Towards an Adaptive and Federated Testbed for AI Research in Africa

1st Alex Mwotil*
Department of Computer Science
Makerere University
Kampala, Uganda
alex.mwotil@mak.ac.ug

2nd Innocent Ndubuisi-Obi Jr*
Paul G. Allen School of Computer Science and Engineering
University of Washington
Seattle, USA
innoobi@cs.washington.edu

3rd Kurtis Heimerl
Paul G. Allen School of Computer Science and Engineering
University of Washington
Seattle, USA
kheimerl@cs.washington.edu

4th Benjamin Kanagwa
Department of Networks
Makerere University
Kampala, Uganda
benjamin.kanagwa@mak.ac.ug

5th Engineer Bainomugisha
Department of Computer Science
Makerere University
Kampala, Uganda
baino@mak.ac.ug

***Abstract*—In a low-resource setting, unique challenges such as limited infrastructure, high network costs, and unpredictable network connectivity can hinder the setup and operation of research testbeds. Often, researchers rely on local and limited institutional resources or distant, less accessible platforms in other continents. However, a federated research testbed can provide an expanded resource pool with unified administrative features, promoting collaboration and contextual research in areas such as Artificial Intelligence (AI) and protocol design that align with local needs. In this paper, we propose a federated compute systems testbed spanning Research and Education Networks in Africa that is adaptive to the challenges unique to computing in Africa. Five key design considerations for such a testbed are presented: *Adaptive*, *Ubuntu*, *Replicability*, *Ease-of-use* and *Secure and Policy-driven*. The main ideas put forth in this paper are that a federated and topology-aware network and systems testbed anchored in NRENs can lower the costs of AI research and serve as a practical workaround for the prevalent skill and infrastructural gaps in Africa.**

***Index Terms*—federation, testbed, low-resource, research, education**

I. Introduction

Artificial Intelligence (AI) has attracted research interest in low-resource settings like Africa, where its potential aligns closely with ongoing continental initiatives to uplift education, healthcare [33], and broader economic landscapes. AI provides new avenues to address some of the most pressing local contextual challenges, such as improving access to quality education, improved production in agriculture, environmental sustainability, optimizing healthcare delivery [33], and fostering economic development [23]. AI-driven solutions can enhance teaching and learning resources, provide tailored health diagnostics, and support predictive analytics in agriculture and commerce, all of which have direct and positive implications for socio-economic growth [4]. AI fields - such as machine learning, natural language processing, computer vision, and data science - require reliable and vast resources that include compute, storage and network to process and ultimately serve applications to users. There are a number of challenges and opportunities for African researchers in this domain:

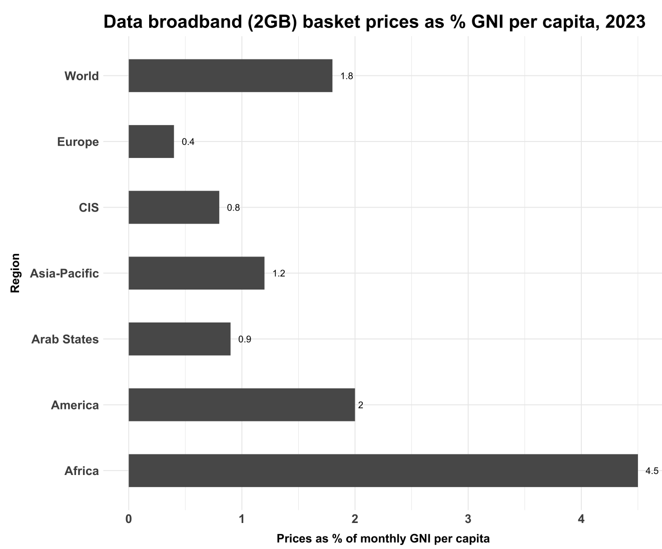

Fig. 1. Adapted from *The State of broadband 2024: Leveraging AI for Universal Connectivity* [12]

A. Significant Cost Barriers to Meaningful Use

Sub-Saharan Africa's (SSA) digital infrastructure coverage, access, and quality still lag in comparison to other regions. In the Oxford Insight AI Readiness Index 2024, of the nine global regions reported, SSA ranked last [21]. In 2023, the cost of 2 gigabyte of mobile and fixed-broadband Internet as a percentage of monthly per-capita Gross National Income (GNI) was 4.5% and 14.8% respectively [9], [12]. The recommendation level by the United Nations Broadband Commission is 2

* Authors contributed equally to this work.

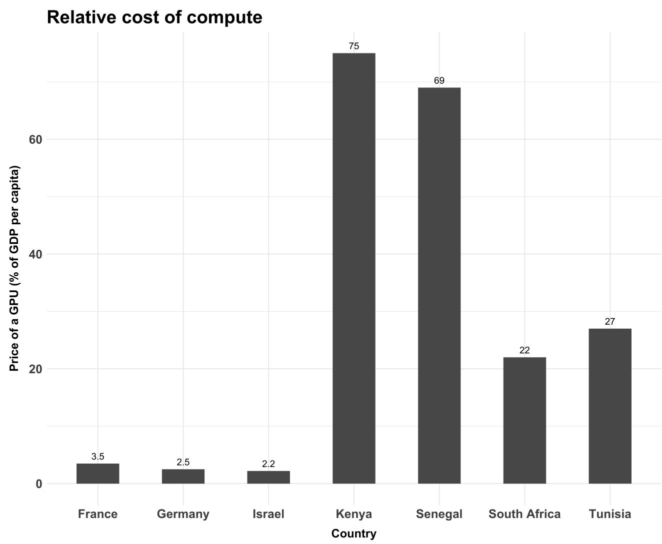

Fig. 2. Adapted from *AI in Africa: The state and needs of the ecosystem* [3]

percent. Additionally, AI researchers face unique challenges that exacerbate existing cost barriers. In Figure 2, the price of a GPU in the South African and Kenyan economies respectively represents 22% and 75% of GDP per capita making it 9 and 31 times more expensive than in high-income countries [3]. In this cost environment and with limited access to specialized processing hardware, many rely on old, CPU-based resources like local workstations, servers, and laptops [23] to do AI research.

B. The Need for Resilient Cloud Infrastructure

Cloud environments require highly resilient infrastructure characterized by low-latency backbones and stable power supplies. However, low-resource settings are mired in the challenges of inconsistent power supply [6], fiber cuts, and poor network inter-connectivity. For instance, Angola, Uganda, and Cameroon experience approximately 3,000, 1,800, and 400 hours of national grid power outages annually, respectively [15]. The lack of redundant network paths and the high likelihood of fiber cuts, often caused by poor route planning and construction activities [5], can render infrastructure unusable. These issues make it extremely difficult for cloud providers to operationalize new regions or sites within such settings. Therefore, researchers reply on slower computing options which stretches even basic AI tasks over days. To overcome these constraints, some researchers rely on hosted Notebook services like Google Colab [1] for ML training needs, which, while valuable, come with stringent restrictions on compute time, memory, and bandwidth that restrict the scope and efficiency of projects. Additionally, because these platforms are hosted on public cloud infrastructure with minimal presence in Africa, researchers encounter latency issues and limited upstream availability.

[1] https://colab.google/

C. Accessibility and Capacity Gaps

Significant disparities in talent and capacity continues to contribute to an ‘AI divide’ between the Global North and the Global South [42]. Many organizations like DeepLearning Indaba, Data Science Africa, and Women in Machine Learning and Data Science has attempted to fill this capacity and skills gap on the continent [4], [24], [29]. For the more skilled researchers with access to on-prem hardware-accelerated clusters, additional challenges exist. These are usually limited in capacity and suffer from fragmentation, leading to long queue times and delayed access to necessary resources. Additionally, knowledge gaps in configuring and maintaining these systems can result in underutilized infrastructure, leaving valuable resources idle or inefficiently managed, affecting the progress of AI research across the continent and necessitates the urgent need for accessible, affordable, and efficient solutions specifically tailored to the needs of African researchers.

D. Federation and Resource Sharing

Federation of resources through consolidation of available compute infrastructure across Africa can provide avenues for collaboration and resource sharing on AI projects. Similar initiatives, such as the Pacific Research Platform [37], CloudLab [16], Chameleon [27], Fabric [8], PlanetLab [14] and EdgeNet [36], have addressed the computing needs of researchers worldwide, enabling institutions to contribute resources to a shared cloud pool. A federation allows users to provision and execute research experiments across domains like cloud computing, distributed systems, security, AI, and network measurements. However, these research testbeds are predominantly concentrated in regions like the United States, Europe, and Asia. The reasons for this uneven distribution are complex, with cost being a primary factor. Acquiring, operating, and maintaining such infrastructure — let alone contributing it to a shared resource pool — is expensive. In Africa, this challenge of federation is exacerbated by the lack of policy harmonization across governments related to data access and privacy, information flow, responsible generation, use, and storage of data.

E. NRENs as a Backbone for Federated AI Compute

The *African network* is an interesting use-case in the design of federated research testbeds for AI, particularly due to its reliability issues: frequent outages, often caused by power failures, cable cuts, and network misconfigurations, combined with a lack of robust infrastructure (e.g., redundant links), limit consistent network availability. Moreover, circuitous, triangular [13], and asymmetric routing paths within the continent contribute considerable latency to even continental international communication. Although Internet eXchange Points (IXPs) and remote peering are meant to enhance local connectivity, recent studies suggest that these are yet to achieve the intended impact [28]. How do we understand or design testbeds that are aware and characteristic of these settings? National Research and Education Networks (NRENs)in Africa provide a good backbone for a federated cluster of AI compute.

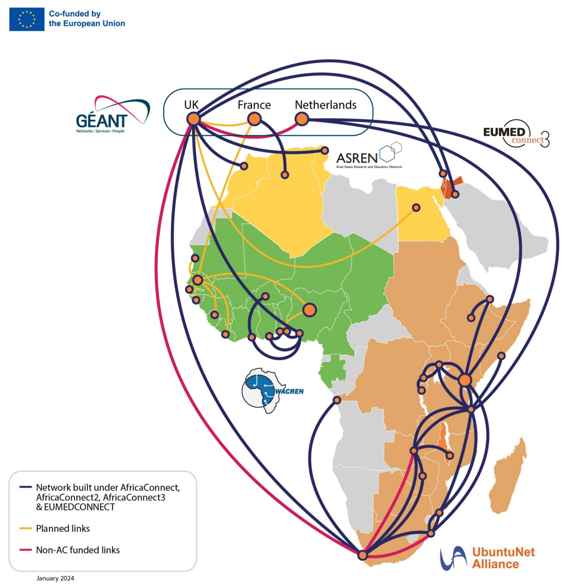

Fig. 3. National research and education networks are interconnected at regional and global levels to enable universities, research institutions and educational organizations to collaborate. Adapted from the AfricaConnect3 project connectivity map [34].

NRENs are specialized not-for-profit networks that provide dedicated connectivity and supporting services tailored to needs of the academic and research community [18]. These networks facilitate reliable, high-performance data exchange, enabling collaboration across universities, research institutions, and educational organizations. NRENs are interconnected through Regional Research and Education Networks (RRENs), which in turn connect to the global research and education network via multiple Points of Presence (PoPs) around the world as shown in Figure 3 for Africa. This structure creates an international infrastructure that supports large-scale data transfers, shared computing resources, and collaborative projects. Through these interconnections, NRENs can also provide essential backbone support for experiments and innovations in fields such as AI, climate science, and genomics.

In Africa, 38 NRENs [34] currently operate, covering over 70% of the continent. These NRENs vary widely in maturity, influenced by factors such as network capacity, existence, and range of services offered. Despite their collective commitment, only a few NRENs in Africa currently provide operational AI research infrastructure, and those that do are characterized by diverse configurations in terms of size, availability, and usage. For example, the Kenya Education Network (KENET) manages a GPU-based research cluster with just a minimal 4-node setup, which is heavily strained, often requiring reservation requests to wait several days before access [1]. In other regions, while some AI infrastructure exists, it remains underutilized due to gaps in technical expertise, rendering valuable resources idle and ineffective.

Given these challenges, research testbeds designed for Africa must effectively account for and adapt to the conditions facing African networks such as poor network path quality, high-bandwidth delay products, intermittent connectivity, unstable power grids, and diverse network configurations. These testbeds must treat adaptive routing, redundancy planning, and dynamic resource management as first-class features. They should benefit from localized caching, decentralized compute nodes, and protocols that minimize dependency on unstable path and circuitous routes. By embedding these contextual factors into testbed infrastructure, we can create a platform capable of efficiently and sustainably supporting AI research that aligns with the specific needs and use cases of African researchers. Thus the key idea put forth in this paper is the following: *A federated and topology-aware networking and systems testbed anchored in NRENs can lower the costs of AI research and serve as a practical workaround for the prevalent skill and infrastructural gaps in the region.* We describe the requirements and propose an abstract model for such a pan-African research testbed that is both aware of and responsive to the challenges inherent in low-resource settings. The remainder of the paper is organized as follows: Section II reviews related work in the field, Section III details the design requirements for the testbed, Section IV presents the proposed architecture, and Section VI concludes the paper while discussing directions for immediate and future research.

II. Related Work

Testbeds are integral to any research domain, serving as controlled environments where theoretical concepts and practical implementations can be tested and validated [26]. In Computer Science, research testbeds offer a representative platform for performing experiments, enabling researchers to evaluate systems, protocols, and applications under various conditions [32]. These experiments often have far-reaching implications, such as in the development of new technologies and innovations. These implications shape how systems are designed, deployed, and optimized in real-world scenarios. By replicating complex environments, testbeds bridge the gap between theoretical research and practical application, to influence the next generation of computing and network technologies [17]. The rapid growth of compute-intensive AI workloads has necessitated significant adaptations and resource enhancements to support these applications in these testbeds. A number of research testbeds that have now evolved to support AI workloads exist and can be broadly characterized into: generic, federated, and hybrid testbeds.

A. Generic

Designed to provide physical or virtualized computing and network resources to simulate real-world environments, these testbeds provide a foundation to test and evaluate hardware, software, and systems. Availability of infrastructure is the baseline for any research testbed. The capabilities of the testbed are dependent on its goals and the availability of resources to support these experiments. Scalability, performance and resource allocation are important attributes of these testbeds. The operators provide an interface and tools to configure the hardware, network topologies and compute environments as per user needs. The infrastructure typically includes on-premise physical hardware, virtual machines, cloud instances, or hybrid environments. Most research testbeds provide some form of generic infrastructure to the users. For example, GENI (Global Environment for Network Innovations) [**?**] now transitioned to FABRIC [8], Grid'5000 [11], CloudLab [16], Chameleon [27] provide enormous compute, network and some dedicated AI resources. In Kenya, KENET's GPU testbed [1] provides virtual GPU resources to researchers. The testbeds provision and manage hardware resources using a combination of provider software, and open-source platforms like OpenStack and Kubernetes. It should be noted that these testbeds have evolved to additionally support AI workloads. While many testbeds aim to cater to multiple fields, creating a one-size-fits-all platform is complex. This complexity is due to the diverse nature of research needs, ranging from high-performance computing for AI workloads to distributed systems and networking experiments, each demanding specialized configurations and capabilities. For example, PEERING [35] aims to support a tested that allows researchers to carry out experiments on the Internet's BGP routing systems, while testbeds like FABRIC [8], Grid'5000 [11], CloudLab [16] are primarily designed to support networking, distributed computing, and storage research.

B. Federated

Institutions with a common research agenda and collaborations can opt to contribute all or part of their resources to build a federated research testbed. Government support for testbeds focuses on in-country federation of resources typically hosted at research and education institutions. Federated infrastructures enable large-scale experiments and allow researchers to address complex challenges that demand substantial computational and networking resources. Computer Science is inherently global, and resource federation across borders unlocks new opportunities for innovation and experimentation. CloudLab [16], Chameleon [27], and EdgeNet [36] demonstrate the importance of federated testbeds in enabling research to global scales. Beyond compute, research has highlighted the role organizations like NRENs can play in federation of DNS resolvers to improve latency and resilience of DNS resolution and increase the adoption of standards [10]. In addition, federated setups reduce the duplication of efforts, enhance reproducibility, and foster interdisciplinary research by integrating global resources into cohesive, collaborative environments. Cost-sharing, redundancy, and fail-over are essential benefits of a federated research testbed.

C. Hybrid

Hybrid research testbeds are dynamic, borderless distributed resource pools designed to cater for multiple research domains and geographies. By combining physical and virtual resources across a wide range of institutions and regions, these testbeds enable a high degree of flexibility and scalability. They are particularly valuable for interdisciplinary research, as they offer access to diverse infrastructures such as cloud computing, high-performance clusters, and specialized hardware like GPUs. These testbeds bridge gaps between localized testbeds and global federations, fostering cross-regional collaborations. Examples like PlanetLab [14], which spans multiple continents, and EdgeNet [36], with its decentralized model, highlight the potential of hybrid testbeds to integrate heterogeneous environments, making them accessible to researchers worldwide. Their borderless nature not only enhances resource efficiency but also drives innovation across geographical and disciplinary boundaries.

III. Design Requirements for a low-resource Federated AI Testbed

Low-resource settings like Africa present unique challenges and opportunities for the design of hybrid or federated research testbeds. Limited infrastructure often prevents the establishment of local testbeds, let alone contributions to global federations. Moreover, skill shortages in testbed setup and management lead to under-utilization of available resources. Network limitations, such as high latency and frequent outages, further complicate integration into broader federated systems. Despite these obstacles, the collaborative potential of resource federation is evident, allowing institutions to share limited assets and bridge gaps in research capacity. Without addressing these challenges, Africa risks being excluded from advancements in critical fields such as AI, distributed systems, and Internet development, leaving innovations tailored to high-resource environments inaccessible. Federated research testbeds, designed with African realities in mind, can democratize access to computing resources and drive inclusive global progress. In this section, we describe 5 key design considerations for building a federated testbed to support particularly AI research in low-resource settings:

A. DC1: Adaptive

A federated testbed in a low-resource setting should be adaptive to the volatile conditions of the environment. For the platform to be adaptive, it must be aware of the *topology*, *costs*, and *resources* that support the experiments to ensure seamless operation under changing conditions:

1) Topology-aware: Low-resource settings, such as those in Africa, face frequent infrastructure challenges, including bandwidth limitations, power outages, and fiber cuts. These factors can dynamically alter the federation's topology, disrupting access to resources and experiments. In extreme cases, network partitions may isolate portions of the federation, breaking communication and interrupting experiments. A federated testbed should be topology-aware - implements real-time monitoring of resource availability and network connectivity, allowing for dynamic reconfiguration to ensure continuity.

2) Cost-aware: Network inefficiencies, such as "weird routing" observed in Africa, can lead to increased operational expenses and degraded performance of a federated testbed. These routing anomalies can result in higher latency, reduced throughput, and unpredictable network communication costs. A federated testbed should be aware of these costs and strive to minimize them as much as possible, for example through optimized resource allocation and communication. Coupled with real-time network monitoring, the goal is to analyze and address routing and resource allocation inefficiencies. Furthermore, resource management policies should prioritize equitable and efficient use of shared infrastructure, ensuring that cost constraints do not hinder participation or experimentation.

3) Resource-aware: A federated testbed in a low-resource setting is inherently heterogeneous, reflecting variations in available resources, user requirements, and application demands. For instance, AI experiments requiring accelerator hardware, such as GPUs or TPUs, may face uneven distribution or scarcity across locations. Consider a machine learning application designed to assess plant disease incidence and severity using images [31]. This application involves multiple stages — data collection, processing, modeling, deployment, inference, and maintenance — each with distinct computational and user-specific needs. The modeling stage demands high compute power, often necessitating accelerators, while data-related stages require substantial storage capacity and may involve localization to comply with data residency or regulatory requirements at the country level. Institutions across NRENs could also participate collaboratively at different stages of this pipeline. A federated testbed must therefore be capable of dynamically generating configurations that align with the specific demands of an application, including compute, storage, and localization needs, while also accommodating user-specific preferences and workflows. This can provide for optimal usage of the available limited resources in these settings.

B. DC2: Ubuntu - Collaborative and Cooperative

A low-resource federated testbed should embody the philosophy of Ubuntu - *I am because we are*. This philosophy emphasizes collective responsibility and mutual support, to foster collaboration among research and education institutions. By contributing their resources (computing and human) to the federation, no matter how limited, institutions can create a shared pool that benefits all participants. Such a spirit of cooperation not only strengthens the federation but also ensures that even the most resource-constrained organizations can access the tools and infrastructure for their experiments. This collective approach leverages unity to overcome individual limitations, and drive innovation across the federation.

The testbed should integrate tools that simplify the process of registering new institutional resources into the platform. These tools should streamline onboarding, enabling institutions to contribute compute, storage, and networking resources with minimal technical expertise. Features such as automated configuration, user-friendly interfaces, and adaptive integration protocols can lower barriers to participation, ensuring wider adoption. By easing registration, the testbed can rapidly scale its resource pool, enhancing its capacity to support diverse research initiatives while fostering inclusivity within the federation.

Incentive programs, such as offering more user credits to contributing institutions, can boost participation in federated testbeds. These credits could provide prioritized access to shared resources or additional computing time, regardless of the scale of resources an institution contributes. This approach should encourage more stakeholders to join the federation, with increased participation strengthening the overall resource pool and hence utility to all users.

C. DC3: Replicability

Replication allows for validation of experiments and ensures that they are reproducible across different environments. An experiment profile, which specifies the resources and configurations required to reproduce results, is an important feature of a testbed. This enables users to rerun experiments under consistent conditions and verify outcomes. However, in a low-resource federated testbed, achieving replication is challenging due to variations in resource availability and connectivity across the federation. These differences can introduce inconsistencies, such as deviations in performance metrics or computational results, depending on the topology and resource constraints at the time of execution. Federated testbeds should incorporate adaptive replication mechanisms that account for such variations, ensuring that replication is as robust as possible despite the limitations of the testbed environment.

D. DC4: Ease of Use

In low-resource settings, technical expertise is scarce and simplicity is an important design consideration of a testbed. A testbed platform must provide user-friendly interfaces and intuitive navigation to ensure accessibility for researchers. A federated testbed provides a unified platform reducing the complexities of interacting with disparate systems. This allows users to focus on their research rather than the underlying infrastructure. The simplicity of the system should extend to both the user interface and the process of contributing resources. Federation should create a seamless experience, where institutions can easily register their resources and researchers can quickly access the tools they need. Constructing clear, straightforward workflows that simplify resource management, experiment setup, and result retrieval is important. This reduces the learning curve and enables users to maximize the utility of the testbed without being hindered by the additional technical barriers.

E. DC5: Secure and Policy-driven

The platform should provides a structured means of managing multiple tenants each with potential overlapping resource allocation need and access control/security requirements. Utilizing identity federations, such as federated identity management systems (e.g., Shibboleth or OpenID Connect), can ease user on-boarding by allowing users to authenticate using their existing institutional credentials. There should also be a means to ensure that resources allocated to a tenant can be provably isolated for those given to others. As a federated system extends to multiple governance domains, it is essential to ensure that local and global constraints in compute, access, and information flow can be guaranteed. This is particularly valuable in federated environments, where researchers may come from diverse institutions with different administrative systems.

The platform will need a governance framework that defines roles and responsibilities for managing resources, resolving conflicts, and maintaining transparency in decision-making. These are essential to ensure the long-term sustainability and equitable operation of the testbed. A participatory governance model involving stakeholders from member institutions can foster trust, inclusivity, and shared ownership of the platform. Policies on resource contributions, data access, and experiment prioritization should also be clearly defined to support fair usage and collaboration among diverse users.

It is also important to monitor ongoing experiments to ensure compliance with acceptable use, access control, and information flow policies. Policy enforcement systems should be designed to detect violations in real time, flagging any resource abuse or misuse of testbed capabilities, while ensuring the seamless operation of research experiments. Effective federation management relies on clear organization, simplified user access, and comprehensive monitoring to foster an environment where researchers can use the platform effectively and efficiently.

The research and education community in Africa is well-organized and collaborative. Annual conferences bring members together to present and discuss issues of shared interest. The community also engages in numerous capacity-building activities, including technical workshops and meetings, which serve as platforms for knowledge exchange. One of the most impactful initiatives is direct engineering assistance, where NRENs collaborate by sharing expertise and resources to strengthen their networks and systems. The federated testbed can leverage this community for ongoing development, management, and support, including onboarding and training in platform usage.

IV. Proposed Architecture

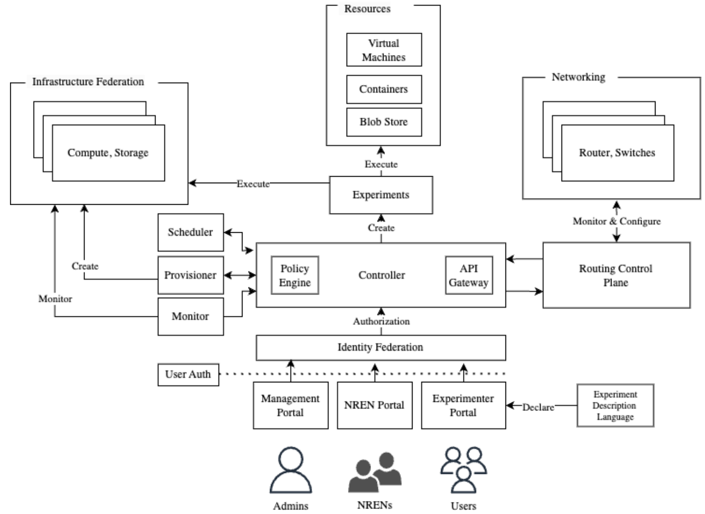

Fig. 4. A high-level architecture of the federated testbed for low-resource settings.

Designing a federated testbed for low-resource settings is a multi-disciplinary effort that involves both technical and human considerations to ensure the platform is utilized to its full potential. Such a testbed must be cognizant of the specific challenges faced in low-resource environments, adapting dynamically to the ever-changing conditions. It should be extensible, allowing for the simple integration of new resource pools as they become available. As with any experiment research platform, ensuring the ability to replicate experiments is essential for validating results and fostering collaboration. Managing a diverse set of users and resources calls for additional features, such as identity management systems and comprehensive monitoring tools, to ensure fair usage, compliance, and maximum benefit for all users. A well-designed federated testbed should be adaptive, scalable, and capable of facilitating collaboration across a range of domains and borders. The proposed architecture for the federated testbed is composed of seven components that work cohesively to provide an adaptive, and user-friendly platform:

- ***Routing Control Plane***: The routing control plane maintains routing information related to transit and peering links and also forwarding information related to tenant networks. It performs dynamic routing based on a number of criteria from shortest path to *cheapest* path. The routing control plane acts to keep all traffic where possible within the NREN fabric: either through peering or dedicated bilateral private network interconnects (PNIs) between the NRENs and RRENs.
- ***Workload Description Language***: This component allows users to design and specify experiments in a declarative format. When the controller and provisioner deploy resources for the workload, it leverage the specification and automated synthesis to find a configuration of resources that is satisfiable for that workload and existing workloads. This component would provide uses with push-button solver-aider tools for defining policies and constraints for their workloads.
- ***Policy Engine***: This component includes tooling for defining, assigning, and enforcement of acceptable use, access control and information flow policies. It tracks the status of both the infrastructure and ongoing experiments.

It collects metrics on resource availability, usage, and connectivity while providing real-time alerts for any anomalies or disruptions. This ensures the platform can adapt dynamically to changing conditions.

- ***Experiment Description Language***: The experiment description language is a component that should specify the resource needs and configurations for the expected workload. It should provides a highly opinionated API for describing the resource needs (memory, storage, region, etc) for a particular experiment on the testbed.
- ***Controller***: The controller is the central coordinator that handles the configuration and scheduling operations of the platform. It allocates resources, schedules experiments based on user requirements and available infrastructure (from monitoring information), and resolves conflicts to optimize platform utilization.
- ***Provisioner***: This component manages the physical and virtual resources that constitute the testbed. It handles the on-boarding of new resources, ensuring seamless integration and interoperability with existing components. This layer abstracts hardware specifics while maintaining performance and resource awareness.
- ***Portal***: The users access the testbed through a set of portals depending on their use case of the testbed, and the relevant access control policies are applied accordingly. The use of federated identity management will expedite user on-boarding to the platform. The three main user portals include:
  - Experimenter: This portal provides users with the front-end for core interaction with the experiments. It will allow the user to create new, and manage existing experiments.
  - NREN: The NRENs contribute resources to the federation, and this portal provides management features to support this function. The portal can be used to scale and monitor the federation infrastructure provided by the NREN.
  - Admin: The administrative portal is used by the federation operators to manage and monitor all components of the testbed.

This modular architecture ensures scalability, adaptability, and efficiency, making it suitable for the unique challenges of low-resource settings. Each component plays a role in enabling a resilient and collaborative research testbed environment.

V. Use Cases and Motivating Examples

To motivate the need for the testbed proposed in this paper, we provide two use cases of such a testbed and how it could responds to existing challenges in Africa.

1) **Policy-aware Scholarly Data Commons as Testbed for Pan-African Data Challenges.** Today, there is a gap in locally-relevant data that can be used to build AI tools and models relevant to the African context [2], [7], [23], [40]. This proposed testbed can play the coordinating role of a *data commons* [3], [20], [22] responding to calls for collaborative platforms for data collection and sharing. Similar to existing solutions such as Dataverse[2] in Harvard, it can play a role of coordinating access to high-quality data that can be used by research institutions in model training. Additionally, current national policies on cross-border data-sharing complicate the use of AI and other data-intensive systems by restricting flow of data across national borders. In some cases, these policies require the use of "domestic mirrors" which require that access to data must need local policy constraints [19]. The increasing trend of internet fragmentation and data localization will only further complicate issues around data sovereignty [19], [23], [38], [41]. NRENs are well positioned to experiment and create mechanisms for data sharing and use between themselves. Such mechanisms can be supported by our proposed testbed and can not only support sharing of data between NRENs and the individuals and institutions they support, but also inform conversations around data harmonization in Africa.

2) **Shared Compute, Storage, and Networking for Availability and Affordability.** As we previously mentioned, the high costs of GPUs and cloud computing are significant barriers to large-scale data science and AI research in Africa [3]. Additionally, low broadband penetration, high compute cost, and dollar-denominated egress fees associated to high data shuffling to multi-node AI workloads can have significant costs implications for African users [3], [30], [39]. Additionally the risk to resilience of unstable energy grids call for workloads that can be distributed. Maintaining high-performance computing clusters with high up-times can be cost exorbitant and carbon-positive as diesel generators become the most reliable method of ensuring up-times. By distributing storage and compute across NRENs, our proposed research testbed can leverage more resilient compute infrastructures and affordable and lower latency storage nodes. Additionally, leveraging existing inter-NREN connectivity fabrics and peering, our proposed tested can overcome cost barriers associated to deploying AI workloads using more expansive transit providers or dollar-denominated egress fees in available public clouds. As AI training costs and platform fees continue to rise [3], our platform can help keep AI research and use within the reach of African researchers by building resilient local capacity.

[2]https://dataverse.harvard.edu

[3]By one measure, it is 1,000 times more capable than its predecessor, GPT-2. But training GPT-3 cost, by some estimates, almost $5 million [25]. In 2021, it was estimated that when models need to be trained for specific tasks and can cost more than $50,000, paid to cloud computing companies to rent their computers and programs [25].

VI. Conclusion & Future Work

Low-resource settings, such as Africa, face unique challenges including constrained resources, limited technical expertise, and interconnectivity hurdles, which make participation in federated testbeds particularly difficult. As a result, researchers often rely on localized institutional testbeds with limited capacity or turn to distant ones based in the USA, Europe, and Asia. This geographic disparity raises critical questions: how can researchers in these settings effectively conduct experiments on pressing local issues such as AI and protocol design that directly impacts them?

To address this gap, this work proposes a Pan-African research testbed tailored to support emerging AI applications and broader computational research needs. We describe five critical design considerations for such a platform, taking into account the inherent volatility and resource constraints of the deployment environment. Furthermore, a high-level architecture is presented to provide a blueprint for implementation.

Looking ahead, our focus will be on contextualizing this proposal to align with the needs of African research and education institutions. These institutions, which form a well-organized network spanning over 70% of African countries, are natural candidates for collaboration given their existing national research and education networks. Additionally, this proposal builds upon our current work on designing and implementing a distributed systems and network measurement research testbed. Insights from this foundational work will contribute valuable monitoring capabilities to the envisioned Pan-African platform.

Acknowledgment

The authors thank all contributions especially the funding from the Internet Society Foundation in supporting this vision.

References

[1] Addressing hpc accessibility in kenya: the kenet gpu cluster's role in advancing scientific research and collaboration, 2024.

[2] AI4D, SIDA, IDR, and Genesis. Ai in africa: The state and needs of the ecosystem - data.

[3] AI4D, SIDS, IDR, and Genesis. Ai in africa: The state and needs of the ecosystem - compute.

[4] Emmanuel Ogiemwonyi Arakpogun, Ziad Elsahn, Femi Olan, and Farid Elsahn. Artificial intelligence in africa: Challenges and opportunities. *The Fourth Industrial Revolution: Implementation of Artificial Intelligence for Growing Business Success*, 2021.

[5] Ben Asante and Teddy Hayford-Acquah. Fiber cut impacts and the scientific analysis of its financial losses. *IOSR J. Electron. Commun. Eng.*, 16:34–51, 2021.

[6] Timothy King Avordeh, Adam Salifu, Christopher Quaidoo, and Rockson Opare-Boateng. Impact of power outages: Unveiling their influence on micro, small, and medium-sized enterprises and poverty in sub-saharan africa-an in-depth literature review. *Heliyon*, 2024.

[7] Olusesan Michael Awoleye. Reconfiguring data infrastructure ecosystem in africa: A primer toward digital sovereignty. *ArXiv*, abs/2109.14186, 2021.

[8] Ilya Baldin, Anita Nikolich, Jim Griffioen, Inder Monga, Kuang-Ching Wang, Tom Lehman, Paul Ruth, and Ewa Deelman. Fabric: A national-scale programmable experimental network infrastructure. *IEEE Internet Computing*, 23:38–47, 2019.

[9] World Bank. Digital transformation drives development in africa.

[10] Leandro M. Bertholdo, Ralph Holtz, Roland van Rijswijk-Deij, Lisandro Granville†‡, and Cristian Hesselman. An anycast federation for dns resolvers, 2023.

[11] Raphael Bolze, Franck Cappello, Eddy Caron, Michel J. Daydé, Frédéric Desprez, Emmanuel Jeannot, Yvon Jégou, Stéphane Lanteri, Julien Leduc, Nouredine Melab, Guillaume Mornet, Raymond Namyst, Pascale Vicat-Blanc Primet, Benjamin Quétier, Olivier Richard, El-Ghazali Talbi, and Iréa Touche. Grid'5000: A large scale and highly reconfigurable experimental grid testbed. *The International Journal of High Performance Computing Applications*, 20:481 – 494, 2006.

[12] International Telecommunications Union Broadband Commission. The state of broadband: Leveraging ai for universal connectivity.

[13] Josiah Chavula, Nick Feamster, Antoine B. Bagula, and Hussein Suleman. Quantifying the effects of circuitous routes on the latency of intra-africa internet traffic: A study of research and education networks. In *International Conference on e-Infrastructure and e-Services for Developing Countries*, 2014.

[14] Brent N. Chun, David E. Culler, Timothy Roscoe, Andy C. Bavier, Larry L. Peterson, Mike Wawrzoniak, and Mic Bowman. Planetlab: an overlay testbed for broad-coverage services. *Comput. Commun. Rev.*, 33:3–12, 2003.

[15] Matthew A Cole, Robert JR Elliott, Giovanni Occhiali, and Eric Strobl. Power outages and firm performance in sub-saharan africa. *Journal of Development Economics*, 134:150–159, 2018.

[16] Dmitry Duplyakin, Robert Ricci, Aleksander Maricq, Gary Wong, Jonathon Duerig, Eric Eide, Leigh B. Stoller, Mike Hibler, David Johnson, Kirk Webb, Aditya Akella, Kuang-Ching Wang, Glenn Ricart, Lawrence H. Landweber, Chip Elliott, Michael Zink, Emmanuel Cecchet, Snigdhaswin Kar, and Prabodh Mishra. The design and operation of cloudlab. In *USENIX Annual Technical Conference*, 2019.

[17] Sarah Edwards, Xuan Liu, and Niky Riga. Creating repeatable computer science and networking experiments on shared, public testbeds. *ACM SIGOPS Oper. Syst. Rev.*, 49:90–99, 2015.

[18] Michael Foley. The role and status of national research and education networks (nrens) in africa. 2016.

[19] Google. Bridging the gap: Ai's potential in africa's data desert. *The Atlantic*, 2023.

[20] IndiaAi. https://archive.ics.uci.edu/about.

[21] Oxford Insights. Government ai readiness index 2024.

[22] UC Irvine. Uci machine learning repository.

[23] Sountongnoma Martial Anicet KIEMDE and Ahmed Dooguy Kora. The challenges facing the development of ai in africa. *2020 IEEE International Conference on Advent Trends in Multidisciplinary Research and Innovation (ICATMRI)*, pages 1–6, 2020.

[24] Sountongnoma Martial Anicet KIEMDE and Ahmed Dooguy Kora. The challenges facing the development of ai in africa. *2020 IEEE International Conference on Advent Trends in Multidisciplinary Research and Innovation (ICATMRI)*, pages 1–6, 2020.

[25] Will Knight. Ai's smarts now come with a big price tag, 2021.

[26] Mikael Lindvall, Ioana Rus, Paolo Donzelli, Atif M. Memon, Marvin V. Zelkowitz, Aysu Betin Can, Tevfik Bultan, Christopher Ackermann, Bettina Anders, Sima Asgari, Victor R. Basili, Lorin Hochstein, Jörg Fellmann, Forrest Shull, Roseanne Tesoriero Tvedt, Daniel Pech, and Daniel Hirschbach. Experimenting with software testbeds for evaluating new technologies. *Empirical Software Engineering*, 12:417–444, 2007.

[27] Joe Mambretti, Jim Hao Chen, and Fei Yeh. Next generation clouds, the chameleon cloud testbed, and software defined networking (sdn). *2015 International Conference on Cloud Computing Research and Innovation (ICCCRI)*, pages 73–79, 2015.

[28] Fabrício M. Mazzola, Pedro de B. Marcos, Ignacio Castro, Matthew J. Luckie, and Marinho P. Barcellos. On the latency impact of remote peering. In *Passive and Active Network Measurement Conference*, 2022.

[29] Ibomoiye Domor Mienye, Yanxia Sun, and Emmanuel Ileberi. Artificial intelligence and sustainable development in africa: A comprehensive review. *Machine Learning with Applications*, 2024.

[30] Rose M. Mutiso. Ai in africa: Basics over buzz. *Science*, 383 6690:eado8276, 2024.

[31] Ernest Mwebaze and Godliver Owomugisha. Machine learning for plant disease incidence and severity measurements from leaf images. *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 158–163, 2016.

[32] Lucas Nussbaum. Testbeds in computer science. 2017.

[33] Ayomide Owoyemi, Joshua Owoyemi, Adenekan Osiyemi, and Andy Boyd. Artificial intelligence for healthcare in africa. *Frontiers in Digital Health*, 2, 2020.

[34] The AC3 Project. The africaconnect3 (ac3) project connectivity map 2024, 2024.

[35] Brandon Schlinker, Todd Arnold, Ítalo F. S. Cunha, and Ethan Katz-Bassett. Peering: virtualizing bgp at the edge for research. *Proceedings of the 15th International Conference on Emerging Networking Experiments And Technologies*, 2019.

[36] Berat Can Senel, Maxime Mouchet, Justin Cappos, Olivier Fourmaux, Timur Friedman, and Rick McGeer. Edgenet: A multi-tenant and multi-provider edge cloud. *Proceedings of the 4th International Workshop on Edge Systems, Analytics and Networking*, 2021.

[37] Larry Smarr, Camille Crittenden, Thomas A. DeFanti, John Graham, Dmitry Mishin, Richard Moore, Philip M. Papadopoulos, and Frank Würthwein. The pacific research platform: Making high-speed networking a reality for the scientist. *Proceedings of the Practice and Experience on Advanced Research Computing*, 2018.

[38] Folashadé Soulé. Digital sovereignty in africa: Moving beyond local data ownership. *Centre for International and Governance Innovation*, 2024.

[39] Abdullahi Tsanni. What africa needs to do to become a major ai player, 2024.

[40] Tyler Venske. Navigating digital sovereignty in africa: A review of key challenges and constraints. Technical report, Good Governance Africa, 2023.

[41] John Walubengo. Africa left behind: Lack of local data hinders ai effectiveness.

[42] Danni Yu. The 'ai divide' between the global north and the global south.