우 (C 비트가 1인 경우), 몫에 대한 가변길이 압축데이터에 대한 일진 디코딩을 수행하여 고정길이 나머지와의 연결시킴으로써 원본 데이터를 복원하게 된다. 만약 미압축 데이터인 경우(C 비트가 0인 경우)는, 입력데이터 중 C 비트와 T 비트를 제거한 뒤 원본 데이터의 몫과 나머지의 순서를 복원하는 방법으로 간단히 디코딩 동작을 마무리하게 된다. 압축율이 1보다 매우 크기 때문에, 디코더가 한 번에 입력 받은 데이터 내에 T 비트가 다수개가 존재하게 된다. 또한 현재 디코딩 중인 데이터의 T 비트가 다음 입력 패킷에 존재하는 경우도 발생할 수 있다. 따라서 인코더와 디코더 내에는 이와 같은 상황을 처리할 수 있도록 적절한 레지스터링을 통해서 오류 없는 데이터 패킷/디패킷 동작을 수행할 수 있도록 하드웨어를 설계하고 검증 하였다.

IV. 실험결과

1. 실험환경

1.1 사이클 수준의 NoC 시뮬레이터

제안하는 방법의 성능향상 정도를 평가하기 위해서 본 논문에서는 사이클 수준의 NoC 시뮬레이터를 C 언어로 제작하였다. 제작된 시뮬레이터는 개별 구성 블록들을 하드웨어와 유사하게 독립적으로 구현함으로써 클럭 사이클 단위로 제안된 NoC 구조를 실험할 수 있도록 설계되었다. 내부 IP 코어들은 랜덤으로 데이터 패킷을 발생시키고, 이 패킷들은 제안된 압축 알고리즘에 의해서 인코딩되어 온 칩 네트워크를 통해서 리시버 IP 코어로 전달된다. 전달된 데이터 패킷은 제안된 디코딩 알고리즘에 의해서 디코

딩되어 리시버 IP 코어에 전달된다. 이때 네트워크 레이턴시는 패킷의 생성 시점에서, 온 칩 네트워크 상에서의 버퍼링 및 라우팅을 통해서 최종적으로 압축해제까지 걸리는 클럭 사이클로써 측정되었다.

2. 실험결과

2.1 제안하는 구조의 네트워크 레이턴시

제안하는 구조의 성능을 평가하기 위한 NoC 구조로서, 기본적인 8x8 메쉬 구조를 이용하였다. 개별 IP 코어와 NI와의 인터페이스와 온 칩 네트워크 상의 물리적 링크의 대역폭은 모두 4바이트로 가정하였다. 실험을 위해서 전부 1,000,000개의 랜덤 데이터를 발생시켰고, 개별 데이터의 크기 또한 최소 4바이트에서 32바이트 사이에서 4바이트 단위로 랜덤하게 생성하였다. 정해진 수의 데이터가 모두 리시버에 전달될 때 까지 데이터 생성, 데이터 압축, 패킷화, 네트워크, 디패킷화, 압축해제를 반복실험하면서 개별 데이터의 전송 시간을 클럭 사이클로 계산하여 평균을 측정하였다.

표 1. 제안하는 구조의 성능평가 결과

Table 1. Experimental results

패킷 삽입율	네트워크 레이턴시 (클럭사이클)	
	8x8 메쉬 기반 기존의 NoC	제안하는 NoC (8x8 메쉬)
0.1	35.3	29.9
0.2	37.2	30.8
0.3	39.8	32.0
0.4	103.2	80.4
평균	53.9	43.3

표1은 제안하는 압축 기술을 이용한 온 칩 네트워크의 성능향상을 설명하는 표이다. 패킷 삽입율 (한 사이클 당 각각의 라우팅 노드에 삽입되는 패킷의 개수, 예를 들어 8x8 메쉬 구조의 경우 0.4의 패킷 삽입율은 전체 16개의 라우팅 노드 중 평균적으로 한 사이클 당 6.4개의 노드가 패킷을 입력 받고 있음을 의미한다.) 모든 경우에 대해서 제안하는 방법의 네트워크 레이턴시가 매우 적은 것을 알 수 있다. 평균적으로 19.7%의 네트워크 레이턴시를 향상시킬 수 있음을 증명하였다.

2.2 제안하는 구조의 하드웨어 오버헤드

제안하는 압축기술은 상대적으로 매우 간단한 알고리즘인 Golomb-Rice 구조를 이용하기 때문에

하드웨어 면적이 매우 작은 수준이라고 할 수 있다. 본 논문에서는 제안하는 구조의 하드웨어 오버헤드를 평가하기 위해서 제안하는 구조를 Verilog HDL을 이용하여 RTL (register transfer level) 설계를 수행하여 Synopsys의 Design Vision을 이용하여 게이트 수준으로 합성함으로써 면적과 타이밍 오버헤드를 계산하였다. 합성에 사용된 라이브러리는 CMOS 0.13 마이크로 공정을 이용하였다. 하드웨어 면적은 2-입력 낸드 게이트 기준으로 인코더 732개, 디코더 827개의 게이트 수를 보였으며, 크리티컬 패스 딜레이는 인코더 3.3ns, 디코더 3.4ns 로 측정되었다.

V. 결 론

본 논문은 효율적인 하드웨어 기반의 무손실 데이터 압축 기술을 이용하여 NoC 구조의 네트워크 레이턴시를 향상 시킬 수 있는 새로운 온 칩 네트워크 구조를 제안하였다. 무손실 압축 기술을 위하여 Golomb-Rice 알고리즘에 대한 하드웨어 구조를 제안하고 이를 구현하여 실험을 통해 검증을 증명하였고, 사이클 기반의 시뮬레이터를 제작하여 제안하는 구조의 효율성을 실험하였다. 제안된 구조는 기존의 모든 NoC 토폴로지에 NI 입출력에 대한 간단한 추가를 통해서 구현이 가능한 장점이 있다. 실험 결과 제안하는 압축 기술을 적용한 NoC 구조는 적은 하드웨어 오버헤드를 통해서 온 칩 네트워크 성능을 매우 크게 향상 시킬 수 있음을 증명하였다.

참고문헌

- [1] J. Kim, D. Park, T. Theocharides, N. Vijaykrishnan, and C. Das, "A low latency router supporting adaptivity for On-Chip Interconnects", Proc. of Design Automation Conference, pp. 559-564, Jun. 2005.
- [2] P. Guerrier and A. Greiner, "A generic architecture for On-Chip packet-switched interconnections", Proc. of Design and Test in Europe, pp. 250-256, Mar. 2000.
- [3] S. Kumar et al., "A network on chip architecture and design methodology", Proc. International Symposium on VLSI, pp. 117-124, 2002.
- [4] F. Karim et al, "An interconnect architecture for networking systems on chips", IEEE Micro, Vol.22, No.5, pp. 36-45, 2002.
- [5] F. Moraes, N. Calazans, L. Moller, and L. Ost, "HERMES: An infrastructure for low area overhead packet-switching networks on chip", the VLSI integration, Vol.38, pp. 69-93, 2004.
- [6] N. Easley and L. -S. Phe, "High-level power analysis for on-chip networks", Proc. of CASES, pp. 104-115, Nov. 2004.
- [7] C. Erika, K. F. Lima, C. Maico, A. Pedro, M. Paulo, A. Alexandre, L. Marcelo, "A high-fault-coverage approach for the test of data, control and handshake interconnects in mesh networks-on-chip", IEEE Trans. Computers, Vol.57, No.9, pp. 1202-1215, 2008.
- [8] J. Duato and T. M. Prinkston, "A general theory for deadlock-free adaptive routing using a mixed set of resources", IEEE Trans. on Parallel Distributive Systems, Vol.12, No.12, pp. 1219-1235, 2001.
- [9] A. Jalabert, S. Murali, L. Benini and G. D. Micheli, "Xpipes compiler: a tool for instantiating application specific networks on chip", Proc. of Design and Test in Europe, pp. 884-889, 2004.
- [10] R. F. Rice, "Some practical universal noiseless coding techniques", Jet Propulsion Laboratory, California Institute of Technology, Pasadena, Technical report, pp. 79-22, 1979.

저 자 소 개

김 홍 식



1997년 2월 : 연세대학교 전기공학과 학사.
1999년 8월 : 연세대학교 전기 및 컴퓨터공학과 석사.
2004년 8월 : 연세대학교 전기전자공학과 박사.

2007년~현재, 연세대학교 연구교수.
관심분야 : NoC, DFT, 데이터압축, 그래픽 가속기.

Email : hongsik@yonsei.ac.kr

김 현 진



1997년 2월 : 연세대학교 전기공학과 학사.
1999년 8월 : 연세대학교 전기 및 컴퓨터공학과 석사.
현재, 연세대학교 전기전자공학과 박사과정.

관심분야 : IDS, DVS.

Email : wkhong@daegu.ac.kr

홍 원 기



1995년 2월 : 연세대학교 전산과학과 학사.
1997년 2월 : 연세대학교 컴퓨터과학과 석사.
2001년 8월 : 연세대학교 컴퓨터과학과 박사.

현재, 대구대학교 정보통신공학부 조교수.
관심분야 : 임베디드시스템, NoC설계방법론.
Email : wkhong@daegu.ac.kr

강 성 호



1986년 2월 : 서울대학교 제어공학과 학사.
1988년 2월 : University of Texas at Austin 전기 및 컴퓨터공학과 석사.
1992년 2월 : University of Texas at Austin 전기 및 컴퓨터공학과 박사.

현재, 연세대학교 전기전자공학과 교수.

관심분야 : SoC 설계 및 테스트.

Email : shkang@yonsei.ac.kr