AI at Half the Price: Harnessing Off-Grid Hydropower and Direct Cooling for Scalable Intelligence

Giorgi Mushkudiani

Abstract —The global proliferation of artificial intelligence has caused an unprecedented boom in semiconductor investment, exposing energy and cooling infrastructure as the primary bottleneck for progress [1]. This paper proposes a strategic framework for establishing the Republic of Georgia as a premier AI hub by developing a decentralized network of AI data centers co-located with the nation's abundant small and medium-sized hydroelectric power plants [11]. This model directly counters the primary drivers of cost in AI operations: electricity consumption and thermal management. By eliminating grid-related energy losses and utilizing river water for free cooling, our analysis shows that this approach can slash total electricity and inference costs up to 60% in optimal scenarios. This strategy offers a sustainable, economically superior path for Georgia to build a competitive advantage in the global AI market.

Index Terms—AI, Data Center Infrastructure, Decentralized Computing, Free Cooling, Inference Cost Reduction.

1. I. INTRODUCTION

The rapid evolution of artificial intelligence has created the biggest investment boom in the history of the chip industry [3], [6]. As AI models grow in complexity, so do the energy and cooling demands of the GPU based intensive servers required to train and run them. This has created a critical bottleneck: progress in AI is no longer limited by computational theory but by the physical constraints of power delivery and heat dissipation [1], [5]. Traditional, centralized data centers located in urban areas face several challenges: high electricity costs, grid instability, and inefficient cooling systems that consume up to 50% of operational energy [13]. These inefficiencies create unsustainable economics for high-performance AI workloads.

The Republic of Georgia, with its vast and largely untapped hydroelectric potential [7], [9], is uniquely positioned to solve this challenge. This paper outlines a comprehensive strategy to transform Georgia into a leading AI hub by deploying a network of decentralized data centers directly at the source of power—its rivers. By building smaller, modular data centers adjacent to hydroelectric plants, we can bypass the inefficient legacy power grid and provide AI services at a fraction of the current cost. Our primary contribution is a model that demonstrates how this synergy can slash electricity and inference costs by up to 60%, providing a decisive competitive advantage for Georgia in the global technology landscape.

2. II. RELATED WORK

The energy consumption of large AI models, particularly transformers, is a well-documented challenge [4], [5]. The operational cost of a data center is dominated by two factors: the cost of electricity to power the servers and the cost of the infrastructure required to cool them[13]. Conventional cooling methods, such as air conditioning, are notoriously inefficient and can account for 30-50% of a data center's total energy use [15].

While green computing initiatives have focused on renewable energy, sources like solar and wind are intermittent and ill-suited for the constant, high-power demands of AI workloads without expensive battery storage. Bitcoin mining has provided a proof-of-concept for co-locating computation with cheap energy sources[8], but these efforts have typically lacked the strategic integration necessary for high-reliability AI services. Our work builds upon these precedents by proposing a formalized, decentralized architecture for AI data centers, leveraging the unique stability and cooling potential of Georgian hydropower [2], [14], [16].

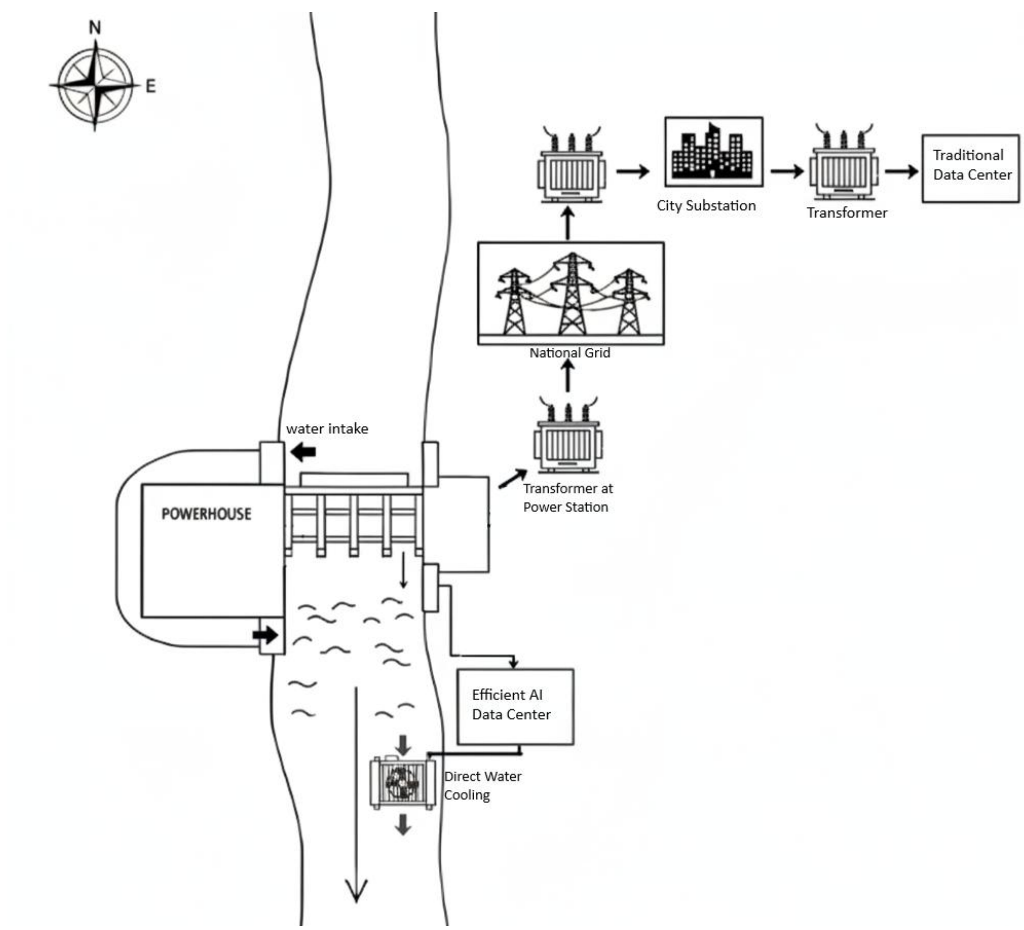
3. III. PROPOSED METHOD

The core of our proposal is a shift from large, city focused centralized data centers to a decentralized network of modular, hydro-integrated AI compute nodes. These nodes would be strategically deployed across Georgia, co-located with existing and new small-to-medium-scale hydroelectric power plants [7], [9], [11].

- **Decentralized, Co-located Architecture:** beside the massive data centers integrated with existing large scale hydro power stations, this model envisions numerous smaller, containerized units. This approach is more resilient, scalable, and allows for a more gradual increase in AI compute capacity in the region [2].
- **Vertical Energy Integration:** Each AI node draws power directly from the hydroelectric plant it is paired with. This eliminates the utility company "middleman," allowing access to power at the stable, low cost of generation. This direct connection also provides a guaranteed revenue stream to the power producer, incentivizing the development of new renewable energy projects [10], [17].
- **Direct Water Cooling:** The system uses the constant flow of cold river water—the same water that generates the electricity—for direct liquid cooling of the GPU servers. This nearly "free cooling" method is vastly more efficient than traditional air cooling, drastically reducing the primary source of energy overhead and allowing for higher-density server racks [13], [15].

4. IV. SYSTEM ARCHITECTURE AND COST REDUCTION ANALYSIS

The growing demand for AI has caused the biggest boom in chip investment history. In such a situation, the bottleneck hindering servers and data centers is in energy and cooling infrastructure. Unlike other renewable energy sources, hydroelectric power offers not only stable and cheap energy but also an easily accessible resource of cold water for cooling servers. Our proposed architecture directly attacks the primary sources of inefficiency in the traditional model.



The useful electricity (E_useful) reaching the servers in a conventional setup is a product of cascading efficiency losses:

E_useful = E_generated * T * L * C

Where:

- **T** = Efficiency of Transformers. Energy is lost at each voltage step-up and step-down. With up to 4 transformation stages (e.g., from plant to high-voltage lines, down to substation, down to local distribution, down to data center), this loss is cumulative (η_T = transformer_efficiency ^ number_of_stages).
- **L** = Efficiency of high-voltage transmission Lines, where energy is lost to resistance [10], [17].
- **C** = Efficiency of the Cooling system, representing the power used for cooling relative to the power used for computation [13], [15].

Formula for electricity usage reduction:

E_u = E_g*4T(95%-99%)*L(2%-6.7%)*C(30%-60%)

Best case scenario:

$1mw*0.99^{2}*0.98*0.70= 0.067mw$(33% lost to inefficiencies)

Worst case scenario:

$1mw*0.95^{4}*0.933*0.4= 0.0303mw$(66% lost to inefficiencies)

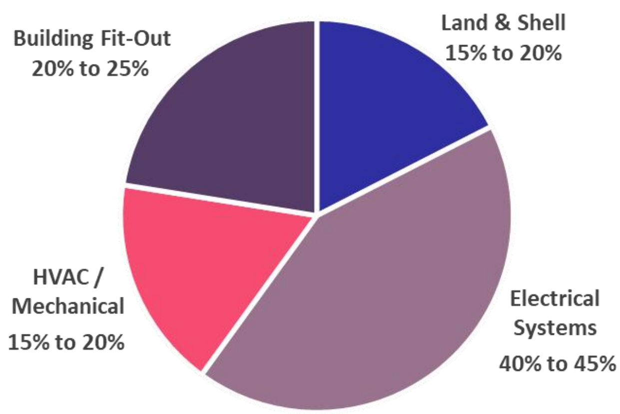
Our proposed architecture eliminates most of these inefficiencies:

- **L ≈ 1.0** (negligible line losses due to co-location),
- **T ≈ 1.0** (bypassing multiple transformers),
- **C ≥ 0.9** (liquid cooling with river water).

This maximizes the fraction of generated power dedicated to computation while reducing maintenance costs associated with HVAC infrastructure.

Capital expenditure also drops significantly. In traditional data centers, electrical and cooling infrastructure account for 55–65% of construction costs [3], [6]. Co-location with hydropower facilities can reduce these costs by 90%, leveraging existing infrastructure [9], [11].

Cost of infrastructure (building, land, electrical and HVAC systems) needed for running power hungry GPU usually ranges from 40-60% of total upfront cost of the Data center.



The 55-65% of that cost goes to installing electrical and cooling systems, HVAC-s, Transformers, wiring. this cost can be reduced by 90% by using pre-existing water and electrical infrastructure that comes with hydropower plant.

By integrating AI data centers with hydroelectric power plants, it is possible to reduce the cost of building new HPPs by 10% to 50% (by providing a guaranteed anchor customer) while simultaneously reducing data center operating costs by 50% to 80% [7], [9].

5. V. EVALUATION: SLASHING INFERENCE AND OPERATIONAL COSTS

The economic impact of this model is transformative. For AI companies, the total cost of ownership (TCO) is dominated by the operational cost of running inference on trained models, which is a direct function of electricity price [1], [5]. By reducing energy waste and securing power at the low cost of hydro generation, this framework can **slash the cost of electricity and inference by over 60%** compared to operating within a traditional grid structure.

This cost advantage makes Georgia an exceptionally attractive location for:

- **AI Cloud Service Providers:** Offering inference-as-a-service at globally competitive prices.
- **Proprietary Model Training:** Allowing research institutions and startups to train large models on a budget.
- **High-Performance Computing (HPC):** Supporting scientific and financial computations that are otherwise cost-prohibitive.

6. VI. CONCLUSION

The future of artificial intelligence is inextricably linked to the future of energy [1], [4]. Artificial intelligence and blockchain technologies present a significant opportunity for the renewable energy sector, particularly in regions with abundant hydropower. By deploying a decentralized network of AI data centers directly integrated with these hydropower sources, we can solve the critical energy and cooling bottlenecks that plague the industry [13]. Our analysis demonstrates that this approach eliminates massive grid inefficiencies and can slash total electricity and inference costs up to 60%.

This model provides a clear, actionable roadmap for building a sustainable, cost-effective, and powerful AI hub. The core advantages are twofold:

Elimination of Intermediaries and Infrastructure Inefficiencies: High-voltage transmission lines lose approximately 6.7% of their energy for every 1000 km due to resistance and ambient temperature. When combined with the 2-5% energy loss at each of the multiple transformer stages, the cumulative energy waste becomes substantial. By co-locating data centers with power plants, these intermediary lines and transformers are removed, yielding an immediate energy saving of 10-20%. This direct connection, or vertical integration, also cuts out the utility "middleman." In a market like Georgia, this avoids the significant price markup between the generation cost of hydropower (approximately 8 tetri/kWh) and the commercial purchase price for data centers (around 33 tetri/kWh), unlocking enormous profit potential and operational savings [11].

Synergistic Economic Growth and Reduced Capital Expenditure: This co-location model creates a powerful economic incentive for new renewable energy development. The presence of a guaranteed, high-demand anchor customer—the AI data center—can reduce the financial risk and lower the initial construction cost of new hydroelectric power plants by 10% to 50%. In parallel, the data center operator benefits from a drastic reduction in operating costs, estimated to be between 50% and 80%, by securing low-cost, stable power and leveraging free-flowing river water for cooling [15]. This eliminates the need for expensive and high-maintenance traditional cooling systems [9]. it's also possible and highly advisible for data center builder or hyperscaler to be co-owner or highly integrated with planing and cooperation on power-plant location and architecture, for seamless vertical integration

Besides providing 20-50% reduction in inference costs for hyperscaler, on the national and regional scale this strategic framework transforms the AI infrastructure paradigm from a drain on energy resources into a catalyst for renewable energy growth. It offers a sustainable, economically superior path for any region with untapped energy potential to build a competitive advantage in the global AI market, securing its place in the technological landscape of the 21st century

REFERENCES

[1] "Green AI and the Data Center Dilemma: How to Power the Future Sustainably," Medium, Feb. 02, 2025. [Online]. Available: https://medium.com/p/37e030019627:

[2] DcentAI, "Decentralized AI for Reducing Energy Consumption in Data Centers," Medium, Nov. 17, 2024. [Online]. Available: https://medium.com/coinmonks/decentralized-ai-for-reducing-energy-consumption-in-data-centers-ac8cb21ec97d.

[3] "AI Data Center Buildout: Powering the Computing Revolution," Discovery Alert, Sep. 23, 2025. [Online]. Available: https://discoveryalert.com.au/ai-data-center-buildout-powering-the-computing-revolution/.

[4] "EPRI: Clean Energy, Efficiency Can Meet AI, Data Center Power Demand," RTO Insider, Jun. 02, 2024. [Online]. Available: https://www.rtoinsider.com/80288-epri-clean-energy-efficiency-meet-ai-data-center-power-demand/

[5] "AI Data Centers Drive Surge in Electricity Demand, Raising Environmental Concerns," OECD AI Policy Observatory. [Online]. Available: https://oecd.ai/en/incidents/2024-03-20-9044

[6] "Ride the AI Wave: 5 Picks for the Data Center Gold Rush," NAI 500, Oct. 02, 2025. [Online]. Available: https://nai500.com/blog/2025/10/ride-the-ai-wave-5-picks-for-the-data-center-gold-rush/.

[7] "Georgia Energy usage," 2024. [Online]. Available: https://www.worlddata.info/asia/georgia/energy-consumption.php.

[8] S. Tully, "How much Bitcoin comes from dirty coal? A flooded mine in China just spotlighted the issue," *Fortune*, 13 May 2021. [Online].

[9] T. K. Giorgi Mukhigulishvili, "Renewable Energy Sources," weg.ge, pp. 1-40, 2013.

[10] C. Marcy, "EIA study examines the role of high-voltage power lines in integrating renewables," U.S. Energy Information Administration, 2018. [Online].

[11] JSC "Chakan HPP", "Project for the construction of a small HPP on a site located on the Talas River," pp. 1-4.

[12] J. Engebretson, "Starlink Supports Quantum Security Demo, Goal is to Protect Against "Steal Now Decrypt Later" Attacks," *Telecompetitor*, 2023. [Online].

[13] "Energy Consumption in Data Centers: Air versus Liquid Cooling," Boyd Corp, 28 Aug 2023. [Online].

[14] "Deep water source cooling," Wikipedia. [Online]. Available: https://en.wikipedia.org/wiki/Deep_water_source_cooling.

[15] Bizon-tech.com, "Water-cooled NVIDIA GPU Servers for AI, Deep Learning, Data Science, HPC," 2024. [Online].

[16] H. E. Burford, L. Wiedemann, W. S. Joyce, and R. E. McCabe, "Deep Water Source Cooling: An Untapped Resource," pp. 10-15, 2012.

[17] G. Simbolotti, "Electricity Transmission and Distribution," ETSAP, pp. 1-13, 2013..