

Thesis Project Portfolio

Fortifying the Cloud using Big Data Computing

(Technical Report)

Exploring the Opacity of the Cloud

(STS Research Paper)

An Undergraduate Thesis

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Dagim Tekle

Spring 2024

Department of Computer Science

Table of Contents

1. Sociotechnical Synthesis
2. Fortifying the Cloud using Big Data Computing
3. Exploring the Opacity of the Cloud
4. Thesis Prospectus

Sociotechnical Synthesis

Secure and Sustainable Cloud

“If possible, so far as it depends on you, live peaceably with all.” - St. Paul, Romans 12:18 ESV

“Brothers and sisters, I do not consider myself yet to have taken hold of it. But one thing I do: Forgetting what is behind and straining toward what is ahead,” - St. Paul, Philippians 3:13 NIV

When contemplating the meaning of *the good life* or how to *live well* in the twenty-first century, I find inspiration in the teachings of Jesus of Nazareth and the writings of his followers. This perspective guides my reflective ethical inquiries, particularly through a Christian lens that emphasizes ongoing perpetual refinement in our pursuits, and living well (with others). As cloud computing gains traction, challenges arise concerning security and sustainability. In the technical section of my thesis, I reflected on my internship experience where I researched how to improve security log querying through cloud-native, scalable indexing and querying strategies. Improving the security postures of the cloud, we can take another step towards improving the privacy and overall well-being of users. In the sociotechnical section, I explored the responses of key stakeholders to the electric crisis triggered by the growing presence of data centers in Northern Virginia. By raising awareness among residents, this portion seeks to empower them with knowledge to actively engage with the evolving landscape of cloud infrastructure.

Starting with the technical portion, I explored the evolution of cloud computing from its conceptualization in 1961 to its widespread adoption by major providers such as Amazon, Microsoft, and Google. Despite its prevalence in modern society, concerns about privacy and

security persist among users. To address these concerns, robust access monitoring and efficient querying strategies are essential. Leveraging Lucene on Amazon Web Services (AWS), I examined the feasibility of a scalable indexing and querying strategy. While this approach simplifies deployment, challenges such as Lambda's time, memory, and concurrency limits cause index file proliferation and unresponsive subsequent querying. To overcome these limitations, future work involves exploring alternative AWS running environments like AWS EMR Serverless to enhance scalability and performance.

In the STS portion, I examined the concept of cloud computing and its impact on society, with a focus on Virginia's data center infrastructure. Despite the cloud's integration with everyday life, there seems to be a widespread lack of understanding of the cloud. By analyzing various sociotechnical theories, such as the Social Construction of Technology and Engineering as Social Experimentation, I aimed to reveal the hiddenness of cloud infrastructure and how it affects residents, businesses, and policymakers. Through an exploration of relevant groups such as Dominion Energy, the Virginia Economic Development Partnership (VEDP), Amazon Web Services (AWS), and the Piedmont Environmental Council (PEC), I identified diverse interests and interpretations of cloud technology's impact on Virginia's economy, electricity supply, and ecology. My findings emphasize the importance of increasing technosocial awareness and encourage stakeholders to make more informed decisions about how to live well with emerging technologies.

By considering the sociotechnical aspects alongside the technical components of cloud computing, I gained a comprehensive understanding of the rapid growth of the cloud market and its broad societal implications. The technical analysis offered valuable insights into scalable indexing and querying strategies, highlighting privacy and security challenges as cloud usage

increases. Meanwhile, the STS perspective uncovered the holistic impact of this technology, such as Virginia's data center infrastructure, which involves complex interactions between technological progress and stakeholders like energy providers, policymakers, and local communities. This awareness emphasizes the necessity for responsible, informed decision-making, balancing technological advancement with ethical considerations and sustainable practices. By combining an understanding of societal dynamics with technical expertise, we can proactively address the forthcoming challenges of the cloud's rapid expansion and maximize the benefits of these emerging technologies.