Infrastructure Architecture Map

Document: High-level diagram and description of the "World Bank's" IT infrastructure (on-prem, cloud, hybrid model)

- **Purpose:** To provide a clear, high-level visual representation and accompanying narrative describing the overall IT infrastructure architecture of the World Bank. This map illustrates how on-premises data centers, private cloud environments, and public cloud services are structured and interconnected to support the Bank's global operations.
- **Strategic Approach and Hybrid Model:** The World Bank operates a complex, global IT environment. Its infrastructure strategy embraces a **hybrid cloud model**. This approach is common for large financial institutions that must balance the need to modernize and innovate with the requirement to maintain control over critical legacy systems and highly sensitive data. The hybrid model allows WDB to:
    - Leverage public cloud services (IaaS, PaaS, SaaS) for agility, scalability, cost-efficiency, and access to advanced technologies (e.g., AI/ML, big data analytics).
    - Utilize private cloud environments for workloads requiring greater control, specific configurations, or dedicated resources, often serving as a stepping stone for cloud adoption or for hosting internal development and testing platforms.
    - Maintain robust on-premises data centers for core legacy systems (e.g., mainframe applications, certain critical transaction processing systems), systems with unique hardware dependencies, or data that is subject to stringent residency or security constraints that are perceived to be best met on-prem.

Key Architectural Components and Zones

This balancing act between **legacy and modernity** is a key characteristic of WDB's infrastructure. While new services and applications are often designed with a **cloud-first** or cloud-native approach, significant investment exists in on-premises infrastructure that continues to be vital for core operations. The architecture map must clearly depict this co-existence and the strategies for managing integration and security across these diverse environments.

The WDB's infrastructure is organized into several key architectural zones:

1. **On-Premises Data Centers:**

   - **Locations:** WDB operates multiple geographically dispersed data centers. These include a primary data center at its Headquarters (e.g., Washington D.C. or London), and one or more regional data centers that may also serve as disaster recovery (DR) sites for critical systems.
   - **Systems Hosted:** These data centers typically host core banking systems, mainframe applications (if applicable), enterprise resource planning (ERP) systems, critical financial databases, and other systems requiring high levels of physical security and control.

2. **Private Cloud Environment:**

   - **Technology:** WDB maintains a private cloud infrastructure, likely based on virtualization technologies such as VMware vSphere or an OpenStack deployment.
   - **Services Hosted:** This environment is used for hosting internal applications, development and testing environments, virtual desktop infrastructure (VDI), and potentially some production workloads that benefit from cloud-like agility but require greater control than public cloud offerings.

3. **Public Cloud Presence (Multi-Cloud Strategy):**

   - **Providers:** WDB likely employs a multi-cloud strategy, utilizing services from major providers such as Amazon Web Services (AWS), Microsoft Azure, and Google Cloud Platform (GCP) to avoid vendor lock-in and leverage best-of-breed services.
   - **Regions:** Services are deployed in multiple geographic regions to support global operations, provide high availability, and meet data residency requirements.
   - **Services Used:** A wide range of services are consumed, including:
     - **IaaS:** Virtual machines, storage, networking for scalable compute capacity.
     - **PaaS:** Database services, application development platforms, container orchestration (e.g., Kubernetes), serverless computing for building and deploying modern applications.
     - **SaaS:** Specific business applications such as Customer Relationship Management (CRM), Human Resources Management (HRM), collaboration suites (e.g., Microsoft 365), and security services.
     - **Data Analytics Platforms:** Cloud-based data warehouses, data lakes, and AI/ML services for economic research and business intelligence.

4. **Edge Computing:**

   - **Connectivity and Integration:** Secure and reliable connectivity between these diverse environments is paramount:

    - **Wide Area Network (WAN):** High-speed, dedicated WAN links (e.g., MPLS) connect major on-premises data centers and large regional offices.
    - **Cloud Connectivity:** Secure connections to public cloud providers are established via dedicated interconnects (e.g., AWS Direct Connect, Azure ExpressRoute) or secure VPNs over the internet.
    - **Virtual Private Networks (VPNs):** Used for secure remote access for employees, site-to-site connections between smaller offices and data centers, and inter-cloud connectivity between different public cloud environments.
    - **API Gateways:** Centralized management of Application Programming Interfaces (APIs) facilitates secure and controlled data exchange and service integration between internal systems, cloud services, and potentially with external partners.
    - **Identity and Access Management (IAM):** A federated IAM system aims to provide consistent identity and access control across on-premises and cloud environments.
- **High-Level Infrastructure Diagram:** A conceptual diagram (inspired by blueprinting principles) would accompany this description. This diagram would not be a detailed network schematic but would visually represent:
    - The major zones: HQ Data Center, Regional DR Data Center, Private Cloud, AWS Cloud Region(s), Azure Cloud Region(s), GCP Cloud Region(s).
    - Key types of workloads or services typically hosted in each zone.
    - Major connectivity paths between these zones (WAN links, Cloud Interconnects, VPNs).
    - The placement of key shared services like central IAM, DNS, and security monitoring.
- **Security Considerations:** The architecture incorporates security at multiple layers:
    - **Perimeter Security:** For on-premises data centers (firewalls, IDS/IPS).
    - **Cloud-Native Security:** Leveraging security tools and services provided by cloud providers (e.g., security groups, network ACLs, cloud WAFs, key management services).
    - **Consistent Security Policies:** Striving to apply consistent security policies and standards across the hybrid environment, managed through centralized security governance.
    - **Data Security:** Encryption of data at rest and in transit, data loss prevention (DLP) technologies.

- **Security Monitoring:** Centralized logging and security information and event management (SIEM) to provide visibility across the hybrid landscape.