

2024 IEEE VLSI Review

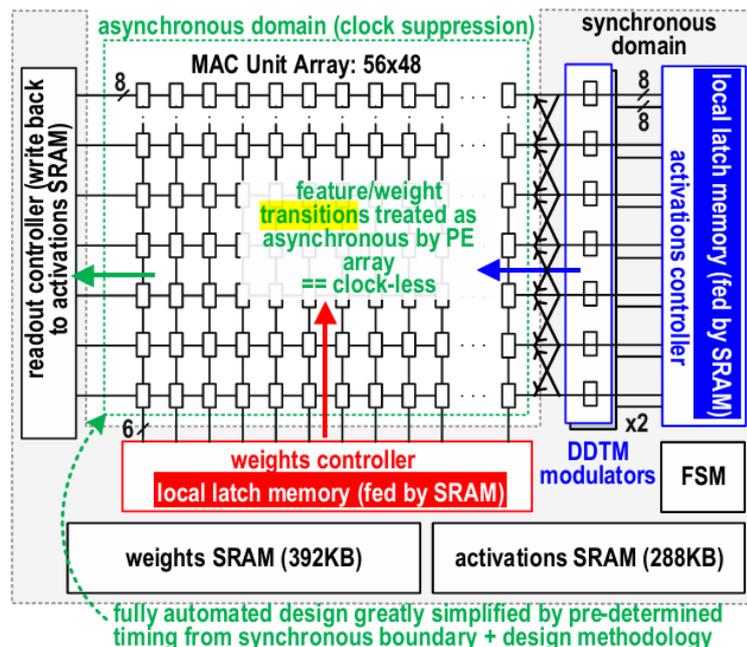
한양대학교 신소재공학과 석박통합과정 송충석

Session 3 AI/ML Accelerators and CiM

이번 2024 IEEE CICC의 Session C3은 AI/ML Accelerators and CiM라는 주제로 총 5편의 논문이 발표되었다. 본 review에서는 3-1, 3-4, 3-5를 리뷰하고자 한다.

#3-1 논문에서는 Dyadic Digital Transition Modulation (DDTM) 이라는 새로운 데이터 표현 방식을 기반으로 한 DNN 가속기 구조를 발표하였다. 디지털 기반 인공신경망을 위한 transition density를 통해 저전력 아키텍처를 목표로 하고 있으며 비동기식 카운터를 이용하여 MAC 연산을 간단히 하고 높은 에너지 효율성을 이루었다. 비동기식 카운터를 사용하기 때문에 클럭기반으로 동작하지 않고, pseudo-sparsity를 적용하여 MAC 연산기의 activity를 줄였다. 해당 논문은 40nm 공정을 이용하여 100TOPS/W 이상의 성능을 기록하였다.

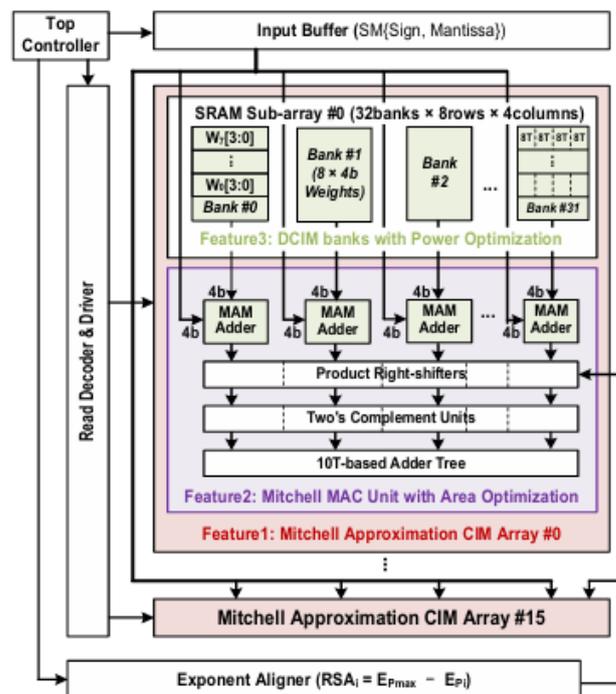
본 논문의 아키텍처에 이용된 비동기식 MAC array는 카운터의 동작을 98.4배 감소시켜 2bit 연산에 368.9 TOPS/W의 성능을 나타내었다. 효율을 높이기 위해 설계된 비동기식 MAC 연산기를 통해 하락된 정확도는 재학습을 통해 복구하게 된다.



[그림 3-1] 논문 3-1에서 제안한 전체 아키텍처

#3-4 논문에서는 Mitchell's Approximate Multiply (MAM) 기술을 적용하여 에너지와 면적 효율을 증가시킨 디지털 기반 CIM을 발표하였다. MAM은 기존에 CIM에서 사용하는 bit-serial 기반 곱셈방식으로 MAC 연산을 하지만, 본 논문은 더욱 효율적인 방법을 이용하여 부동소수점 연산을 구현하였다. MAM 방식은 복잡한 부동소수점 곱셈을 간단한 부동소수점 덧셈으로 근사화하는 방법으로 n bit 연산을 위해 n번의 cycle이 필요한 CIM 곱셈을 한번에 처리함으로써 처리량이 n 배가 증가하게 된다.

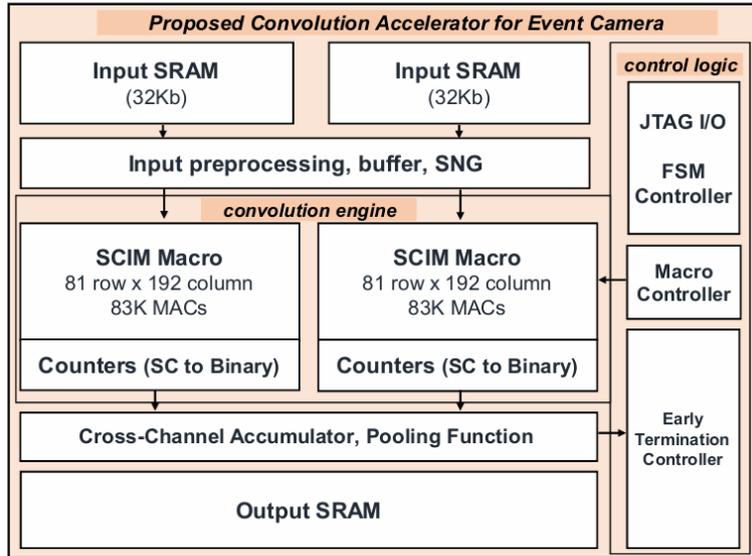
본 논문의 디지털 기반 CIM 매크로는 BF16에서 65.15TFLOPS/W의 에너지 효율과 3.04 TFLOPS/mm²의 면적 효율을 나타내었다. VGG16과 ResNet20 모델에서 CIFAR-10 데이터셋을 이용해 각각 90.07%, 91.71%의 정확도를 보였고, ResNet50 모델에서 ImageNet 데이터셋으로 80.52%의 정확도를 달성했다.



[그림 3-4] 논문 3-4에서 제안한 전체 아키텍처

#3-5 논문에서는 이벤트 카메라를 위한 ADC-less CIM 기반 컨볼루션 가속기를 발표하였다. CIM 메모리에 binary data를 저장하고 stochastic number generator (SNG)를 사용하여 실시간으로 probability 비트로 변환함으로써 기존 방식보다 10배 이상 적은 저장 공간을 사용하고 각 weight 마다 32개의 작은 MAC 연산기를 내장하고, 불필요한 0 계산을 건너뛰는 방식으로 효율성을 높였다.

본 논문에서 발표한 아키텍처는 485 TOPS/W의 에너지 효율과 278-514 Mevent/s의 처리량을 보여주며 기존의 이벤트 카메라를 활용한 가속기보다 60배 이상의 처리량을 기록하였다. 비 이벤트 모드에서도 다른 CIM 아키텍처보다 2~3배 높은 에너지 효율성을 보여주었으며 처리량은 10배 이상을 달성하였다.



[그림 3-5] 논문 3-5에서 제안한 전체 아키텍처

저자정보



송충석 석박통합과정 대학원생

- 소속 : 한양대학교
- 연구분야 : 딥러닝 가속기 설계
- 이메일 : scs940430@naver.com
- 홈페이지 : <https://sites.google.com/site/dsjeonglab1>

2024 IEEE VLSI Review

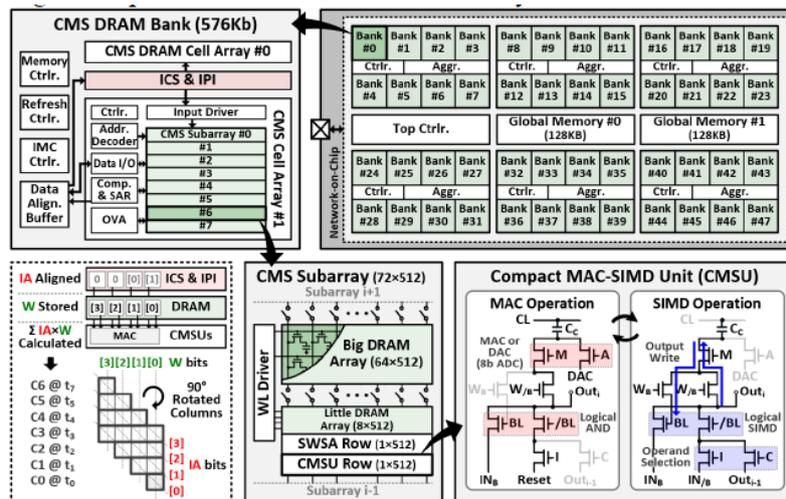
한양대학교 신소재공학과 석박통합과정 송충석

Session 20 Processing for AI

이번 2024 IEEE CICC의 Session C20은 Processing for AI 라는 주제로 총 4편의 논문이 발표되었다. 본 review에서는 20-1, 20-2, 20-3을 리뷰하고자 한다.

#20-1 논문에서는 Dymond라는 새로운 1T1C DRAM 기반 CIM을 발표했다. Dymond는 column방향으로 accumulation을 하는 column addition (CA) 데이터플로우를 사용하고 높은 메모리 사용 효율성과 에너지 효율성을 기록하였다. LSB쪽 연산에 사용되는 LSB-CA는 ADC 사용을 최소화 하여 에너지 효율을 높이고, MSB쪽 연산에 사용되는 MSB-CA는 signal enhanced MAC과 signal shifted ADC를 통해 SQNR을 향상시켰다. 또한 switchable sense amplifier를 사용하여 저전력 기반 CIM 연산을 구현하였다.

28nm CMOS 기반으로 제작된 본 논문의 Dymond는 6.48mm^2 면적에 27Mb DRAM 메모리를 implementation하여 최대 27.2 TOPS/W의 에너지 효율성을 달성하였으며 ResNet, BERT, GPT-2 와 같은 최신모델에서도 성능을 입증하였다. 기존의 eDRAM 기반 CIM과 비교하여 Dymond는 1.8배 높은 효율성을 보였고, figure of merit은 7.9배 높였다.

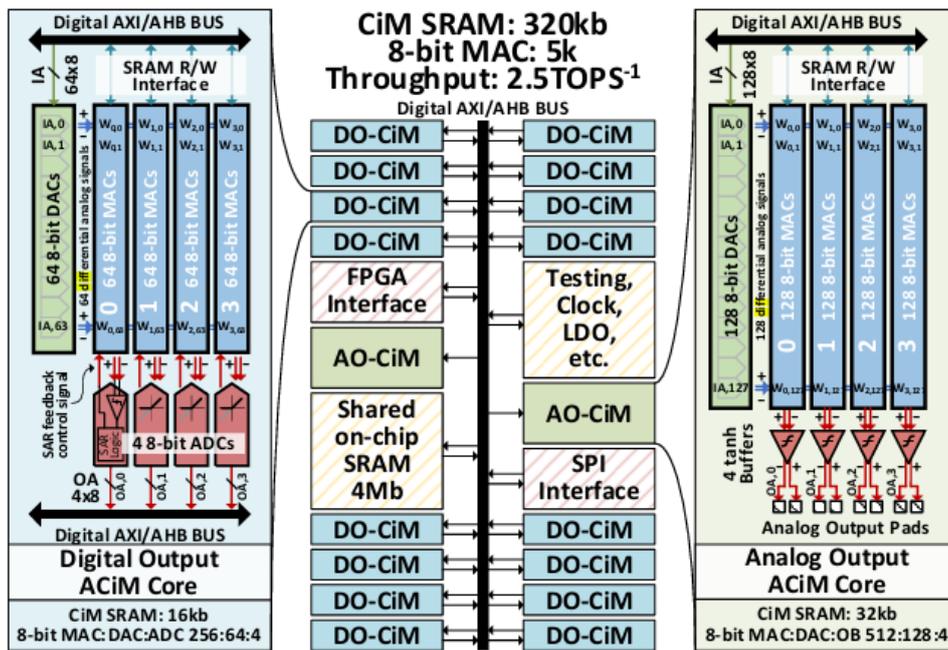


[그림 20-1] 논문 20-1에서 제안한 전체 아키텍처

#20-2 논문에서는 아날로그 컴퓨팅 방식 CIM에서 발생하는 빈번한 데이터 변환 문제를 해결하기 위해 signed 8bit MAC 연산기를 새롭게 제안하였다. 더불어 ReLU는 SAR-ADC에서 LSB 스킵 방식으로, tanh은 아날로그 버퍼를 통해 구현하여 이를 MAC 연산기에 통

합시켜 효율성을 극대화시켰다.

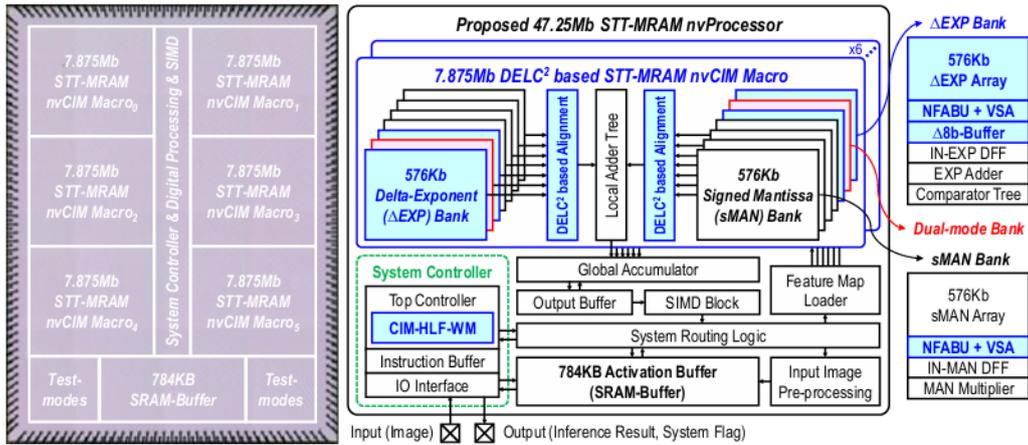
본 논문에서 제안한 프로세서는 디지털 및 아날로그 연산코어를 포함한 하이브리드 형태를 띠고 있으며 2.5TOPS의 throughput을 달성하였다. 64개의 8bit DAC, 16Kb SRAM 셀, 256개의 8bit MAC 유닛 및 4개의 8bit ADC가 통합된 DO-CiM 코어와, 데이터 변환기를 사용하지 않고 tanh 활성화 함수를 사용할 수 있는 AO-CiM 코어를 통해 에너지 효율과 면적 효율을 상승시켰다. 실험 결과, 55TOPS/W의 성능을 보이며, AO-CiM 코어에서는 최대 104.5 TOPS/W 를 기록하였다.



[그림 20-2] 논문 20-2 에서 제안한 전체 아키텍처

#20-3 논문은 22nm CMOS 공정 기반 AI 엣지 프로세서를 제안하였다. 비휘발성 메모리인 STT-MRAM을 사용하여 near memory computing을 구현하여 47.25Mb 메모리의 21.4TFLOPS/W 의 에너지 효율성을 기록하였다. 연산 효율성을 위해 compression을 데이터 손실 없이 구현하였다.

본 논문에서 개발한 프로세서는 428.58us의 짧은 응답속도를 가지며 압축된 데이터를 만들기 위한 하드웨어 오버헤드를 줄이고 SRAM 버퍼 접근을 감소시켜 전체적인 시스템 성능을 향상시켰다. 이러한 방법으로 인해 기존 비휘발성 메모리 프로세서보다 2.55배 이상 높은 시스템 수준 에너지 효율성을 달성했다.



[그림 20-3] 논문 20-3 에서 제안한 전체 아키텍처

저자정보



송충석 석박통합과정 대학원생

- 소속 : 한양대학교
- 연구분야 : 딥러닝 가속기 설계
- 이메일 : scs940430@naver.com
- 홈페이지 : <https://sites.google.com/site/dsjeonglab1>

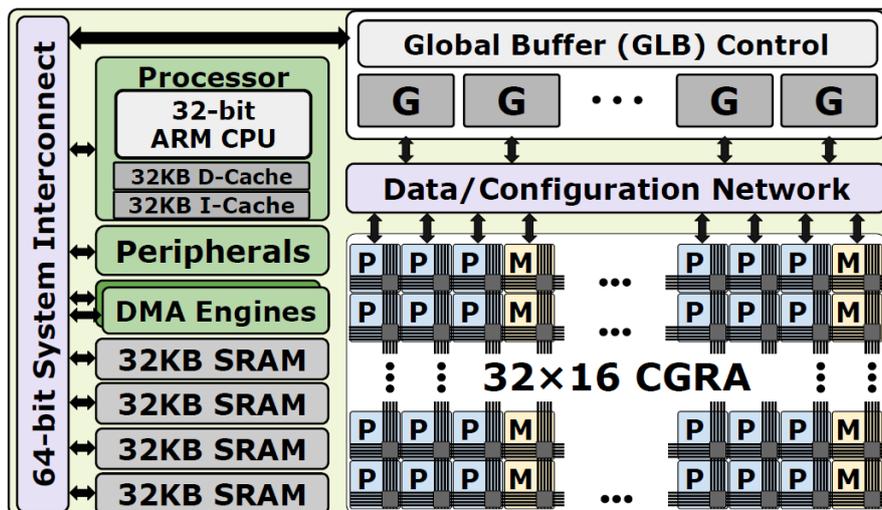
2024 IEEE VLSI Review

KAIST 전기및전자공학부 박사과정 엄소연

Session 7 Processors I

이번 2024 VLSI의 Session C7은 Processor라는 주제로 총 4편의 논문이 발표되었다. 이 세션에서는 dense/sparse application, diffusion model, stochastic analog SAT solver, 그리고 stencil/sparse linear algebra를 위한 하드웨어가 발표되었다. 이번 후기를 통해 4개의 논문에 대해 간략하게 살펴보고자 한다.

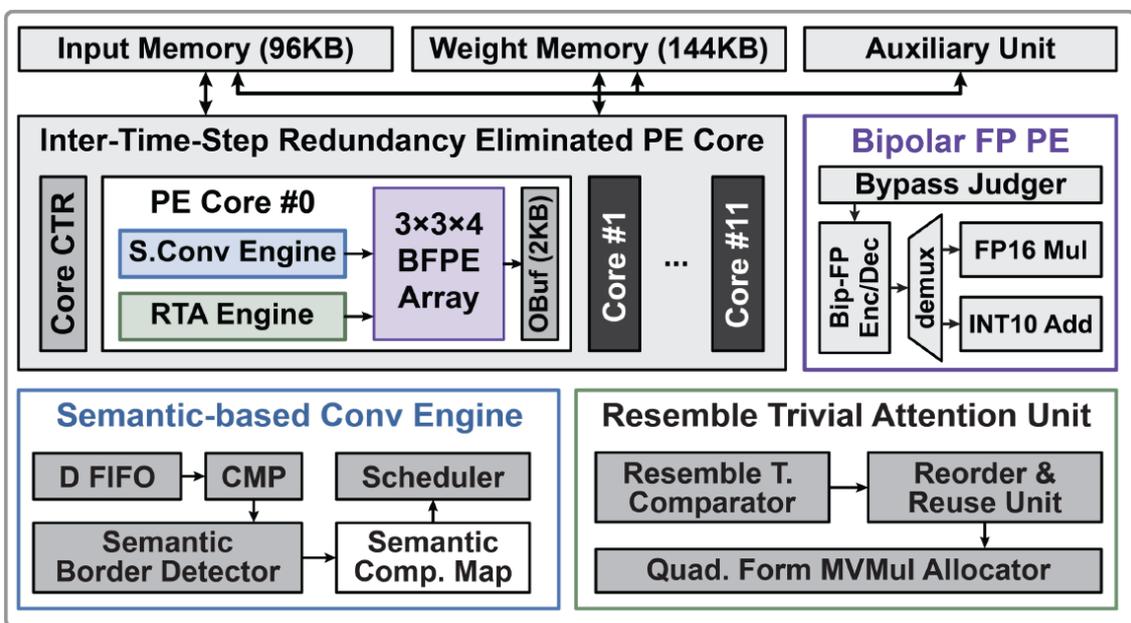
#7-1 이 논문은 Stanford University에서 발표한 연구로, Onyx라는 최초의 완전 프로그래머블 Coarse-Grained Reconfigurable Array (CGRA)를 소개한다. Onyx는 밀집(dense) 및 희소(sparse) 응용 프로그램 모두를 가속할 수 있으며, 임의의 고차원 텐서 연산을 지원한다. Onyx는 압축된 텐서를 처리할 수 있는 메모리 및 계산 프리미티브를 사용하여 불필요한 연산을 제거하고, 이미지 처리 및 머신러닝에서 효율성을 높인다. 28nm 공정에서 제작된 Onyx SoC는 384개의 프로세싱 타일과 128개의 메모리 타일을 사용하며, 메모리 타일은 피버트리(fibertree) 기반 스트리밍 구조로 희소 데이터를 효율적으로 처리한다. 제안된 시스템은 희소 행렬 곱셈에서 CPU 대비 최대 565배, 이미지 처리와 머신러닝에서 각각 76%, 85%의 에너지-지연 제품(EDP) 개선을 달성한다.



[그림 1] Onyx 시스템-온-칩 아키텍처

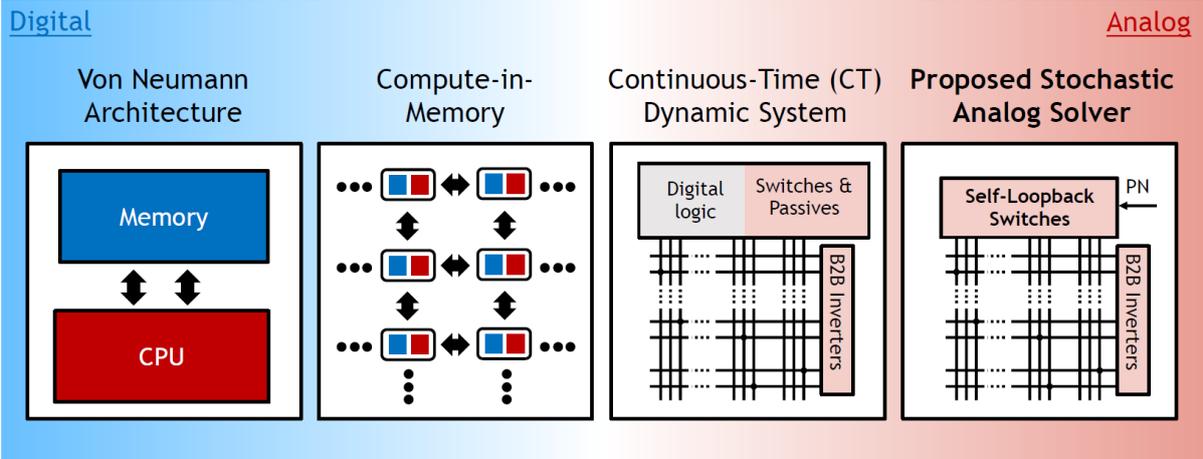
#7-2 이 논문은 Tsinghua University에서 발표한 에너지 효율적인 디퓨전 모델 프로세서를 소개하며, 22nm CMOS 공정에서 제작되었습니다. 이 프로세서는 시간 단계 간의 중복

을 제거하는 방식을 활용하여 성능과 에너지 효율을 크게 향상시킵니다. 주요 특징으로 는 세 가지가 있습니다. 첫째, semantic-segment sparse convolution 엔진을 통해 중요하 지 않은 이미지 부분에서 발생하는 중복된 convolution 연산을 88.5%까지 제거하여, 중 요한 부분만 계산함으로써 연산 효율을 극대화합니다. 둘째, resemble trivial attention exponent inheritance 기법을 통해 어텐션 레이어의 계산 중복을 줄여 16.7배의 연산 효 율을 달성합니다. 셋째, bipolar floating-point multiplier는 mantissa 곱셈의 비효율성을 줄 여 25.4%의 연산량을 절감합니다. 결과적으로 이 프로세서는 CIFAR-10 및 ImageNet 데 이터셋을 대상으로 실험되었으며, 실험 결과 평균 52.01 TFLOPS/W의 에너지 효율을 달 성했습니다. 더불어 이미지 품질 손실은 1% 미만으로 유지되었습니다. 이 diffusion model 프로세서는 convolution layer와 attention layer에서 각각 11.76배, 22.51배의 에너 지 효율 향상을 기록했으며, 기존 가속기 대비 최대 23.14배의 성능 향상을 보여줍니다.



[그림 2] 디퓨전 모델 아키텍처

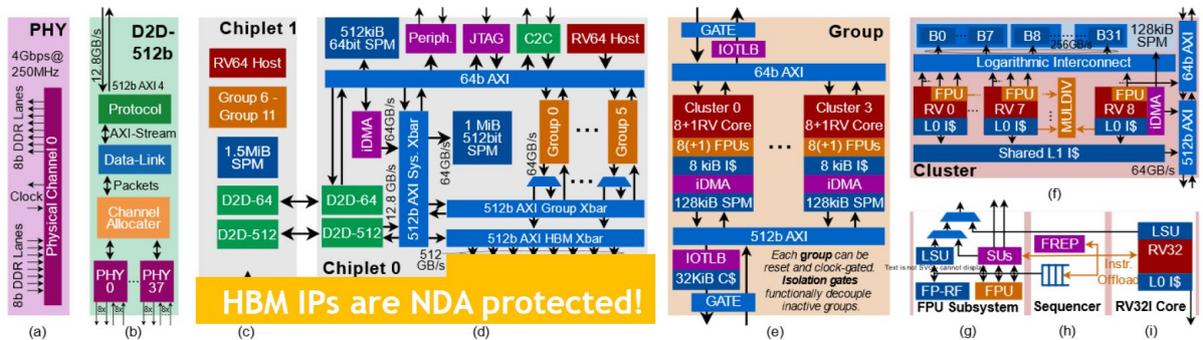
#7-3 이 논문은 65nm CMOS 공정에서 구현된 확률적 아날로그 SAT 솔버를 제시하며, 기존 디지털 및 아날로그 SAT 솔버에 비해 속도와 에너지 효율성을 개선했다. 제안된 솔 버는 병렬 연산을 가능하게 하는 아날로그 오픈 루프 구조를 사용하며, 연속 시간 셀프 루프백 풀업 스위치와 Pseudo Noise(PN) 기반 스크램블링 기법을 도입해 해결 가능성을 높인다. 하이브리드 PN 생성기를 통해 비용을 절감하면서 성능 최적화를 달성했다. 제안 된 솔버는 3-SAT 문제 해결을 위한 크로스바 구조를 기반으로 하며, 각 열은 변수, 각 행은 절을 나타낸다. 3비트 PN 제어 코드를 사용하여 절을 난수화된 상태로 제어하며, 이를 통해 문제 해결 가능성을 높인다. 그 결과, 기존 아날로그 SAT 솔버보다 1000배, 디 지털 SAT 솔버보다 10배 빠른 속도를 기록했으며, 평균 해결 시간은 6.6 μ s였다.



- Long solving time
- High energy consumption
- Energy-efficient
- Long critical path
- Low solvability
- Energy-efficient
- Fast solving speed
- Energy-efficient

[그림 3] SAT Solver 카테고리

#7-4 이 논문은 희소 선형대수와 스텐실 연산에 최적화된 432코어 RISC-V 기반 듀얼 칩렛 시스템 Occamy를 소개한다. Occamy는 FP64, FP32, FP16, FP8 연산을 지원하며, 희소하고 불규칙한 메모리 접근을 효율적으로 처리하기 위해 설계되었다. 각 칩렛은 16GiB HBM2E 메모리 스택과 216개의 RISC-V 코어로 구성되어 있으며, 클러스터당 64 GiB/s의 높은 메모리 대역폭을 제공한다. 이 시스템의 주요 특징은 다양한 precision 연산 코어와 희소 스트리밍 유닛을 사용하여 간접 주소 지정과 희소 데이터 병합을 가속화하는 것이다. 또한, 확장 가능한 지연 허용 아키텍처와 분산 DMA 유닛을 통해 데이터와 제어 트래픽을 효율적으로 처리할 수 있다. 실리콘 테스트 결과, Occamy는 스텐실 연산에서 최대 571 GFLOP/s, 28.1 GFLOP/s/W의 성능을 기록했으며, 희소-밀집 행렬 연산에서 307 GFLOP/s, 희소-희소 연산에서는 187GCOMP/s의 성능을 보여주었다. 이 시스템은 CPU와 GPU 대비 최대 11배 높은 연산 밀도와 에너지 효율을 제공하며, 스텐실 및 희소 연산에서 뛰어난 성능을 보였다.

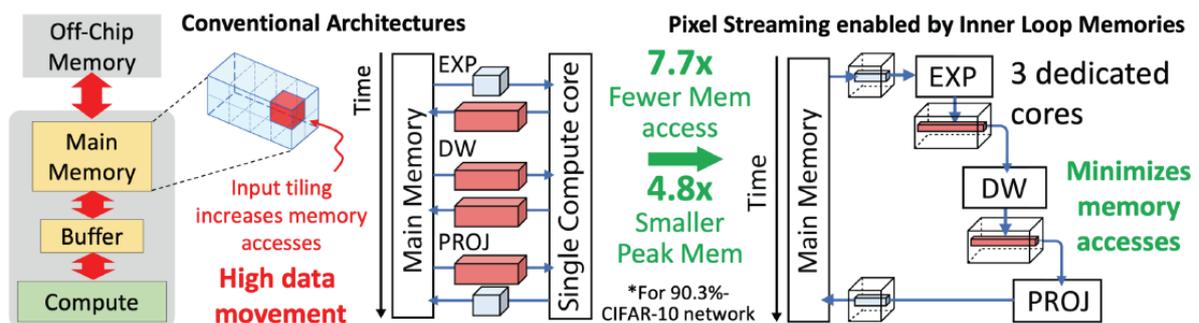


[그림 4] 듀얼 칩렛 아키텍처

Session 28 Processors II

2024 VLSI의 Session C28 또한 Processor라는 주제로 총 4편의 논문이 발표되었다. 이 세션에서는 tinyML 작업을 위한 프로세서, 에너지 효율적인 이중 SoC, 시각적 문맥 이해를 위한 Scene Graph Generation 프로세서, 그리고 고속 광통신을 위한 소프트 디시전 오류 수정 디코더가 발표되었다. 이번 후기를 통해 4개의 논문에 대해 간략하게 살펴보고자 한다.

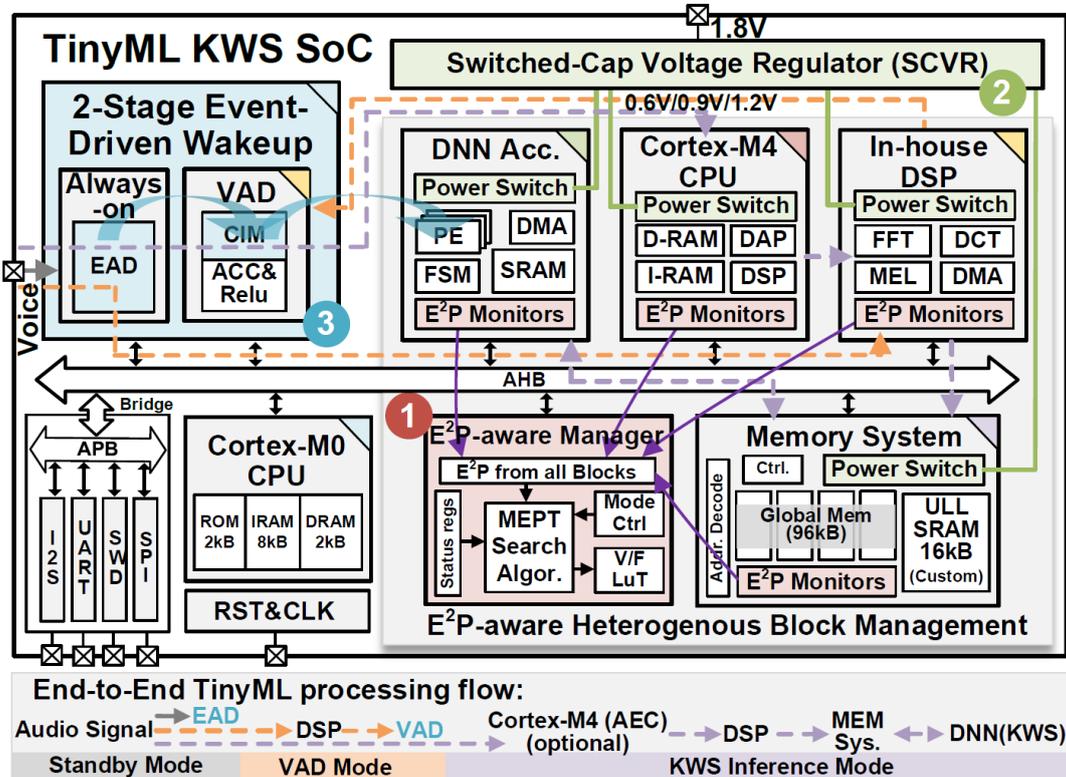
#28-1 Medusa는 28nm CMOS 공정에서 제작된 프로그래머블 8비트 프로세서로, 향상되어 있는 tinyML 작업에서 업계 최고의 추론 에너지를 달성합니다. Medusa는 6T 래치 기반의 Inner Loop Memories (ILMs)를 사용하여 15 fJ/Byte의 낮은 읽기 에너지를 제공하며, Pipelined Pixel Streaming (PPS) 아키텍처를 통해 시스템 레벨의 메모리 접근 에너지를 최대 9.5배 절감합니다. 주요 특징으로는 메모리 집약적인 병목층에서 에너지 소비를 줄이고, ILMs와 통합된 비트 직렬 공급기 및 덧셈 트리를 사용하여 연산 영역과 에너지를 최소화하며, 열 게이팅 및 파이프라인 우회 체계를 통해 불필요한 스위칭 에너지를 줄입니다. Medusa는 CIFAR-10에서 0.83/4.6 $\mu\text{J}/\text{프레임}$ 과 86.2/91.6%의 정확도를 기록하며, 추론 에너지에서 3.4배/4.9배의 향상을 보여줍니다. Google Speech Commands에서는 0.23 $\mu\text{J}/\text{프레임}$, Visual Wake Words에서는 5.0 $\mu\text{J}/\text{프레임}$ 을 달성합니다. Medusa는 전압-주파수 스케일링을 지원하여 에너지와 지연 시간을 최적화할 수 있으며, 전체 메모리 접근 에너지는 시스템 총 에너지의 20%에 불과합니다.



[그림 5] Pipelined Pixel Streaming 아키텍처

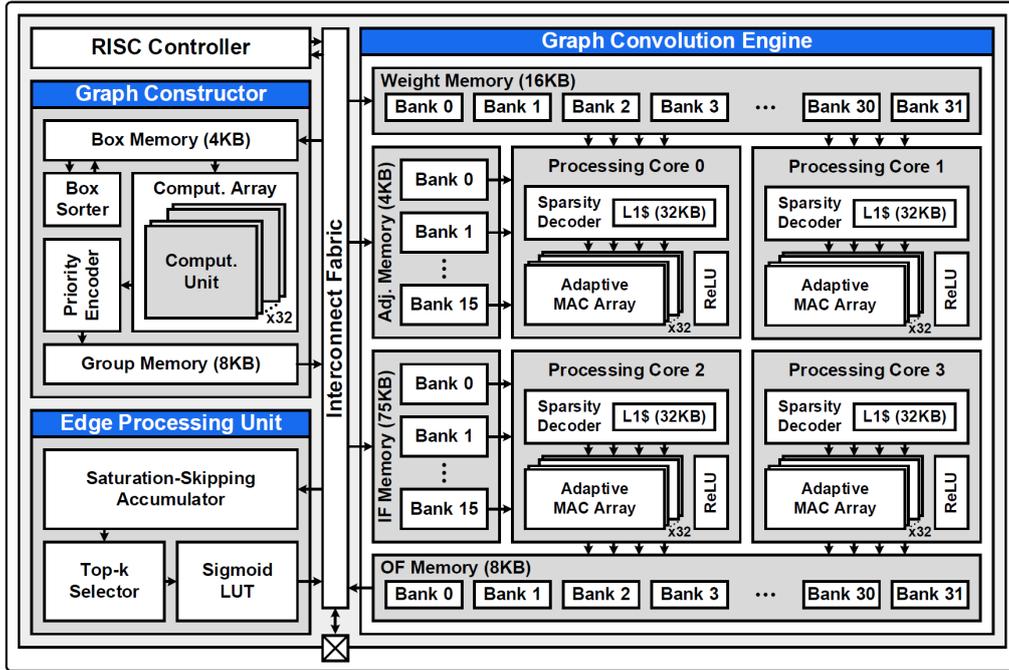
#28-2 이 논문은 Peking University에서 발표한 연구로, 에너지 이벤트 성능(E2P)-인식 관리 기능과 CIM 기반 2단계 이벤트 구동 웨이크업 방식을 갖춘 이중 TinyML SoC를 소개한다. 이 시스템은 리소스가 제한된 엣지 AI 기기에서 효율적으로 동작하도록 설계되었으며, 최소 전력 소비는 3.5 μW 로, 피크 대 대기 전력 비율은 최대 30,000배에 이른다. 특히, E2P 인식 시스템은 다양한 이중 블록의 런타임 상태를 실시간으로 모니터링하고, 계층적 전압 조정 메커니즘을 통해 시스템 수준에서 최소 에너지 지점(MEPsys)을 찾아냄으

로써, 에너지 효율을 28% 이상 향상시킨다. 또한, 이 연구에서는 CIM 기반 2단계 이벤트 구동 웨이크업 방식을 제안하여, 항상 켜져 있어야 하는 모드의 에너지를 87.3% 절감하는 데 성공했다. 이 SoC는 딥러닝 가속기와 같은 다양한 도메인 특화 가속기를 포함한 이중 SoC 구조를 채택하여, 엣지 AI 애플리케이션에서 초저전력 성능을 보여준다.



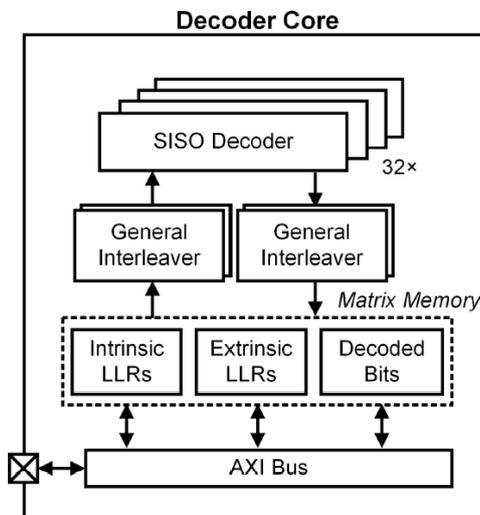
[그림 6] TinyML SoC 아키텍처

#28-3 이 논문은 National Taiwan University에서 발표된 연구로, 시각적 문맥 이해를 위한 최초의 전용 Scene Graph Generation (SGG) 프로세서를 소개한다. SGG는 객체 간 관계를 파악하여 고수준의 시각적 정보를 추출하는 작업으로, 기존 객체 감지보다 훨씬 복잡하다. 본 연구는 알고리즘-아키텍처 공동 최적화를 통해 계산 복잡성을 줄이고, 하이브리드 희소성 인코딩과 재구성 가능한 MAC 배열 구조를 활용해 메모리 사용량을 최소화하고 에너지 효율성을 높였다. 제안된 프로세서는 40nm CMOS 공정에서 제작되었으며, 280fps의 처리 속도와 0.36mJ/frame의 에너지 소비를 달성했다. 이는 GPU 대비 154배 높은 처리 속도와 1,800배 적은 전력 소모를 기록하며, 총 2.7×10^5 배 더 높은 에너지 효율을 보여준다.



[그림 7] Scene Graph Generation 시스템 아키텍처

#28-4 이 논문은 University of Michigan에서 발표한 연구로, 고속 광통신을 위한 소프트 디시전 오픈 전방 오류 수정(oFEC) 디코더를 제안한다. oFEC 디코더는 타일 기반 아키텍처와 차이 기반 소프트 디시전 디코딩을 사용해 성능을 극대화하며, 32개의 파이프라인 SISO BCH 디코더를 통해 256비트 코드워드를 병렬로 처리한다. 또한, 매트릭스 메모리와 일반 인터리버(GIL)를 최적화해 메모리 영역을 83.6% 줄였고, 지연 시간을 33% 줄였다. 이 디코더는 인텔 16 공정에서 제작되었으며, 11.4mm²의 면적을 차지하고 40.2Gbps의 처리량과 17.43pJ/b/iteration의 에너지 효율을 보인다. 기존 FPGA 기반 oFEC 디코더와 비교해 10배 높은 처리량과 낮은 전력 소모를 달성해 차세대 광통신을 위한 솔루션으로 주목받고 있다.



[그림 8] Decoder Core 아키텍처

저자정보



엄소연 박사과정 대학원생

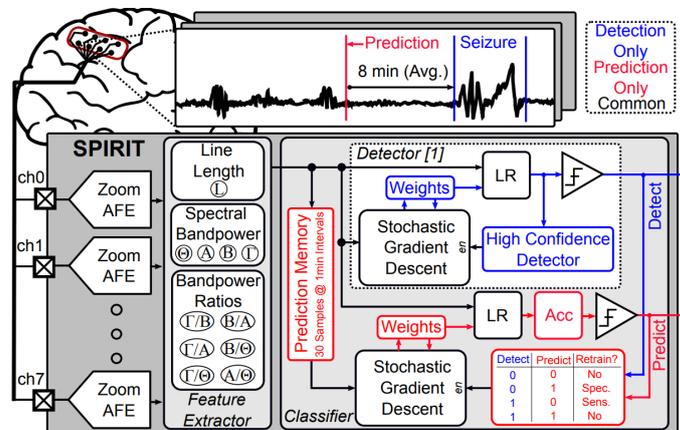
- 소속 : KAIST 전기및전자공학부
- 연구분야 : Computing-In-Memory Processor
- 이메일 : soyeon.um@kaist.ac.kr
- 홈페이지 : <https://ssl.kaist.ac.kr/>

2024 IEEE VLSI Review

DGIST 전기전자컴퓨터공학과 박사과정 위정윤

Session 23 Neural Recording Interfaces

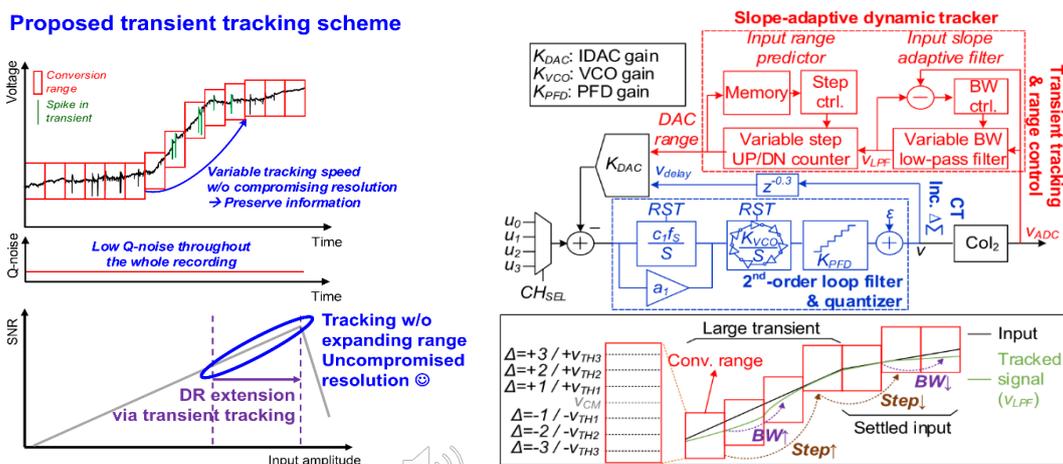
#23-1 본 논문은 칩 내부에서 retraining 및 정확도 조정이 가능한 Stochastic-gradient-descent-based Predictor with Integrated Retraining and In situ accuracy Tuning (SPIRIT) 시스템을 제안했다. (그림 4) 기존 제안된 시스템들은 칩 외부에서 데이터별 training 및 정확도 조정을 수행한 반면, 본 논문은 해당 기술들을 모두 on-chip화 했다는 점에서 높은 완성도를 지닌다. 본 시스템은 크게 analog front-end와 classifier로 나눌 수 있다. 뇌신호 측정을 위한 analog frontend는 coarse conversion을 위한 SAR ADC와 fine conversion을 위한 incremental ADC로 구성되어 있다. Classifier는 일반적인 logistic regression (LR) classifier이 사용되었으며, future weight 조정을 위해 stochastic gradient descent (SGD) 기술이 사용되었다. SPIRIT 시스템은 저전력 incremental ADC를 활용하여 효율적으로 생체 신호를 수집하고, 온라인 학습 알고리즘을 통해 지속적으로 예측 모델을 개선함으로써 실시간 간질 발작 예측이 가능하다는 점에서 기존 기술과 차별성을 갖는다.



[그림 4] 제안된 SPIRIT system level diagram

#23-3 본 논문은 dynamic zoom-and-track 기술을 사용하여 stimulation/motion artifact와 electrode DC offset이 발생한 상황에서도 sub- μ V, 10kHz 대역폭의 신경신호를 안정적으로 기록할 수 있는 Incremental $\Delta\Sigma$ Neural Recording Frontend를 제안했다. (그림 5) 신경신호는 recording front-end의 입력단에 직렬로 연결된 100kohm 상당의 neural 전극을 통해 측정된다. 이때 사용되는 전극은 frontend 입력단에 DC offset를 발생시킨다. 또한 전극과 조직 사이에서 발생하는 electrode-electrolyte impedance (ETI)로 인해 발생하는 motion artifact와 신경 자극으로 인해 발생하는 stimulation artifact는 sub- μ V크기의 신경

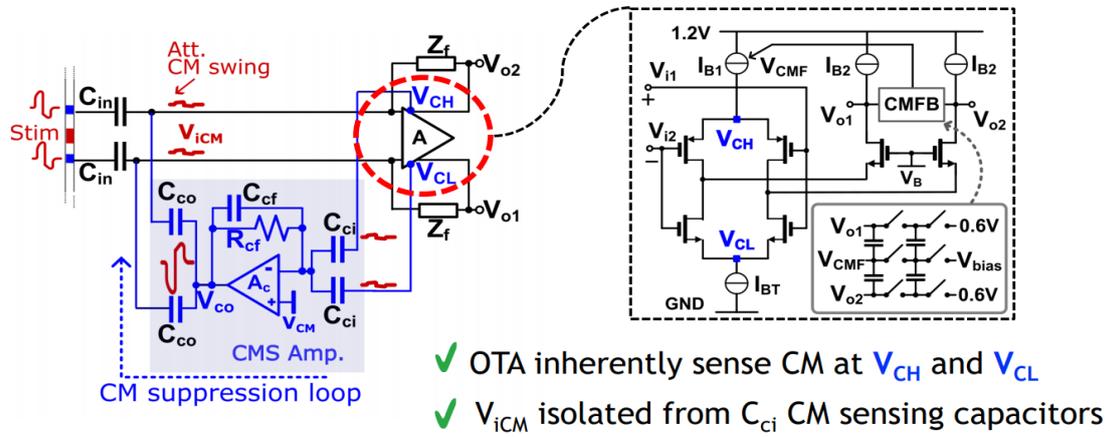
신호에 $<100\text{mV}_{pp}$ 크기의 fast transient 특징을 갖는 DC성분이 함께 기록된다. 이런 상황에서 $\text{sub-}\mu\text{V}$ 의 신경신호를 안정적으로 측정하기 위해서 recording frontend는 최소 100mV_{pp} 의 dynamic range (DR)를 가져야 한다. Interference 제거를 위해 넓은 DR를 구현한 기존 구조들은 quantization noise 레벨을 높이는 단점을 가지고 있다. 예를 들어, pre-amplifier의 gain을 낮춰 신호의 saturation을 방지함으로써 conversion range를 넓히는 구조는 quantizer의 LSB는 고정인 상황에서 amplifier의 gain을 낮춰 frontend의 입력범위를 넓혔기 때문에, input-referred quantization noise는 gain에 반비례하여 증가한다. Coarse ADC로 큰 크기 신호의 transient를 추적한 뒤에, fine ADC로 quantization하는 방법의 경우는 추적하는 순간의 quantization noise는 coarse ADC에 의해 결정되어 큰 크기의 quantization noise를 갖는다. 이러한 기존 기술들의 한계를 극복하기 위해서는 크고 빠른 간섭 신호의 영향을 제거하고 양자화 잡음을 최소화하기 위해서는 넓은 대역폭의 DC servo-loop를 구현하여 DC 신호를 빠르게 추적하고 제거하는 것이 필수적임을 알 수 있다. 본 논문에서는 입력 신호의 slope에 따라 DC servo-loop의 대역폭과 입력범위를 dynamic하게 변경할 수 있는 slope-adaptive dynamic tracker 기술을 제안했다. 본 기술의 내부회로는 그림xx와 같이 대역폭 가변이 가능한 input slope adaptive filter와 입력범위 조절이 가능한 input range predictor로 나누어 확인할 수 있다.



[그림 5] 제안된 Zoom-and-track Incremental $\Delta\Sigma$ Neural Recording Front-end

#23-5 본 논문은 closed-loop neuromodulation 수행 시 발생하는 common-mode interference (CMI)를 제거하기 위한 analog frontend를 제안했다. (그림 6) CMI 발생 시에도 신경신호를 왜곡없이 측정하기 위해서는 보통 넓은 입력범위를 갖는 신경신호 기록 ADC가 필요하다. 하지만 본 논문의 기술은 CMI를 suppression하기 때문에 신경신호 기록 ADC의 설계를 간소화할 수 있다는 장점이 있다. 본 논문은 CMI-tolerant neural AFE를 16채널로 구현하였으며, common-mode suppression에 필요한 amplifier를 8개 채널이 공유함으로써 채널 당 전력소모를 줄였다. 구현된 시스템은 34G의 높은 입력 임피던스를

달성했으며, $3.125\mu\text{W}/\text{Ch}$ 전력소모와 0.006mm^2 의 매우 작은 면적을 달성했다.



[그림 6] 제안된 CMI-tolerant neural AFE

저자정보



위정윤 박사과정 대학원생

- 소속 : DGIST EECS
- 연구분야 : Analog ICs for biomedical application
- 이메일 : wiejung@dgist.ac.kr
- 홈페이지 : <http://ins.dgist.ac.kr>

2024 IEEE VLSI Review

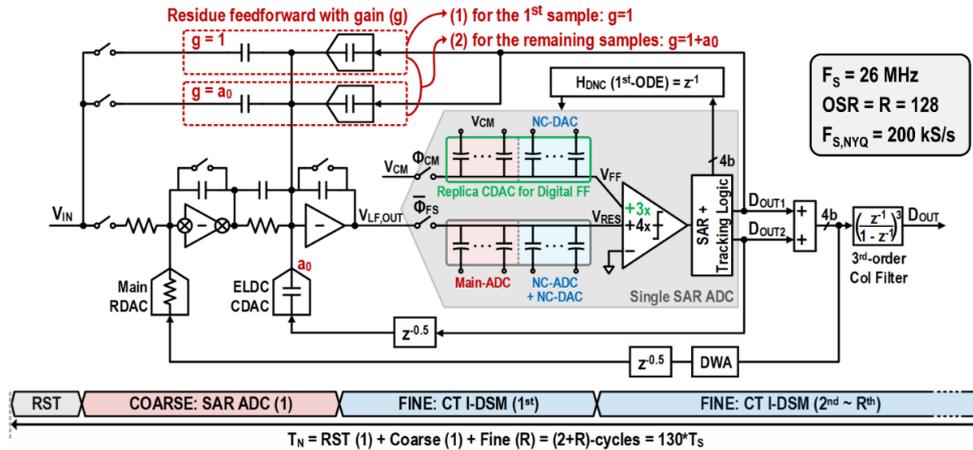
DGIST 전기전자컴퓨터공학과 박사과정 위정윤

Session 18 Data Converter Techniques

최근 제안되고 있는 ADC는 높은 전력효율을 얻기 위해 OSR을 낮추는 방향으로 연구되고 있다. OSR을 낮추면서 동시에 높은 resolution을 보장하기 위해 고차 noise-shaping, pipelined, 그리고 coarse-fine conversion 등 다양한 기술을 응용한 ADC가 제안되고 있다. 최근 noise-shaping 기술들은 전력효율을 높이기 위해 3차 이상의 고차 ADC로 구현되고 있다. 하지만 고속 응용분야에서는 이런 고차 ADC 구현이 어렵기 때문에 pipelined ADC가 빈번히 제안되고 있다. 반면에 센서분야에서는 ADC 출력에 digital 출력을 재구성하기 위한 외부 decimation filter 또는 넓은 면적의 합성 decimation filter를 필요로 한다. 하지만 고차 noise shaping ADC는 on-chip decimation filter 구현에 면적 및 전력 면에서 단점을 가지기 때문에 decimation filter 구현이 용이한 incremental ADC (IADC)가 연구되고 있다. Incremental ADC 내부의 reset 기능으로 인해 decimation filter를 간단하게 구현할 수 있지만, 높은 resolution 구현을 위해서는 매우 높은 oversampling ratio(OSR)에서의 동작이 필요해 높은 power efficiency를 기대하기 어렵다. 이런 IADC의 한계를 극복하기 위해 최근 Nyquist sample rate을 유지하면서 낮은 OSR로도 높은 resolution 구현이 가능한 coarse-fine conversion incremental ADC가 제안되는 추세이다.

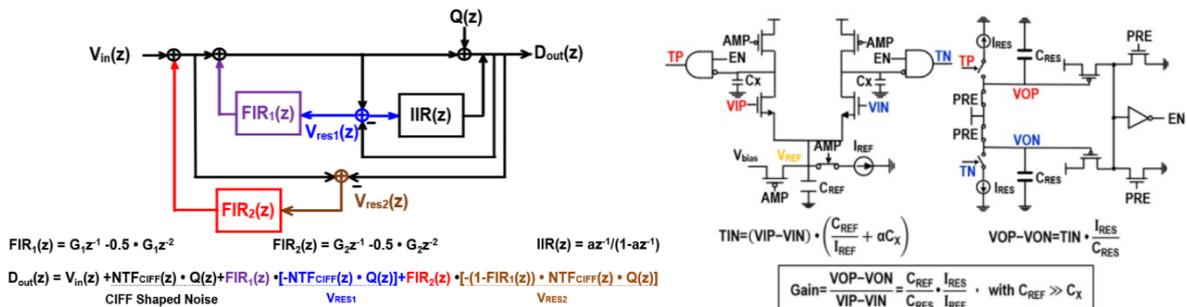
#18-1 본 논문은 tracking-zoom, residue gain switching, 그리고 digital noise-coupling feedforward (NC-FF) 세 가지 기술을 제안함으로써 높은 power efficiency와 resolution이 필요한 센서 인터페이스를 위한 zoom 동작을 탑재한 IADC를 구현했다 (그림1). Tracking-zoom 기술은 coarse conversion의 SAR ADC를 fine conversion에서의 quantizer로 재사용 가능하도록 함으로써 높은 resolution 및 면적 최소화에 기여했다. 해당 기술은 coarse SAR ADC의 sampling rate을 Nyquist rate으로 가져가면서 동시에 signal-to-quantization noise ratio (SQNR) 성능 저하를 피하기 위해 제안된 기술로, fine conversion 시에 추가적으로 coarse code를 업데이트한다. Fine conversion의 디지털 출력을 threshold 및 low pass filtering 시킴으로써, SAR ADC의 OSR의 한계로 인해 저하된 coarse conversion SQNR 성능을 fine conversion에서 DC 성분의 평균을 빼고, 해당 평균으로 coarse code를 업데이트함으로써 1-conversion SQNR 저하를 보상한 기술로 해석할 수 있다. Residue gain switching 기술은 excessive loop delay (ELD) 보상을 위해 추가된

path에서 발생하는 gain 변화를 보상하기 위해 feedforward path의 gain을 ELD path의 gain과 동일하게 설정하는 기술이다. 본 기술을 통해 fine conversion의 signal transfer function (STF)를 1로 맞추으로써 continuous time IADC에서도 coarse, fine conversion에서의 fuzz 문제를 해결했다. 마지막으로 제안된 digital NC-FF기술은 기존 digital NC에 feedforward path를 추가하여 noise transfer function (NTF)의 코너 가변성을 구현했다. 언급된 기술들을 통해 제안된 ADC는 169.8의 $FoM_{S,SNDR}$ 과 174.9의 $FoM_{S,DR}$ 성능을 달성했다.



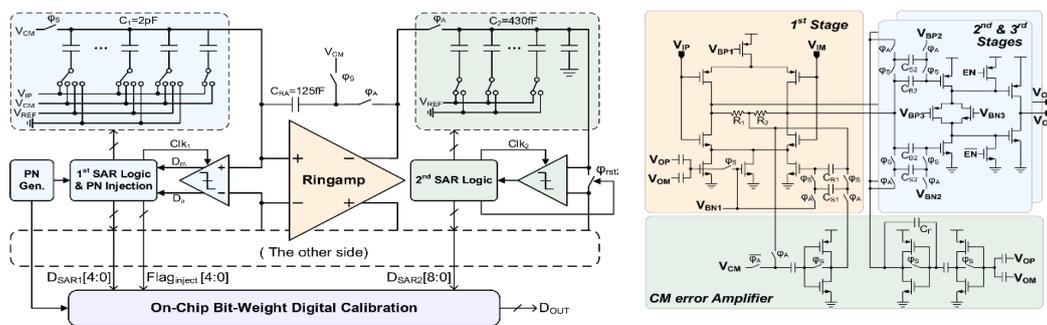
[그림 1] 제안된 continuous-time tracking-zoom incremental ADC

#18-2 본 논문은 error feedback (EF)과 cascade of Integrator with feedforward (CIFF) 기술을 응용한 EF-EF-CIFF 기반 5차 noise shaping 기술을 제안하고, 내부 amplifier를 process, voltage, 그리고 temperature에 강인한 voltage-time-voltage (VTV) 컨버터를 통해 구현함으로써 고성능의 SAR ADC를 구현했다 (그림2). Amplifier의 경우, PVT에 강인하게 설계하기 위해선 추가적인 면적, 전력이 소모된다. 본 논문은 전압을 PVT에 강인한 시간영역으로 변경한 뒤, 이득을 추가하여 전압으로 변경하는 VTV를 제안하여 PVT에 강인하면서도 저전력을 소모하는 ADC 구현에 성공했다. 결과적으로 고차를 구현하여 매우 낮은 OSR로 높은 SNDR를 달성함과 동시에 전력소모를 최소화하여 180.2-FoM_S의 높은 전력효율을 달성했다.



[그림 2] 제안된 5차 NS-SAR ADC 블록도 및 voltage-time-voltage 컨버터

#18-4 본 논문은 switched-capacitor-based ringamp with robust CMFB, auxiliary-latch-based PN-junction window detector, separate-averaging algorithm for bit-weight calibration 세 가지 기술을 제안함으로써 sub-ADC 간 gain error 및 DAC mismatch를 on-chip에서 보정가능한 시스템을 구현했다. (그림 3) 고속 pipelined ADC에서는 전력효율을 높이기 위해 ring 타입의 residue amplifier을 응용하고 있지만, 기존 ringamp들은 안정적인 common-mode (CM) 구현을 위해서 신호 왜곡을 감수하거나 또는 높은 정확도의 제어 clock으로 인한 추가적인 회로를 필요로 했다. 본 논문에서 제안한 switched-capacitor-based ringamp는 1st-stage는 DC coupled fully differential amplifier을 사용하고 2nd-stage는 switched-capacitor를 통해 안정적으로 CM를 바이어싱하는 회로를 구현했다. 이와 더불어 discrete-time CM error amplifier를 추가하고, CMFB에 필요한 path를 줄여 phase margin을 확보한 CMFB를 구현했다. Pipelined ADC의 sub-ADC 간 gain error 및 첫번째 sub-ADC의 DAC mismatch는 ADC 선형성에 큰 영향을 미친다. 본 논문은 pseudo random noise (PN)를 인가하여 bit-weight digital calibration을 수행함으로써 문제를 해결했다. 적절한 타이밍에 PN을 인가하기 위해 auxiliary-latch-based PN-junction window detector를 quantizer에 결합시켰으며, separate-averaging algorithm for bit-weight calibration 기법을 구현하였다. 상기 언급된 세가지 기술을 통해 본 논문은 low frequency 입력에서 175.3-FoM₅를 Nyquist 입력에서 169.7-FoM₅를 달성했다.



[그림 3] 제안된 continuous-time tracking-zoom incremental ADC

저자정보



위정윤 박사과정 대학원생

- 소속 : DGIST EECS
- 연구분야 : Analog ICs for biomedical application
- 이메일 : wiejung@dgist.ac.kr
- 홈페이지 : <http://ins.dgist.ac.kr>

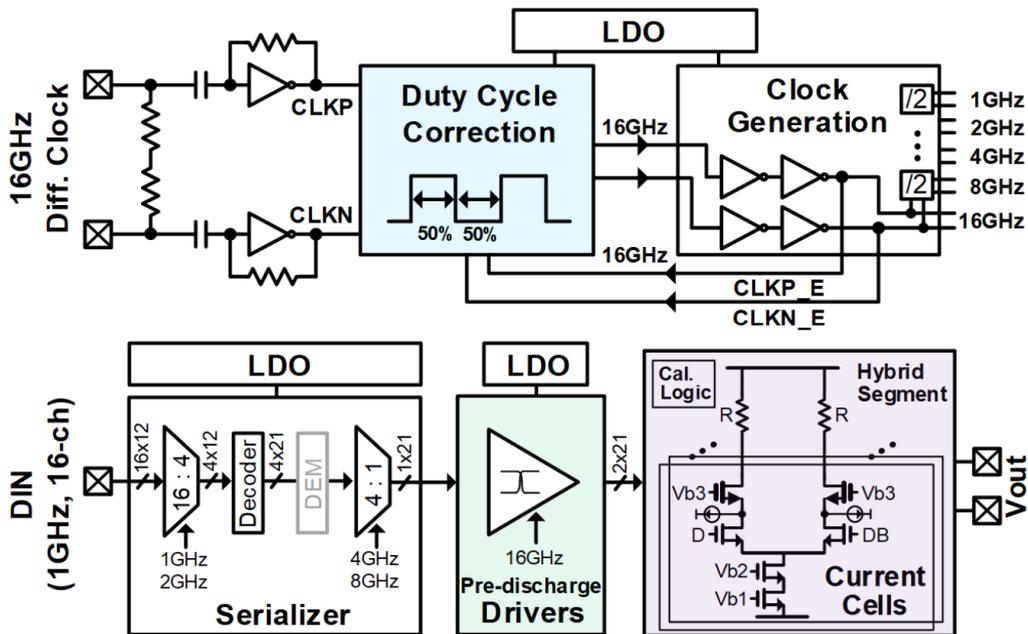
2024 IEEE VLSI Review

고려대학교 반도체시스템공학과 박사과정 김현진

Session 24 High-Speed Data Converters

이번 2024 IEEE VLSI의 Session 24는 High-Speed Data Converters라는 주제로 총 5편의 논문이 발표되었다. 이 세션에서는 RF 통신 어플리케이션에 적합한 고속 ADC 및 DAC와 wireline에 적합한 고속 ADC 구조들이 소개되었고, 주로 고속 동작을 위해서 time-interleaved architecture로 설계될 경우 발생하는 이슈들을 해결하는 방향으로 연구가 진행되었다. 특히, 올해는 추가적인 성능 하락 없이도 background calibration을 구현하여 실용적인 data converter를 설계하는 논문들이 다수 발표되었다는 점이 주목할 만하다.

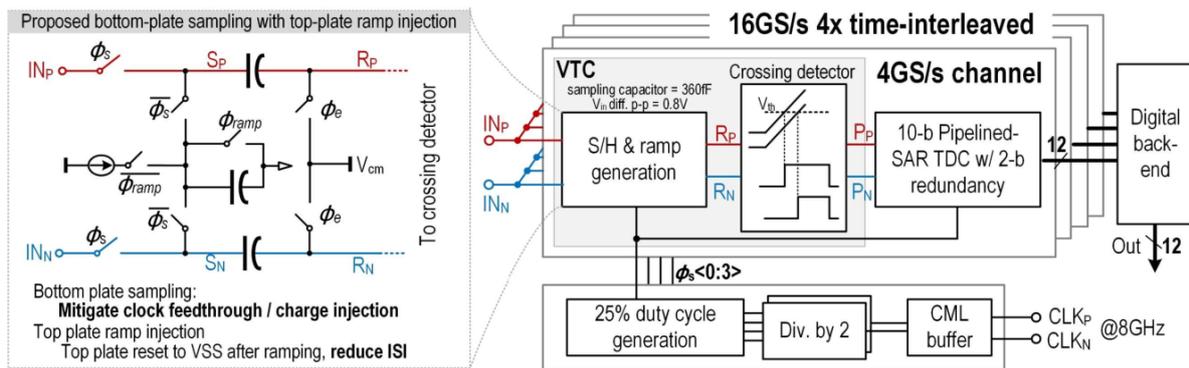
#24-1 본 논문에서는 5G 통신 어플리케이션에 적합한 high-resolution & high-speed DAC 구조를 제안한다. Current-steering 구조를 기반으로 설계되었으며, MSB와 LMSB 사이의 dynamic mismatch로 인하여 성능 저하가 없도록 새로운 hybrid current cell segmentation 방식으로 동작한다. 또한, code 기반의 free switch driver를 통해서 DAC 동작 시 발생하는 데이터 의존성을 제거해 주었다.



[그림 1] 본 논문에서는 소개된 전체적인 RF-DAC의 아키텍처이다.

제안된 회로는 5nm FinFET 공정으로 설계되었으며, -85 dBc의 IMD3와 -163 dBFS/Hz의 노이즈 스펙트럼 밀도를 보여주었다. 또한, 16 GS/s에서 7.1 GHz 출력 주파수로 동작할 때 486mW의 낮은 전력 소모량을 달성하였다.

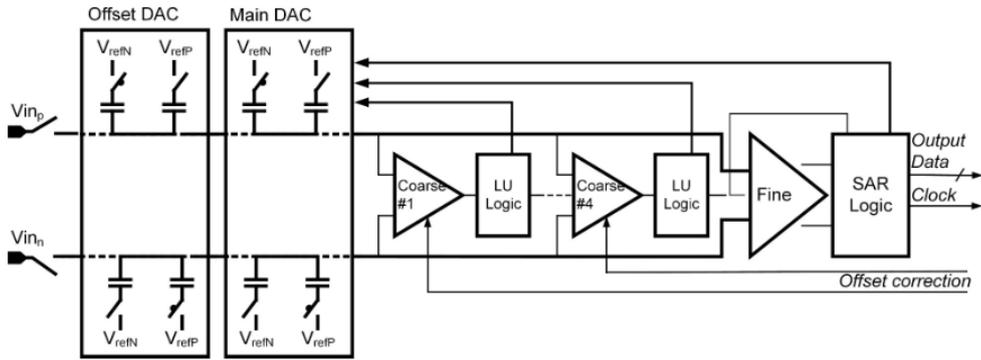
#24-2 본 논문에서는 새로운 구조의 voltage-to-time converter와 background delay offset calibration을 통해서 time-domain ADC의 PVT variation과 연관된 이슈들을 해소하여, 4x time-interleaving 만으로도 16 GS/s 및 10-bit 해상도를 달성한 새로운 RF 어플리케이션용 ADC 아키텍처를 제안한다.



[그림 2] 본 논문에서 제안된 time-interleaved time-domain ADC 구조이다.

먼저, 제안된 time-domain ADC는 SAR 알고리즘을 기반으로 동작하며, SAR 동작에 redundant quantization range를 여유로 두어 radix error와 offset에 의해 야기되는 quantization residue를 보상하였다. 하지만, 고해상도 동작을 위해서는 추가된 redundant stage들의 delay offset이 calibration 되어야한다. 따라서, delay offset을 적분한 integrator 출력을 이용하여 DTC를 제어하는 방식으로 background offset calibration을 진행하였다. 또한, 새로운 bottom-plate sampling 기반의 VTC 구조로 ADC의 추가적인 정확도 향상을 달성하였다. 측정 결과, ADC는 Nyquist 주파수에서 55.93 dB의 SFDR과 44.48 dB의 SNDR을 달성했으며, 94.2 mW를 소비하여, 기존 논문들 대비 높은 FoM 성능을 달성하였다.

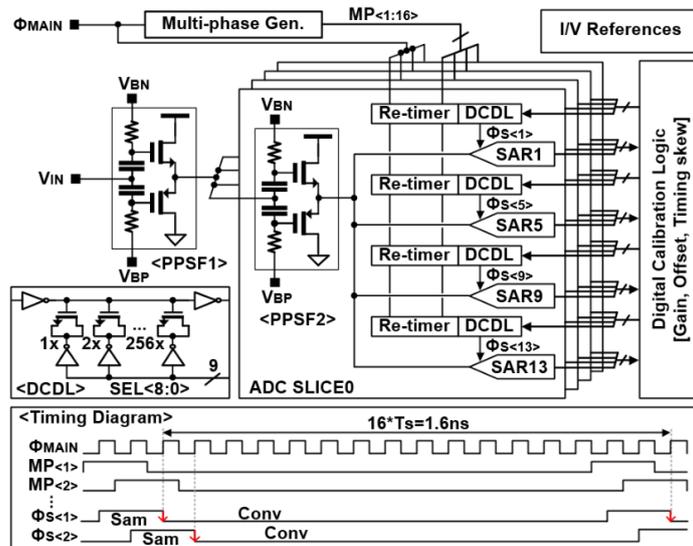
#24-3 본 논문에서는 partial loop unrolled SAR ADC 회로를 sub ADC 요소로 사용하는 64-way time-interleaved ADC 아키텍처를 소개한다. 전통적인 SAR ADC 구조는 비교기와 SAR 로직들의 순차적인 동작 방식으로 time-interleaved ADC 구조에서 사용되기에 느리다.



[그림 3] 본 논문에서 sub ADC로 사용된 partial loop unrolled SAR ADC 구조이다.

따라서, 논문에서는 [그림 3]과 같이 loop unrolled ADC를 MSB의 결정 동작에 사용하여 sub ADC의 전체적인 동작 속도를 증가시켰다. 하지만, ADC 회로에서는 두개 이상의 비교기를 사용할 경우, 비교기 간의 offset을 calibration해야하는 이슈가 발생한다. 본 논문에서는 redundancy를 사용하여, loop unrolled 구간과 SAR ADC 구간에서의 DAC 값을 일치시켜서, 추가적인 하드웨어 없이도 같은 아날로그 전압을 통해서 calibration을 진행하는 새로운 calibration 방식을 제안하였다. 60 GS/s 동작 속도를 감안하여 아날로그 front-end 회로 또한 신중하게 설계되었고, 논문에서 소개된 ADC는 기존에 발표된 30 GS/s 이상의 time-interleaved ADC 구조들과 비교하였을 때 향상된 FoM 성능을 달성하였다.

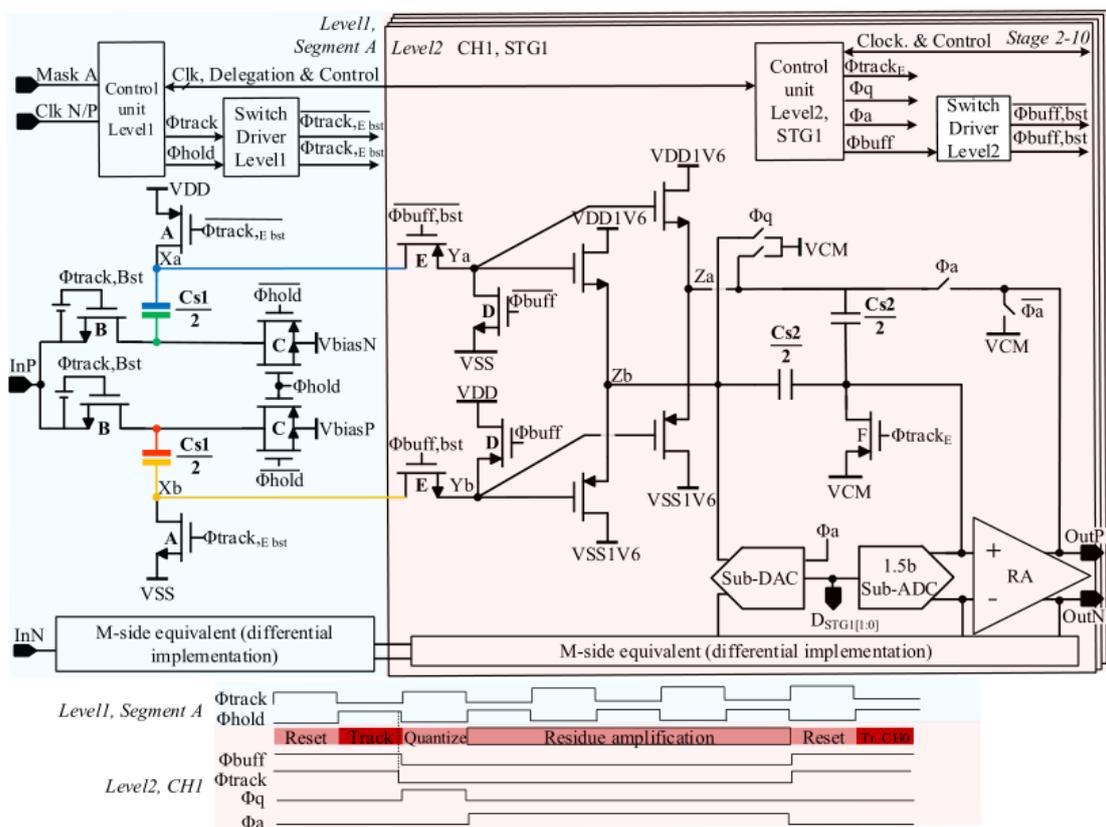
#24-4 본 논문에서는 background timing skew calibration을 적용시킨 over-Nyquist 12-비트 10 GS/s time-interleaved ADC 구조를 소개한다. 제안된 background calibration 방식은 auto-correlation 알고리즘을 통해서 absolute timing skew 에러를 계산하고 보정하며, 짝수 및 홀수 채널을 구분하여 동시에 calibration하여 3단계의 보정만으로 16개의 SAR ADC의 timing skew를 전부 calibration한다.



[그림 4] 본 논문에서 제안된 전체적인 time-interleaved SAR ADC 구조이다.

제안된 ADC 칩은 5nm FinFET 공정으로 제작되었으며, 5GHz 및 9GHz 입력 주파수에서 각각 50.2dB 및 46.2dB의 SNDR를 달성하였다. 본 논문에서 소개된 ADC는 추가적인 하드웨어 없이 디지털 도메인에서 timing skew를 calibration 하였으므로, 386mW의 비교적 낮은 전력 소비량을 달성하였다.

#24-5 본 논문에서는 10 GS/s 이상의 RF 어플리케이션용 ADC 구조를 소개한다. 회로는 hierarchical time-interleaved ADC 구조를 기반으로 설계되었고, front-end의 signal 및 clock buffering은 고속에서도 높은 해상도를 달성하는 것을 목표로 집중적으로 분석 및 설계되었다.



[그림 5] 본 논문에서 제안한 전반적인 time-interleaved ADC 구조이다.

논문에서 제안한 wide-band-linear signal buffering 기술과 효율적인 hierarchical ADC architecture 를 통해서 low noise (>56 dB SNDR) 성능과 low distortion (>68 dB SFDR) 성능을 달성하였고, 이는 기존의 비슷한 RF 어플리케이션용 ADC들의 성능들 보다 높다.

저자정보



명예기자 김현진

- 소속 : 고려대학교 반도체시스템공학과 박사과정
 - 연구분야 : PMIC & ADC & Ising Machines
 - 이메일 : jamespul@korea.ac.kr
 - 홈페이지 : <https://kilby.korea.ac.kr>
-

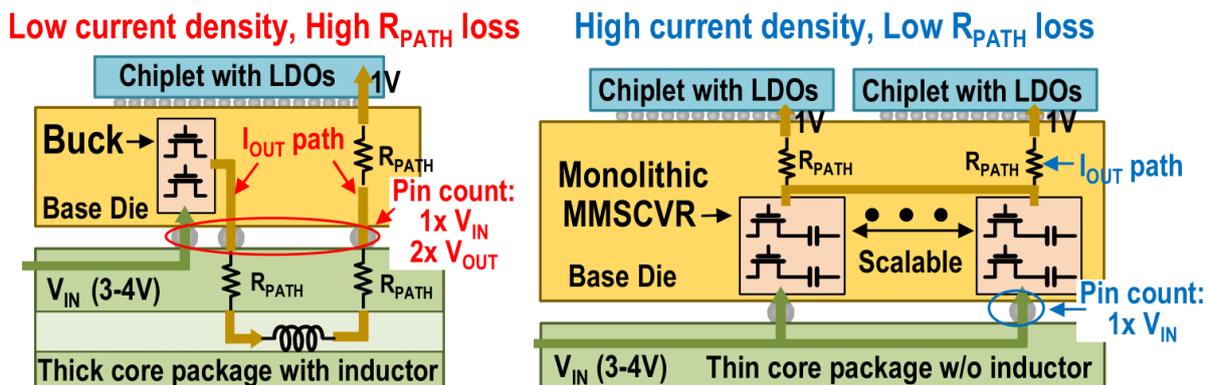
2024 IEEE VLSI Review

고려대학교 반도체시스템공학과 박사과정 김현진

Session 17 Power Management Techniques

이번 2024 IEEE VLSI의 Session 17은 Power Management Techniques라는 주제로 총 5편의 논문이 발표되었다. 이 세션에서는 switched capacitor voltage regulator, wireless power converter, low dropout regulator 등 다양한 구조의 power converter들이 소개되었고, 주로 각각의 converter들의 어플리케이션에서 겪고 있는 중요한 디자인 이슈들을 해결하는 방향으로 연구가 진행되었다. 특히, 올해는 이전 학회들에 비해서 integrated voltage regulator 논문들이 다수 발표되었고, 창의적인 DC-DC converter 구조 변화로 기존의 computing architecture들의 문제를 해결하려 했다는 점이 주목할 만하다.

#17-1 본 논문에서는 16 nm CMOS 공정의 scalable design을 기반으로 한 monolithic multi-stage modular switched capacitor voltage regulator (MMSCVR)를 제안한다. 제안된 MMSCVR 회로는 7A까지 출력되며, 이중 3D-IC 어플리케이션에서 수직 전력 전달 방식으로 동작하기 위해서 논문에서는 self-timed deadtime generator와 안전한 startup 동작 또한 제안하였다. MMSCVR 회로는 4:1 또는 3:1 전압 강하를 선택해서 동작할 수 있도록 모듈식으로 설계되었고, vertical scalable level shifter를 통해서 최적화된 게이트 드라이빙을 수행한다.

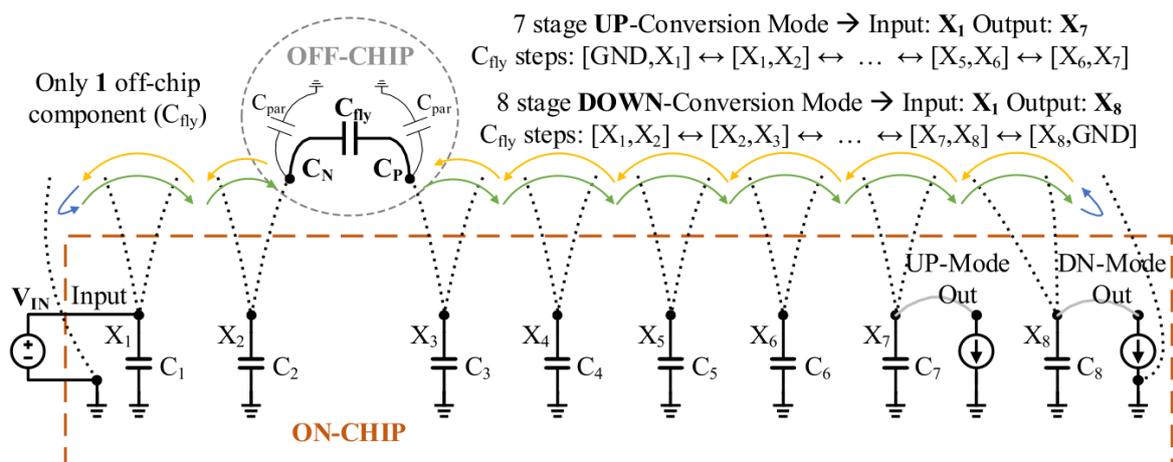


[그림 1] 본 논문에서는 기존의 base-die buck 변환기 대비 적은 pin count와 높은 current density, 낮은 dc 저항 손실을 발생시킨다.

본 논문에서 제안된 switched-capacitor 회로는 3:1 및 4:1의 전압 강하에도 90.6%와 87.3%의 높은 전력 전달 효율을 달성하였다. 하지만, 본 논문에서처럼 DC-DC converter

가 소수의 고정된 전압 강하 비율로 동작하게 된다면, 자유로운 dynamic voltage scaling 동작에서는 종래의 continuously-scalable한 switched-capacitor DC-DC 변환기들 대비 전력 전달 효율이 크게 감소하게 된다는 단점이 존재한다. 또한, 전체적인 시스템적인 분석 말고도 논문에서 소개된 변환기 회로 자체가 종래의 고정된 3:1 및 4:1 전압 강하 switched-capacitor DC-DC 변환기들 대비 어떠한 장점이 어느정도 있는지에 대한 정량적인 분석이 필요하다고 생각된다.

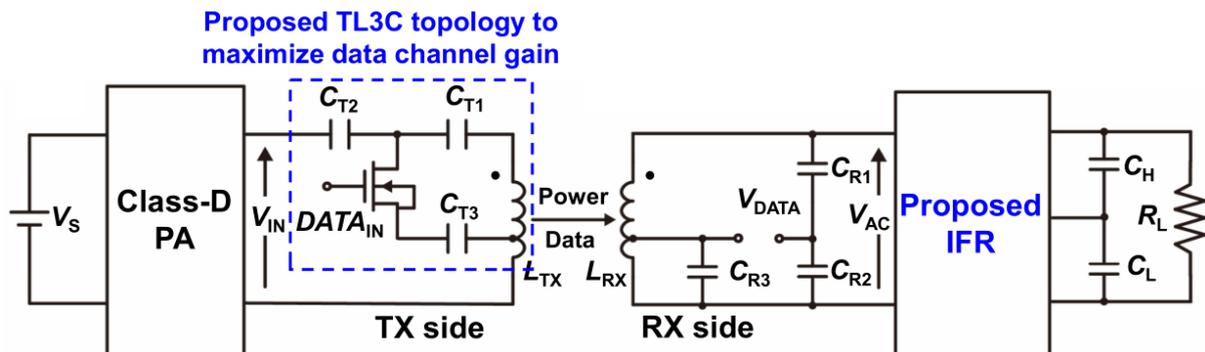
#17-2 본 논문에서는 μW 급 센서 어플리케이션을 위한 고전압 single-capacitor switched DC-DC up/down 변환기를 제안한다. 기존 고전압 승압용 Dickson charge-pump 구조는 변환기 stage가 많아지고 출력 전류량이 적을수록 전반적인 전력변환효율이 감소하므로, μW 출력 전력 수준으로 100V 이상의 전압을 생성할 때 1%의 전력전달효율 밖에 가질 수 없었다. 따라서, 제안된 회로는 기존 고전압 switched-capacitor들과 다르게 하나의 0.1 μF off-chip 커패시터를 사용하여, on-chip 커패시터들로 구성된 전압 사다리를 오르내리면서 전력을 효율적으로 전달한다.



[그림 2] 본 논문에서 제안된 switched-capacitor DC-DC 변환기 구조이다.

변환기는 [그림 2]와 같은 구조를 통해서 승압과 강압 변환을 진행하고, 승압 모드에서는 175 V 출력에서 62%의 효율을, 강압 모드에서는 184 V 입력에서 43%의 효율을 달성하였다. 특히, 논문에서는 승압 변환 효율 성능은 기존의 100V 이상의 출력전압과 nA 출력 부하를 가진 변환기들보다 16배 높다고 밝혔다. 논문에서 소개된 회로는 on-chip 커패시터들을 사용해서 off-chip의 parasitic capacitance로 인한 손실을 낮춘 새로운 구조이므로, on-chip 커패시턴스와 off-chip 커패시터의 parasitic capacitance 값의 비율에 따른 전력변환효율이 변화 분석이 필요하다고 생각된다.

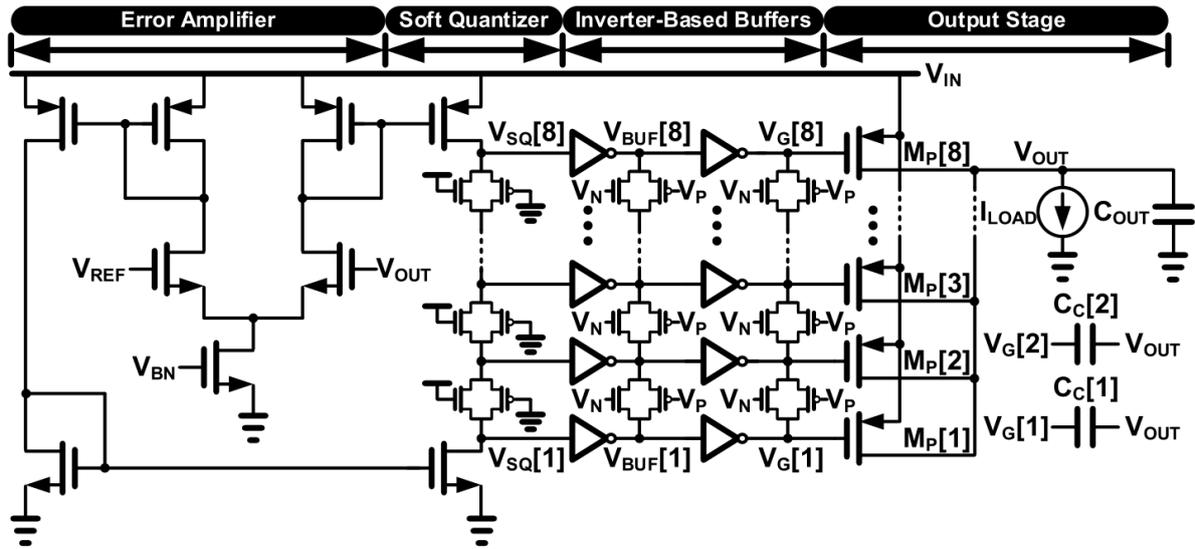
#17-3 본 논문에서는 bridge inverter의 기초적인 성분과 harmonic 성분을 이용하여 전력과 데이터를 각각 전송하는 무선 전력 및 데이터 전송 시스템을 제안하였다. 종래의 full bridge rectifier는 간섭 전압을 일으켜 data flipping 이슈를 초래할 수 있다. 따라서, 본 논문에서는 active 소자인 트랜지스터들을 사용하여, interference-free 정류기(IFR)를 구현하였고, 이러한 이슈를 해결하였다. 또한, 전력 채널 이득을 유지하면서, 데이터 채널 이득을 극대화하기 위해 TL3C 토폴로지를 제안하였다.



[그림 3] 본 논문에서 제안된 전체적인 WPDT 구조이다.

제안된 IFR은 5개의 phase로 동작하면서 interference voltage ratio를 -17.0 dB에서 -45.2 dB로 크게 낮췄다. 칩은 180 nm CMOS 공정으로 제작되었으며, 6.78 MHz로 동작하고, 82 mW의 부하 전력을 공급하면서 동시에 4.0 Mb/s의 속도로 데이터를 보낼 수 있었다. 하지만, 제안된 IFR 구조는 커패시터와 MOSFET 스위치들을 직렬연결하는 방식으로 다양한 phase를 생성하므로, 회로의 전체적인 전력 밀도 및 전력 전달 효율을 어느정도 감소시킬 수밖에 없다. 따라서, 논문에서처럼 passive full bridge rectifier와만 비교하기 보다는 최신 WPDT 구조와의 비교 분석도 필요하다고 생각된다. 또한, 실제 측정 파형에서는 V_{AC} 전압이 ringing하면서 논문의 이론적인 그림과 형태가 불일치하는 phase들이 있는데, 이에 대한 원인 설명 또한 필요하다고 생각된다.

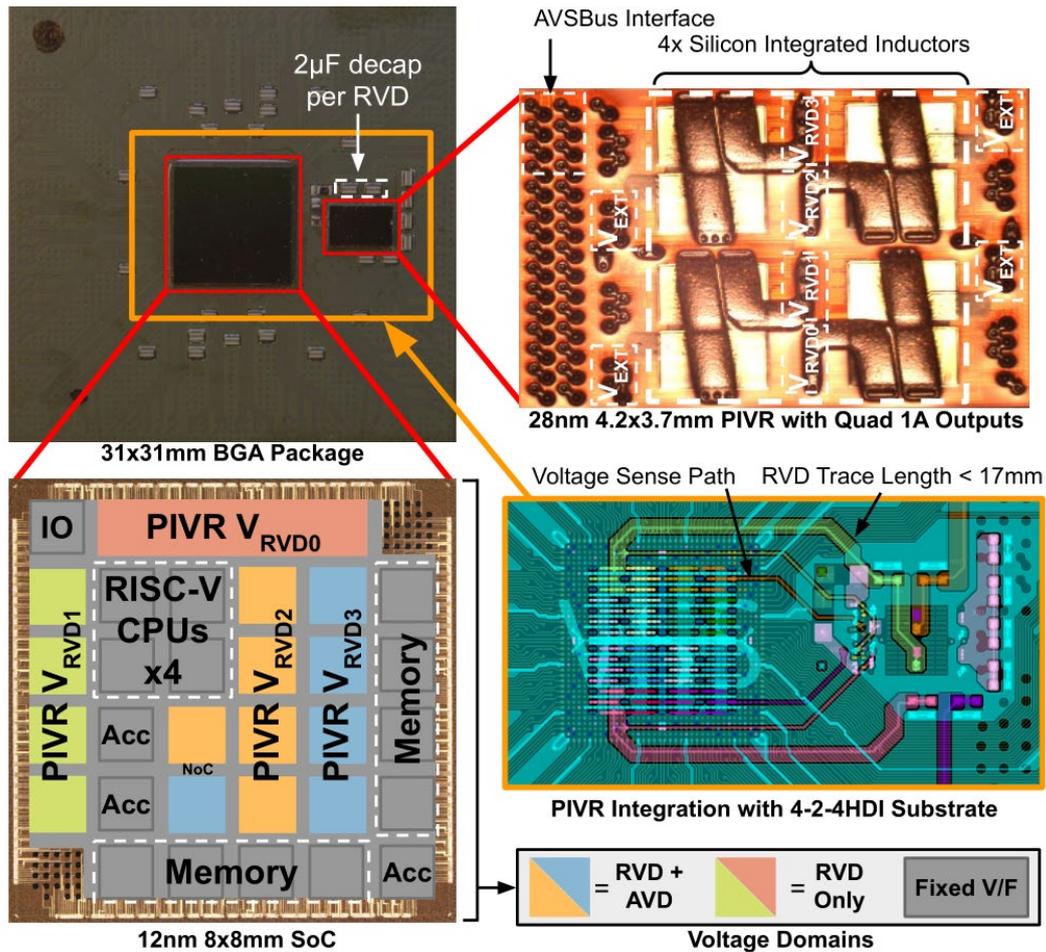
#17-4 본 논문에서는 soft-switching 동작을 통해서 아날로그 low dropout regulator (LDO)와 디지털 LDO의 장점을 결합한 새로운 구조의 LDO를 소개한다. 제안된 LDO는 아날로그 OTA의 출력을 transmission gate로 구성된 soft quantizer를 통해서 전압으로 변환하고, inverter 기반의 버퍼를 통해서 power transistor를 키거나 끄면서 출력을 regulation하게 설계하였다.



[그림 4] 본 논문에서 제안된 LDO 회로도조이다.

[그림 4]은 본 논문에서 제시된 LDO의 회로도를 보여준다. 논문에서 제시된 구조는 65 nm 공정을 사용하였고, 1 V 입력 전압에서 31.3 A/mm²의 전류 밀도, 0.6 V 입력 전압에서 99.99%의 전류 효율을 달성하였다. 하지만, 첫번째 stage는 종래의 error amplifier를 사용하므로 낮은 전압에서의 transient response 성능이 감소하는 구조이고, 어플리케이션 특성상 앞단의 inductive converter의 동작 주파수가 1 MHz 이상임에도 PSR이 1 MHz 이하에서 감소하는 추세를 보인다는 부분이 아쉽다고 생각된다. 또한, transmission gate의 저항 값과 error amplifier의 current 출력이 곱해지면서 전압을 생성하므로 PVT variation에 따른 성능 변화가 예상된다. 마지막으로 buffer들 사이에서 trip point variation이 발생할 때 생기는 동작 이슈나, steady-state condition에서 buffer들이 trip point에 멈춰 있는 상태로 short-through 전류가 발생하는 경우가 있을지에 대한 추가적인 설명이 필요하다. Digital 동작을 사용하여 analog LDO의 성능을 향상시킨 또다른 논문으로 ISSCC'20에 발표된 Ring-Amplifier-Based LDO 논문이 이미 있으므로, 해당 논문과의 성능 비교 또한 필요하다.

#17-5 본 논문에서는 패키지에 집적된 buck voltage regulator(PIVR)와 standard-cell 기반의 디지털 LDO를 통합한 새로운 dynamic-voltage-and-frequency-scaling (DVFS) 아키텍처를 소개하였다. 이 DVFS 아키텍처는 22코어 SoC 어플리케이션에서 개별 코어를 400 ns settling time의 속도로 제어하였다.



[그림 5] 본 논문에서 제안한 22코어 SoC 어플리케이션용 DVFS 아키텍처이다.

백 컨버터는 silicon integrated 인덕터를 통해서 패키지에 집적하였고, 309 mW/mm^2 의 전력 밀도를 달성하였다. 아키텍처는 [그림 5]와 같은 실제 SoC 어플리케이션에서 측정되었고, DVFS를 사용하지 않은 기본값 대비 최대 23%의 전력을 절감하였으며, 논문에서는 이러한 결과를 통해서 제안된 아키텍처가 기존 방식들 대비 가장 높은 효율의 전력 공급 아키텍처임을 밝혔다.

저자정보



명예기자 김현진

- 소 속 : 고려대학교 반도체시스템공학과 박사과정
- 연구분야 : PMIC & ADC & Ising Machines
- 이 메 일 : jamespul@korea.ac.kr
- 홈페이지 : <https://kilby.korea.ac.kr>

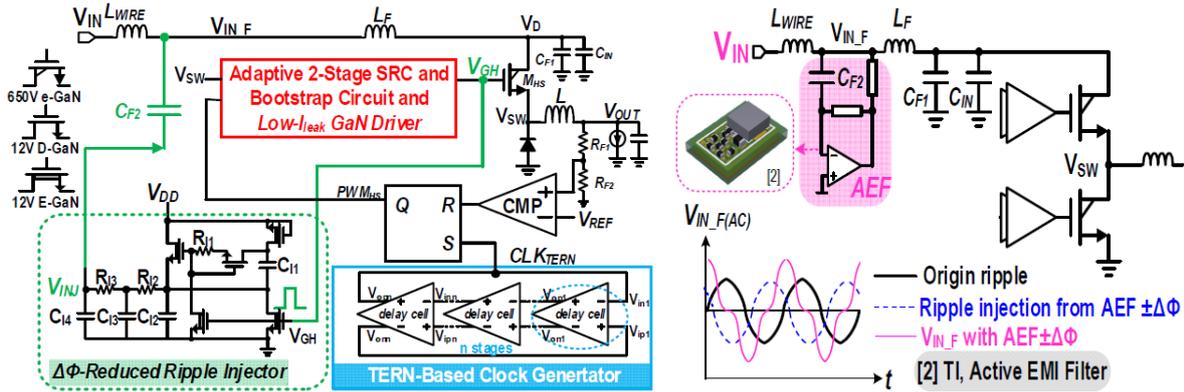
2024 IEEE VLSI Review

KAIST 전기 및 전자공학부 박사과정 박수연

Session 2 Power at High Voltage and Current

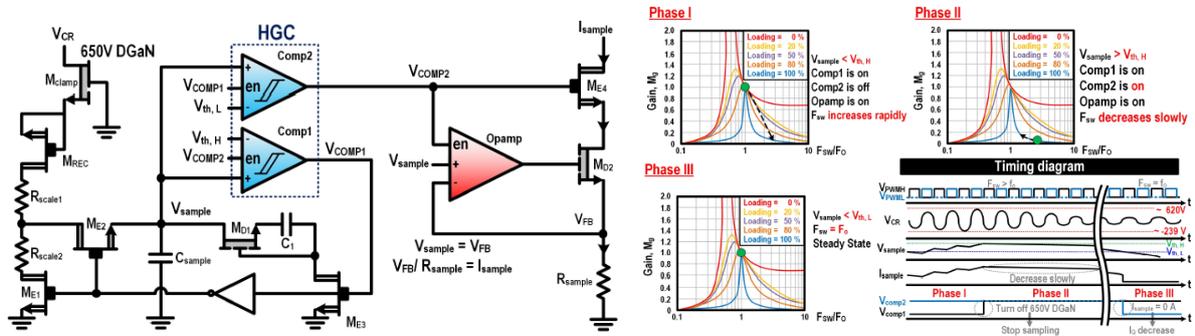
2024 VLSI **Session 2 “Power at High Voltage and Current”**에서는 1) 차량용 CISPR25 규격을 만족시키는 GaN-Buck의 낮은 EMI 달성 기법과 작은 누설 전류를 갖는 Gate driver, 2) 스타트업 초기 공진에 의한 over-voltage를 감소시킨 소프트-스타트업 LLC 컨버터, 3) 고속 하프 브리지 스위치의 LV-Cap Stack 방식의 차지 펌프 Gate driver, 4) 인풋 common 전압 transient에 immune한 전류 센서의 전압과 온도 변화에 따른 저항 값 변화 calibration, 5) trans-inductor를 활용한 High-Step down 컨버터의 고속 transient response가 발표되었다. 본 리뷰는 각 논문의 innovation을 소개한다.

#2-1 차량용 PMIC는 주변 전자기기로의 전자기 신호 간섭을 방지하기 위한 CISPR EMI 규격이 존재한다. 스위칭 레귤레이터 특성상 필연적으로 발생하는 스위칭 노드와 입력에서 발생하는 EMI를 제거하기 위해서 C2-1 논문은 입력 노드의 EMI 감소 방법으로 입력 신호와 동기화된 High-Side GaN의 Gate 신호를 입력으로 받아 입력 신호와 상쇄할 반전된 신호의 위상 차이 $\Delta\phi$ 를 감소시켰다고 주장한다. 스위칭 노드 전압의 EMI는 eGaN의 gate instability 특성을 활용한 TERN (Trapping-Effect Random Number) 클락 생성기로 spread-spectrum을 구현하여 감소시켰다. HS 스위치의 턴-오프 시 발생하는 Drain 전압의 spike는 EMI를 증가시키는 또다른 원인이며, 이를 감소시키기 위해 Adaptive 2-Stage Slew Rate Control 기법이 제안되었고, High Side switch의 gate node의 방전 slew를 조절, EMI를 감소시켰다. 또한 E-GaN을 이용하여 bootstrapping 시 V_{GS} 전압의 감소 없이 낮은 leakage를 갖는 Cross-coupled 구조의 gate driver를 제안하여 효율을 개선하였다. 그 결과 6mW의 전력 소모로 Peak EMI 21.51dBuV, 피크 효율 94.5%, 최대 240W출력을 달성하였다.



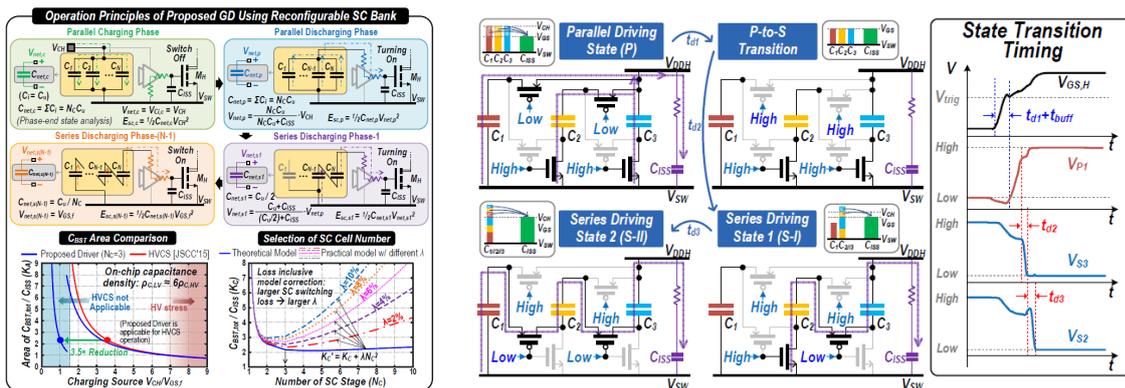
[그림 1] Overall Block Diagram (좌, V_{IN} -Ripple Injector, TERN-Based CLK, 2-Stage SRC)과 총래의 Active Filter 기반 입력 노드 EMI 제거 기법

#2-2 LLC 컨버터는 ZVS와 ZCS가 가능한 특성을 활용해, Front-end에 주로 사용하며 고효율 전력 전달을 가능하게 한다. 본 논문은 LLC 컨버터 startup 시 1차측 transformer의 LC 공진으로 인해 Resonant capacitor에 걸리는 과도한 voltage-stress를 방지하기 위해 아래 [그림 1]의 soft-startup clamping 회로를 제안하였다. 제안하는 기법은 샘플링한 Resonant Capacitor (C_R) 전압이 High-threshold ($V_{TH,H}$) 보다 높은 경우 스위칭 주파수를 천천히 감소시키고 (Phase I), $V_{TH,H}$ 보다 낮은 경우 스위칭 주파수를 급격히 증가, Low-threshold ($V_{TH,L}$) 보다 낮은 경우 스위칭 주파수를 공진 주파수로 동작시켜 C_R 전압의 over-voltage를 방지한다. 그 결과 제안하는 soft-start clamping technique을 적용하지 않은 경우와 비교하여 start-up 시 걸린 2.51kV, -2.15kV의 큰 V_{CR} 전압이 620V와 -239V로 대폭 감소하는 효과를 얻었다. 효율 면으로는 Energy Star와 80 Plus - Titanium 등급의 고효율 컨버터 인증을 받기 위해서 I_Q (Quiescent Current)를 줄이기 위한 Low- I_Q 달성 및 bootstrap cap precharging, High Gate Driving Strength를 갖는 Fast Inverter를 제안하였다. 강한 gate driving strength를 갖는 Fast Inverter에 의해 설계한 LLC 컨버터는 대기 상태에서 즉각적으로 정상상태에 돌입할 수 있으며, 대기 상태에서의 누설 전류를 수 mA에서 60.4uA로 감소시켰다.



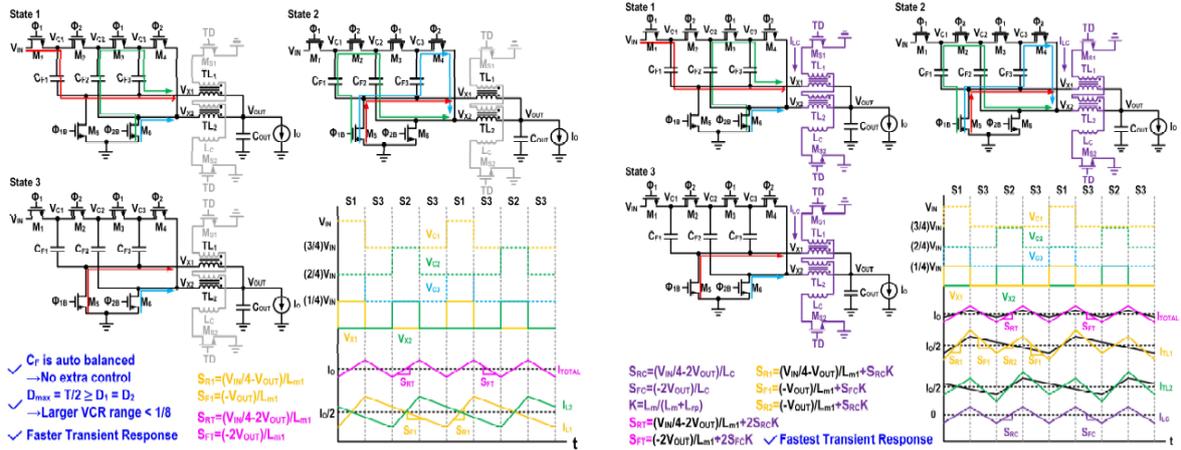
[그림 2] Resonance Capacitor C_R - Soft Startup Clamping 회로와 Phase 별 주파수 조절 방식

#2-3 Half-Bridge(HB) 컨버터의 High-Side 스위치를 구동하기 위한 bootstrap 방법으로, off-chip bootstrap cap을 사용하는 방법은 gate-loop에 발생하는 기생 인덕턴스에 의해 구동 속도 저하, 불안정성, EMI 증가 및 수율 감소의 문제를 가져온다. 이에 대한 방안으로 bootstrap cap을 내장하는 방법이 있으며, 차지 펌프 방식의 gate 구동은 높은 gate-cap 충전 속도로 GaN 소자와 함께 고속 스위칭을 가능하게 하는 장점이 있다. 그러나, 한 번의 커패시터 hard-charging으로 gate capacitance (C_{ISS})를 충전하기 위해서는 높은 전압이 필요하며, 높은 내압 소자 사용은 큰 driving loss를 갖는 문제가 있다. C2-3 논문은 큰 내압 소자 기반의 차지 펌프 Gate driver를 낮은 내압의 소자를 Stacking하여 사용한 high density 차지 펌프 Gate driving 방식을 제안한다. 기존의 차지 펌프 Gate driver는 switched capacitor의 output impedance에 의해 이상적인 전압까지 출력 전압을 레귤레이션할 수 없는 문제가 있지만, 제안하는 LV-Stacking 방식의 차지 펌프 Gate driver를 사용할 경우 LV-Cap-stack을 더 쌓아 V_{GS} 전압을 이상적인 Full turn-on 전압까지 올릴 수 있는 장점이 있다. 제안하는 bootstrap 방식의 GD는 $7.7nC/mm^2$ 의 밀도를 가지며, 이를 적용한 하프 브리지 컨버터는 6MHz의 스위칭 주파수로 동작한다.



[그림 3] 제안하는 LV-Cap Stacking 방식의 multi-step 차지 펌프 Gate Driver 원리 및 동작

#2-4 전류-Path에 series 저항을 연결한 Current sensor는 active 소자를 사용한 센싱 방법 대비 정밀한 값을 얻을 수 있는 장점이 있다. 그러나, 저항 자체의 전압에 따른 저항 값 변화, 온도 변화에 따른 저항 값 변화는 정확한 전류 센싱에 오차를 발생시키는 주 원인이다. 또한, 스위칭 레귤레이터의 스위칭 노드와 연결된 저항은 양단 전압을 센싱할 때 스위칭 시 입력 common 전압의 변동으로 offset 제거용 HV chopper 스위치의 입력 common 전압에 대한 빠른 Tracking이 가능해야한다. C2-4 논문은 입력 common 전압의 빠른 변화를 Tracking하는 High-Voltage chopper를 제안하며, 전압과 온도에 따른 저항 값 변화를 각각 dual-slope ADC 기반의 12-bit 메모리를 갖는 trimming과 1-bit 2차 델타 시그마 모듈레이터와 128-TAP FIR DAC 피드백으로 calibration하는 CCIA를 제안하였다. 그 결과 종래의 sensor와 비교했을 때 가장 낮은 '0.1mΩ'의 낮은 저항으로 ±100A의

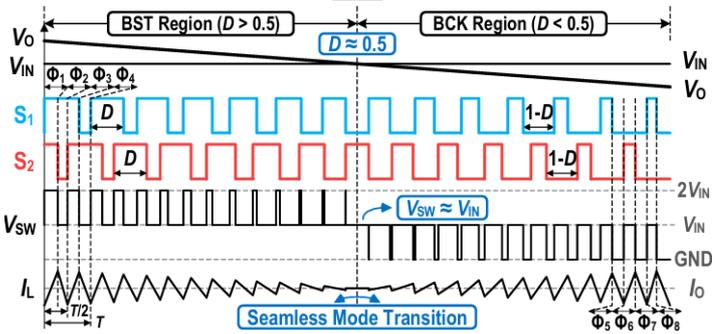
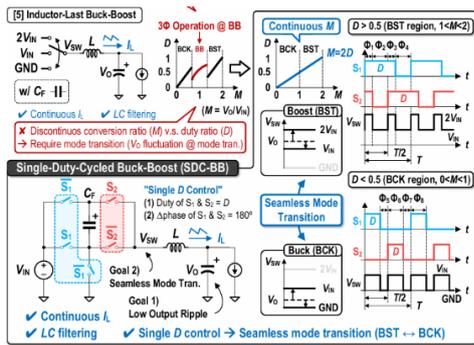


[그림 5] 제안하는 컨버터의 Steady state 동작과 Transient 상황의 트랜스-인덕터 제어 기법

Session 21 Power Converters

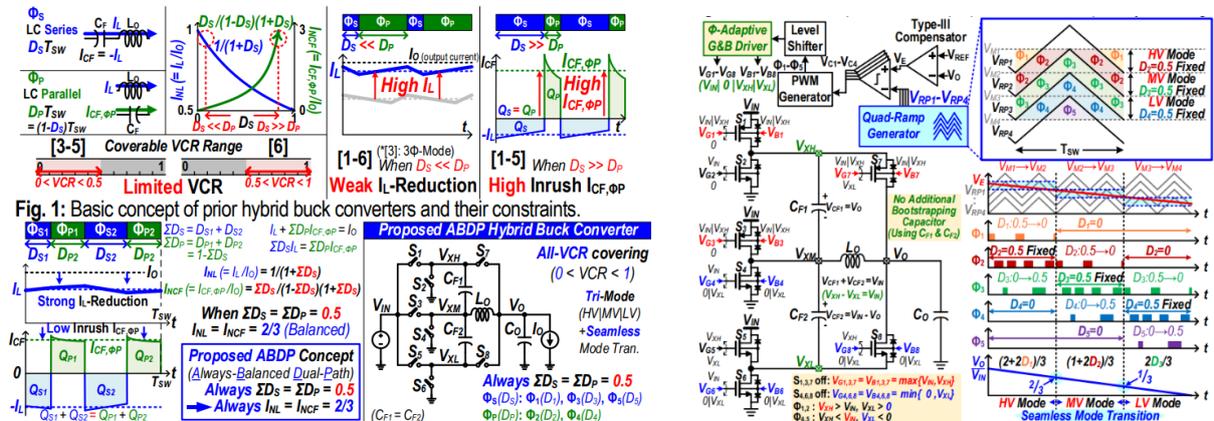
2024 VLSI Session 21의 “Power Converters”에서는 모바일 OLED 디스플레이용 Low-Ripple Buck-boost Converter, DVS용 All-VCR Dual-Path Buck Converter, Substrate Conduction loss를 감소시킨 고 입력전압 IVR chiplet, High Step Down Converter의 짧은 on-time을 개선한 Hybrid Topology 논문이 발표되었다.

#21-1 모바일 OLED 디스플레이의 ELVDD 전원은 디스플레이 Pixel LED의 전류(밝기)를 결정하는 주 전압원이며, 디스플레이의 플리커 현상을 줄이기 위해 ELVDD PMIC 설계 시 낮은 출력 전압 리플을 요구한다. 저휘도에서 고휘도까지 지원하는 디스플레이는 종래의 Boost 컨버터가 아닌 Buck과 Boost 영역을 모두 커버하는 Buck-Boost PMIC를 요구하게 되었고, ELVDD PMIC로서 종래의 백-부스트 컨버터와 Dual-Path Buck-Boost Converter는 출력 전압으로의 불연속적 전류와 capacitor의 Inrush 전류 전달로 인해 큰 전압 리플을 발생시켜 디스플레이의 플리커 현상을 심화시킨다. C21-1 논문은 아래 그림 1과 같이 항상 연속적이며 Inductive 전류만을 전달하는 “Inductor Last” 구조와 3-Level Converter와 같은 Single-Duty Cycle로 Buck에서 Boost 영역까지 컨버터를 제어한다. VCR=2D의 컨버터를 0~D, D~2D의 구간으로 나누어 스위칭 노드 전압 스윙을 V_{IN} 으로 감소시켜 낮은 출력 전압 리플을 갖는다. Voltage mode 기반의 동작으로 DVS 시 seamless transition이 가능하며 $V_{IN}=3.5\sim 4.5V$, $V_O=4V$ 의 transient 상황에서 2.1mV 이하의 출력 전압 리플을 달성하였다.



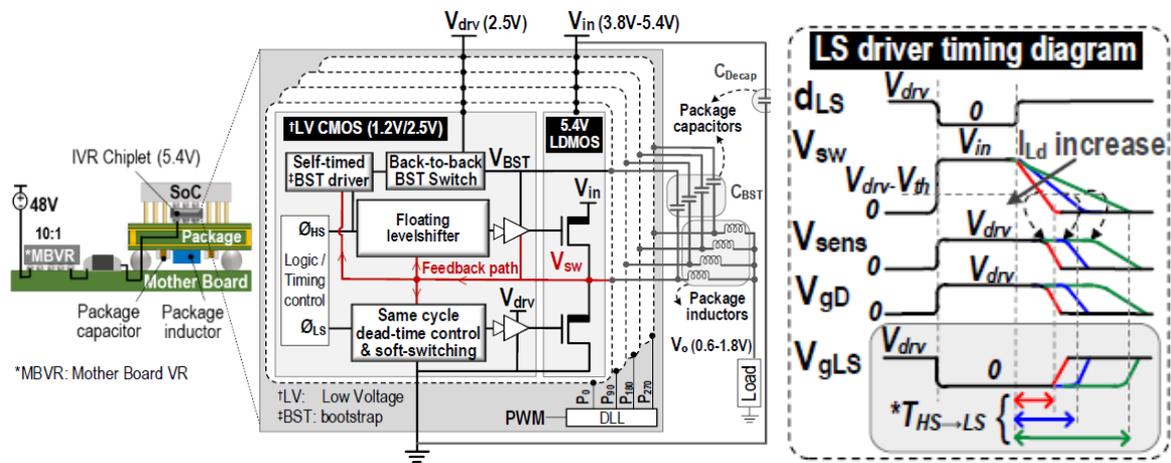
[그림 1] 낮은 출력 전압 리플을 달성한 Inductor-Last Buck-Boost의 Single-Duty Cycle Control

#21-2 인덕터에 기생하는 직렬 저항은 스위칭 레귤레이터로 큰 부하 전류를 공급할 때 효율을 감소시키는 주 원인이다. 최신 소형화 트렌드에 맞는 작은 부피의 인덕터는 기생 직렬 저항이 매우 크기 때문에 부하 공급량이 커짐에 따라 큰 전력 손실을 발생시킨다. 기존의 하이브리드 컨버터는 L-C의 직/병렬 연결 시간을 조절하여 인덕터 전류의 DC 값을 감소, 효율을 개선한다. 그러나, 해당 방법은 필연적으로 L-C의 직렬 혹은 병렬 연결 시간이 줄어들면 인덕터 전류의 DC 값 감소 효과가 줄어들거나, Capacitive한 Inrush 전류로 인한 극심한 효율 감소가 나타나는 특징이 있다. C21-2 논문은 L-C의 직/병렬 연결 시간의 합을 모든 VCR에서 주기의 절반 시간으로 동일하게 제어하여, VCR (i.e., Duty)에 관계없이 항상 인덕터 전류와 Capacitive 전류의 DC 값을 부하 전류의 2/3만큼 감소시킨다. 그 결과 작은 사이즈의 인덕터를 사용하면서, 모든 VCR에서 큰 부하를 공급할 시 발생하는 Conduction loss를 감소시킨다. 해당 제어는 4개의 Level로 구분된 Quad-Ramp Generator로 LV(0~1/3), MV(1/3~2/3), HV(2/3~1) VCR 영역의 사용을 Seamless하게 구현하였으며, 설계한 PMIC는 칩과 수동 소자를 포함하여 2.7mm³의 부피로 최대 2.5A의 부하를 공급, 94.2%의 피크 효율을 갖는다.



[그림 2] VCR에 무관한 인덕터 DC 전류, 커패시터 전류 Reduction 및 Quad-Ramp 기반 구현

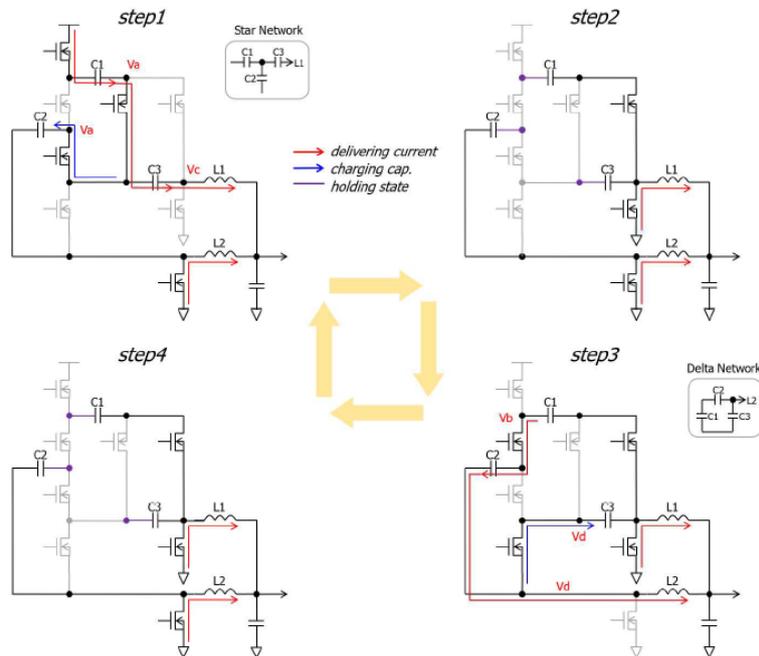
#21-3 고성능 서버 컴퓨팅 및 GPU는 연산 시 매우 큰 전류를 필요로 하며, 빠른 출력 전압의 recovery를 위해 낮은 출력 임피던스를 갖는 Multiphase Buck 컨버터를 사용한다. 그러나 Power-Supply Chain(PSC)의 끝단에 위치한 SoC는 낮은 입력 전압으로 (예, 1.8V) 전력을 공급할 시 PSC의 앞단으로 갈수록 증가하는 고전류에 의한 전력 손실이 두드러지며, 이 때 증가한 인풋 파워는 높은 전류를 더 증가시킨다. 특히 Voltage Regulator(VR) 연결부에서의 기생 저항에 의한 효율 감소와 VR의 직렬 연결 시 발생하는 구조적 효율 감소는 전력 누수의 주 원인이다. C21-3 논문은 종래의 1.8V 대비 3배 증가한 5.4V 입력 전압에서 0.6~1.8V를 출력 전압을 레귤레이션하는 3차원 집적(3D) Multiphase Buck 컨버터를 제안, Planar한 VR 연결부를 Z축-3차원 전력 공급 방식으로 변경하여 연결부에서의 전력 손실을 감소시켰다. 또한, 패키지 인덕터와 커패시터를 사용하는 10MHz의 높은 주파수의 IVR을 NMOS로 구현할 때 필요한 Bootstrapping의 Deadtime을 확보하는 회로를 설계하여 안정적인 동작을 보장하였다. 스위칭 노드 전압의 변화를 관찰한 Deadtime 설정으로 LS 스위치의 부하-적응형 Deadtime이 가능하다. 그 결과 10MHz의 고속 스위칭을 달성하며 LS 스위치의 ZVS로 고주파수로 동작 시 발생하는 스위칭 손실을 개선하였다. 발표한 PMIC는 9.3A/mm²의 전류 밀도를 가지며 최대 80A의 부하를 공급하며 94.5%의 피크 효율을 갖는다.



[그림 3] 제안하는 3D 집적 Multiphase Buck Converter와 Load Adaptive Deadtime 제어

#21-4 Power-supply chain의 주 효율 감소 원인 중 하나는 VR의 직렬 연결에 의한 Cascade efficiency drop이다. 48V Bus로부터 곧바로 SoC 전원을 공급하는 High-Step down 컨버터는 VR의 직렬 연결에 의한 효율 감소를 Single-stage로 구현하여 효율을 증가시키는 장점이 있다. 그러나, Buck (=Step-down) 컨버터의 높은 입력전압에 의해 낮아진 Duty는 SoC의 V_{DD} 마진과 함께 높아진 스위칭 주파수에 의해 물리적으로 짧은 on-time을 가지며 부하 전류의 과도 응답, DVS 제어 시 출력 전압의 큰 under-overshoot을 발생시킬 수 있다. C21-4 논문은 48V to 0.6~2V 전원 공급을 하는 Hybrid 컨버터를 제안하며, 종래의 Buck 컨버터 대비 5x, 8x Duty expansion을 가능하게 하는 Start-Delta 스위

칭 토폴로지를 제안하였다. 또한, 저면적 게이트 드라이빙용 플로팅 전원, 라인-레귤레이션 성능 개선, 디지털 Circuit의 비선형성 개선 아이디어를 소개하였으며, 2A/2us의 부하 응답 시 최소 45mV의 under/overshoot을 달성하였다.



[그림 4] 제안하는 5x Duty extended 48V to 1V 컨버터의 Star-Delta 스위칭 동작

저자정보



명예기자 박수연

- 소속 : KAIST 전기 및 전자공학부 박사과정
- 연구분야 : Power Management IC 설계
- 이메일 : tndjs12221@kaist.ac.kr
- 홈페이지 : <https://www.icdesignlab.net>

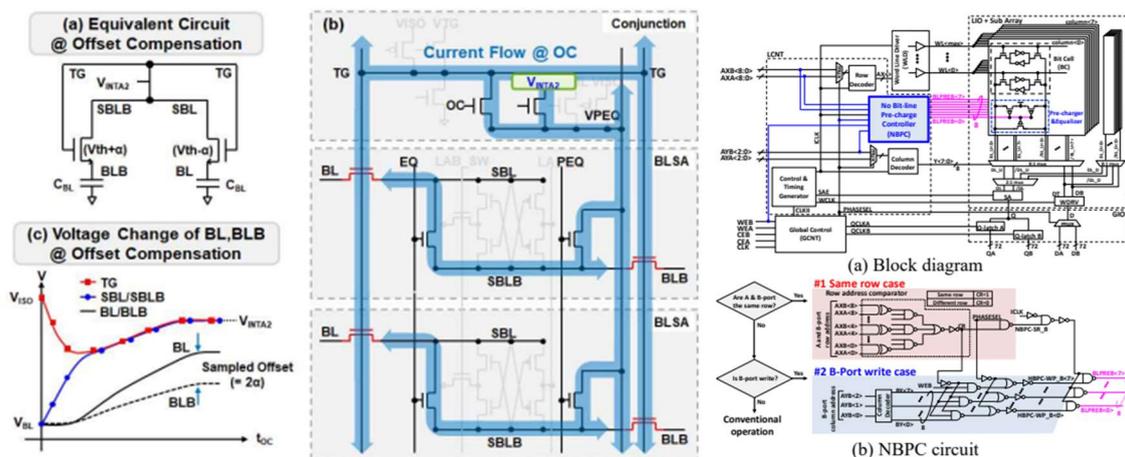
2024 IEEE VLSI Review

서울대학교 전기정보공학부 박사과정 박현준

Session 16 Memory Circuits

Session 16은 최신 반도체기술을 기반으로 한 다양한 고성능 메모리 설계 및 회로 최적화 기법들이 소개되었다. 이 세션에서는 SRAM, MRAM 등 다양한 메모리 기술이 고속 동작, 저전력 소모, 그리고 높은 신뢰성을 확보하기 위한 방법론을 제시하며, 특히 Automotive grade 1 및 HPC 애플리케이션을 목표로 한 논문들이 소개되었다.

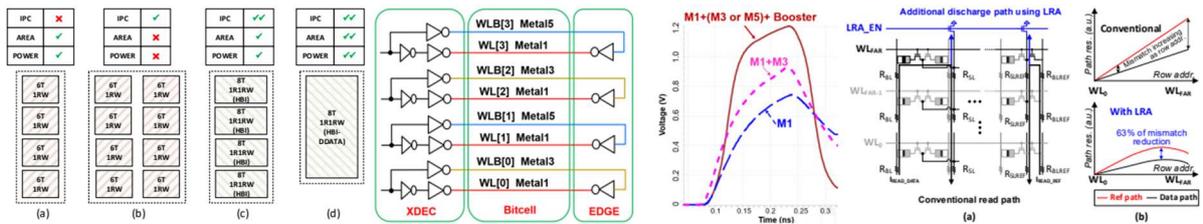
#16-1 이 논문은 저전압 DRAM을 위한 OC-CTPS BLSA를 제안한다. 제안된 비트 라인 센스 앰프(BLSA)는 14nm DRAM 공정으로 구현되었으며, 오프셋 보정을 통해 트랜지스터의 V_{th} 불일치를 보정함으로써 충전 전달(CT)의 정확성을 높였다. 기존 BLSA 방식에서 공정 변동으로 인한 V_{th} 오프셋은 충전 전달의 불안정성을 유발해 불량 비율(FBR)을 증가시키는 문제가 있었으나, OC-CTPS 방식은 이를 다이오드 연결을 통해 해결하였다. 이는 추가적인 회로 크기 증가 없이 BLSA의 성능을 개선할 수 있게 해준다. 제안된 OC-CTPS BLSA는 2.75ns의 오프셋 보정 시간(t_{OC})과 1.5ns의 충전 전달 시간(t_{CT})을 통해 안정적인 감지 기능을 제공한다. 실험 결과, -25°C 와 100°C 에서 각각 250ps와 500ps의 3시그마 창을 확보하였으며, 이는 다양한 온도 조건에서도 전압 임계값 변화를 안정적으로 추적할 수 있음을 보여준다. 또한, 0.75V의 낮은 동작 전압에서도 기존 BLSA보다 94% 향상된 불량 비율을 기록하며, 저전압 환경에서의 안정성을 증명하였다. 결론적으로, 이 논문에서 제안된 OC-CTPS BLSA는 저전력, 저전압 환경에서 DRAM의 신뢰성을 크게 향상시킬 수 있는 효과적인 솔루션을 제공한다.



[그림 2] (좌) 16-1 (우) 16-2

#16-2 이 논문은 3nm Fin-FET 기술을 기반으로 한 19.87 Mbit/mm²의 비트 밀도를 달성한 2RW(2-읽기/쓰기) 의사 듀얼 포트(PDP) 6T SRAM 설계를 제안한다. 이 설계는 성능을

개선하기 위해 고저항 와이어 추적을 도입하여 다양한 PVT(공정, 전압, 온도) 조건에서 최적의 속도를 유지하고, 순차적 접근을 고려한 동적 전력 감소 기법을 통해 불필요한 비트라인 프리차지를 제거함으로써 동적 전력을 최대 43%까지 절감한다. 이 SRAM은 288-kbit PDP 매크로로 제작되었으며, 2RW 듀얼 포트 SRAM과 비교하여 면적 효율성을 극대화하였다. 6비트 DAC 아키텍처를 사용하여 데이터를 처리하며, 비트라인 프리차지를 제어하는 새로운 회로 구조를 통해 전력 소모를 크게 줄일 수 있었다. 특히, 동일 행(row)을 순차적으로 접근할 때 1차 프리차지를 생략하여 전력 절감을 실현했다. 또한, 고저항 와이어가 포함된 회로에서 발생할 수 있는 타이밍 문제를 해결하기 위해 비트라인 프리차지 타이밍 최적화를 도입하였다. 이를 통해 다양한 PVT 조건에서도 안정적인 동작을 보장하며, SRAM의 성능과 전력 효율성을 향상시켰다. 제안된 SRAM은 43%의 비트라인 전력 절감을 달성했으며, 1.64 GHz에서 동작할 수 있다. 최종적으로, 제안된 3nm PDP SRAM은 경쟁 기술에 비해 더 높은 비트 밀도와 전력 효율성을 제공하여 차세대 SoC 및 FPGA 설계에 적합한 솔루션을 제시한다.



[그림 2] (좌) 16-3 (중) 16-4 (우) 16-5

#16-3 이 논문은 3nm 기술에서 7GHz 이상의 성능을 제공하는 1R-1RW(1 읽기-1 쓰기/쓰기) 고대역폭 SRAM을 Arm HPC(High Performance Computing) 프로세서에 통합한 설계를 소개한다. 기존 8T-1R1W 메모리 아키텍처를 개선하여 추가적인 읽기 포트를 제공함으로써 1R1RW 기능을 구현하였으며, 이를 통해 L1-데이터 캐시에서 읽기 대역폭을 두 배로 증가시키고 Instruction Per Cycle을 1% 이상 향상시켰다. 또한, 이 새로운 HBI(High Bandwidth Instance) 메모리 아키텍처는 CPU의 물리적 설계에서 라우팅 혼잡과 지연을 줄여 13%의 면적 감소와 10-15ps의 라우팅 지연 감소를 달성하였다. 제안된 HBI-1R1RW 메모리는 기존 6T-1RW 메모리와 비교했을 때 33% 더 작은 L1-데이터 캐시 면적을 구현했으며, 두 개의 읽기 포트를 통해 읽기 성능을 개선하였다. 이 아키텍처는 word line 길이를 절반으로 줄이고, 두 개의 메모리 뱅크를 동시에 접근할 수 있는 구조를 사용하여 메모리 접근 시간을 최적화했다. 또한, RC 불일치 문제를 해결하기 위해 이중 word line 기술을 사용하여 두 번째 포트에서의 성능을 개선하였으며, 고속 다이내믹 멀티플렉서와 센싱 회로를 적용하여 첫 번째 포트에서도 빠른 읽기 속도를 보장하였다. HBI-Ddata 메모리는 네 개의 HBI-1R1RW 뱅크를 통합하여 CPU 설계에서 핀 접근성을 개선하고 라우팅 혼잡을 완화하였다. 이를 통해 CPU 성능 향상과 함께 전력 효율성을

극대화하였다. 또한, 백엔드 금속을 사용하여 내부 메모리 라우팅의 저항과 결합 커패시턴스를 줄임으로써 메모리 성능을 최적화하였다. 3nm 공정으로 제작된 테스트 칩에서는 7GHz 이상의 주파수와 11.2 Mbit/mm²의 8T 기반 SRAM 비트 밀도를 기록하였으며, 0.945V에서 108ps의 빠른 메모리 접근 시간(0.75V, TT에서 85°C 기준)을 달성하였다. 최종적으로, HBI-1R1RW 메모리는 Arm HPC 프로세서의 L1-데이터 캐시에서 사용되며, 높은 성능과 효율성을 제공하여 CPU의 전반적인 주파수를 향상시킬 수 있음을 입증하였다.

#16-4 이 논문은 3nm FinFET 기술을 사용하여 1024x640 및 2048x640 크기의 단일 포트 고속 멀티뱅크 SRAM 매크로를 설계한 내용을 다룬다. 이 설계는 고성능 컴퓨팅(HPC) 애플리케이션을 대상으로 하며, single-port 6T 비트셀을 기반으로 하여 최고 3.3GHz에서 작동하는 SRAM 매크로를 구현하였다. 또한, 0.55V-1.35V의 넓은 전압 범위에서 동작이 가능하여 DVFS와 같은 전력 효율성 개선 기술을 지원한다. 제안된 SRAM 설계에서는 워드라인, 글로벌 클록, 글로벌 비트라인 부스팅 기술을 통해 속도를 개선하고, 디코더 신호용 스플릿 드라이버를 도입하여 속도를 37% 향상시켰다. 또한, 읽기 보조 회로를 통해 최대 전압(V_{max})을 1.35V로 높여 성능을 최적화했다. 이 모든 기술을 적용한 결과, 1V 100°C에서 3.3GHz의 성능을 달성하였다. 이 설계는 멀티뱅크 구조로 구현되었으며, 이를 통해 워드라인과 X-Dir 신호 전송의 효율성을 높였다. 부스팅 회로와 라우팅 최적화를 통해 리피터 추가 없이 신호 전송 속도를 개선하였다. 특히, Y방향에서는 워드라인과 비트라인 신호를 개선하기 위해 보조 인버터와 글로벌 신호 부스터를 사용했으며, X방향에서는 글로벌 클록(GCK) 지연과 X 프리디코드 신호를 최적화했다. 결과적으로, 이 논문에서 제시한 회로 기술은 전력 효율성을 개선하고, SRAM의 면적 효율성을 극대화하였다. 이를 통해 기존의 멀티뱅크 SRAM 설계보다 더 큰 매크로 크기를 구현하면서도 밀도×주파수/전력 지표에서 최고의 성능을 달성했다.

#16-5 이 논문은 14nm FinFET 공정에서 구현된 128Mb eMRAM 설계를 제안하며, 고온에서의 읽기 마진 저하 문제를 해결하기 위한 회로 기술을 소개한다. LRA를 통해 데이터와 참조 읽기 경로 사이의 저항 불일치를 최소화하고, 온도 추적 회로를 사용하여 참조 저항을 적응적으로 변경하여 고온에서도 신뢰할 수 있는 읽기 동작을 보장한다. 이 설계를 적용한 256-IO MRAM 매크로는 17.88Mb/mm²의 고밀도를 달성하였으며, 0.60V 코어 전압에서 -40°C에서 160°C까지의 온도 범위에서 11ns의 읽기 속도를 성공적으로 구현하였다. STT-MRAM(스핀 전이 토크 MRAM)**은 높은 내구성, 유지력, 빠르고 저전력 쓰기 성능 등의 장점이 있지만, 터널 자기 저항 비율(TMR)이 고온에서 저하되는 문제가 있다. 이는 데이터 '0'과 '1' 사이의 저항 차이가 감소하여 읽기 마진이 줄어드는 문제를 초래한다. 이 논문에서는 이러한 문제를 해결하기 위해 오프셀 누설 전류와 저항 불일치를 줄이고, 고온에서도 신뢰할 수 있는 읽기 성능을 유지하는 방법을 제안한다. LRA 데이터와 참조 경로 간의 저항 불일치를 완화하여 읽기 속도와 전류를 증가시키며, TTC는 온도

변화에 따라 참조 저항을 적응적으로 조정하여 TMR 저하와 저항 불일치를 보상한다. 또한, NWL 생성기를 도입하여 선택되지 않은 셀에서 발생하는 누설 전류를 제거하고 접근 트랜지스터의 신뢰성을 개선하였다. 128Mb MRAM 매크로는 7.16mm²의 면적을 차지하며, 이전 연구에 비해 40% 향상된 읽기 마진을 기록하였다.

저자정보



박현준 박사과정 대학원생

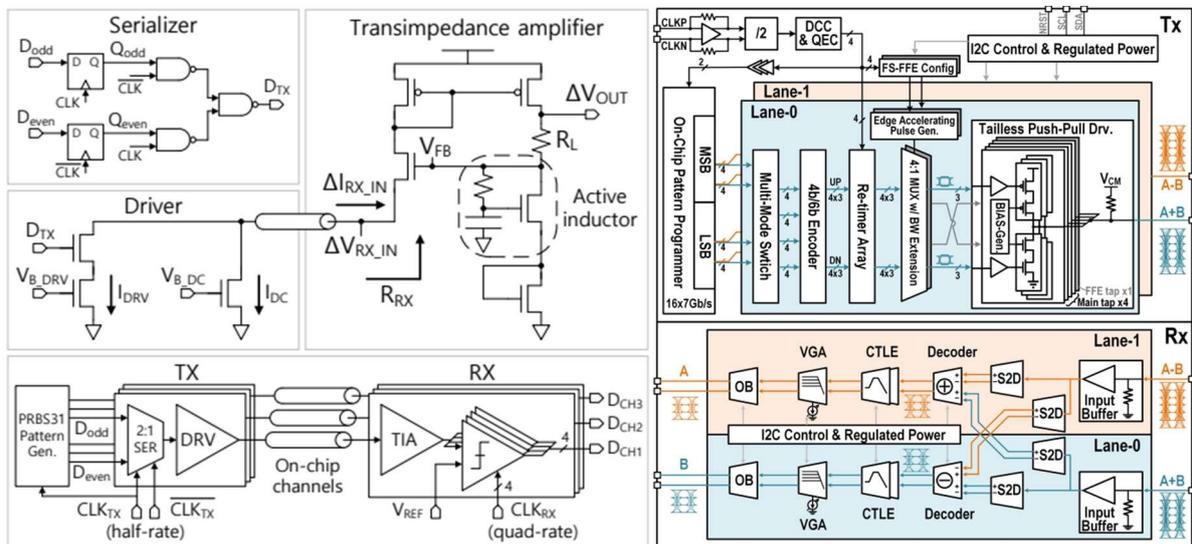
- 소속 : 서울대학교
- 연구분야 : HBM, Chord Signaling, Information Theory
- 이메일 : spp098@snu.ac.kr
- 홈페이지 : <https://sites.google.com/view/wschoi?pli=1>

2024 IEEE VLSI Review

서울대학교 전기정보공학부 박사과정 박현준

Session 1 Wireline Circuits

Session 1은 저전력 고속 Wireline 통신 기술에 초점을 맞춘 4개의 혁신적인 논문이 발표되었다. 이 세션을 통해 Wireline 기술에 요구되는 최신 동향과 발전된 설계 및 테스트 방법에 대한 통찰력을 얻을 수 있다.

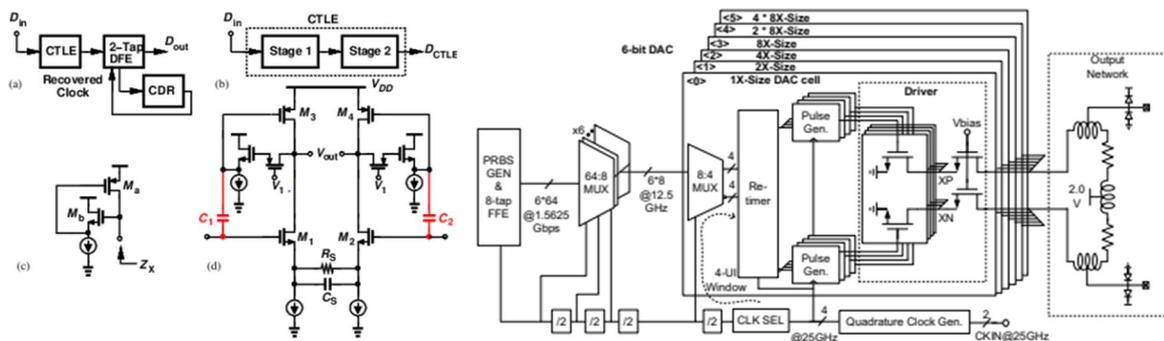


[그림 1] (좌) 1-1 (우) 1-2

#1-1 이 논문은 고밀도, 단거리 온칩 interconnect에서 crosstalk 문제를 해결할 수 있는 single ended current mode transceiver를 제안한다. 이 트랜시버는 XTC(crosstalk cancellation)를 통해 고속 데이터 전송 시 발생하는 신호 간섭 문제를 해결한다. 기존 방식에서는 shielding을 사용하지만, 이는 대역폭 밀도를 감소시킨다. 제안된 트랜시버는 current mode signaling과 RX termination 저항 최적화를 통해 별도의 오버헤드 없이 크로스토크를 효과적으로 상쇄하여, 13.3Tb/s/mm의 높은 edge density에도 불구하고 246fJ/bit의 매우 낮은 에너지 효율성을 달성하였다.

#1-2 이 논문은 2×56 Gb/s 단일 종단 직교 PAM-7 트랜시버를 제안하며, 채널 독립적 크로스토크 상쇄를 위한 인코더 기반 크로스토크 상쇄(EB-XTC) 기법을 적용하였다. 제안된 트랜시버는 두 개의 상관되지 않은 PAM-4 신호를 직교 PAM-7 신호로 변환하여 밀접하게 결합된 차동 채널에서 발생하는 강한 크로스토크 잡음을 제거한다. 이 트랜시버

는 **인코딩 송신기(TX)**와 **디코딩 수신기(RX)**로 구성되며, TX는 크로스토크 면역성을 갖춘 PAM-7 신호를 전송하고, RX는 PAM-7 신호를 다시 PAM-4 신호로 복원한다. 본 연구의 주요 기여는 채널 독립적인 크로스토크 상쇄를 실현했다는 점이다. 이는 기존의 PAM-4 신호가 크로스토크에 민감하다는 한계를 해결하는 새로운 접근법으로, PAM-4 신호를 PAM-7 신호로 인코딩함으로써 크로스토크를 제거할 수 있었다. 또한, 송신기에서 5개의 슬라이스로 이루어진 tailless push-pull 드라이버를 사용하여 신호의 전압 스윙을 0.6 Vpp로 높여 충분한 SNR을 확보하였고, 수신기에서 고주파 증폭 및 조정을 통해 신호 복원을 보장하였다. 측정 결과는, 56 Gb/s/pin의 전송 속도를 유지하면서 크로스토크를 효과적으로 제거한 것이 입증되었다. EB-XTC를 적용했을 때 크로스토크가 완전히 제거되었고, 채널 1에서 0.3 UI의 수평 눈 개방(horizontal eye opening)과 67 mV의 수직 눈 개방(vertical eye opening)을 기록했다. 또한, 제안된 PAM-7 신호 및 EB-XTC 방식은 채널 비대칭성에도 불구하고 완전한 크로스토크 제거와 채널 독립적 성능을 유지했다.



[그림 1] (좌) 1-3 (우) 1-4

#1-3 이 논문은 56 Gb/s NRZ 수신기를 제안하며, 저전력과 소형 면적을 목표로 새로운 아키텍처와 회로 기술을 적용하였다. 28nm CMOS 기술로 구현된 수신기는 25 dB 이상의 채널 손실을 가지는 환경에서 56 Gb/s의 전송 속도를 지원하며, 비트오류율(BER)이 10^{-12} 이하를 달성하였다. 이 수신기는 기존의 PAM-4 방식에 비해 전력 소모가 크지 않으면서도 높은 성능을 유지한다. 논문에서 제시된 설계는 풀레이트(full-rate) 방식을 채택하여, 반레이트나 쿼터레이트 아키텍처보다 더 낮은 전력 소비를 구현했다. 이는 DFE와 CTLE에서의 부하가 감소하고, CDR의 입력 커패시턴스가 작아지면서 가능해졌다. 컴팩트한 플로어플랜 덕분에 인터커넥트 길이가 짧아져 버퍼가 필요 없고, 이는 추가적인 전력 절감으로 이어진다.

#1-4 이 논문은 200-Gb/s PAM-4 송신기를 제안하며, 1.6-Vppd 출력 스윙과 클록 스쿠보정 기능을 12nm FinFET 기술에서 구현한 설계를 다룬다. 데이터 센터 간의 고속 전송을 위한 광 전송 기술이 중요해짐에 따라, 기존의 SiGe 기반 브리지 드라이버는 높은 전

력 소모와 CMOS ASIC과의 호환성 문제를 가지고 있다. 이를 해결하기 위해, 제안된 송신기는 CMOS 기반으로 고출력 스윙을 직접 생성하여 통합 밀도를 높이고 전력 효율성을 개선한다. 송신기는 6비트 디지털-아날로그 변환기(DAC) 아키텍처를 채택하며, 8탭 디지털 FFE와 25-GHz Quadrature 클록을 사용하여 고속 데이터를 처리한다. 펄스 생성기는 두 단계의 동적 논리 구조를 사용하여 지터를 감소시켰고, QCG는 클록 스큐를 감지하고, 오류를 보정하여 정확한 데이터 전송을 가능하게 했다. 결과적으로, 이 송신기는 3.32 pJ/bit의 전력 효율과 0.116 mm²의 소형 면적으로 높은 출력 스윙을 제공하며, 고속 통신을 위한 데이터 센터 및 실리콘 포토닉스 응용에 적합한 솔루션을 제시한다.

저자정보



박현준 박사과정 대학원생

- 소속 : 서울대학교
- 연구분야 : HBM, Chord Signaling, Information Theory
- 이메일 : spp098@snu.ac.kr
- 홈페이지 : <https://sites.google.com/view/wschoi?pli=1>

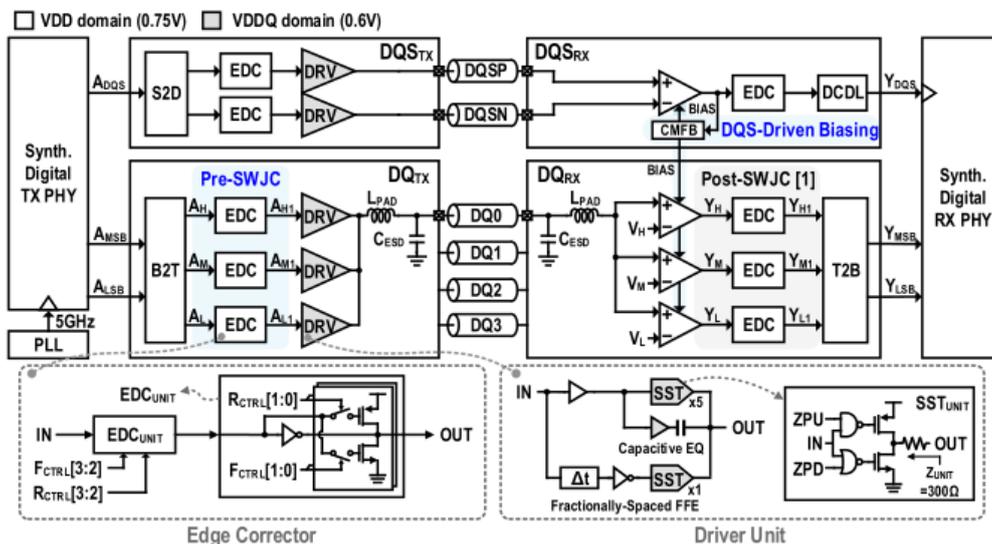
2024 IEEE VLSI Review

KAIST 전기및전자공학부 박사과정 임규완

Session 22 Wireline Circuits II

“Wireline Circuits” 라는 주제로 만들어진 이번 VLSI의 22번 session에서는 총 5편의 논문이 발표되었다. 5편 중 2편은 high-speed optical communication을 주제로 하고, 3편은 memory 및 pcie 등에 쓰이는 PAM-4 TRx에 관한 연구 결과를 제시하였다. 이번 리뷰에서는 switching jitter compensation technique을 통해 timing margin을 개선한 22-3번 논문에 대해 살펴보겠다.

22-3 본 논문은 equalization과 별개로 발생하는 switching jitter에 주목하여, 이를 switching jitter compensation (SWJC)를 사용해 보상함으로써 timing margin을 0.26UI에서 0.39UI까지 늘리는 방법을 제시하였다.



[그림 1] 제안하는 PAM-4 transceiver 구조

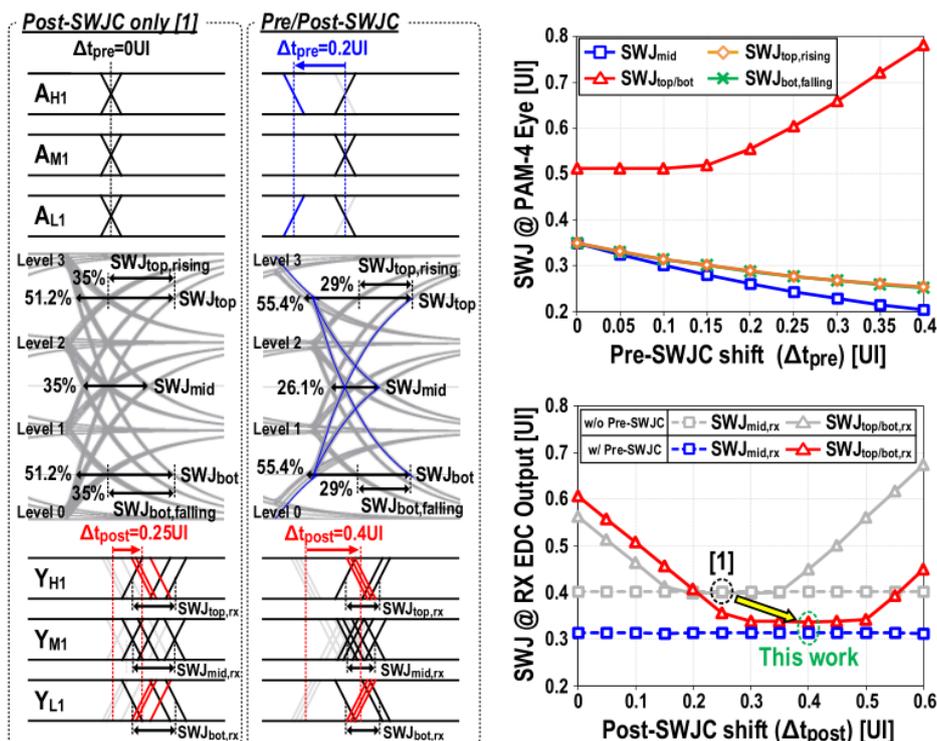
[그림 1]은 논문에서 제안하는 PAM-4 transceiver의 구조를 보여준다. 이의 transmitter 및 receiver에 적용된 edge corrector (EDC)가 논문에서 제안하는 주요한 아이디어이다.

이는 rising 및 falling edge의 delay를 개별적으로 4-bit control code를 이용하여 control 할 수 있고, 이의 resolution은 4ps이며 tuning range는 60ps이다.

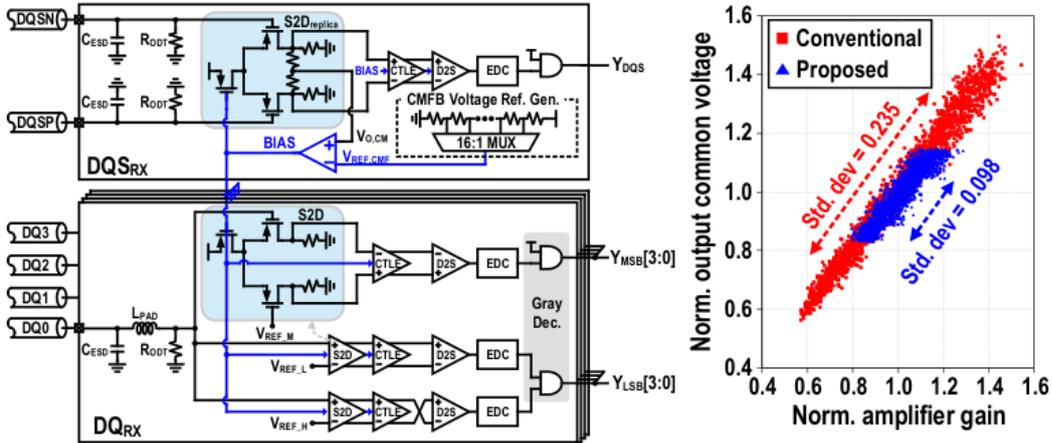
Transmitter은 source-series termination (SST) driver 구조를 기본으로 하여, fractionally-spaced feedforward equalization 및 capacitive peaking equalization을 이용하였다. 여기서 EDC는 main driver 전에 삽입되어 (pre-SWJC) rising/falling delay를 조절한다.

Receiver은 제안하는 work에서 새로이 제안된 DQS-driven biasing amplifier를 이용하여 CMOS level thermometer code로 입력 신호를 변환하고, 다시 EDC를 통하여 이의 timing margin을 늘린 후 (post-SWJC) thermometer-to-binary 변환을 거쳐 deserialize 된다.

[그림 2]는 제안하는 EDC를 통한 Tx단에서의 pre-SWJC 및 RX단에서의 post-SWJC의 효과를 보여준다. Post-SWJC는 이미 이전 work에서 제안된 바 있으며, 이는 Rx에서 bottom과 top의 SWJ를 줄여주는 효과가 있지만, middle의 SWJ는 줄일 수 없다. 제안하는 work에서는, 중앙의 SWJ를 줄이기 위해, AH1과 AL1 (thermometer-coded driver input)의 falling (AH1) edge와 rising (AL1) edge를 앞으로 당겨, level 3에서 falling하는 3개의 transition과 level 0에서 rising하는 3개의 transition을 일찍 일어나게 한다. 이를 통해, PAM-4의 middle eye의 timing margin이 그림 2에 보는 것처럼 증가하게 된다. 이러한 timing margin 개선 효과는 pre-SWJC shift가 클수록 개선되지만, 0.15UI이상의 shift를 발생시킬 경우, top 및 bottom의 SWJ에 영향을 주게 되어 이 이상의 shift는 사용할 수 없다.



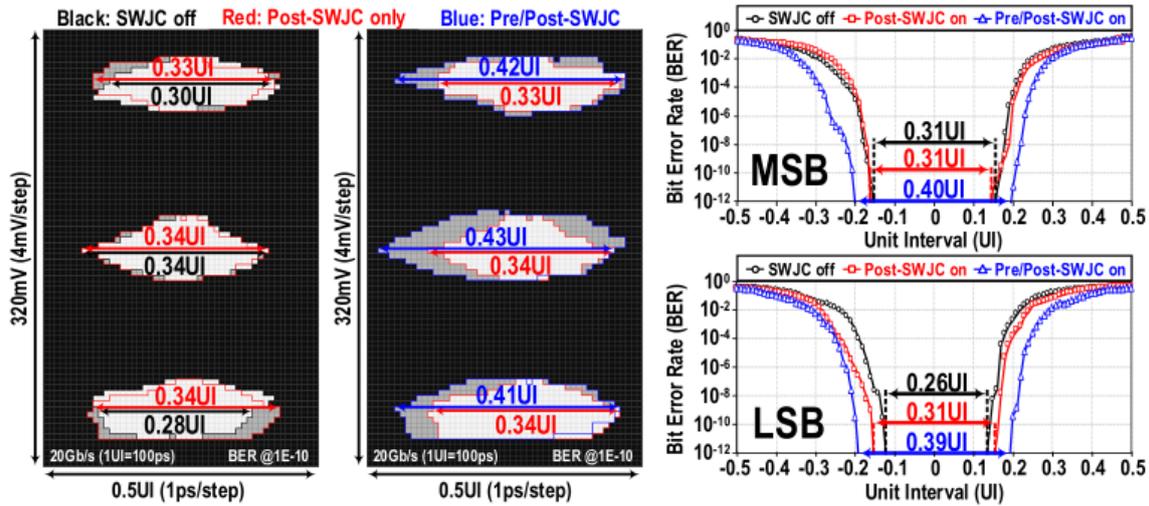
[그림 2] 제안하는 Pre/Post SWJC



[그림 3] 제안하는 DQS-driven biasing 및 이의 효과

[그림 3]은 제안하는 work에서 사용한 DQS-driven biasing 및 이의 효과를 보여준다. 제안하는 work은 differential clock의 common-mode 전압이 거의 변화하지 않는다는 점을 활용하여, 두 clock의 common mode를 활용하여 S2D amplifier의 bias를 설정하고, 이를 다른 single-ended lane의 S2D에 동일하게 사용하여 각 amplifier의 bias를 설정하였다. 이러한 방식을 통해, 각 S2D의 bias는 [그림 3]의 오른쪽처럼 variation이 줄어들게 된다.

[그림 4]는 20Gb/s에서 측정된 RX eye와 bathtub curve를 보여준다. 먼저 RX eye를 살펴 보면, 제안하는 SWJC로 인해 RX eye width가 0.33UI/0.34UI/0.34UI에서 0.42UI/0.43UI/0.41UI로 증가한 것을 확인할 수 있으며, bathtub curve에서는 제안하는 pre/post SWJC가 10^{-12} 이하의 BER에서 타이밍 마진을 0.31UI/0.26UI에서 0.4UI/0.39UI로 충분히 증가시키는 것을 확인할 수 있다.

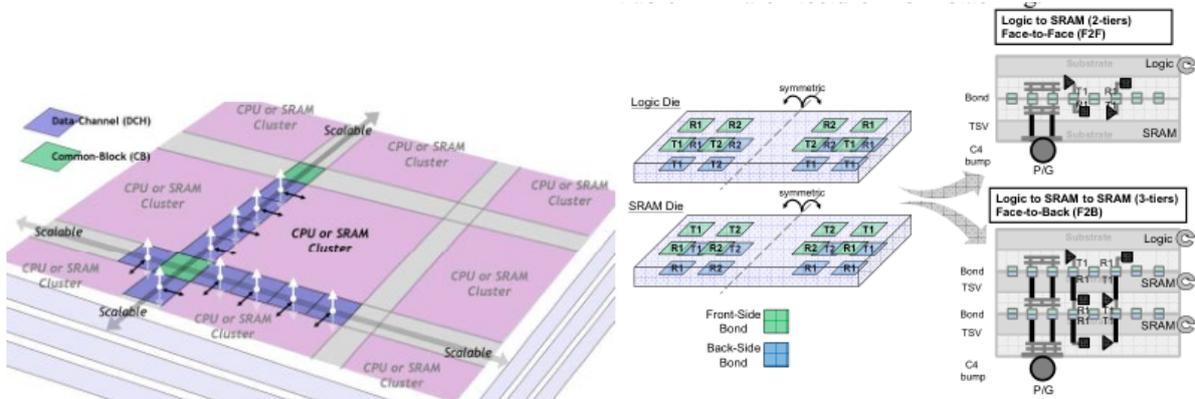


[그림 4] 측정된 Rx eye 및 bathtub curve

Session 14 Very High-speed Wireline

“Very High-speed Wireline”이라는 주제로 만들어진 이번 VLSI의 14번 session에서는 총 5편의 논문이 발표되었다. 5편 중 3편의 논문은 die-to-die (D2D) extremely short-reach (XSR) serial interface를 주제로 하고, 2편은 high-speed optical communication을 주제로 연구 결과를 제시하였다. 이번 리뷰에서는 9um pitch 3D package를 사용해 매우 높은 면적당 data rate을 달성한 14-1번 논문에 대해 살펴보겠다.

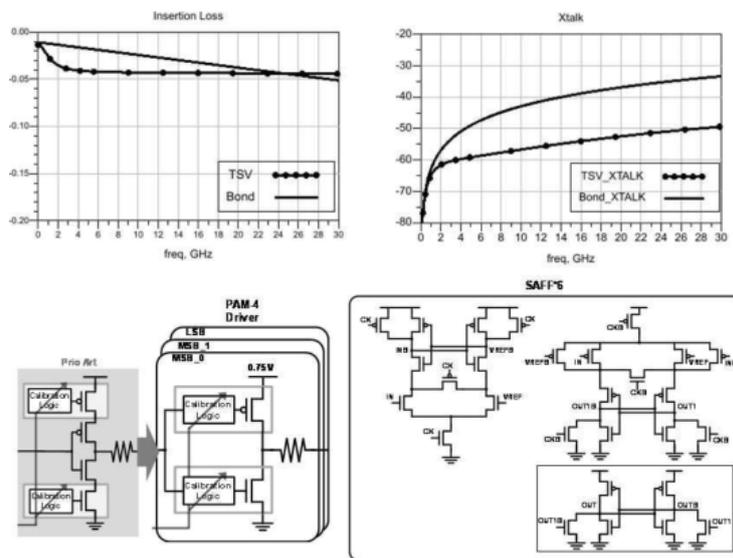
14-1



[그림 1] (왼쪽) scalable D2D 구조 with 3D stacking, (오른쪽) 3D stacking을 위한 bond/TSV 구조

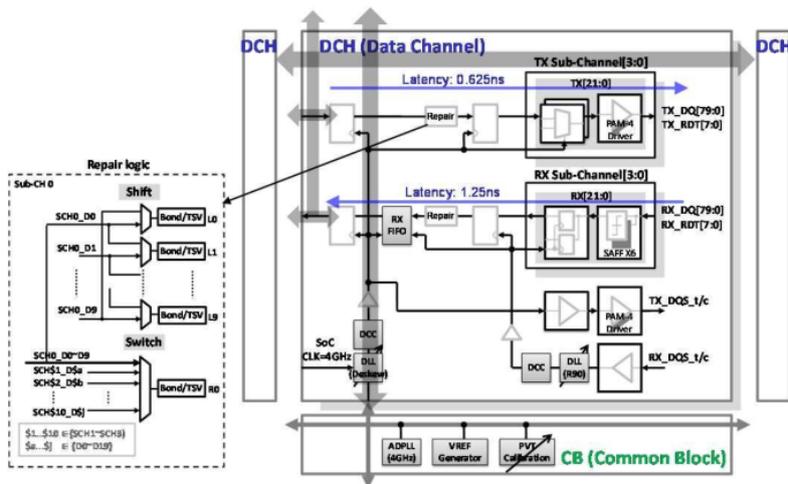
본 논문은 5nm 공정으로 제작된 CPU (logic) die와 6nm 공정으로 제작된 SRAM die를 9um pitch를 가지는 bond와 through-silicon-via (TSV)를 이용하여 3D stacking 하여 높은 interconnect density를 달성하고 면적당 데이터 전송 효율을 극대화한 D2D interface를 제안한다.

[그림 1]은 논문에서 제안한 scalable한 D2D 구조와 이를 위한 3D stacking 구조를 보여준다. 제안하는 구조에서 각 die는 위/아래에 각각 bond가 위치하여 3D stacking을 통한 연결이 가능하게 되어 있고, [그림 1]의 오른쪽과 같이 face-to-face (F2F)로 연결되거나, face-to-back (F2B)로 연결되고, F2B로 연결된 경우는 face와 back 간의 연결은 TSV로 구성된다.



[그림 2] (위) Bond/TSV loss & crosstalk, (아래) PAM-4 driver & Receiver sense-amp FF (SAFF)

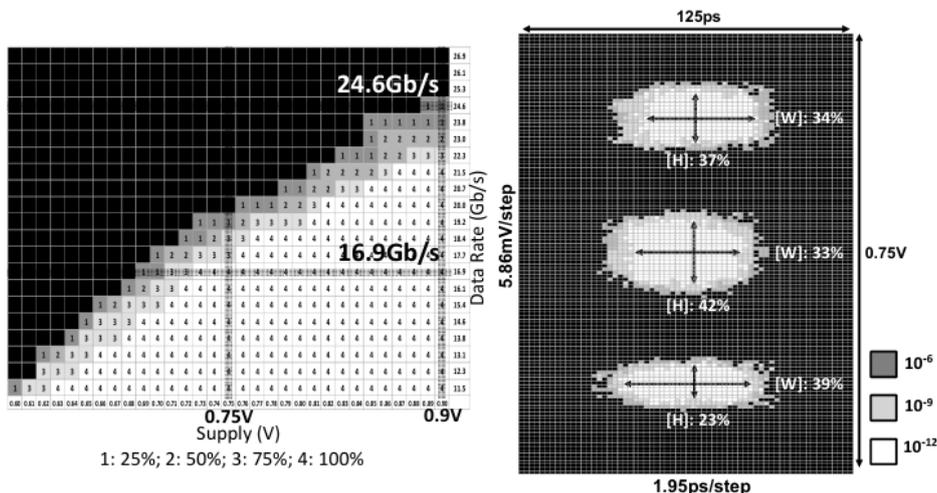
[그림 2]는 bond 및 TSV의 loss와 crosstalk (위)와 PAM-4 driver과 receiver에 쓰이는 strongarm latch 기반 flip-flop (SAFF)를 보여준다. Driver의 경우, cascode device 없이 compact하게 설계하기 위해 calibration을 driving MOSFET gate side로 옮겼고, SAFF는 낮은 reference부터 높은 reference까지의 입력 level에 대응하기 위해 p-type 및 n-type stage가 병렬로 이루어진 구조를 사용하였다.



[그림 3] data channel (DCH) 및 common block (CB) 구조

[그림 3]은 제안하는 data channel (DCH) 및 common block (CB)의 구조를 보여준다. 제안하는 DCH는 RX에서 90도 phase shift를 제공하는 DLL을 통해 data를 center에서 sampling 할 수 있게 하고, TX에서는 TXFIFO 대신 deskew DLL을 이용하여 SoC에서 DCH 까지의 data 전달을 동기화할 수 있게 했다.

또한, 제안하는 work에서는 'repair logic' 을 제안하여 defect가 있는 bond가 생길 경우 이 bond를 통하는 lane을 redundant lane으로 우회할 수 있게 하여 동일 cluster에 있는 10개의 lane에 대한 defect를 수리할 수 있게 한다.



[그림 4] 측정된 shmoo plot 및 PAM-4 eye contour plot

[그림 4]는 측정된 shmoo plot 및 PAM-4 eye를 보여준다. shmoo plot은 active lane이 25%~100% 일 때의 supply에 따른 data rate을 보여준다. PAM-4 eye plot은 built-in self-

test circuit을 통해 high/middle/low 패턴을 적절히 인가하여 측정한 값으로, 10-12 BER 기준으로 125ps의 34%/33%/39%의 eye width와, 250mV의 37%/42%/23%의 eye height을 보였다.

제안하는 work은, data rate을 높이기 위한 특별한 circuit technique을 사용했기 보다는, 9um의 아주 작은 pitch를 가지는 bond를 통한 3D stacking을 통해 면적당 data 전송 효율을 극대화했다고 요약할 수 있다.

저자정보



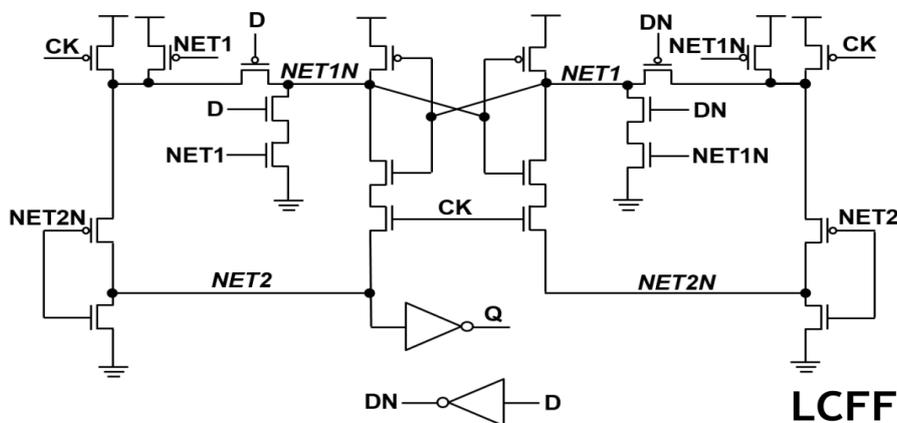
명예기자 임규완

- 소 속 : KAIST 전기및전자공학과 박사과정
 - 연구분야 : Display Driver IC, Readout IC
 - 이 메 일 : limkw@kaist.ac.kr
 - 홈페이지 : <https://ICdesignlab.net>
-

#25-2 모바일 및 웨어러블 어플리케이션의 성장함에 따라 저전력 동작이 중요한 요구 사항이 되었다. Flip-flop은 ASIC design에서 기초적인 sequential element로, ASIC의 전체 전력의 20% 이상을 차지하고 있다. Flip-flop의 전력 소모를 줄이기 위한 방안은 static operation, contention-free transitions, minimize clock power가 있는데, 이를 위해 이전에 ACFF, TCFF, 26TSPC, REFF 등과 같은 flip-flop의 구조가 제안되었다. 하지만 ACFF와 TCFF 구조는 contention의 문제로 yield issue가 발생하며, 26TSPC와 REFF 구조는 diffusion breaks (DB)로 인한 area penalty를 고려하지 않았다. 따라서 본 논문은 path-sharing과 minimal DBs를 통해 compact area를 가지며, true single-phase clock (TSPC), conditional capture (CC), contention-free (CF) 구조를 통합한 flip-flop을 제안한다.

제안한 low-power conditional capture flip-flop (LCFF) 구조는 기존의 transmission gate를 이용한 flip-flop이 아닌 bubble pushing을 통해 static CMOS flip-flop을 구현하여 하나의 clock phase만 사용하는 TSPC 구조를 달성하였다. 또한, flip-flop 내 master와 slave latch에 존재하는 complementary inputs을 이용한 CC 구조로 저전력으로 동작이 가능하도록 하였다. 마지막으로, contention-free를 위해 같은 polarity를 가진 circuit을 merge하였고, 짝수 drain/source 노드를 연결시킴으로써 DB를 최소화하였다. 제안한 구조는 이러한 path-sharing을 통해 flip-flop의 TR count와 DB를 감소시켰다.

본 논문에서는 TSPC, CC, CF 구조 및 TR과 DB의 사용을 최소화하여 기존에 제안한 flip-flop들과 비교하여 면적과 전력 측면에서 유리하다는 결과를 보여주고 있다. Post-layout simulation 결과에 따르면, 제안한 LCFF 구조는 기존의 FF들보다 7% ~ 24% 면적이 작다. 또한, conventional transmission gate flip-flop (TGFF)와 비교했을 때, 전체 전력과 누설 전력이 평균적으로 각각 63%, 26% 줄었다는 결과를 보여주고 있다. 그리고 Yield test 및 race immunity test를 위해 7nm FinFET Samsung test chip을 이용한 결과, LCFF는 0.33V의 최소 전압에서 정상적으로 동작하였고, path sharing으로 인해 PVT variation에 insensitive한 race immunity를 보여주고 있다.



저자정보



신현우

- 소 속 : KAIST 전기및전자공학과 박사과정
 - 연구분야 : High Speed ADC
 - 이 메 일 : shin6223@kaist.ac.kr
 - 홈페이지 : <https://msicl.kaist.ac.kr>
-

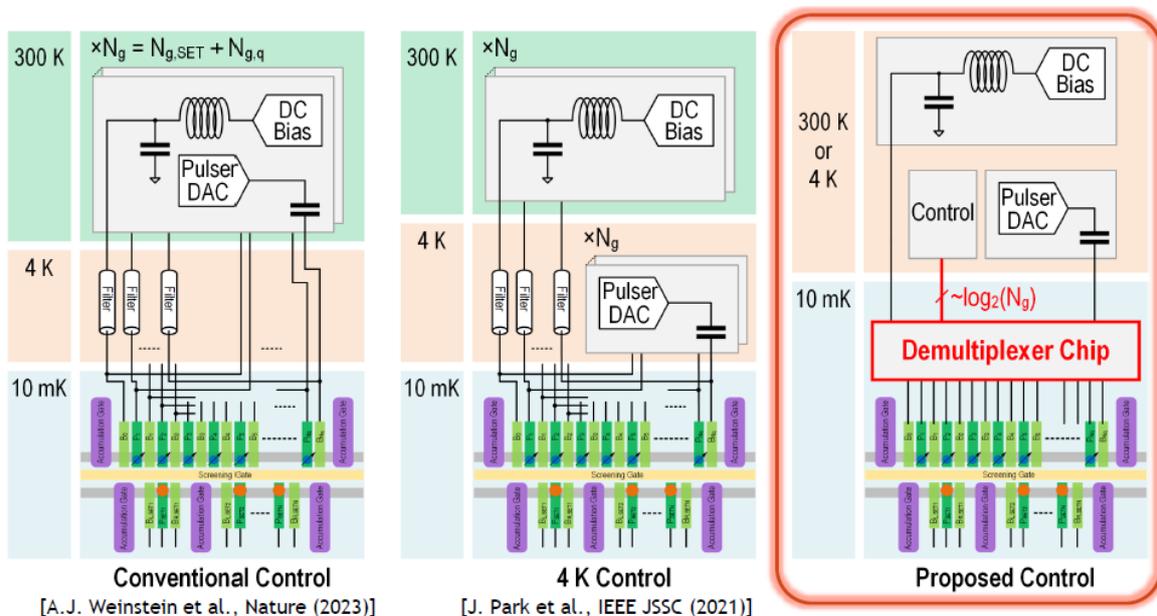
2024 IEEE VLSI Review

경북대학교 전자전기공학부 박사과정 여성일

Session 26 Analog Techniques II

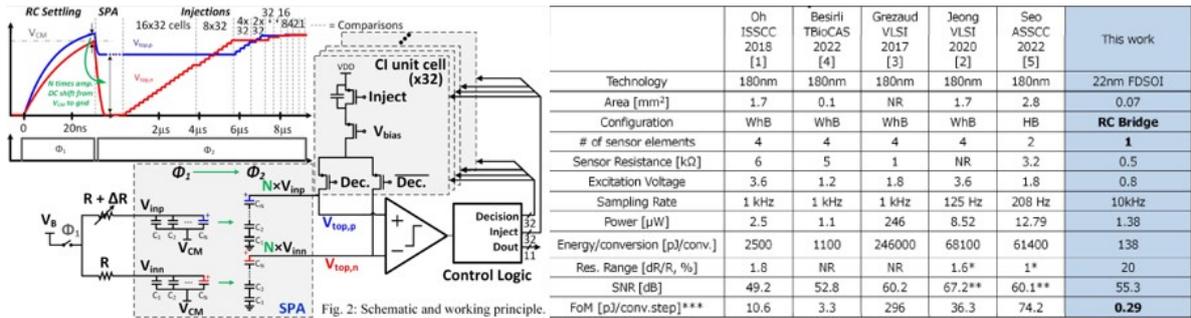
이번 VLSI 2024의 Session 26은 Analog Techniques 라는 주제로 총 4편의 논문이 발표되었다. 온도 센서나 crystal oscillator 같은 다양한 아날로그 회로에 관련된 논문이 소개되었다.

#26-1 이 논문은 Intel에서 발표한 논문으로 양자 컴퓨팅 회로를 위한 극저온 Demultiplexer에 관련된 논문이다. 기존 양자 컴퓨팅 회로에서 qubit의 control을 위해 개별적인 DC bias와 제어가 필요하였다. 이러한 방법은 대규모 qubit 배열을 제어할 때 배선 및 노이즈 문제가 발생할 수 있다는 단점이 있다. 본 논문에서는 이러한 단점을 개선하기 위해 qubit과 같은 mK영역에서 동작하는 Demultiplexer를 제안하고 있다. 그림 1에서 볼 수 있듯 기존 제어 방식과 달리 하나의 DC 전압 입력만으로 최대 64개의 qubit 배열을 제어할 수 있고 qubit 과 같은 PCB에 통합되어 동작하기에 노이즈 특성 또한 향상시킬 수 있었다.



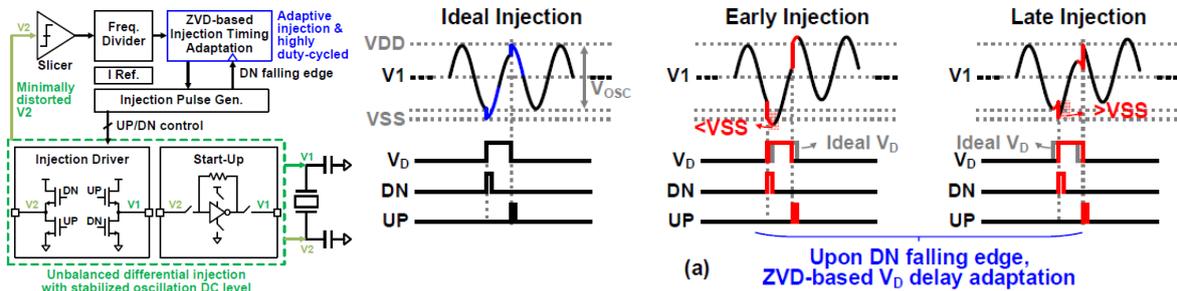
[그림 1] 26-1에서 제안하는 Qubit device control 방식 및 비교

#26-3 이 논문은 ETH Zurich에서 발표한 논문으로 척추 수술 후 척추의 가해지는 힘을 모니터링 할 수 있는 Bridge-to-Digital converter에 관련된 논문이다. 이 회로는 체내에 삽입되기 때문에 전력 소모를 줄이는 것이 중요한 요인이다. 이 논문은 그림 2에서 볼 수 있듯 Wheatstone Bridge 대신 RC bridge를 이용하여 전력소모를 줄일 수 있었으며 이를 통해 기존 논문 대비 가장 뛰어난 FoM을 얻을 수 있었다.



[그림 2] 제안하는 구조의 schematic 및 comparison table

#26-4 이 논문은 Southern University of Science and Technology에서 발표한 논문으로 Low power crystal oscillator(XO)에 관련된 논문이다. crystal oscillator는 최소한의 전력을 사용하면서 PVT variation에도 일정한 주파수를 만들어 내는 것을 목표로 하고 있다. 기존의 crystal oscillator들은 이를 위해 주파수의 peak와 valley에서만 전력을 공급하는 pulse-injection-based XO를 사용하였는데 이러한 방식은 pulse injection 주파수를 Nth-order sub-harmonic frequency로 낮추게 되면 전력 소모를 크게 줄일 수 있다. 이 논문은 16th-order sub-harmonic frequency를 사용하였고 그림 3에서 볼 수 있듯이 ZVD-based timing adaptation circuit을 이용하여 injection 주파수의 정확도를 높이고 전력 소모를 더욱 줄일 수 있었다.



[그림 3] 제안하는 구조의 동작 원리

저자정보



명예기자 여성일

소 속 : 경북대학교 전자전기공학부 박사과정

연구분야 : DC-DC Converter

이 메 일 : sungil1020@knu.ac.kr

홈페이지 : <https://sites.google.com/view/icslab>
