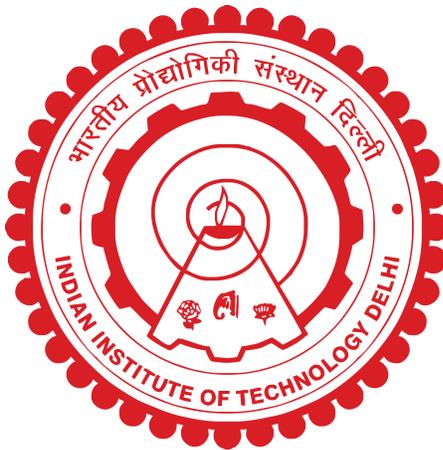


**MODELING AND ENERGY OPTIMIZATION  
OF MULTICORE PROCESSOR-CACHE  
SYSTEMS WITH EMERGING  
TECHNOLOGIES**

**DIVYA PRANEETHA RAVIPATI**



**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING  
INDIAN INSTITUTE OF TECHNOLOGY DELHI  
JULY 2024**

© Indian Institute of Technology Delhi (IITD), New Delhi, 2024

# **MODELING AND ENERGY OPTIMIZATION OF MULTICORE PROCESSOR-CACHE SYSTEMS WITH EMERGING TECHNOLOGIES**

by

**DIVYA PRANEETHA RAVIPATI**

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

Submitted

in fulfillment of the requirements of the degree of Doctor of Philosophy

to the



**INDIAN INSTITUTE OF TECHNOLOGY DELHI**

**JULY 2024**

DEDICATED TO

*My mother & father*

*Smt. Mannava Sampurna Kumari*

*Shri Ravipati Ramaiah (Late)*

# Certificate

This is to certify that the thesis titled “**Modeling And Energy Optimization Of Multicore Processor-Cache Systems With Emerging Technologies**” being submitted by **Ravipati Divya Praneetha** for the award of **Doctor of Philosophy** in Computer Science and Engineering is a record of bonafide work carried out by her under my guidance and supervision in the Department of Computer Science and Engineering, Indian Institute of Technology Delhi. The work presented in this thesis has not been submitted elsewhere, either in part or full, for the award of any other degree or diploma unless otherwise stated explicitly.

Certain works included in this thesis involved collaboration or use of measurement data obtained by other students, which have been explicitly specified/acknowledged in the corresponding chapters.

**Preeti Ranjan Panda**

Professor

Department of Computer Science and Engg.

Indian Institute of Technology Delhi

New Delhi- 110016

# Acknowledgments

*“Gratitude can transform common days into thanksgiving, turn routine jobs into joy and change ordinary opportunities into blessings.”*

—**William Arthur Ward**

Every Ph.D. journey is distinct, with unique experiences, moments of joy, and challenges. In the end, it is the culmination of every fragment that gives the journey its sense of fulfilment. I take this opportunity to express my sincere gratitude to all the individuals who have been instrumental in making my Ph.D. journey truly memorable. Their contributions have not only enriched my academic experience but have also played a pivotal role in shaping me into a better version of myself.

The most invaluable blessing bestowed upon me during research journey was the opportunity to work under the mentorship of **Prof. Preeti Ranjan Panda**. I am forever grateful to Preeti for the profound manner in which he guided me through my research expedition. Under his mentorship, I not only learned to navigate current challenges but also to discern future research opportunities, fostering a mindset that extends beyond immediate expectations. Prof. Preeti took commendable initiatives to mentor our MARG team on research career directions. The team lunches and outings provided us with an avenue to have discussions well beyond research. These discussions helped me in moulding my thought process and approach towards life. He also recognizes that life unfolds during the Ph.D. journey, affording me the time to reflect and return to research with an open mind.

I extend my thanks to my research collaborators, **Prof. Hussam Amrouch** and **Victor M. Van Santen**, for providing the measurement data needed and valuable feedback through out my research work. I am thankful to Prof. Hussam for inviting me to work at the University of Stuttgart, Germany. This research visit facilitated extensive discussions with Prof. Hussam and Victor, opening doors to explore new research problems. The insightful technical discussions with Victor throughout my journey have been invaluable, enhancing my ability to perceive possibilities amidst challenges and improving my technical writing.

My sincere thanks to my student research committee members, **Prof. Kolin Paul**, **Prof. Subodh Vishnu Sharma**, and **Prof. Abhisek Dixit** for their invaluable timely feedback. I thank my course instructors Prof. Kolin Paul, Prof. Subodh Kumar, Prof. Smruti Ranjan Sarangi, and Prof. Preeti Ranjan Panda for enriching my knowledge and skills on the subjects. I extend my sincere appreciation to Prof. Anshul Kumar, Prof. M. Balakrishnan, Prof. Vireshwar Kumar, and Prof. Preeti Ranjan Panda for entrusting me with Teaching Assistant

(TA) responsibilities, enhancing my skills and contributing to my academic and professional journey.

A Ph.D. provides the opportunity to connect with minds sharing similar vigor and passion. I am grateful for the new bonds forged during my Ph.D. My sincere thanks to Sakshi Tiwari who stood by me in my ups and downs, encouraged me, and made my Ph.D. journey beautiful. I am thankful to Rajesh Kedia, Hadi Brais, and Neetu Jindal for the countless discussions, both technical and non-technical, which helped me in evolving my thought process. Being a night owl, I extend special thanks to Sakshi and Hadi for accompanying me and creating countless cherishing memories through many discussions, late-night coffees, and early morning breakfasts. I extend my thanks to Aritra Bagchi who was always there a seat away for rejuvenating, intellectual, and motivating discussions. I also thank Garima Modi for regularly checking on me when I got COVID, and for many cheering discussions. Special thanks to my colleagues, Ayushi Agarwal, Naman Jain, Shailja Pandey, Lokesh Siddhu, Sreenivasulu, Shubhi, Priyanka Singla, Hameedah Sultan, Shubhani, Ramanuj Goel, Diksha, Shubhankar Singh, Shikha, Sandeep Kumar, Dishant, Vinayak, Sidharth Ranjan, Divyanjali, Dharamjeet, Radhika, Ankita, Chahat Bansal, Nivedita, Chandrika Bhardwaj, Shubrima Gosh, Shivendra Singh Parihar, Shubham Gupta, Swetaki Chatterjee, and Himanshu Gandhi for their constant support, encouraging spirit, intellectual discussions, and informal interactions. I also thank Telugu committee, IIT Delhi, who had taken many efforts to bring the telugu community together and made attempts to not miss home during festivals by organizing Ugadi festival (Telugu New Year) every year. I am thankful to Venkat, Hari krishna, Sudhakar, Venkatappa Reddy, Soumya, Surya, Siva, and Venkatesh for their support and rejuvenating conversations. I extend my thanks to my neighbours, Puja and Mythili for taking care of me whenever I fell ill.

I also thank our lab and office admins Vandana, Hemanth Prasad, Rekha, and Suresh for their support at the work place and taking care of many administrative hassles.

No words can truly capture the depth of gratitude I hold for the unwavering support, understanding, and belief my family has bestowed upon me. I am thankful to **Prof. Ram Babu Kodali**, my father-in-law, who recognized my capabilities during our initial interactions and encouraged me to pursue Ph.D., which I had postponed for several years due to social pressure. I extend my heartfelt thanks to my husband, **Dr. Rahul Babu Kodali**, for cheering me up during my challenging phases, broadening my perspective and cognitively impacting my thought process. Your belief in me, even during moments of self-doubt, has been a constant source of strength. My deepest thanks and boundless love to my daughter, **Ruthvika Janaki Kodali**. This thesis is as much a reflection of our shared journey as it is of my academic endeavors. Your remarkable understanding and support at such a young age towards

accomplishing my Ph.D. are truly commendable. I promise to compensate for the sacrifices made.

My greatest blessings came from my father, **Late Ravipati Ramaiah**, who always believed in my capabilities, instilled in us that hard work inevitably leads to success, encouraged me to push the boundaries and venture beyond the comfort zone. I believe that many of my qualities are inspired by his profound influence. I am profoundly grateful to my mother, **Mannava Sampurna Kumari** (Retd. Teacher), for her unwavering belief and support. Due to pollution concerns in Delhi, we decided to leave our four-month-old daughter with her during the early years of my Ph.D. Despite her old age, she has taken care of our child since kid's birth. Words cannot adequately justify or express the depth of her sacrifices and love. I extend my deepest gratitude to my brother, **Mahesh Babu Ravipati**, for his unwavering encouragement, emotional support, and being my pillar of strength. He played a pivotal role in raising my daughter with immense love, becoming her most trusted confidant. I am also grateful to **Shadow**, our pet, for bringing joy into our lives.

I extend my thanks to my extended family, cousins (Krishna Chaitanya, Ruhi, Krishna Priya, Mohan Krishna, Sravya, Suchendra Nath, Aruna, Vamsi Krishna, Devi, Teja, Deepthi, Prasanna, Aswini), friends (Kiran, Hareesh, Sravanthi, Chandana, Vamsi, Teja, Amulya, Kishore), and well wishers (Kishore Uncle, Vani Aunty, Sathyavathi Aunty, Prof. Anjaiah Devineni) for being there for me any time I need to talk and cheering me. While it is impossible to name everyone individually, your collective influence has been invaluable. My deepest thanks to all those who have significantly contributed to shaping and positively impacting my academic pursuit.

**Ravipati Divya Praneetha**



# Abstract

Technology advancements and shrinking transistor feature size have improved system performance and efficiency. However, with the transistor feature size shrinkage racing towards sub-nm regime, a multitude of challenges such as slowdown of Moore’s law, halt of Dennard scaling, significant increase in leakage power consumption, and quantum effects have emerged, particularly at smaller geometries. To address these challenges, researchers and engineers are exploring new transistors and technologies. In this thesis, we explore two such emerging technologies, negative capacitance FinFET (NC-FinFET/ NCFET) and cryogenic computing, at system-level. Researchers have investigated the advantages of NCFET and cryogenic transistors at device-level and circuit-level. The research findings revealed that the unique characteristics of NCFET offer higher frequency of operation without the need to increase voltage along with reducing the leakage current. Similarly, the examination of cryogenic transistors at device- and circuit-levels has unveiled their potential to enhance system performance and mitigate leakage power concerns through operation at low temperatures. Both of these technologies address the challenge of rising leakage current in smaller geometries while simultaneously enhancing performance without the need for voltage increment. Recognizing the potential advantages of these two technologies, we make an attempt to bridge the research gap by exploring their system-level implications.

To comprehensively explore the trade-offs between performance and energy efficiency in CPU-memory systems at smaller geometries and with new technologies, it is essential to revise the processor and cache models used by instruction-level simulators. CACTI and McPAT are popular tools for system-level architectural studies to estimate power, performance and area of the system. However, the tools are primarily designed for CMOS technology using the data obtained from various projections. Furthermore, the models use various approximations at higher geometries which do not align with the current trends for the FinFET technology at lower geometries.

For the first time, we make an effort to revise the CACTI and McPAT models to suit the newer technologies, while respecting the overall modeling methodology of the tools for

FinFET-/NCFET-based system estimates. Moreover, for the first time, we integrate the transistor characteristics from a 14 nanometer commercial FinFET technology within the tools. First, we revise CACTI models to support FinFET and NCFET technologies. We use the developed tool (**FN-CACTI**) to assess the energy efficiency of NCFET-based caches compared to FinFET-based caches. Additionally, we identify the optimal voltage to minimize cache energy consumption for FinFET- and NCFET-based caches at various access rates and cache sizes. Our investigations show that the optimal voltage for NCFET-based caches spans a range of voltages depending on the cache access rate, while the optimal voltage is within the lower range of applicable voltages for FinFET-based caches. As the next step towards revising the tools to estimate system-level delay, power and area estimates, we update the McPAT models. We first integrate FN-CACTI with our modeling tool, **FN-McPAT**, to derive estimates for FinFET- and NCFET-based memory structures. Then, we synthesize the BOOM CPU core with FinFET and NCFET technologies to revise the core component models in FN-McPAT. We use FN-CACTI and FN-McPAT to investigate the performance improvements and energy-efficiency of NCFET-based and FinFET-based systems. Our investigations provide novel insights into energy consumption of NCFET-based caches on systems running various workloads.

Motivated by our findings on the new trends in energy consumption behavior of NCFET-based caches, we present the first work towards optimizing energy in NCFET-based caches with minimal impact on performance. We leverage the unique characteristics offered by NCFETs and propose a dynamic voltage scaling policy, **CAPE**. Along with an approach that is suitable for both NCFET- and FinFET-based caches, we also introduce a novel metric to capture cache criticality. Our experimental evaluations indicate that the CAPE policy achieves 19.2% more last-level cache (LLC) energy savings compared to operating at the maximum available voltage.

Cryogenic circuits have applications in fields such as quantum computing, particle detectors, and magnetic resonance imaging. While cryogenic logic circuits are being addressed by the research community, there is limited work on designing with larger cryogenic systems at lower temperatures (10 K). Moreover, there is no tool to estimate delay, power and area of cryogenic systems at lower geometries. As a first step towards modeling cryogenic multi-core systems, we model Cryo-CACTI for cryogenic caches due to their vital role in performance improvement and significant contribution to both area and power of the processor. Using Cryo-CACTI, we provide our insights on efficiency of cryogenic caches and propose new research directions.

## संक्षेप

प्रौद्योगिकी उन्नति और ट्रांजिस्टर फीचर आकार में कमी ने सिस्टम प्रदर्शन और दक्षता में सुधार किया है। हालाँकि, ट्रांजिस्टर फीचर आकार में कमी के साथ सब-एनएम शासन की ओर बढ़ रहा है, मूर के नियम की मंदी, डेनार्ड स्केलिंग का रुकना, लीकेज पावर खपत में उल्लेखनीय वृद्धि और क्रांति प्रभाव जैसी कई चुनौतियाँ सामने आई हैं, खासकर छोटी ज्यामिति पर। इन चुनौतियों का समाधान करने के लिए, शोधकर्ता और इंजीनियर नए ट्रांजिस्टर और तकनीक की खोज कर रहे हैं। इस थीसिस में, हम सिस्टम-स्तर पर दो ऐसी उभरती हुई तकनीकों, नेगेटिव कैपेसिटेंस फिनफेट (एनसी-फिनफेट/एनसीएफईटी) और क्रायोजेनिक कंप्यूटिंग का पता लगाते हैं। शोधकर्ताओं ने डिवाइस-स्तर और सर्किट-स्तर पर एनसीएफईटी और क्रायोजेनिक ट्रांजिस्टर के लाभों की जाँच की है। शोध निष्कर्षों से पता चला है कि एनसीएफईटी की अनूठी विशेषताएँ लीकेज करंट को कम करने के साथ-साथ वोल्टेज को बढ़ाए बिना संचालन की उच्च आवृत्ति प्रदान करती हैं। इसी तरह, डिवाइस और सर्किट-स्तर पर क्रायोजेनिक ट्रांजिस्टर की जाँच ने सिस्टम प्रदर्शन को बढ़ाने और कम तापमान पर संचालन के माध्यम से लीकेज पावर चिंताओं को कम करने की उनकी क्षमता का खुलासा किया है। ये दोनों ही तकनीकें छोटी ज्यामिति में बढ़ते लीकेज करंट की चुनौती का समाधान करती हैं, साथ ही वोल्टेज वृद्धि की आवश्यकता के बिना प्रदर्शन को बढ़ाती हैं। इन दोनों तकनीकों के संभावित लाभों को पहचानते हुए, हम उनके सिस्टम-स्तरीय निहितार्थों की खोज करके शोध अंतर को पाटने का प्रयास करते हैं।

छोटी ज्यामिति पर और नई प्रौद्योगिकियों के साथ CPU-मेमोरी सिस्टम में प्रदर्शन और ऊर्जा दक्षता के बीच व्यापार-नापसंद का व्यापक रूप से पता लगाने के लिए, निर्देश-स्तर सिमुलेटर द्वारा उपयोग किए जाने वाले प्रोसेसर और कैश मॉडल को संशोधित करना आवश्यक है। CACTI और McPAT सिस्टम-स्तरीय आर्किटेक्चरल अध्ययनों के लिए लोकप्रिय उपकरण हैं, जो सिस्टम की शक्ति, प्रदर्शन और क्षेत्र का अनुमान लगाते हैं। हालाँकि, उपकरण मुख्य रूप से विभिन्न प्रक्षेपणों से प्राप्त डेटा का उपयोग करके CMOS तकनीक के लिए डिज़ाइन किए गए हैं। इसके अलावा, मॉडल उच्च ज्यामिति पर विभिन्न सन्निकटन का उपयोग करते हैं जो कम ज्यामिति पर FinFET तकनीक के लिए वर्तमान रुझानों के साथ सरेखित नहीं होते हैं।

पहली बार, हम CACTI और McPAT मॉडल को नई तकनीकों के अनुरूप संशोधित करने का प्रयास करते हैं, जबकि FinFET-/NCFET-आधारित सिस्टम अनुमानों के लिए उपकरणों की समग्र मॉडलिंग पद्धति का सम्मान करते हैं। इसके अलावा, पहली बार, हम उपकरणों के भीतर 14 नैनोमीटर वाणिज्यिक FinFET तकनीक से ट्रांजिस्टर विशेषताओं को एकीकृत करते हैं। सबसे पहले, हम FinFET और NCFET तकनीकों का समर्थन करने के लिए CACTI मॉडल को संशोधित करते हैं। हम FinFET-आधारित कैश की तुलना में NCFET-आधारित कैश की ऊर्जा दक्षता का आकलन करने

के लिए विकसित उपकरण (FN-CACTI) का उपयोग करते हैं। इसके अतिरिक्त, हम विभिन्न एक्सेस दरों और कैश आकारों पर FinFET- और NCFET-आधारित कैश के लिए कैश ऊर्जा खपत को कम करने के लिए इष्टतम वोल्टेज की पहचान करते हैं। हमारी जाँच से पता चलता है कि NCFET-आधारित कैश के लिए इष्टतम वोल्टेज कैश एक्सेस दर के आधार पर वोल्टेज की एक सीमा तक फैला हुआ है, जबकि इष्टतम वोल्टेज FinFET-आधारित कैश के लिए लागू वोल्टेज की निचली सीमा के भीतर है। सिस्टम-स्तरीय विलंब, शक्ति और क्षेत्र अनुमानों का अनुमान लगाने के लिए उपकरणों को संशोधित करने की दिशा में अगले कदम के रूप में, हम McPAT मॉडल को अपडेट करते हैं। हम सबसे पहले FN-CACTI को अपने मॉडलिंग टूल, FN-McPAT के साथ एकीकृत करते हैं, ताकि FinFET- और NCFET-आधारित मेमोरी संरचनाओं के लिए अनुमान प्राप्त किए जा सकें। फिर, हम FN-McPAT में कोर घटक मॉडल को संशोधित करने के लिए FinFET और NCFET तकनीकों के साथ BOOM CPU कोर को संश्लेषित करते हैं। हम NCFET-आधारित और FinFET-आधारित सिस्टम के प्रदर्शन सुधारों और ऊर्जा-दक्षता की जांच करने के लिए FN-CACTI और FN-McPAT का उपयोग करते हैं। हमारी जांच विभिन्न कार्यभार चलाने वाले सिस्टम पर NCFET-आधारित कैश की ऊर्जा खपत में नई अंतर्दृष्टि प्रदान करती है।

NCFET-आधारित कैश के ऊर्जा उपभोग व्यवहार में नए रुझानों पर हमारे निष्कर्षों से प्रेरित होकर, हम प्रदर्शन पर न्यूनतम प्रभाव के साथ NCFET-आधारित कैश में ऊर्जा के अनुकूलन की दिशा में पहला काम प्रस्तुत करते हैं। हम NCFET द्वारा पेश की गई अनूठी विशेषताओं का लाभ उठाते हैं और एक गतिशील वोल्टेज स्केलिंग नीति, CAPE का प्रस्ताव करते हैं। NCFET- और FinFET-आधारित कैश दोनों के लिए उपयुक्त दृष्टिकोण के साथ, हम कैश क्रिटिकलिटी को कैप्चर करने के लिए एक नया मीट्रिक भी पेश करते हैं। हमारे प्रायोगिक मूल्यांकन से संकेत मिलता है कि CAPE नीति अधिकतम उपलब्ध वोल्टेज पर संचालन की तुलना में 19.2% अधिक अंतिम-स्तरीय कैश (LLC) ऊर्जा बचत प्राप्त करती है।

क्रायोजेनिक सर्किट का उपयोग क्वांटम कंप्यूटिंग, पार्टिकल डिटेक्टर और मैग्नेटिक रेजोनेंस इमेजिंग जैसे क्षेत्रों में किया जाता है। जबकि क्रायोजेनिक लॉजिक सर्किट पर शोध समुदाय द्वारा विचार किया जा रहा है, कम तापमान (10 K) पर बड़े क्रायोजेनिक सिस्टम के साथ डिजाइनिंग पर सीमित काम है। इसके अलावा, कम ज्यामिति पर क्रायोजेनिक सिस्टम की देरी, शक्ति और क्षेत्र का अनुमान लगाने के लिए कोई उपकरण नहीं है। क्रायोजेनिक मल्टी-कोर सिस्टम मॉडलिंग की दिशा में पहले कदम के रूप में, हम प्रदर्शन सुधार में उनकी महत्वपूर्ण भूमिका और प्रोसेसर के क्षेत्र और शक्ति दोनों में महत्वपूर्ण योगदान के कारण क्रायोजेनिक कैश के लिए डीपक्रायो-सीएसीटीआई (डीसी-सीएसीटीआई) मॉडल बनाते हैं। डीसी-सीएसीटीआई का उपयोग करके, हम क्रायोजेनिक कैश की दक्षता पर अपनी अंतर्दृष्टि प्रदान करते हैं और नए शोध दिशा-निर्देश प्रस्तावित करते हैं।

# Contents

<b>Certificate</b>	<b>ii</b>
<b>Acknowledgments</b>	<b>iii</b>
<b>Abstract</b>	<b>vii</b>
<b>List of Figures</b>	<b>xv</b>
<b>List of Tables</b>	<b>xxi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Slowdown of Moore’s Law and the End of Dennard’s Scaling . . . . .	1
1.2 Need for Emerging Technologies . . . . .	2
1.2.1 NC-FinFET Technology . . . . .	2
1.2.2 Cryogenic computing . . . . .	3
1.3 Need for Emerging Technology Based Models . . . . .	3
1.4 Existing System Models and their Limitations . . . . .	4
1.4.1 Memory models . . . . .	4
1.4.2 Core models . . . . .	5
1.4.3 Cryogenic system models . . . . .	5
1.5 Generic strategy to update system models . . . . .	6
1.6 Contributions . . . . .	7
1.6.1 Contribution 1 . . . . .	8
1.6.2 Contribution 2 . . . . .	9
1.6.3 Contribution 3 . . . . .	9
1.6.4 Contribution 4 . . . . .	10
1.7 Outline of the Thesis . . . . .	10

<b>2</b>	<b>Background</b>	<b>11</b>
2.1	NCFET - Innovation in Transistor Technology . . . . .	11
2.1.1	Related work . . . . .	14
2.2	Exploring Cryogenics - The Chilling Frontier . . . . .	16
2.2.1	Related work - Cryogenic System at 77 K . . . . .	16
2.3	Caching Fundamentals and the CACTI Model . . . . .	18
2.3.1	Memory hierarchy . . . . .	18
2.3.2	Cache structure . . . . .	20
2.3.3	CACTI . . . . .	21
2.4	Architectural Power Modeling Tool: McPAT . . . . .	23
2.4.1	Heirarchical Model . . . . .	26
<b>3</b>	<b>FN-CACTI: Advanced CACTI for FinFET and NC-FinFET Technologies</b>	<b>29</b>
3.1	Introduction . . . . .	29
3.2	Related Work . . . . .	31
3.3	FinFET and NC-FinFET Device Modeling . . . . .	34
3.3.1	Transistor Modeling for Technology Parameter Extraction . . . . .	34
3.3.2	FinFET Calibration in TCAD . . . . .	35
3.3.3	NC-FinFET Calibration in TCAD . . . . .	38
3.4	Developing FN-CACTI: An Advanced CACTI Model for FinFETs and NC-FinFETs . . . . .	40
3.4.1	Enhancing FinFET Modeling . . . . .	40
3.4.2	Incorporating NC-FinFET Characteristics . . . . .	43
3.5	Experimental Studies . . . . .	44
3.5.1	Experimental Setup . . . . .	44
3.5.2	Impact of Voltage Specific Device Parameters . . . . .	45
3.5.3	Impact of Modeling Changes . . . . .	46
3.5.4	Combined Impact of Voltage-Dependent Parameters and Modeling Changes . . . . .	48
3.5.5	FinFET Caches versus NC-FinFET Caches . . . . .	51
3.6	Application of the Proposed Model . . . . .	53
3.7	Summary . . . . .	56
<b>4</b>	<b>Performance and Energy Studies on NC-FinFET Cache-based Systems with FN-McPAT</b>	<b>59</b>
4.1	Introduction . . . . .	59

4.2	Related Work . . . . .	62
4.2.1	NC-FinFET Research . . . . .	62
4.2.2	McPAT Development towards FinFET . . . . .	62
4.2.3	Objectives of our work . . . . .	63
4.3	FinFET and NC-FinFET Technology Calibration . . . . .	64
4.3.1	FinFET and NC-FinFET Transistor Calibration . . . . .	65
4.3.2	FinFET and NC-FinFET Standard Cell Libraries . . . . .	65
4.3.3	FinFET and NC-FinFET BOOM Processor . . . . .	66
4.4	FN-McPAT: McPAT for FinFET-based and NC-FinFET-based multi-core systems	67
4.4.1	Updating FN-CACTI . . . . .	69
4.4.2	Updating McPAT for FinFET-based and NC-FinFET-based multi-core systems . . . . .	69
4.5	Experimental Studies . . . . .	70
4.5.1	Experimental Setup . . . . .	71
4.5.2	Evaluation of FinFET and NC-FinFET-based systems . . . . .	75
4.5.3	Evaluation of FinFET-based LLC and NC-FinFET-based LLC at various voltages . . . . .	80
4.6	Summary . . . . .	86
<b>5</b>	<b>CAPE: Criticality-Aware Performance and Energy Optimization Policy for NCFET-based Caches</b>	<b>87</b>
5.1	Introduction . . . . .	87
5.1.1	Power Management in NCFET-based Caches . . . . .	88
5.1.2	Cache Criticality Factor - An accurate Cache Performance Metric . . . . .	90
5.1.3	Novel Contributions . . . . .	91
5.2	Background . . . . .	91
5.2.1	NCFET Caches . . . . .	91
5.2.2	Cache voltage control . . . . .	93
5.3	Related Work . . . . .	93
5.3.1	Energy Optimization in Caches . . . . .	93
5.3.2	NCFET-based Logic Circuits . . . . .	95
5.4	The <i>CAPE</i> Voltage Scaling Strategy . . . . .	96
5.4.1	Our Proposed Novel Metric . . . . .	96
5.4.2	<i>CAPE</i> policy . . . . .	98
5.4.3	Computing CCF . . . . .	102
5.5	Experiments and Results . . . . .	103

5.5.1	Experimental Setup . . . . .	103
5.5.2	Baseline Policies . . . . .	104
5.5.3	Deciding Thresholds . . . . .	105
5.5.4	Voltage Decisions . . . . .	106
5.5.5	NCFET-based LLC results evaluation . . . . .	109
5.5.6	Evaluating adaptability of CAPE to new workloads (Test set) . . . . .	110
5.5.7	Adaptability of CAPE to different cache sizes and system configurations . . . . .	112
5.5.8	FinFET-based LLC results evaluation . . . . .	117
5.5.9	Sensitivity analysis of the thresholds . . . . .	118
5.5.10	Overhead Analysis . . . . .	120
5.6	Summary . . . . .	121
<b>6</b>	<b>Cryo-CACTI: Cryogenic-Aware CACTI for Cache Modeling down to 10K in Advanced 7nm FinFETs</b>	<b>123</b>
6.1	Introduction . . . . .	123
6.1.1	Cryogenic Cache Exploration with CACTI . . . . .	124
6.2	Related Work . . . . .	125
6.2.1	Transistor and Circuit Research at Cryogenic Temperatures . . . . .	126
6.2.2	System-level works at cryogenic temperatures . . . . .	126
6.2.3	Limitations of Existing Work . . . . .	127
6.3	Cryogenic FinFET Transistors . . . . .	127
6.3.1	Semiconductor Physics at Cryogenic Temperatures . . . . .	127
6.3.2	FinFET Transistors at cryogenic 77K . . . . .	128
6.3.3	FinFETs at low cryogenic 10K . . . . .	128
6.3.4	Cryogenic ASAP7 Transistor Modeling - With Scaling . . . . .	129
6.4	Cryogenic SRAM Memory Banks (Caches) . . . . .	131
6.4.1	Cryogenic SRAM with planar MOSFET . . . . .	131
6.4.2	Cryogenic SRAM with FinFET . . . . .	132
6.5	CACTI for Cryogenic Caches . . . . .	133
6.5.1	Transistor Parameters Update . . . . .	134
6.5.2	Wire Models Update . . . . .	135
6.6	Experimental Results . . . . .	136
6.6.1	Experimental Setup . . . . .	136
6.6.2	Evaluation of cryogenic caches . . . . .	136
6.7	Applications with Cryogenic caches . . . . .	142
6.7.1	Impact on performance and cache energy due to operating voltage . . . . .	143

6.7.2	Exploring iso-delay configuration . . . . .	144
6.7.3	Exploring wire types . . . . .	146
6.8	Summary . . . . .	147
<b>7</b>	<b>Conclusion and Future Work</b>	<b>149</b>
7.1	Summary of Contributions . . . . .	149
7.2	Future Research Directions . . . . .	152
	<b>Bibliography</b>	<b>159</b>
	<b>Publications</b>	<b>175</b>
	<b>Biography</b>	<b>177</b>



# List of Figures

1.1	Generic approach for updating system-level metrics estimation models . . . . .	6
1.2	An overview of the contributions . . . . .	8
2.1	3D structure of NCFET and its electrical representation . . . . .	12
2.2	ON-current and OFF-current for FinFET and NCFET at varying voltages . . .	13
2.3	ON-current and OFF-current for FinFET and NCFET at various FE thickness ( $t_{fe}$ ) [87]. . . . .	13
2.4	Cache hierarchy . . . . .	19
2.5	Cache structure . . . . .	20
2.6	CACTI framework . . . . .	21
2.7	An example cache layout organization . . . . .	22
2.8	McPAT hierarchical modeling . . . . .	24
2.9	McPAT framework . . . . .	25
3.1	Transistor structures . . . . .	32
3.2	a) 3-D FinFET structure. b) FinFET cross-section and the geometry for the baseline FinFET and NC-FinFET as well as the NC-FinFET NC related parameters. . . . .	36
3.3	a) FinFET $I_{DS}$ calibration showing a close match between TCAD simulation (lines) and Intel 14 nm FinFET measurement data (symbols) [84] w.r.t. different $V_{GS}$ biases b) FinFET $I_{DS}$ calibration w.r.t. different $V_{DS}$ biases . . .	37
3.4	Evaluation of NC-FinFET, i.e., studying the impact of replacing the high- $\kappa$ $\text{HfO}_2$ in the FinFET with a ferroelectric layer. Highlighted in orange arrows are the improvements in the FinFET drain current ( $I_{DS}$ ) and increase in the total gate capacitance ( $C_{GG}$ ). The green arrows show how NC-FinFET can maintain the same current despite reducing $V_{GS}$ from 0.7 V to 0.55 V. . . . .	39
3.5	Novel contributions of FN-CACTI . . . . .	41

3.6	Various metrics corresponding to different parameter changes at 0.8 V, for 4 MB, 8W, 4B cache configuration . . . . .	46
3.7	Access times of NC-FinFET normalized to that of FinFET. Dashed lines correspond to not updating $C_g$ according to NC-FinFET while solid lines correspond to $C_g$ values extracted from TCAD simulations. . . . .	47
3.8	Normalized cache metrics for various cache sizes (all proposed changes combined) . . . . .	49
3.9	Comparing NC-FinFET caches with FinFET caches for various metrics and for various cache sizes. . . . .	51
3.10	Energy per access for 4 MB, 8W, 8B cache. The marker indicates the lowest energy point. (a) Shows that the voltage corresponding to lowest energy for FinFET is limited within 0.4 V or 0.5 V. (b) Shows that for NC-FinFET, the voltage corresponding to the lowest energy decreases with an increase in the cache access rate and ranges from 0.8 V till 0.4 V. For various cache access rates and voltages, the total energy per access for NC-FinFET is always lower than FinFET. . . . .	55
3.11	Energy consumption (at minimum energy point) of NC-FinFET-based caches normalized to FinFET-based caches for various access rates and sizes of cache. . . . .	56
4.1	High-level view of processor core components in McPAT. Memory structures and regular logic are in green. . . . .	60
4.2	Overview of transistor modelcard calibration, cell library characterization and BOOM processor synthesis. . . . .	64
4.3	Propagation Delay of BOOM Processor. NC-FinFET can maintain <1 ns even at 0.3 V, while FinFET exhibits 1.54 ns delay even at 0.5 V. . . . .	66
4.4	Total Power Consumption of BOOM Processor. NC-FinFET consumes more power at all voltages. Each bar captures the processor power while operating at the respective voltage and the maximum sustainable frequency at the voltage. Note, that frequency of NC-FinFET processor is higher at 0.7V (see Fig. 4.3). For a comparison at the same frequency, follow the iso-performance arrow. . . . .	67
4.5	An overview of the FN-McPAT tool . . . . .	68
4.6	Comparing relative power contribution of core subdomains with the results published by Haj-Yihia et al. [44] . . . . .	73
4.7	Comparing relative power contribution of the multi-core system with the results published by Cheng et al. [27] . . . . .	74

4.8	Comparing relative energy contribution of the LLC with the results published by Chakraborty et al. [21] . . . . .	74
4.9	Energy contribution of the processor components to the total energy of the processor. Absolute energy in milli Joules (mJ) is provided above the bars. . . .	76
4.10	Evaluation metrics of NC-FinFET-based system normalized with FinFET-based system for various SPEC2017 benchmarks. a) Execution time. b) Energy consumption. c) EDP . . . . .	79
4.11	LLC Energy of various SPEC CPU2017 benchmarks for FinFET-based LLC and NC-FinFET-based LLC . . . . .	80
4.12	Evaluation metrics of the system with LLC operating at various voltages normalized with those of the system with LLC operating at 0.7 V: (i) system with FinFET-based LLC (ii) system with NC-FinFET-based LLC . . . . .	82
4.13	System EDP with LLC operating at various voltages normalized with the EDP of the system with LLC operating at 0.7 V: (i) system with FinFET-based LLC (ii) system with NC-FinFET-based LLC . . . . .	83
4.14	Comparing multi-core system with NC-FinFET-based LLC against the multi-core system with FinFET-based LLC at various voltages (NC-FinFET/FinFET): (a) performance; (b) multi-core system energy; and (c) EDP. . . . .	85
5.1	16 MB cache energy savings at various voltages compared to operating the cache at $V_{max}$ (0.7 V). Refer Table 5.2 for the workload mixes. . . . .	88
5.2	Normalized evaluation metrics with 2 MB NCFET-based cache operating at various voltages. a) perlbench b) lbm c) imagick . . . . .	90
5.3	Dynamic energy per access and leakage power of 2 MB FinFET-based and NCFET-based caches operating at voltages ranging from 0.3 V to 0.7 V . . . .	92
5.4	Leakage power of 2 MB NCFET-based cache operating at voltages ranging from 0.3 V to 0.7 V . . . . .	93
5.5	Correlation between performance (IPC) and various cache metrics for SPEC2017 benchmarks. a) perlbench b) lbm . . . . .	96
5.6	Skeleton of <i>CAPE</i> strategy . . . . .	98
5.7	<i>CAPE</i> : Actions on low criticality . . . . .	99
5.8	<i>CAPE</i> : Actions on high criticality . . . . .	101
5.9	Voltages selected by baseline policies and <i>CAPE</i> for various workloads . . . .	107
5.10	Normalized NCFET-based LLC Energy savings, Slowdown, EDP savings, and Throughput degradation with the baseline policies and <i>CAPE</i> in percentages compared to operating the LLC at 0.7 V . . . . .	108

5.11	Normalized LLC Energy savings, Slowdown, EDP savings, and Throughput degradation with the baseline policies and CAPE in percentages compared to operating the LLC at 0.7 V on a new set of workloads - Test set . . . . .	111
5.12	Normalized LLC Energy savings, Slowdown, EDP savings, and Throughput degradation with the baseline policies and CAPE in percentages compared to operating the LLC at 0.7 V on 4 cores, 4 MB LLC system configuration . . . . .	113
5.13	Normalized LLC Energy savings, Slowdown, EDP savings, and Throughput degradation with the baseline policies and CAPE in percentages compared to operating the LLC at 0.7 V on 4 cores, 8 MB LLC system configuration . . . . .	114
5.14	Normalized FinFET-based LLC Energy savings, Slowdown, EDP savings, and Throughput degradation with the baseline policies and CAPE in percentages compared to operating the LLC at 0.7 V on 8 cores, 8 MB LLC system configuration . . . . .	115
5.15	Normalized FinFET-based LLC Energy savings, Slowdown, EDP savings, and Throughput degradation with the baseline policies and CAPE in percentages compared to operating the LLC at 0.7 V on 16 cores, 32 MB LLC system configuration . . . . .	116
5.16	Normalized FinFET-based LLC Energy savings, performance degradation, and EDP savings of the baseline policies and CAPE in percentages compared to operating the LLC at 0.7 V . . . . .	117
5.17	Sensitivity analysis of the thresholds (a) Variation in percentage slowdown at various interval sizes and criticality thresholds ( <i>criticality_th</i> ) while fixing interval threshold ( <i>interval_th</i> ) to 20, (b) Variation in percentage slowdown at various interval sizes and interval thresholds ( <i>interval_th</i> ) while fixing criticality threshold ( <i>criticality_th</i> ) to 0.2, and (c) Variation in percentage slowdown at various interval thresholds and criticality thresholds while fixing interval size to 100 $\mu$ s. . . . .	119
6.1	On-wafer lakeshore's cryogenic probe station for cooling to 10 K . . . . .	129
6.2	Measured transfer characteristics of 5 nm p-FinFET and n-FinFET from 300 K down to 10 K. . . . .	130
6.3	Cryo-CACTI model. . . . .	133
6.4	Access time variations with voltage for various caches a) L1 caches with 2 Banks, 4 Ways, 1 Read and 1 Write ports. b) LLC with 8 Banks, 8 Ways, 1 Read-Write port. . . . .	138

6.5	Access time variations due to various cache sizes and banks for 8 way set-associative cache. . . . .	139
6.6	Impact on cache access time, dynamic energy per access, and leakage power due to low-swing wires and global wires for various cache sizes and banks with 8-way set-associativity . . . . .	141
6.7	Variation in system evaluation metrics for system with 8 cores with 8MB LLC a) Performance b) LLC Energy . . . . .	143
6.8	Performance variations with various cache configurations at 10 K normalized to the performance with 8 MB LLC at 300 K. . . . .	143
6.9	LLC energy variations with various cache configurations at 10 K normalized to the LLC energy of 8 MB LLC at 300 K. Absolute values in mJ is shown above the bars. . . . .	145
6.10	Evaluation metrics variations with various LLC cache sizes and wire types (Low-swing (ls) and Global (g) wires) at 10K compared to those of 8MB LLC ls wire model at 300K a) Performance b) LLC Energy. Absolute energy in milli joules (mJ) in given above the bars for the baseline cache configuration. .	146
7.1	Voltage and frequency possibilities for CMOS-based and NCFET-based systems	152
7.2	Power consumption with various voltage and frequency combinations for CMOS-based system . . . . .	153
7.3	Power consumption with V7-F7 combination for CMOS-based and NCFTE-based systems . . . . .	154



# List of Tables

3.1	Various Configurations Explored . . . . .	45
3.2	Absolute values of cache metrics for FinFET and NC-FinFET-based 4MB, 8W, 4B cache . . . . .	50
3.3	Access rates for various SPEC2017 benchmarks for different cache levels . . .	54
4.1	System configuration . . . . .	71
4.2	SPEC CPU2017 benchmark mixes . . . . .	72
5.1	System configuration . . . . .	103
5.2	SPEC CPU2017 benchmark mixes . . . . .	105
5.3	SPEC CPU2017 benchmark mixes - Test set . . . . .	110
6.1	Calibration of ASAP7 7 nm FinFET with our 5 nm FinFET Measurements . . .	129
6.2	Cache configurations under exploration . . . . .	135
6.3	System configuration . . . . .	142