

Diss. ETH No. 30177

# Circuit Design for IoTs and In-Memory Computing

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES

(Dr. sc. ETH Zurich)

presented by

GIORGIO CRISTIANO

MSc. Nanotechnologies for ICTs, Politecnico di Torino

born on 09.07.1996

accepted on the recommendation of

Prof. Dr. Taekwang Jang, examiner

Prof. Dr. Mingoo Seok, co-examiner

Dr. Abu Sebastian, co-examiner

2024

# Abstract

There has been a clear trend towards smaller, cheaper, and more abundant devices in the last decades. The next computing platform, represented by Internet of Things (IoT) nodes, forgoes the limits imposed by human accessibility to achieve a much more aggressive miniaturization. This allows IoT nodes to be placed directly at the source of information, where they can sense, process, and transmit data. However, the stark area reduction poses different design challenges and tradeoffs that must be correctly addressed. The much smaller form factor heavily impacts the battery capacity that can be fit on those nodes. Hence, IoT circuits must be designed for high energy efficiency during operation and very low idle leakage. This dissertation proposes to address challenges across all the main operations of an IoT node.

A low-power wake-up timer that can achieve a single digit temperature coefficient (TC) with one point calibration only is proposed. Such work, fabricated in 180nm technology, can achieve  $8.7\text{ppm}/^\circ\text{C}$  TC while consuming  $694\text{nW}$ . Additionally, a low noise phase locked loop (PLL) is shown based on extracting phase error information at the output frequency rate. It has been designed in 22nm FD-SOI technology and achieves a simulated integrated jitter of  $53.3\text{fs}$  while consuming  $6.4\text{mW}$ , leading to an FoM of  $-257.3\text{dB}$ . Finally, a  $512\times 512$  in-memory compute core (IMC) based on phase change memories (PCMs) is presented. We also propose a time-based ADC with improved linearity thanks to a controlled charge removal process through charge injection cells and an input modulation scheme that, together with the ADC, can compute an 8b input in just 38ns. The tile has been designed in 14nm FinFet technology and achieves a

simulated peak performance of  $13.8TOPS$  at an energy efficiency of  $19.6TOPS/W$  and area efficiency of  $8.5TOPS/mm^2$ .

# Sommario

Negli ultimi decenni si è assistito a una chiara tendenza verso dispositivi piú piccoli, piú economici, e quindi piú abbondanti. La successiva piattaforma di computazione, i cosiddetti nodi Internet of Things (IoT), non sono costretti dai limiti imposti dall'accessibilità umana e riescono a raggiungere livelli di miniaturizzazione molto piú estremi. Per questo, i nodi IoT possono essere installati direttamente alla sorgente dell'informazione, dove possono misurare, processare e trasmettere i dati. Tuttavia, la netta riduzione dell'area pone diversi problemi di progettazione e di sfide progettuali e compromessi che devono essere affrontati correttamente. Una dimensione molto piccola influisce pesantemente sulla capacità della batteria che può essere inserita in questi nodi. Pertanto, i circuiti IoT devono essere progettati per garantire un'elevata efficienza energetica durante il funzionamento e una perdita di energia molto bassa. Questa tesi si propone di affrontare le sfide relative alle operazioni principali di un nodo IoT.

Viene proposto un timer a basso consumo che può raggiungere un coefficiente di temperatura (TC) a una cifra con un solo punto di calibrazione. Tale lavoro, realizzato in tecnologia a 180 nm, può raggiungere  $8.7ppm/^{\circ}C$  TC con un consumo di  $694nW$ . Inoltre, viene stato presentato un Phase Locked Loop (PLL) a basso rumore basato sull'estrazione di informazioni sull'errore di fase alla frequenza di uscita. È stato progettato in tecnologia FDSOI a 22 nm e raggiunge un jitter integrato simulato di  $53.3fs$  con un consumo di  $6.4mW$ , avendo una FoM di  $-257.3dB$ . Infine, un nucleo per In-Memory Computing (IMC) basato su Phase Change Memories (PCM) di dimensione  $512 \times 512$ . Proponiamo anche un ADC basato sul tempo con una linearità migliorata grazie a un processo di rimozione controllata

della carica di rimozione controllata della carica attraverso celle di iniezione di carica e uno schema di modulazione dell'ingresso che, insieme all'ADC, può calcolare un ingresso di 8b in soli 38ns. Questo chip é stato progettato in tecnologia FinFet a 14nm, e raggiunge una performance di picco simulata di  $13.8TOPS$  ad una efficienza energetica di  $19.6TOPS/W$  ed una efficienza spaziale di  $8.5TOPS/mm^2$ .