Kofax Equitrac

Print Release High Availability

Version: 6.1.0

Date: 2020-06-23

Table of Contents

Equitrac Print Release High Availability

Overview

This guide details print release high availability for Kofax Equitrac®. This guide highlights the installation and configuration of a highly available printer release environment via Equitrac DCE server.

This guide assumes that you have already created, configured and tested your Network Load Balancer (NLB) appliance. For information on setting up NLB, refer to your NLB vendor documentation.

What is print release high availability?

High Availability (HA) print environments are designed to provide full-time system availability. High availability systems typically have redundant hardware and software that makes the system available in the event of a failure, and help distribute workload. To ensure that HA systems avoid having single points-of-failure, any hardware or software component that can fail has a redundant component of the same type.

When failures occur, the processes performed by the failed component are moved (or failed over) to the redundant or backup component. This process resets system-wide resources, recovers partial or failed transactions, and restores the system to normal as quickly and as seamlessly as possible. A highly available system is almost transparent to the users.

HA for Print Release allows users to release jobs at the MFP even in the event of a failure between the MFP and DCE. The user can still release jobs at an MFP when DCE is unavailable.

Print release HA configuration workflow

To set up highly available print release solution, do the following:

1. Configure a virtual service on your NLB appliance for DCE with a Virtual IP (VIP). See Configure a virtual service.
2. Create a DNS record that resolves to the VIP for the NLB. See Create a DNS record.
3. Add a loopback adapter to each DCE in your HA DCE deployment configured with the same IP Address as the VIP for the NLB. See Add a loopback adapter.

   Note If DRE is installed on the same server as DCE, also see Add the IP address variable.

4. Install DCE on multiple servers in DCE HA mode. See Install DCE in an high availability setup.

Network load balancer

With a load balancing solution, such as Windows Server Network Load Balancing, several VMs are configured identically and the load balancer distributes service requests across each of the VMs fairly evenly. This process reduces the risk of any single VM becoming overloaded. Using load balancing is an effective way to eliminate VM downtime because VMs can be individually rotated and serviced without taking the service offline. However, load balancing only works with identical VMs, which have no shared or centralized data.

Load balancing is an effective way of increasing the availability of critical applications. When server failure instances are detected, they are seamlessly replaced when the traffic is automatically redistributed to servers that are still running. Not only does load balancing lead to high availability it also facilitates incremental scalability. Network load balancing facilitates higher levels of fault tolerance within service applications.

The configuration for the Network Load Balancer (NLB) appliance varies depending on the vendor and type of appliance, and must be managed by the end customer’s internal IT administrator. The generic requirements for each protocol configured on the NLB appliance are outlined.

Load balancing for print release

In a highly available print release environment, the workflow is uninterrupted by having multiple DCEs distributed across different servers connected to a Network Load Balancer (NLB) to distribute the workload.

Network load balancing uses multiple independent servers that have the same setup, but do not work together as in a cluster setup. The NLB forwards requests to either one DCE server or another, but one server does not use the other server’s resources. Also, one resource does not share its state with other resources. In an NLB setup, all resources run at the same time, and a management layer distributes the workload across them. This process reduces the risk of any single server becoming overloaded.

The recommended load balancing method used is Layer 4 in a direct server return (DSR)/N-Path/direct routing configuration. Layer 4 load balancing uses information defined at the networking transport layer as the basis for deciding how to distribute client requests across a group of servers.

It is very important that the source IP Addresses are preserved. That is, the EQ DCE service must see the request originating from the individual MFP IP Address and not the NLB appliance.

Layer 4 load balancing forwards traffic to a specific server based upon the selected port or service. The NLB appliance sits between the MFP and the DCE server and on port 2939 for Ricoh and Lexmark for requests from the MFP, and then decides which server to send them on to.

The configuration for the NLB varies depending on the vendor and type of appliance, and must be managed by the end customer’s internal IT administrator. Refer to your the vendor’s documentation for specific NLB appliance requirements, and your Microsoft documentation for general NLB setup.

Configure a virtual service

A virtual service is required on your NLB appliance for DCE with a Virtual IP (VIP) assigned to it. To configure a virtual service on the NLB appliance, do the following. Consult your NLB appliance vendor for support.

- Configure the Virtual Service IP Address (VIP).
- Set the Ports to 2939 (for Lexmark and Ricoh)
- Set the Protocol to TCP.
- Set the Load balancing Forwarding Method to Direct Routing (i.e. layer 4/direct routing/direct server return/N-Path).Ensure the Persistent checkbox is not selected.
- Set the Check Port for server/service online to 2939.

Create a DNS record

On the DNS server, create a hostname and corresponding “Host (A)” record for the virtual DCE that matches the Virtual IP (VIP) for the NLB. This is needed so that the virtual DCE name resolves to the VIP used on the NLB.

When installing DCE in HA mode, the installer prompts the user to supply a virtual server name. This virtual server name should match the DNS record previously created.

When configuring the MFP devices to connect back to the highly available DCE, the DCE hostname/IP Address should be the VIP for the NLB and not the individual DCE hostname or IP Address.

Add a loopback adapter

Typical Layer 4 NLB deployments require that all servers placed behind a load balanced service have primary and secondary network interface cards (NIC) configured. The primary NIC provides the server with a dedicated, full-time connection to a network. The secondary NIC does not physically connect to the network.

When clients request a service via the NLB appliance, they contact an IP Address/Hostname that is configured on the NLB appliance specifically to listen for requests for that service. This is the Virtual IP (VIP) of the NLB appliance. Since the NLB appliance forwards on these requests directly to the servers offering the service without altering the destination IP Address, the servers themselves must contain at least one NIC assigned with the same IP Address as the VIP. If they do not, then the request from the client is rejected as the servers assume that the request was not intended for them.

It is equally important that the secondary NIC added to each server does not actually connect to the production LAN. This ensures that when any client wishes to connect to the NLB appliance on its VIP, the servers with the secondary NIC also containing the VIP do not respond directly to the clients. This would initiate a direct connection between the client and the server and would avoid sending the traffic via the NLB appliance.

In order to avoid direct client to server connection, the majority of NLB appliance vendors advise to add the secondary NIC as a loopback adapter, as this is a virtual interface that does not physically connect to a network. Refer to your vendor’s documentation for more information.

Add the IP address variable

If you are deploying an HA DCE infrastructure on servers that are also running DRE, additional configuration is required to correct certain undesired network behavior. For standalone DCE servers without DRE running on them, the following IP Address configuration in not required.

If the secondary loopback NIC containing the IP Address of the VIP (for HA DCE) is added to a server that is also running DRE, when a user prints the DRE sends both IP Addresses to CAS. When there are multiple registrations with the same IP Address, users do not see jobs that have been printed to different DREs. For example, if a user prints a job to DRE1, and then another job to DRE2, they only see the job on DRE1. This happens because both DREs have registered the same IP Address (the VIP assigned to the loopback adapter), and it is assumed that since DRE1 has been queried for jobs, DRE2 has also been queried.

In order to correct this behavior, and have DCE only send its production IP Address as part of the service registration message, a system environment variable must be added to each DRE/DCE containing the appropriate IP Address.

1. Go to **Control Panel > All Control Panel Items > System Access** on the DRE/DCE, and select **Advanced system settings**.
2. On the System properties window, click **Environmental Variables**.
3. On the Environment Variables window, click **New** from the **System variables** section.
4. Create the Variable name **EQ_IPADDRESSES** with a Variable value of the production IP Address of your Equitrac DRE/DCE server, and press **OK** and then **OK** again.
5. Repeat these steps above for every DRE/DCE in your deployment.

Configure weak host model commands

Common configuration for network hosts is to be multihomed with multiple network interfaces. A multihomed host provides enhanced connectivity because it can be simultaneously connected to multiple networks, such as an intranet or the Internet. However, since they can be connected to both an intranet and the Internet, services running on multihomed hosts can be vulnerable to attack.

In the weak host model, an IP host (IPv4) can send packets on an interface that is not assigned the source IP address of the packet being sent. This is known as weak host send behavior. An IP host can also receive packets on an interface that is not assigned the destination IP address of the packet being received. This is known as weak host receive behavior.

In order for Dynamic Source Routing (DSR) to work, the weak host model must be enabled on the server’s loopback interface, as well as the interface from which requests are received from.

To configure a multihomed server so the network interfaces can send or receive packets for addresses that they are not assigned, run the following commands from an Administrator command console. Replace the interface name in quotes with the names of your server interfaces.

For the VLAN interface:

```
netsh interface ipv4 set interface "Ethernet 2" weakhostreceive=enabled
```

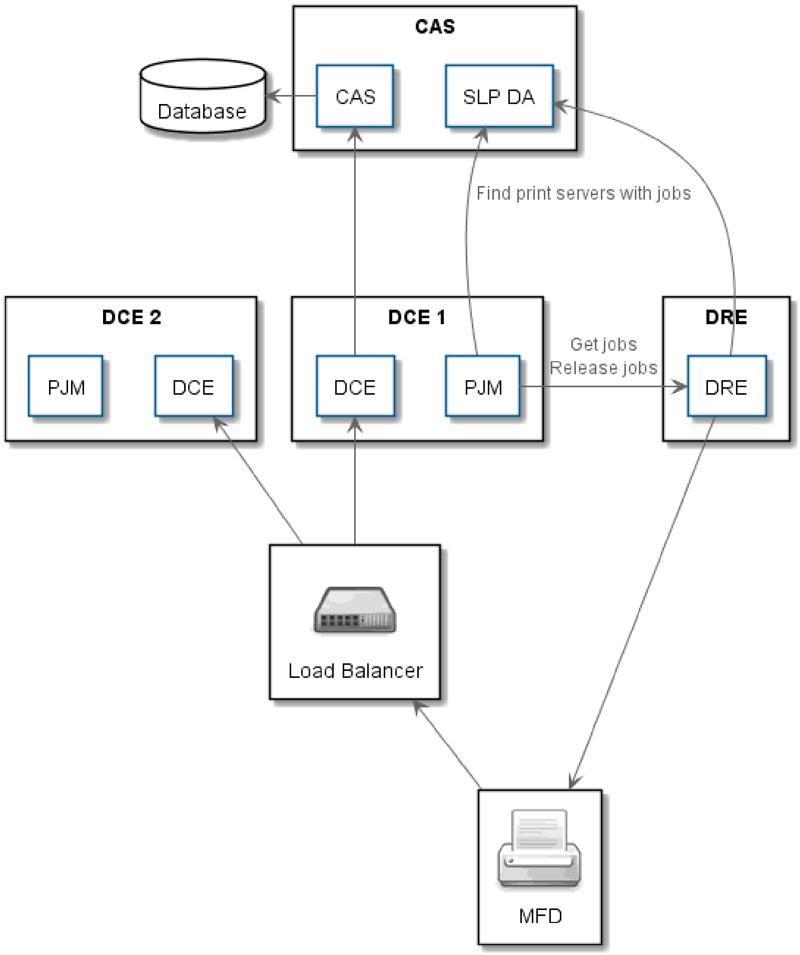
For the Loopback interface:

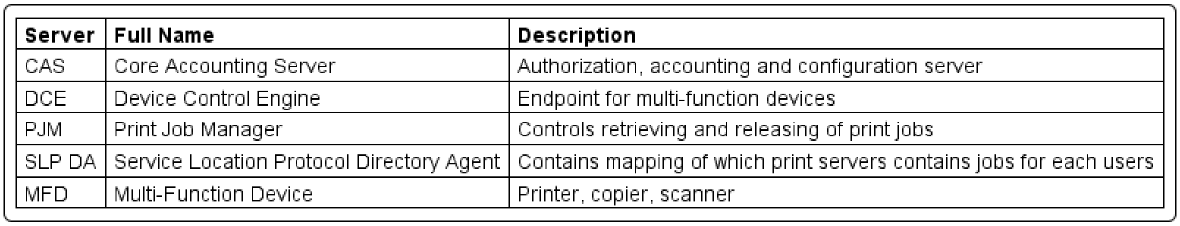
```
netsh interface ipv4 set interface "Ethernnet 3" weakhostreceive=enabled
```

```
netsh interface ipv4 set interface "Ethernet 3" weakhostsend=enabled
```

For detailed description of weak host models refer to the following Microsoft article: https://technet.microsoft.com/en-us/library/ad9db381-1e1b-4077-be1c-bcefb11f1ea8

Print release high availability server deployment





Server	Full Name	Description
CAS	Core Accounting Server	Authorization, accounting and configuration server
DCE	Device Control Engine	Endpoint for multi-function devices
PJM	Print Job Manager	Controls retrieving and releasing of print jobs
SLP DA	Service Location Protocol Directory Agent	Contains mapping of which print servers contains jobs for each users
MFD	Multi-Function Device	Printer, copier, scanner

Install DCE in an high availability setup

You can install multiple DCEs to manage the communication load from release devices.

To install DCE in an HA Setup, do the following:

1. Install ControlSuite with DCE and run the ConfigAssistant.
2. On the **Services** page click the **"…"** button beside **Device Control Engine**, and select **Virtual Site Name** from the list.
3. Enter the **virtual site name** that matches the VIP of the NLB, and click **OK**.
4. On the **Services** page click the **"…"** button beside **Device Control Engine**, and select **Start** from the list.
5. Continue configuring ControlSuite.

Configure certificate pinning

1. Install the first DCE node and configure it.
2. If a self-signed certificate is used (default option), then the DCE's certificate will need to be exported with the private key.
    a. Open the **Windows Certificate Manager** for the Local Machine.
    b. Expand **Personal > Certificates**.
    c. Locate the certificate with the **Friendly name** that matches the **Certificate** name in the Configuration Assistant's Certificate Management.
    d. Right-click the certificate name and select **All Tasks > Export…**
    e. On the Certificate Export Wizard click **Next**.
    f. Select **Yes, export the private key**.
    g. Click **Next** on the Export File Format.
    h. Click the **Password** checkbox and enter a password in the **Password** and **Confirm password** text boxes and press **Next**.
    i. Enter a file name and press **Next**.
    j. Click **Finish**.
3. On the Certificate Management page in the Configuration Assistant, import the certificate from the first DCE node to the remaining DCE nodes.
    a. Select the Equitrac component and select **Import Certificate** from the **Select Action** drop-down list.
    b. Enter the **Filename** and **Password** of the exported DCE certificate, and a friendly name to identify the certificate.
    c. Click **Next**. A Binding Ports window opens updating the certificates. Close the binding ports window when done.

Monitor DCE health in a cluster

After setting up your DCE in a high availability (HA) cluster environment, it is recommended to run a health monitor to verify that DCE is working. If a DCE cluster node fails over, the embedded clients can reconnect to an alternate DCE node in the cluster and continue the user session. In order for this to happen the NLB must first detect that the DCE node is no longer in service. This can be done by using an NLB health monitor for the DCE service.

The DCE service supports a “DCEHealthCheck” URL for active TCP monitoring of the DCE service by the NLB to determine if DCE can respond to the request in a timely manner. The DCE HealthCheck monitor continually pings the DCE nodes on port 2939 and takes the node offline on failure. Once the NLB detects a node failure it stops routing new client connection requests to the failing DCE node. Clients with existing connections to the failing node may have to wait for a connection timeout. An alternative is to configure the NLB to reset existing client connections as soon as the failure is detected, causing the client to request a new connection without waiting for a network timeout.

The following basic configuration settings are required:

- Interval: 15 seconds (how often the monitor sends a request to the DCE node)
- Timeout: 15 seconds (how long the monitor waits for a successful response before taking the failed node offline). These Interval and Timeout values are similar to the DCEs internal timeouts.
- Send String: `GET /DCEHealthCheck HTTP/1.1\r\nConnection: close\r\n`
- Receive String: `HTTP/1.1 200 OK`