IBM

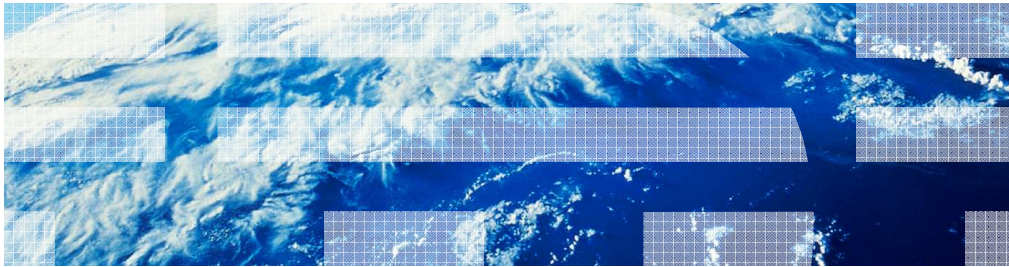# IBM WebSphere Application Server V8.5

Dynamic cluster elasticity

This presentation describes support for dynamic cluster elasticity included in IBM WebSphere Application Server V8.5.

Table of contents

Dynamic cluster elasticity

This presentation will provide an overview of dynamic cluster elasticity, illustrate an example usage scenario involving dynamic cluster elasticity operations, provide a summary of the information covered, and finally provide additional resources for utilizing dynamic cluster elasticity.

Section

Overview

Dynamic cluster elasticity   © 2012 IBM Corporation

The presentation will cover dynamic cluster elasticity in WebSphere Application Server version 8.5.

WAS85_Dynamic_cluster_elasticity.ppt

What is dynamic cluster elasticity?

- In WebSphere Network Deployment 8.5, Intelligent Management allows you to dynamically start, stop, and allocate resources in order to meet current demand for an application running on a dynamic cluster

- The Application Placement Controller is able to minimize the number of running nodes and still meet service policy goals for all dynamic clusters

- When elasticity is enabled, you can define custom operations to:
  - Start node agents of a dynamic cluster when more resources are needed
  - Stop node agents when the resources are no longer necessary

Dynamic cluster elasticity

In WebSphere Network Deployment version 8.5, the new Intelligent Management features enable the dynamic starting, stopping, and allocation of resources in a network deployment cell in order to meet defined service policy goals for an application running on a dynamic cluster.

The Application Placement Controller can be used to minimize the number of running nodes in a dynamic cluster while still meeting established service policy goals for an application. Enable elasticity mode in the Application Placement Controller and create custom elasticity actions to run user-created scripts when an elasticity add or remove operation is triggered. User scripts can accomplish many tasks, like starting a stopped node agent of a dynamic cluster when more resources are needed and stopping the node agent when the resources are no longer necessary.

Elasticity operations

- Add operation:
  – Issued when all servers in a dynamic cluster are started and demand is still not being met
  – Application Placement Controller attempts to consolidate and start all servers on a minimum number of nodes

- Remove operation:
  – Issued when a there is a drop-off in current demand or if the total number of resources are not required to meet the minimum number of instances for a dynamic cluster
  – For example, a custom elasticity action can stop a node agent that is no longer needed

Dynamic cluster elasticity   © 2012 IBM Corporation

The add elasticity operation is triggered when the started servers of a dynamic cluster are unable to meet current demand. Through the add operation, the Application Placement Controller attempts to consolidate and start all existing servers on a minimum number of nodes. This is accomplished through automatic actions and by executing the custom elasticity actions that you have defined under the add operation. Custom add action scripts are triggered before any expansion occurs.

The remove elasticity operation is triggered when there is a drop-off in the demand for an application running on a dynamic cluster and if the total number of active servers is greater than the minimum instances specified by the dynamic cluster. A remove operation is never triggered that can bring a dynamic cluster below the number of minimum active server instances. Any custom remove action script associated with the remove operation is triggered after consolidation of resources has occurred. For example, after a dynamic cluster member whose resources are no longer needed is stopped, a custom elasticity action can stop the node agent that the now stopped dynamic cluster instance was running on because it is no longer needed.

Service policies and elasticity

- Used to determine if demand is being appropriately met for a dynamic cluster
- Consists of:
  - User-defined performance goal
    - Average response time
    - Percent response time
  - Importance Level

Dynamic cluster elasticity   © 2012 IBM Corporation

Service policies are used to determine if demand is being met for a dynamic cluster. A service policy consists of a user-defined performance goal, such as average response time for incoming requests, and importance level of the defined goal. The importance level is selected from one of seven predefined values from lowest to highest.

Once a service policy's performance goal is no longer being met, an add elasticity operation is triggered to increase the resources that are available to the dynamic cluster.

Section

Usage scenario

Dynamic cluster elasticity

This scenario demonstrates a method to implement dynamic cluster elasticity.

IBM

Starting a stopped node agent

- Setup:
  - Dynamic cluster: dynamic_cluster_1
  - Members: node_1 – node agent is started
    node_2 – node agent is stopped
- The application placement controller triggers the add elasticity operation due to an increase in demand for resources on dynamic_cluster_1
- A custom elasticity action is started that does these actions:
  - Looks for any stopped node agents in dynamic_cluster_1
  - Starts up one of the stopped node agents it finds (in this case, node_2)
- Once the node agent is started, the application placement controller is able to start the dynamic cluster server instance on node_2, opening up more resources to handle the increased demand

A dynamic cluster named dynamic_cluster_1 has two members: node_1 and node_2. The node agent for node_1 is currently started and the dynamic cluster server instance is running. The node agent for node_2 is currently stopped, and therefore the dynamic cluster server instance on node_2 is stopped as well.

An add elasticity operation is triggered when the service policy goals are no longer being met for an application on dynamic_cluster_1. A custom elasticity action is started to search for any nodes that are members of dynamic_cluster_1 that have stopped node agents. Once one of those nodes is found, node_2 in this case, the script can start it and the application placement controller is able to start the dynamic cluster server instance on node_2 and the server begins to help processing the increased demand.

Writing scripts for custom elasticity actions

- WebSphere Network Deployment passes these parameters to any scripts for custom elasticity actions
  - cell: The cell name
  - node: The name of the node the elasticity operation was issued for
  - dynamiccluster: The name of the dynamic cluster that needs more resources
- Custom scripts must be placed on the file system of the node that the Application Placement Controller is running on

Dynamic cluster elasticity   © 2012 IBM Corporation

WebSphere Network Deployment Intelligent Management provides these parameters to the scripts started in custom elasticity actions: the cell name, the name of the node the add operation was issued for, and the name of the dynamic cluster that is requesting additional resources.

Any custom scripts that you create must be placed in the file system of the node that the Application Placement controller process is running on.

Example custom script

- You can print out the parameters passed to the script with a shell script:

```
#!/bin/sh
echo "The cell is "
echo $cell
echo ". The dynamic cluster requesting resources is "
echo $dynamiccluster
echo ". And the node involved in the request is "
echo $node
```

- For information about writing non-java custom action scripts, see the "Creating health policy custom actions" topic in the Information Center

This is an example script which prints out the parameters that are passed to the custom action script. The parameters passed to the shell script are saved as environment variables called cell, dynamiccluster, and node.

For more information about writing non-java custom action scripts, see the "Creating health policy custom actions" topic in the information center.

Summary

This section contains a summary.

Summary

- Elasticity allows for the dynamic allocation of resources for a dynamic cluster through add and remove elasticity operations
- Custom elasticity actions can be created to:
  - Start a stopped node agent
  - Stop a node agent that is currently running
- Elasticity allows you to maximize performance and minimize cost through automatic actions based on customer demand.

Dynamic cluster elasticity allows for the dynamic allocation of the resources in a cell depending on the current demand for an application running on a dynamic cluster. When the demand for applications which are deployed to a dynamic cluster increases, an add elasticity operation is triggered and cluster members are started in order to meet that increased demand. Conversely, when the demand for those applications decreases, a remove elasticity operation is triggered and cluster members are stopped in order to free up those resources for other applications. The overall goal of dynamic cluster elasticity is to minimize the number of WebSphere nodes in use while still meeting the current demand for applications.

Custom elasticity actions allow users to start any logic when expansion or consolidation occurs. The custom add operation is started before expansion occurs, before starting a dynamic cluster member on a previously unused WebSphere node whose resources are now needed. The custom remove operation is started after consolidation has occurred, after stopping dynamic cluster members on a WebSphere node whose resources are not currently needed. For example, the custom add action can be used to start the node agent on a WebSphere node which is now needed, and the custom remove action can be used to stop a node agent on a WebSphere node which is not currently needed.

The benefit of elasticity to your environment is maximized performance and minimized cost by automatically expanding and contracting based on what matters most: real customer usage.

References

- Elasticity mode
  http://www14.software.ibm.com/webapp/wsbroker/redirect?version=phil&product=was-nd-mp&topic=cwve_elasticity
- Configuring elasticity mode
  http://www14.software.ibm.com/webapp/wsbroker/redirect?version=phil&product=was-nd-mp&topic=twve_cfgelastic
- Creating health policy custom actions
  http://www14.software.ibm.com/webapp/wsbroker/redirect?version=phil&product=was-nd-mp&topic=twve_odhealthaction

Dynamic cluster elasticity

See these references for additional information about dynamic cluster elasticity.

Feedback

Your feedback is valuable

You can help improve the quality of IBM Education Assistant content to better meet your needs by providing feedback.

- Did you find this module useful?

- Did it help you solve a problem or answer a question?

- Do you have suggestions for improvements?

Click to send email feedback:

mailto:iea@us.ibm.com?subject=Feedback_about_WAS85_Dynamic_cluster_elasticity.ppt

This module is also available in PDF format at: ../WAS85_Dynamic_cluster_elasticity.pdf

Dynamic cluster elasticity

You can help improve the quality of IBM Education Assistant content by providing feedback.

Trademarks, disclaimer, and copyright information