Thesis Project Portfolio

The Struggles of Kubernetes Key-Value Stores

(Technical Report)

Data Center Dreams or Devastation: The True Cost of Computing

(STS Research Paper)

An Undergraduate Thesis

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

University of Virginia • Charlottesville, Virginia

In Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

James Draudt

Spring, 2025

Department of Engineering Computer Science

Table of Contents

Sociotechnical Synthesis

Data centers are among the world's largest energy consumers and emerging sources of pollution. They are projected to overtake the entire aviation industry in emissions, significantly contributing to climate change. Data centers are often built unsustainably, and the software they rely on is not optimized (Bilal et al., 2014). The absence of regulation enables data centers to expand rapidly, overlooking their energy consumption and associated pollution. Within data centers, cloud computing technologies—such as Kubernetes clusters, which combine computers to complete tasks—are currently inefficient. If these problems are not fixed, data centers will continue to grow as the need for more computing power grows. This will only increase emissions and inefficiencies, causing irreparable damage to the environment and communities surrounding data centers. Temperatures will rise, severe weather events will increase, and local power and water infrastructure will be pushed to their breaking point.

As cloud computing continues to mature, one pressing question remains: How can we make this more efficient? This question is particularly relevant to Kubernetes clusters, the industry standard for scalable computing tasks. I chose to address this question and explore optimal solutions. Through an in-depth review of Kubernetes documentation and an examination of emerging technologies, I discovered that Kubernetes' database architecture could, and should, be optimized. Currently, Kubernetes relies on a single database called ETCD, which presents several problems: a shortage of maintainers, poor performance at both large and small scales, and challenges related to interaction and monitoring. Acknowledging these issues, I researched possible alternatives to this built-in limitation and identified a technology called KINE. KINE allows Kubernetes clusters to use preconfigured databases, enabling manual optimization based on cluster size, improving monitoring, and supporting modern, high-performance databases. I found YugabyteDB to be an efficient and reliable database option for large-scale clusters. Furthermore, connecting multiple Kubernetes clusters to a single YugabyteDB instance allowed them to collaborate—a capability that usually requires additional technologies to manage.

As data centers become an integral part of modern infrastructure, questions about their social and communal impact are increasingly urgent. To explore these concerns, I examined case studies, expert analyses, and statistical findings. This research revealed the extent of the damage data centers can cause: Sweden's once-renowned energy grid was severely degraded by data center demands (Libertson et al., 2021); communities are forced to live in drought conditions to meet data center resource needs (Sharma, 2024); and others suffer from constant noise pollution from these behemoths (Monserrate, 2022). These findings highlight that data centers are often built with minimal regard for environmental or societal consequences. Data centers currently account for approximately 2% of global greenhouse gas emissions (Bilal et al., 2014), degrade local power and water systems, create massive noise pollution, and are frequently located on land that should be preserved. As computing-intensive industries like AI and data science expand, data centers will continue to grow at an alarming pace. These facilities are largely self-regulated and not bound by uniform codes or sustainable design standards—allowing profit to take precedence over people (Sharma, 2024). This is a serious issue that must be addressed. National, state, and local regulations are needed to ensure that data centers can grow responsibly and become assets rather than liabilities to the communities they inhabit.

I believe I have successfully identified solutions to mitigate the problems caused by data centers. I proposed a regulatory framework modeled after past environmental policies that would protect the climate without hindering economic progress.

Technological improvements, such as using KINE in combination with YugabyteDB, could significantly improve efficiency. Enhanced cloud computing efficiency would reduce the need for new data centers while decreasing the electrical and water demands of existing ones. However, to create any long term solution further research is needed in both technical and policy domains. Potential legislation for data centers should be explored and drafted. Studies should be conducted to better understand the broader impacts of data centers, and research into cloud computing optimizations must be prioritized. In particular, we need comparative studies to determine which databases are best suited for Kubernetes clusters of various sizes.

I would like to thank Professor Wylie for her guidance throughout my STS research. I am also grateful to Zeta Associates for providing an internship experience that encouraged me to follow my curiosity and pursue this Kubernetes optimization project.