Deploying Liferay Digital Experience Platform in Microsoft Azure

Table of Contents

Introduction

Liferay Digital Experience Platform (DXP) can be deployed into a variety of infrastructures and cloud-based environments. The Liferay Engineering and Global Services teams have extensive experience deploying and managing infrastructure in cloud environments, including Microsoft Azure. Their accumulated experience and best practices are described in this comprehensive reference guide for Liferay Digital Experience Platform (DXP).

This document provides an initial environment for Azure deployments and can be altered depending on your specific requirements for fault tolerance, scalability and other infrastructure and non-functional requirements. The reference architecture is based on the base reference architecture as described in the Deployment Checklist with Azure-specific technologies applied.

For additional hands on support, the Liferay Global Services team also has a specialized Go Live package that can help with your pre-production tuning and configuration.

Reference Architecture

The selection of an appropriate architecture is one of the first decisions in your deployment path. To select an appropriate architecture, you must consider:

- **Information Security:** securing the hardware and sensitive information against malicious attacks and intrusions
- **Performance:** supporting the desired number of total users, concurrent transactions, etc. to a level that matches your requirements
- **Fault Tolerance:** maintaining uptime during unexpected failure or scheduled maintenance
- **Flexibility and Scalability:** designing an expandable architecture to support additional features and users without significant redesign

Overview

The reference architecture depicted in Figure 1 appears complex, but it provides high levels of fault tolerance and flexibility.

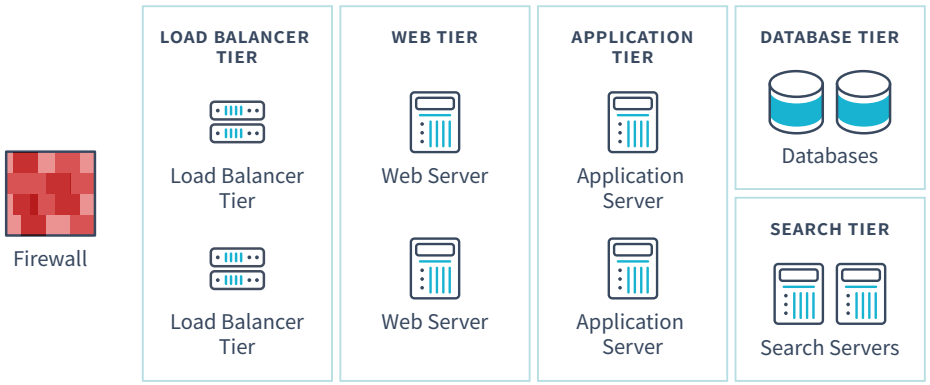

Figure 1 - Liferay DXP Reference Architecture

The architecture contains the following tiers:

- **Firewall**
  - Provides intrusion detection and prevention
- **Load Balancer Tier**
  - Ensures smooth distribution of load between multiple web server resources (and underlying application servers)
- **Web Server Tier**
  - Delivers static content elements like images, rich media, CSS files, etc.
  - Provides integration modules to single sign on solutions like CA Netegrity, Oracle Identity, etc.
  - Handles custom routing needs
- **Application Tier**
  - Hosts Liferay DXP on supported application servers like Tomcat, JBoss, GlassFish, Oracle AS, Oracle/BEA Weblogic and IBM Websphere*
- **Database Tier**
  - Hosts Liferay supported database servers like MySQL, Oracle, MS SQL, IBM DB2 and PostgreSQL*

- **Search**
  - Hosts Liferay supported search servers like Elasticsearch or Solr*

Please see Liferay DXP Compatibility Matrix for complete list of supported application servers)

Sizing

The hardware deployed within each tier varies depending on the type of transactions. We will use Liferay Engineering's benchmarking environment as a hardware specification guide:

- **Firewall**

  Azure Network Security Groups

- **Load Balancer Tier**

  Azure Load Balancer or Azure Application Gateway

- **Web Tier**

  Primarily forwards traffic to the underlying application servers. However, it also provides caching, compression and other capabilities.

  2 – Azure VM size F4s v2 (4 core Intel Xeon Platinum 8168, 8 GB memory) *

  Azure CDN

- **Application Tier**

  Represents the workhorse of the architecture.

  2 – Azure VM size F16s v2 (16 core Intel Xeon Platinum 8168, 32 GB memory)*

- **Database Tier**

  2 – Azure Database for MySQL (8 core Intel Xeon E5-2673 v4, Memory Optimized)*

  Azure Files

- **Search Tier**

  3 - Azure VM size E8 v3 (8 core Intel Xeon E5-2673 v4, 64 GB memory)

* Note that these sizes are subject to change by the Azure team. Please verify current sizes in Azure. These sizes were chosen to best match machine sizes specified in the Liferay Deployment Checklist.

azure.microsoft.com/en-us/pricing/details/virtual-machines

azure.microsoft.com/en-us/pricing/details/mysql

Implementation Details

Firewall

Azure Network Security Groups allows access restrictions on inbound and outbound ports on VMs. This may be used to isolate the entire infrastructure from the public internet and isolate layers such as the web tier and application tier according to security requirements.

Load Balancer

Azure Load Balancer allows traffic to be distributed in an even manner between servers. The load balancer utilizes the source ip and port, destination ip and port and protocol type to maintain sticky sessions. The load balancer may be configured to use any combination of these five attributes to determine session stickiness.

Azure also offers the Application Gateway (AAG) that can be used to load balance between application servers. AAG allows for sticky sessions (Session Affinity) using a cookie. It should be noted that this cookie should not overlap with any cookies set by Liferay or the underlying application server (e.g. JSESSIONID) as it will be set and controlled by AAG. AAG offers additional capabilities that may be useful including more advanced routing options. See this blog post for more information.

Web Tier

The web tier in Azure consists of the Azure Content Delivery Network (CDN) and Azure Virtual Machines (VM). The Azure CDN is responsible for delivering static content, while the VMs are deployed with a web server such as Apache or Nginx.

Apache, Nginx or any of the other Web Server implementations listed in the Liferay DXP Compatibility Matrix can be deployed to Azure's Virtual Machines normally. These web servers can be configured to perform the desired work (compression, caching, etc.) and then forward requests to the appropriate application server in the next tier. Liferay DXP should be configured accordingly by disabling the unnecessary servlet filters in Liferay (e.g. gzip compression filter).

The reference architecture prefers to utilize separate VMs for this tier in order to preserve resources from being competed over, but serving the Web Servers on the same VM as the application server is also a viable option. Note that such a strategy will result in the web and application tiers sharing computing resources and affect your system's performance, making it not ideal for all applications and loads.

You may also choose not to deploy separate web servers. However, doing so may put additional load on the JVMs hosting Liferay DXP for operations like gzip.

As with most architectural choices, there are advantages and disadvantages to each approach. Load tests and performance tuning will help to determine the best option for your solution and requirements.

Please see the Liferay DXP Compatibility Matrix for a complete list of supported operating systems which can be used in the Azure VM instances.

Application Tier

This tier consists of Azure VMs with an application server where Liferay DXP is deployed. Azure users can choose from a variety of Java application servers and operating systems when deploying Liferay, such as Windows Server, RHEL and Ubuntu Server. Additionally, the Azure platform offers a range of virtual machine sizes upon which you may host your chosen OS and application server to suit your requirements. Custom images can also be created to help bootstrap deployments and ensure consistency across multiple VMs.

Please see the Liferay DXP Compatibility Matrix for a complete list of supported application servers and operating systems which can be used in the Azure VM instances.

Database Tier

Azure offers managed database solutions utilizing Microsoft SQL Server, MySQL, MariaDB or PostgreSQL. The Azure Database services offer a low barrier of entry for users looking to deploy highly resilient and scalable database tiers.

Those looking to use other database platforms (e.g. Oracle, etc) may rely upon other Azure partners for available machine images. Users can choose from existing Databases supported by Microsoft Azure or deploy their own implementation. You are free to choose any database platform, assuming the platform is supported according to the Liferay Support Matrix.

Please see the Liferay DXP Compatibility Matrix for a complete list of supported database platforms.

Search Tier

For this tier, we need to set up an Elasticsearch (ES) cluster compatible with DXP. It's recommended to use ES 6.5.x with DXP 7.2 and 6.1.x with DXP 7.1 (Fix Pack 42 or newer).

Liferay and Elastic recommend at least a three-node cluster so as to provide resiliency in case of random failures. Due to potential split-brain issues, it is not recommended to use a two-node cluster. For a three-node cluster, Elasticsearch recommends setting the discovery.zen.minimum_master_nodes property to "2". For more information, please consult the node settings documentation on Elasticsearch.

It's recommended to use a similar JDK for running ES as used for running DXP. At minimum, the JDK major version should match and have the same vendor. This requirement stems from the fact that DXP uses the TCP endpoint of ES to communicate and in some edge cases, the JDK versions on each end may play a role in how the communication is handled.

As with the application tier, VM instances of ES should be created in several availability zones (within one region) to increase the level of resiliency against data center outages.

For details on how to install and setup ES, please check the official documents provided by its creators.

Liferay Specific Considerations

File Storage

Liferay utilizes shared disk storage for its document management capabilities. In non-public cloud deployments, customers tend to provision network attached storage (NAS) or SAN drives mounted via NFS or similar technologies. Azure Files provides similar features.

Using the SMB protocol, Azure Files provides multi-VM shared file access as an operating system mount point. Once you have configured your VMs with Azure Files, you may configure Liferay's document repository to use the appropriate mount path.

Search

In Liferay DXP, Liferay ships an embedded Elasticsearch search engine - the Elasticsearch engine runs in the same JVM as Liferay DXP. Although this solution is great for having out-of-the-box search in Liferay, it will not officially be supported by Liferay for production use, only for development. For production usage, Liferay will only support the use of the Elasticsearch search engine running outside of the DXP JVM (i.e. 1 JVM for Liferay DXP and a separate JVM for the Elasticsearch search engine).

Search engines benefit heavily from caching and their JVM memory profiles are substantially different from a JVM focused on serving content and web views (e.g. Liferay JVM). For these reasons, the two applications should always be kept separate in production environments.

The search engine JVM can run on the same Azure instance as the Liferay JVM; however, this will cause the two processes to compete for the same resources. For heavy search usage, Liferay strongly advises deploying not only to a separate process but to a separate Azure instance to provide dedicated CPU capacity.

Liferay continues to support Solr for Liferay DXP, which can be configured in Azure using one or more VM instances. Please see Solr documentation for detailed deployment instructions.

For those who are deploying Liferay DXP, Liferay recommends deploying Elasticsearch, although Solr remains a supported option.

Cloud Architecture Considerations

Security

As mentioned in the implementation section (see Firewall), the basic means of securing the Liferay deployment in Azure are Virtual Networks, security groups and role-based access control (RBAC) for Azure resources. There are, however, additional measures which can be taken if higher security standards are required.

SSL/TLS

The entrypoint into the Azure infrastructure for users may be an Azure Load Balancer. Azure Load Balancer does not offer SSL termination, it simply forwards traffic to be handled by the backend system. SSL should be terminated at the web tier in this case.

Alternatively, the user may enter through the Azure Application Gateway (AAG). This service offers SSL termination and can forward unencrypted traffic to the backend web or application servers. Azure Application Gateway manages the certificates for communication between the browser and Azure Application Gateway. AAG also offers routing based on the URL, allowing for more complex capabilities such as transparently utilizing a CDN for certain resources.

Some security requirements may require encryption on all communications including communication between the load balancer and the web tier or the web tier and the application tier. Azure offers Azure Key Vault to create and manage these certificates.

Data Encryption

Azure offers encryption of the data at rest for many of its data services. This includes, but is not limited to:

- Encrypted VM disks
- SQL Database and Azure Database for MySQL
- Azure Storage (Including Azure Files)

All these services use Azure Key Vault to store the encryption keys.

Please see the latest Azure documentation to see if a particular Azure service offers encryption options.

Autoscaling

Failure scenarios should be planned (e.g. application not responding, loss of server, loss of data center, etc.) and factored into the high availability and fault-tolerance plan for the application.

Autoscaling is a common strategy used in cloud environments to improve fault tolerance and provide dynamic resource allocation. Azure offers Virtual Machine Scale Sets which allows administrators to manage sets of VMs that will automatically start and stop based on various factors including load or failures. Failure scenarios should be planned (e.g. application not responding, loss of server, loss of data center, etc.) and factored into the high availability and fault tolerance plan for the application.

We will not discuss best practices and how to set up autoscaling in Azure. You may consult the appropriate Azure documentation for guidelines. Liferay is fully compatible with autoscaling architectures, assuming you have the appropriate subscription levels and have deployed Liferay Connected Services as part of your Liferay platform.

High Availability and Disaster Recovery

To support highly-available applications and deployments Azure offers multiple, independent replication regions and isolated availability zones within those regions. Azure's load balancing services allow administrators to set up cross-zone clusters that can ensure that the application is resilient against hardware or network interruptions.

Azure's Files storage solution and Managed SQL Server services also support managed geo-replication across regions and provide a higher SLA.

Azure also provides managed and scriptable backup solutions for VMs, such as Azure Backup and Site Recovery.

Automated Setup

A major part of recovering a DXP system from a disaster is the recreation of the entire DXP system, including the underlying system infrastructure and all its applications. This can be done by carefully documenting all steps in the set up of the infrastructure and applications. However, Liferay recommends fully automating the construction of the entire DXP system, as well as the on-going deployment of application development. With little to no manual intervention, disaster recovery is then simply executing the automated infrastructure reconstruction along with deploying the latest application. This approach in building infrastructure automatically is known as "Infrastructure as Code" or "Configuration as Code."

There are tools that can help with automating the building of infrastructure and deploying applications, such as:

- Azure Resource Manager
- Terraform
- Sceptre

It is important to perform all successive updates (i.e. done after initial installation) using the automated tool, to make sure the stack is recovered exactly to the configuration from before the failure.

Backup Schedule and Replication

Backing up data (or content) properly is an important piece in planning for a successful data recovery procedure. DXP has three sets of data which need to be backed up and recovered after a failure:

- Database
- Document library file store
- Elasticsearch index

Please refer to Azure documentation on backups for each of their services. As a best practice, plan on synchronizing the backups to all three sets of data, so that they remain as consistent as possible.

Putting It Altogether

With both the scheduled backups and data recovery procedures in place, it is important to perform a disaster recovery test. The end goal of this test is to reconstruct a fully functional DXP system from the various backups, using the data recovery procedures. Of course, this should be done in an environment closely matching the production environment, but not the actual production environment (unless downtime is acceptable, or if the project has not been debuted). Additionally, commit to testing disaster recovery regularly.

Networking (Liferay Clustering)

When Liferay DXP is deployed in Azure, it will run inside your Virtual Network. The virtualized networking infrastructure is very similar to physical networks. However, it is important to note that multicast is not available in the Azure Virtual Network.

Liferay implements clustering within its Cluster Link feature. By default, Cluster Link uses multicast for discovery and communication between cluster members. In Azure, Cluster Link can be configured to use JDBC for member discovery and TCP for communication between cluster members.

Cloud vs Bare Metal Performance

Special attention should be paid toward vocabulary used when describing Azure VM performance. With bare metal machines, a CPU refers to the physical chip which typically contains multiple cores. Each core may contain one or more hardware threads. Within the OS and the Azure platform, a core refers to a single physical hardware thread.

Azure's D14 v2 instances are deployed with 16 cores from Intel Xeon ES-2673 v3 chips. This means the VM utilizes 16 hardware threads. Each physical ES-2673 v3 chip contains 12 hyper-threaded cores or 24 hardware threads. Thus, D14 v2 instances provide roughly the equivalent CPU performance of 66% of a physical ES-2673 v3 CPU.

Virtualization will also have a certain degree of overhead, compared to bare metal. For instance, while the D14 v2 instance size provides 66% of the threads from a physical CPU, in reality, it offers less than 66% of bare metal performance. You should factor this into your calculation when performing capacity planning.

Summary

In the preceding sections, we outlined the steps and considerations to design a fully fault-tolerant Liferay DXP deployment for the Azure cloud platform. The described architecture builds a solid foundation for future growth.

Disclaimer

Liferay can only give you an initial tuning recommendation based on benchmarks that have been performed on the core Liferay DXP product. It is up to you as system architects and business analysts to come up with the utilization scenarios that your system will need to service the required amount of users.

It is your responsibility to run the appropriate load tests on your system before production deployment, so that you can identify significant bottlenecks due to custom applications/portlets or other unforeseen system and network issues, and implement appropriate configurations.

Please use this document and the Liferay DXP Deployment Checklist as guides in sizing your system and procuring your hardware.

Moving Forward

Liferay DXP Cloud

If your team would rather focus on critical business needs instead of infrastructure maintenance, Liferay DXP Cloud is a Platform as a Service (PaaS) tailored for Liferay DXP that will help you focus on what matters, saving IT resources for the highest business priorities. For more information, visit liferay.com/dxp-cloud or contact sales@liferay.com.

Liferay Global Services

Learn how Liferay's Global Services team can support your Liferay DXP project with a Go Live consultation. Contact sales@liferay.com for more information.

Liferay makes software that helps companies create digital experiences on web, mobile and connected devices. Our platform is open source, which makes it more reliable, innovative and secure. We try to leave a positive mark on the world through business and technology. Hundreds of organizations in financial services, healthcare, government, insurance, retail, manufacturing and multiple other industries use Liferay. Visit us at liferay.com.

191009