Using a Stovepipe Configuration to achieve Content Navigator and Content Platform Engine High Availability

Abstract

Content Platform Engine (CPE) has two transport layers: Web Services and EJB. All client requests enter CPE through one of these two transport layers.

- Web Services

  The listener exposes the full functionality of the CPE server through a standard Web services API and uses a load balancer for high availability.

- EJB

  The Java EE session beans reside in the EJB tier of the application server. All clients of this layer must perform a JAAS login prior to sending a request to one of the EJBs. The recommended approach when using the EJB transport for high availability is to use application server clusters. A hardware load balancer cannot be used between application servers that are using the EJB transport because EJB uses long lived sessions and only the first EJB lookup call from an application actually goes through the hardware load balancer. All subsequent EJB calls circumvent the hardware load balancer.

With the EJB transport, the recommended approach to creating highly available ECM applications is by creating clusters of application servers that use technology such as WebSphere ND because this approach provides the simplest and most effective method of balancing workloads, session replication and achieving transparent failover.

If using a server cluster technology such as WebSphere ND is not possible, then high availability can be achieved using a "Stovepipe" configuration in which each server in the application tier is directed to a specific CPE instance.

The purpose of this document is to provide the high-level steps for configuring a stovepipe environment for use by IBM FileNet Content Platform Engine (CPE) and IBM Content Navigator (ICN). This document does not provide details on how to configure hardware load balancers for distributing load or for monitoring the availability of components in the stovepipe configuration.

The high availability approach described above has inherent limitations and drawbacks. To provide a comprehensive and transparent highly available solution, IBM strongly recommends the use of an application server clustering solution such as WebSphere Application Server Network Deployment.

Contents

Basic Stovepipe Configuration

Figure 1 illustrates the basic stovepipe configuration. User traffic goes to hardware load balancers that send work to either of the Content Navigator instances. Each Content Navigator instance is configured to communicate with a specific Content Platform Engine instance. In Figure 1, ICN1 is configured to communicate with CPE1, and ICN2 is configured to communicate with CPE2.

- If ICN1 or CPE1 becomes unavailable, the load balancers must redirect all user traffic to ICN2. Any users of ICN1 will have to start a new session on ICN2, and potentially will have to reprocess any in-progress work.

- If ICN2 or CPE2 becomes unavailable, the load balancers must redirect all user traffic to ICN1. Any users of ICN2 will have to start a new session on ICN1, and potentially will have to reprocess any in-progress work.

Users

Hardware load balancers

ICN Servers
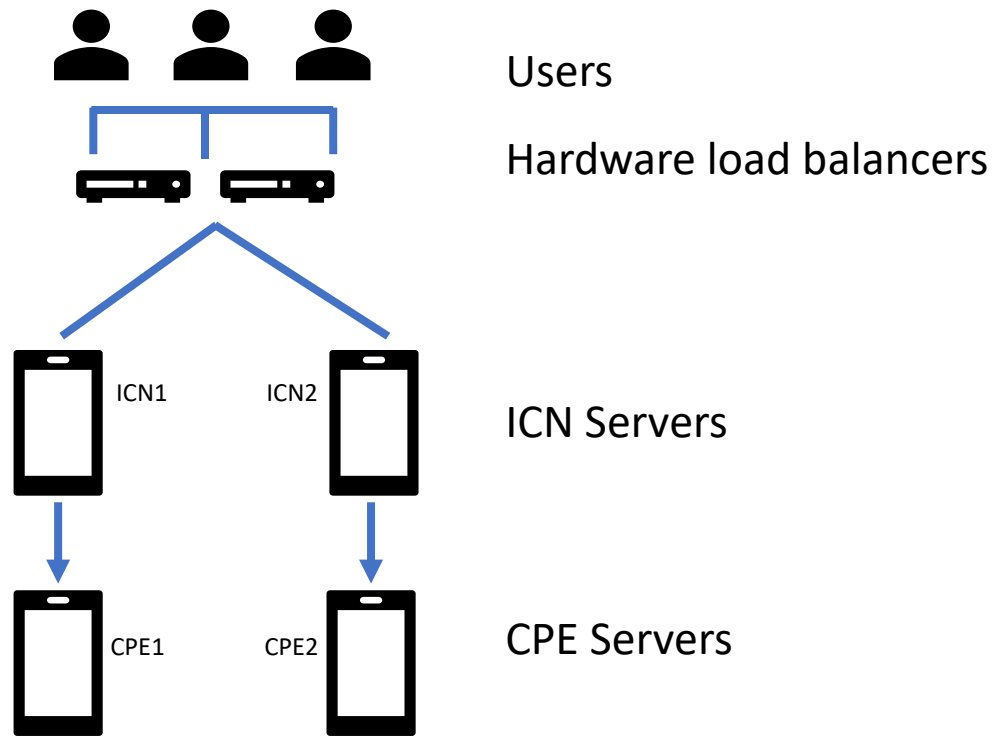
ICN1   ICN2

CPE Servers

CPE1   CPE2

Figure 1: Basic Stovepipe Configuration

Using Non-EJB Applications in a Stovepipe Configuration

Applications that use the .Net API or Web Services to access the CPE servers, including Content Search Services, continue to use a load balancer to connect to all the CPE servers in the environment. Figure 2 illustrates how these components connect to the CPE servers.

The behavior of these applications when a CPE fails is the same in both a traditional HA configuration and when a stovepipe configuration is used.

As a best practice, ensure these applications are not dependent on a specific CPE server being available, either by using ambient sessions or, if the code is running on the CPE server, by using a localhost alias.
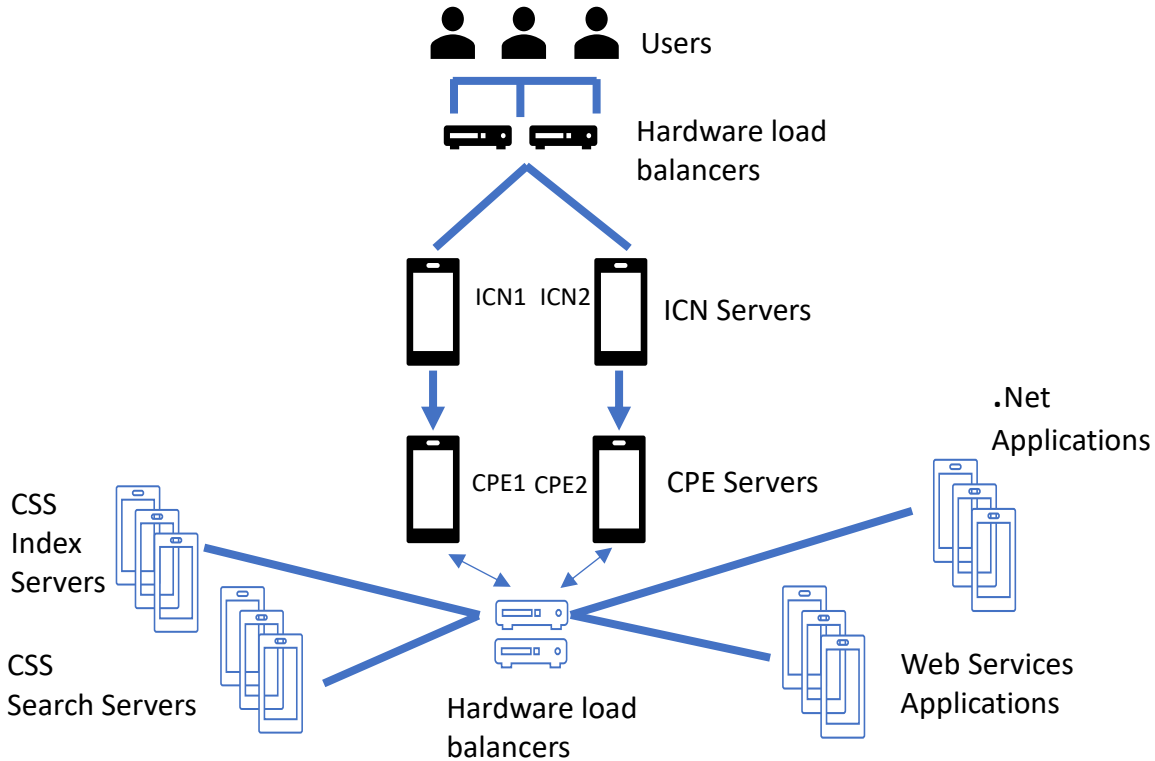
Figure 2: Adding non-EJB Applications

Using Content Search Services (CSS)

Configuring Content Search Services for high-availability is not affected by the architecture used for CPE high availability.

At a minimum deploy two CSS servers running the indexing and search services as a starting point and add additional servers in the future as requirements dictate.

Use a different server name and port for each server. If you use the same server name and port as a previously installed server, the new server installation overwrites the existing server.

Review the information on affinity groups to determine the best configuration of CSS servers for your environment:
https://www.ibm.com/support/knowledgecenter/SSNW2F_5.5.0/com.ibm.p8.ce.admin.tasks.doc/cbr/csscbr_affinitygroups.htm.

Using Workflow in a Stovepipe Configuration

Since connection points are used to identify workflow systems, workflows are always aware of all the CPE servers in the environment irrespective of whether a stovepipe or a traditional HA configuration is used. However, if a CPE fails in a stovepipe configuration, it might take the system a little longer to

recognize that a server has become unavailable and redirect the work to one of the remaining CPE servers.

It is possible to disable workflow running on a specific CPE server, but exercising this option in a stove pipe configuration.

Only a few workflow operations, such as transferring workflows and flushing the cache, that update metadata require inter-server communication. The default server inactive detection time is four minutes and you are strongly recommended not to change this value since the functions such as transferring workflows should, in general, be scheduled for non-production hours.

However, the heartbeat parameters, if needed can be reduced as low as one minute.

To alter the server detection time

1. Log on to ACCE

2. Within an object store, navigate to Administration > Workflow System

3. Click on the Advanced Tab

4. Click on All Servers (See Figure 3)

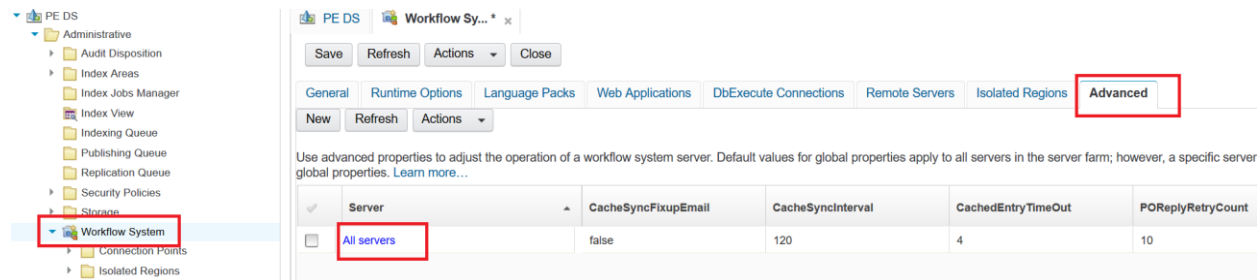   The Edit Server Instance Property dialog displays.

Figure 3

5. On the Edit Server Instance Property Dialog (See Figure 4), review the list of properties. If they are not listed already, add the following properties and assign appropriate values:

   - ServerHeartBeatIntervalInSecods
   - ServerInaciveInterval
   - AllowOverrideFarmMonitorAttributes

   The minimum value that can be set for each of the properties is 60 seconds, but as with any tuning exercise, balance the need for a shorter heartbeat with the impact on CPU and database resource consumption. Be aware that setting the heartbeat too low could result in a server being marked offline as a result of transient issues such as minor variations in clock time or a temporary network slow down, rather than a server really being unavailable.
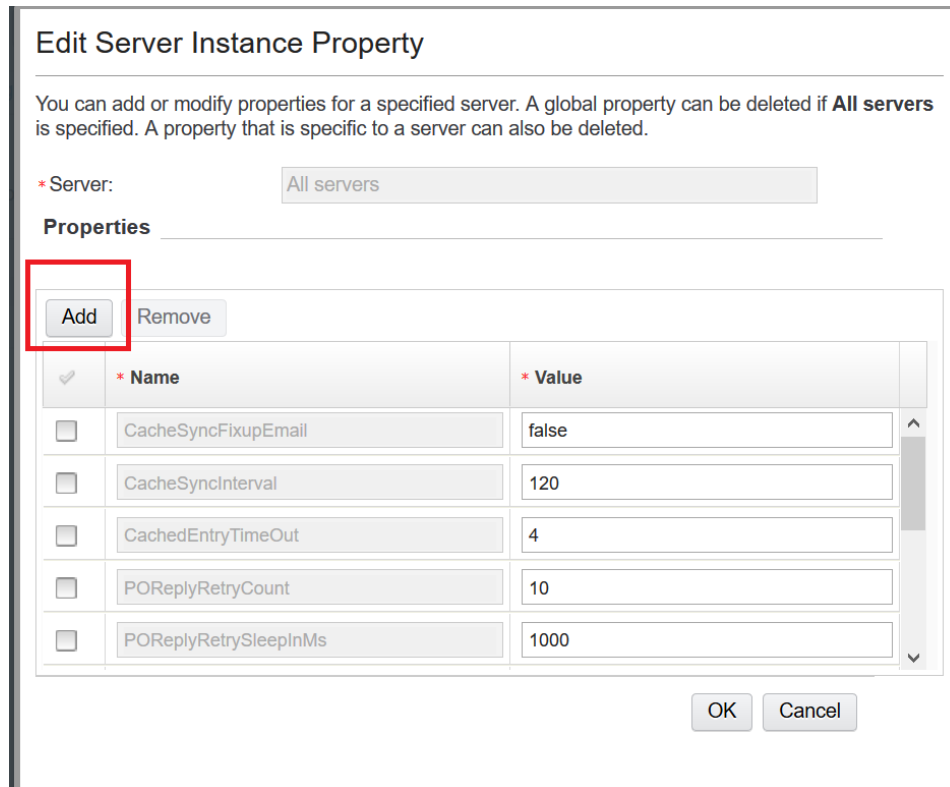
Figure 4 – Updating Process Server Properties

When implementing component steps, ensure the code uses ambient session context and does not need to interact with a specific CPE server. For more information on using component manager, refer to the following tech note: http://www.ibm.com/support/docview.wss?uid=swg27043131.

Using Hosts Files to Identify Servers

The hosts file on each server is used extensively in this stovepipe topology. A generic virtual host name is used for each application in the IBM ECM solution. The virtual host name and virtual IP address resolution are controlled in the "hosts" file on each server.

For example, the mapping from the application tier (ICN) to the tiers below is accomplished via virtual host names that are mapped in the host files on each application server to the physical IP address of the specific CPE server in the tiers below.

Configuring the Load Balancers

The hardware load balancer must be configured with monitors (i.e., availability checks) for the CPE and ICN servers within each stovepipe. If a monitor check fails for a particular server, access to the application instance (ICN) that depends on the stovepipe containing that server must be blocked until the unresponsive server is back online.

IBM ECM deployments typically have a variety of applications that interact with the Content Platform Engine servers. If a Content Platform Engine server becomes unavailable, all application instances

depending on that server will be blocked – that is, the load balancer must not send requests to those application instances.

If an application instance (ICN) becomes unavailable, only that application instance must be blocked from further user requests, while the CPE associated with it can remain online and can continue to process requests that are coming via non-EJB applications.

Requirements

To configure a stovepipe environment, you need

- A hardware load balancer, along with virtual IP addresses and host names configured for all CPE and ICN instances.

  The load balancer must use sticky sessions based on the client source IP address.

  The CPE and ICN components must be configured to use the virtual IP address and virtual host name when connecting. To connect to the CPE and ICN components, the CPE and ICN URLs must use the virtual host name and IP address.

  For example:

  http://<CPE VIP> or <CPE virtual HostName>:port/FileNet/Engine

  http://<CPE VIP> or <CPE virtual HostName>:port/acce

  http://<ICN VIP> or <ICN virtual HostName>:port/navigator

  http://<ICN VIP> or <ICN virtual HostName>:port/navigator/Ping

- Each ICN server must be configured to use the specific CPE server in its respective stovepipe via the host file entries that map the virtual host name to the CPE server that resides in that stovepipe.

  For example, each ICN server host file must contain:

  <CPE server Alias Name>   <CPE server IP address for that particular stovepipe's CPE>

Installation and Configuration

Following is an outline of the steps for installing and configuring the Content Navigator and Content Platform Engine in a stovepipe with WebSphere Application Server:

1. Install WebSphere Application Server following the instructions in the IBM Knowledge Center for WebSphere.

   Ensure that the version of WebSphere Application Server is compatible with the levels of Content Navigator and CPE that you want to use.

   Configure all CPE WebSphere instance profiles across all stovepipes such that they use the same bootstrap port number.

2. Install and configure Content Platform Engine on each of the CPE servers following the instructions provided in the FileNet P8 Platform Knowledge Center.

3. Configure each CPE server with identical configuration settings for Lightweight Directory Access Protocol (LDAP) and JDBC data source settings.

4. Deploy the CPE application on each CPE server.

   All CPEs must be deployed with the same bootstrapped EAR file.

5. Complete the CPE post-deployment steps documented in the Knowledge Center.

6. Navigate to the Administrative Console for Content Platform Engine (ACCE).

   Use the CPE virtual IP or host name and port as the <CPE_server and port> in the ACCE URL.

7. Create the FileNet P8 Domain.

8. Use ACCE to configure one or more Object Stores as needed.

9. Install and configure ICN on each of the ICN servers following the instructions for configuring an IBM FileNet Content Manager Repository in the Content Navigator Knowledge Center.

   Each ICN server in each respective stovepipe must have identical configuration settings for Lightweight Directory Access Protocol (LDAP) and JDBC data source settings.

10. Deploy the IBM Content Navigator application following the instructions in the Content Navigator Knowledge Center.

11. Navigate to the Content Navigator administration desktop.

    Use the ICN virtual IP or hostname and port as the <host_name and port> in the Content Navigator URL.

12. Create and configure one or more IBM FileNet Content Manager repositories.

    The URL of the CPE repository must be constructed using the virtual IP or hostname and port of the CPE server as shown below:

    iiop://< CPE server Alias Name >:<CPE WAS BOOTSTRAP_ADDRESS port>/FileNet/Engine

13. As needed configure additional ICN desktops and teamspaces.

References

- FileNet P8 Platform Knowledge Center: https://www.ibm.com/support/knowledgecenter/SSNW2F

- Content Navigator Knowledge Center: https://www.ibm.com/support/knowledgecenter/SSNW2F

- WebSphere Application Server Knowledge Center: https://www.ibm.com/support/knowledgecenter/SSEQTP/mapfiles/product_welcome_was.html

- Software Product Compatibility Reports:
  https://www.ibm.com/software/reports/compatibility/clarity/index.html