ForgeRock®

WHITEPAPER

Containerized Directory Services

Table of Contents

Introduction

Digital transformation has put traditional data directory models at risk. The way directories were designed and deployed in the past, combined with the exponential growth of data in the modern era, have slowed the retrieval of identity- related information. This puts an organization's competitiveness, security and growth at risk.

Organizations need to modernize to stay competitive, and part of that modernization comes in the form of deploying a next-generation directory service in a "containerized" deployment model.

Containers are modular, easy-to-deploy software that can be set up in minutes. They require

only a light IT touch. They provide resilience, security, and scalability that go beyond what many organizations have in place today. And, they position organizations to better take advantage of opportunities in the digital future.

What Is a “Legacy” Directory Service?

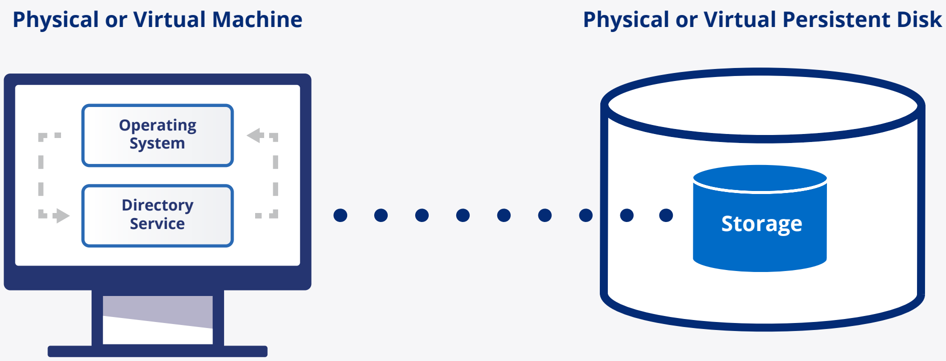
Let’s start with a definition. A directory service is a customizable information store that functions as a single point of truth for the organization. It’s where users and their identity-related information reside. Its main value is to point resources to the directory in order to authenticate and grant access to users.

Information technology is constantly evolving and so are directory services. Many organizations have deployed their directories on physical servers or virtual machines. They were often deployed with static IP addresses and tied to a particular physical/virtual server within a data center. Storage was to a local or shared disk in the data center.

If the directory was to be copied or replicated, a similar setup in another data center would have to be established (on a physical or virtual machine) and the data had to be copied to other local or shared disks. The setup and replication was and continues to be a manual process requiring a database administrator with specialized knowledge of the environments and technology.

This was adequate for the technology of the time, but, ultimately, it leads to a lack of consistency across the organization. Every directory is managed just a little differently, and silos of arcane knowledge have to be maintained on how to service, replicate, and monitor every instance. Deploying this way is also resource-intensive for both the IT personnel and the infrastructure needed to run it. Many organizations, through mergers or organic growth, find themselves running multi-cloud (AWS, Google Cloud, Azure), and, therefore, find they need to maintain cloud-specific support for each environment.

The Traditional “Legacy” Directory Deployment Model


Physical or Virtual Machine
Physical or Virtual Persistent Disk
Operating System
Directory Service
Storage

Traditional directories are deployed on physical and virtual machines with attached storage. This can lead to high IT overhead and maintenance. It also leads to inconsistency in how directory services are managed among different data centers.

The Challenges of Modernization to the Legacy Directory

Data stores were originally designed and built for an entirely different purpose. Telecommunication organizations were the first to use large-scale directories to store the identity data of their base of phone customers. User data usually consisted of just a phone number and username. Retrieval of this information when it was needed – even if the directory was replicated across global locations – could be done with adequate speed so that there would be little to no latency impact on the applications or the users requesting them.

This quickly became the model for directory services in other industries. Enterprises copied this paradigm and deployed directories in a similar fashion to support the growing needs of their internal workforce users. Their directories were populated with email addresses, instead of phone numbers. Usernames and passwords were used to enforce access

to the network and applications. Other identity information soon followed, such as employee number, department, current manager, title, and other parameters. Even with the schema changes to the directory to support this additional information, speed and retrieval remained within acceptable bounds.

Since disaster recovery and business continuity are necessary components of organizations' enterprise risk strategy, this distributed directory model worked well, allowing data to be replicated across locations in near real time so that a user could quickly authenticate to any network or application against a single "source of truth." Organizations thrived and grew, and the directory became an indispensable part of the organization's identity and access management (IAM) infrastructure.

Enter the Era of Cloud and Big Data

But things changed. First, the cloud came into play. The existing fiber and telecommunication links that organizations used to interconnect global networks gave way to more efficient, larger capacity, "bursty" internet connections. Applications moved to the cloud. Users began to work from everywhere. Bring-your-own-device (BYOD) became the de facto standard. Suddenly, everything relied on internet connectivity.

Then came big data. Where the directory of the past only contained a few user identity-related items, that same directory now had to contend with an onslaught of new business-required data that needed to be stored and retrieved at lightning speed. This included user preferences, user personalization data, application data, Internet of Things (IoT) data, entitlements, enrollments, and more. Things got complex fast.

The Impact of 5G and IoT on Directory Services

5G is here. The broadband wireless network promises to connect everything to everything else – everything from traffic sensors, to connected cars, to smart grid technology is coming to your network! This will have a transformative effect on the way we interact with the world.

The Impact: The real-world impact of this on directory services is that any interaction between your users and these smart connected devices requires lightning-fast retrieval to authenticate across a spectrum of global networks. That means sub-second lookup among billions of entries. This won't be a luxury. Rather, it will be a business requirement for your customers and workforce users to interact effectively with your organization.

It is estimated that by 2025, there will be 152,200 connected IoT devices every minute!*

**IDC Data*

As information scales and business requirements grow, what's at stake is nothing less than the performance of your data.

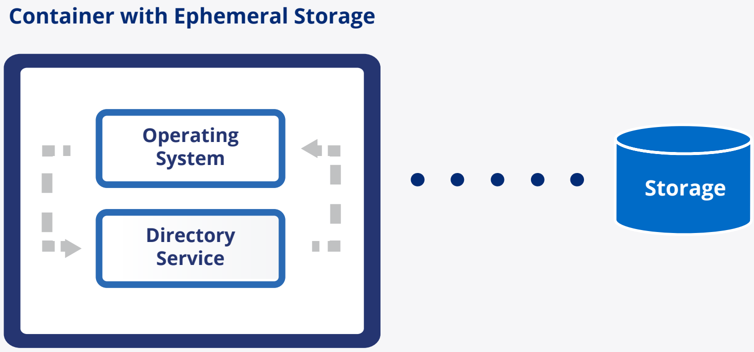
Timely access to the applications and systems your users require is in jeopardy. Latency within and across data centers in authenticating users is already causing disruptions and slowdowns in some organizations. Users can become unproductive, and customers drop out of a digital engagement when they can't immediately get logged on.

Containerization: Enabling a New Directory Services Deployment Paradigm

Containers are a well-tested IT deployment model. Kubernetes and Docker are just a few of the containerization technologies that have been around for years. Millions of business-critical applications are running in containers in data centers across the world.

Containers offer many advantages. They are simple to deploy – just point the compatible software image to the container pod you want to deploy to, and the service can be up and running in a matter of minutes. Containers also offer a consistent approach to deploying, since you don't need to worry about the underlying operating system, machine, or even what cloud to deploy on. And, they offer a level of automation with easier replication and auto-scale than almost any other technology available today. Taken together, containerization is a preferred way to deploy critical services. Containers also use fewer IT resources and run with greater efficiency in the data center.

Containerized Deployment with Storage



Deploying directory services with storage presents a problem: if you collapse your container, you lose your data. ForgeRock uses persistent volumes to help solve this problem.

What's the challenge? Until 2018, Kubernetes did not officially support "stateful sets" of data. If you deployed your directory services within a container pod, you risked potential data loss.

Containers are typically used to run web and microservices that operate in a stateless (no data collection, storage, or retrieval) mode. They can be brought up and collapsed as needed. This flexibility is what makes IT staff favor containers. But it also means that if you want to use containers for directory services, you need a way to keep your data safe. You need to keep the data safe or end up with unrecoverable volumes.

ForgeRock Directory Services with Containerization

ForgeRock Directory Services[1] offers a way to deploy directories using containers that incorporate all the advantages of resiliency, scalability, and flexibility combined with the ability to store and protect persistent data volumes in your environment.

The ForgeRock approach is unique in the industry – we are one of the few vendors that have specifically designed our Directory Services solution to run using containers. In addition to getting all of the great features of a next-generation high-performance, high-availability, internet-scale identity store, you get a solution that you can deploy on your terms, and that works well both within and across your data centers. It's the best of both worlds.

How Do We Do This?

ForgeRock includes a sample DockerFile in the product, so you can get started quickly out of the box (or after downloading).

We provide a Cloud Developer Kit (CDK) that allows you to automate deployment on Kubernetes, Skaffold, and Kustomize. The advantage of the ForgeRock approach is you can choose how and where you deploy the solution: in the cloud, on premises, or or in a hybrid environment. You can use any supported container technology – the solution will run in exactly the same way.

Storage that Ensures Maximum Data Uptime

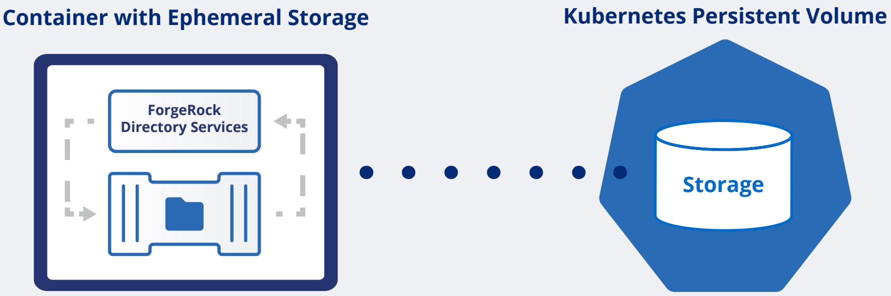
Directory services run in a stateful mode, meaning they generate data sets that need to be preserved. We take these data for your stateless applications and store them in persistent volumes. This way, our Directory Services operate like bare metal or virtual machines (VMs), relying on an external disk or persistent volumes for storage.

Supported persistent volumes include:
> SSDs
> Mounted volumes
> Storage arrays

Not supported
> File shares

The advantage of this approach is that if you need to bring down your Kubernetes environment for any reason, the integrity of the data is preserved. You can create another directory services containerized instance and simply attach it to the named persistent volume. And, given the automated nature of replication, you don't have to worry about manual processes like promoting a replica. Our directory services technology makes sure the data is synchronized across all your volumes.

Containerized Deployment with Storage


Container with Ephemeral Storage
Kubernetes Persistent Volume
ForgeRock Directory Services
Storage

By using persistent volumes with Kubernetes, the underlying container can be collapsed and rebuilt without any impact on the stored data.

[1] Supported from release 7.0.

Security Built In

Organizations need continuous protection against malicious attacks and potential breaches, both from insider threats and the external world. ForgeRock Directory Services provides the necessary encryption to protect the confidentiality and integrity of data at rest and in transit.

Fine-grained access control to your data is an essential. In a traditional database, users can often see everything in a given table, unknowingly exposing sensitive information to potential misuse or privacy implications. ForgeRock restricts viewing only to the specific attribute in the directory a user is authorized to see. We can also restrict privileges to read only, read-write, or write only. We enable personally identifiable information (PII) data masking from logs, so even when data is used for other purposes, like reporting, confidentiality is still protected. We also support High Availability (HA) Proxy Protocol which gives visibility to requesting clients located on the other side of a proxy so that access decisions to important data can be properly made.

Password storage is an area of particular concern for many organizations. Best practices dictate that passwords should not be stored in cleartext but instead be hashed and salted. ForgeRock does this, and we also give the password field

a special computational hash treatment using BCrypt and PBKDF2, among other techniques. We also support the strongest hashing algorithms, like Argon2, to protect again the most sophisticated password cracking attacks. And, as with fine-grained access, a user can reach the password attribute within an authorized directory but only for his or her unique access. Data security cannot be taken for granted. Since multiple applications access the same data directory, having these tools available and in place is an essential requirement for your security architecture.

Automation, Resiliency and Cloud-Scale

With ForgeRock Directory Services, when business requirements dictate that another copy of the directory is needed somewhere (examples: placing the resource closer to users, a seasonal burst in traffic, or disaster recovery), another node can be added quickly. When another copy is no longer needed, it can be removed just as easily.

Automation

The real value of containers lies in automation. Changes made to a configuration can be automatically pushed out to other nodes and clusters. This greatly simplifies the operational aspects of managing your directory services.

Scaling up

Adding a node increases the READ capability of the directory. This is important because, with identity-related information – such as verifying user credentials during login – increasing the READ capacity can greatly speed up authentication. It gets users online faster and gets them the information they need quicker. However, scaling up does not increase the WRITE capacity because it requires data to be copied to the new persistent volume associated with the new node.

Scaling down

When a particular node is no longer needed, it can be collapsed quickly and easily, as if it never existed. The ForgeRock Directory Services software automatically cleans up the remaining servers. You don’t need to repoint replication away from the node that is no longer needed, as it happens automatically. As a matter of good database hygiene, you should delete your persistent volume storage in order to avoid storage and other fees you may incur from your cloud service provider.

Performance

Containers tend to run faster than VMs. They tend to utilize server resources more efficiently. Monitoring container pods and adding new containers or rebuilding existing ones to better support utilization and availability can be done much faster than with VMs. ForgeRock’s unique big data search capabilities utilize indexes and optimized read/write I/O of the disk to return query results in record time.

No service interruption

ForgeRock Directory Services with containers do not require your data to be offline for any period of time. Rolling upgrades are supported. You do not need to go through the cumbersome process of promoting a replica to ensure continuity of service during an outage or planned upgrade.

Conclusion

Organizations need to modernize in order to stay competitive. As the volume of data in legacy directories grows, the ability to deliver business-critical services to users degrades. This directly impacts an organization's employees and customers, and puts them at a disadvantage when they try to make the most of new opportunities.

ForgeRock Containerized Directory Services prepare organizations for the highly connected, ultra-fast business requirements of the future. With automation, directory services work the same way across all your environments and require fewer IT resources. With scalability, new containerized services can be spun up when needed and deployed where they can make the most difference. And for security, fine-grained access ensures that users get what they need and no more. All this ensures that your users are not impacted by latency and slow access, no matter how many applications you have running or where in the world your users happen to be.

About ForgeRock

Follow Us

ForgeRock, the leader in digital identity, delivers modern and comprehensive Identity and Access Management solutions for consumers, employees and things to simply and safely access the connected world. Using ForgeRock, more than a thousand global customer organizations orchestrate, manage, and secure the complete lifecycle of identities from dynamic access controls, governance, APIs, and storing authoritative data – consumable in any cloud or hybrid environment. The company is privately held, and headquartered in San Francisco, California, with offices around the world. For more information and free downloads, visit www.forgerock.com or follow ForgeRock on social media.

ForgeRock