Informatica PowerCenter on the AWS Cloud

Quick Start Reference Deployment Guide

November 2016
([last update](): April 2017)

Informatica Corporation
AWS Quick Start Reference Team

Contents

This Quick Start deployment guide was created by Amazon Web Services (AWS) in partnership with Informatica.

Overview

This Quick Start reference deployment guide provides step-by-step instructions for deploying Informatica PowerCenter on the Amazon Web Services (AWS) Cloud. Quick Starts are automated reference deployments that use AWS CloudFormation templates to launch, configure, and run the AWS compute, network, storage, and other services required to deploy a specific workload on AWS.

Deploying Informatica PowerCenter on the AWS Cloud provides several benefits:

- **Enterprise class data integration.** PowerCenter is an enterprise-proven data integration solution that can process billions of records in the AWS Cloud.

- **Connections to existing data sources and quick onboarding of new data sources and data types**. PowerCenter offers a vast array of connectors, whether you want to connect to on-premises data sources or AWS services such as Amazon Redshift, Amazon Relational Database Service (Amazon RDS), or Amazon Simple Storage Service (Amazon S3).

- **Faster time to insight**. By leveraging your existing on-premises PowerCenter mappings, metadata, and workflows, you can rapidly load data into AWS data services such as Amazon Redshift, delivering the right analytical data to your business stakeholders.

- **Reduced cost of data integration solution.** Data integration is typically compute intensive. You need to increase or decrease the computing power of your data integration platform elastically to meet varying business data loads. On-demand compute bursting in combination with an AWS pay-as-you-go model helps reduce the infrastructure cost.

- **Accelerated data architecture modernization**. If you are planning to modernize your data warehousing initiatives on AWS, PowerCenter's rich functionalities such as metadata-driven data integration, dynamic mappings, SQL conversion mapping, and automatic data validation will help you shorten development cycles and reduce time to market.

- **Delivery of clean, complete, and trustworthy data**. Whether you are offloading or extending on-premises applications to the cloud or fully embracing the cloud, delivering complete, high-quality data is critical. PowerCenter has a long history of helping organizations empower their users with complete, high-quality, actionable data.

- **Meeting business demands by improving developer productivity**. IT departments struggle to deliver trusted data as the amount of data and the demand from business keep growing. To succeed, your organization's data management environment must continue to leverage greater speed and agility in order to delight your customers and outsmart your competitors. PowerCenter's highly visual, easy to use, metadata-driven environment not only enhances the ability to execute on today's projects, but also helps your organization cost-effectively scale future initiatives.

This Quick Start is for users who want to move or offload their applications or data warehouses to AWS. This migration allows them to realize cloud benefits while leveraging their existing data management investment in PowerCenter.

Costs and Licenses

You are responsible for the cost of the AWS services used while running this Quick Start reference deployment. There is no additional cost for using the Quick Start.

The AWS CloudFormation template for this Quick Start includes configuration parameters that you can customize. Some of these settings, such as instance type, will affect the cost of deployment. See the pricing pages for each AWS service you will be using for cost estimates.

This Quick Start requires a license for Informatica PowerCenter. You can use any edition of PowerCenter, depending on the license. PowerCenter is provided in the AWS Marketplace as a bring your own license (BYOL) listing. The two current offerings are [Informatica PowerCenter for Windows (BYOL)](#) and [Informatica PowerCenter for Red Hat Linux (BYOL)](#):

- Release support – PowerCenter versions 10.1.1 and 10.1.0

- Operating system support – Windows Server and Red Hat Enterprise Linux

- Informatica domain DB support – Amazon RDS running Oracle Database

- Informatica license – Bring your own license (BYOL)

You must work with Informatica directly to obtain a license and key.

Architecture

Deploying this Quick Start with the **default parameters** builds the following PowerCenter environment in the AWS Cloud.
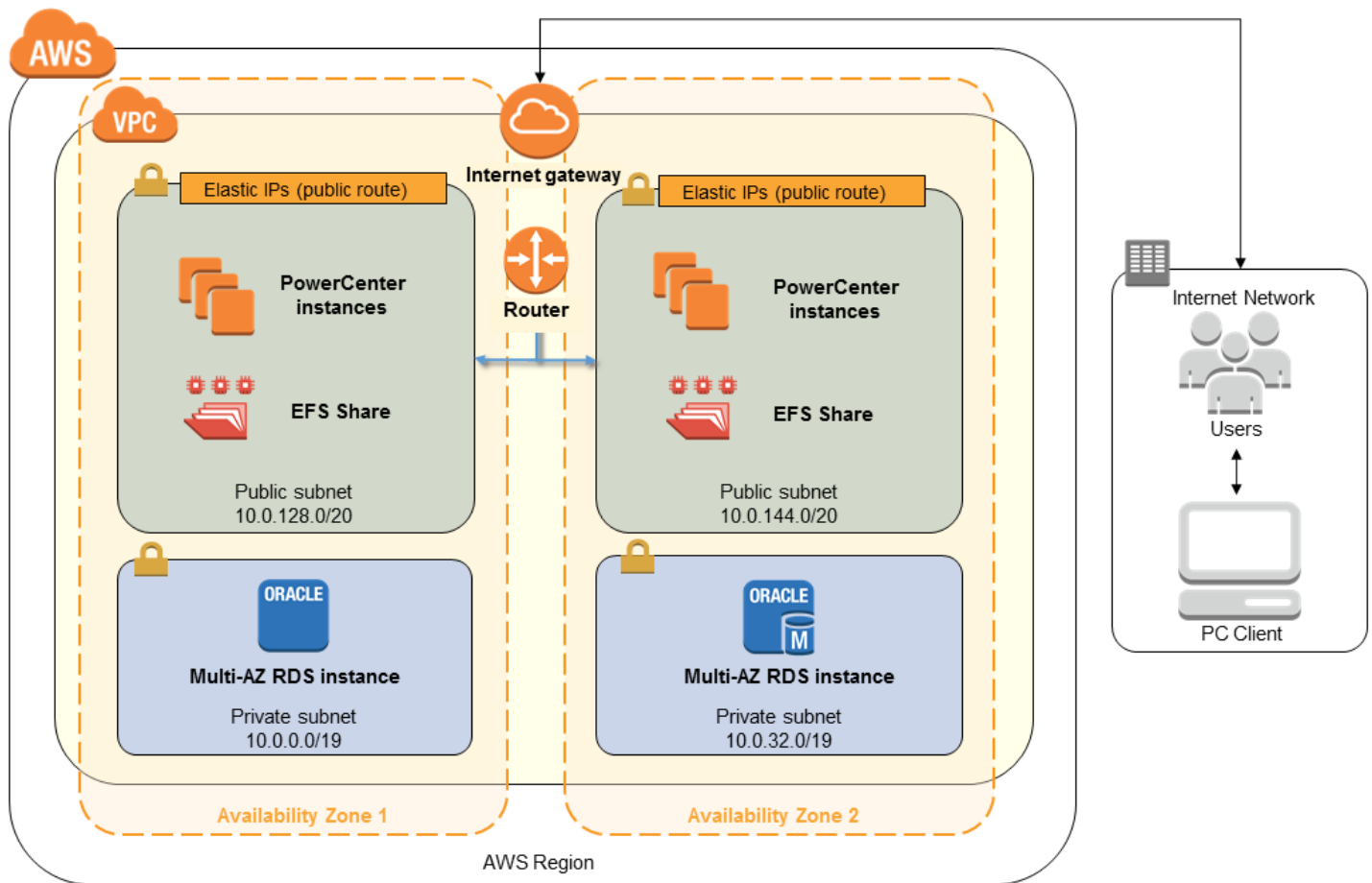
Figure 1: Quick Start PowerCenter architecture on AWS

The Informatica domain runs in the AWS Cloud. The domain includes the following application services:

- **PowerCenter Integration Service.** The PowerCenter Integration Service runs PowerCenter sessions and workflows.

- **PowerCenter Repository Service.** The PowerCenter Repository Service manages the PowerCenter repository. It retrieves, inserts, and updates metadata in the repository database tables. Oracle serves as the PowerCenter repository database.

The PowerCenter services write files, such as cache, source, and target files, to Amazon Elastic Block Store (Amazon EBS). If the **Create Elastic File System** parameter is set to **Yes** (default) on a Linux system, the Quick Start uses Amazon Elastic File System (Amazon EFS) instead for shared storage, if it's available in your region. While the PowerCenter services run on AWS, you can create, configure, and run PowerCenter mappings and workflows from on-premises PowerCenter clients.

Prerequisites

Specialized Knowledge

Before you deploy this Quick Start, we recommend that you become familiar with the following AWS services. (If you are new to AWS, see Getting Started with AWS.)

- Amazon VPC
- Amazon EC2
- Amazon EBS
- Amazon EFS
- Amazon RDS

Technical Requirements

The PowerCenter domain and its nodes can be deployed to an existing virtual private cloud (VPC) or to a new VPC. You need to ensure that connectivity to the required sources and targets are established from the VPC. All nodes are gateway nodes and are configured with optional Elastic IP addresses so that client tools running outside the VPC can connect to the domain.

Make sure that the required number of Elastic IP addresses are available in your AWS account. Also, make sure to enter the IP addresses of all the nodes in the etc\hosts file on the client machine to establish connectivity to the domain. Create additional nodes using a single Amazon Machine Image (AMI) to expand compute capacity of the domain as needed.

Deployment Options

The Quick Start provides two deployment options:

- **Deployment of PowerCenter into a new VPC** (end-to-end deployment) builds a new virtual private cloud (VPC) with public and private subnets, and then deploys PowerCenter into that infrastructure.
- **Deployment of PowerCenter into an existing VPC** provisions PowerCenter components into your existing infrastructure.

The Quick Start provides separate templates for these options. You also have a choice of two operating systems on which Informatica PowerCenter will be deployed: Microsoft Windows Server or Red Hat Enterprise Linux (RHEL).

Deployment Steps

Step 1. Prepare an AWS Account

1. If you don't already have an AWS account, create one at http://aws.amazon.com by following the on-screen instructions.

2. Use the region selector in the navigation bar to choose the AWS Region where you want to deploy PowerCenter on AWS.

3. Create a key pair in your preferred region.

4. If necessary, request a service limit increase for the Amazon EC2 **C4** or **R3** instance type you're planning to use. You might need to do this if you already have an existing deployment that uses this instance type, and you think you might exceed the default limit with this reference deployment.

Step 2. Launch the Quick Start

1. Choose one of the following options to deploy the AWS CloudFormation template into your AWS account. Note that the existing VPC option provides separate templates depending on your operating system.

   Launch Quick Start (for new VPC)   Launch Quick Start (for existing VPC on Windows)   Launch Quick Start (for existing VPC on RHEL)

   The templates are launched in the US West (Oregon) Region by default. You can change the region by using the region selector in the navigation bar.

   Each stack takes approximately two hours to create.

   > **Note:** You are responsible for the cost of the AWS services used while running this Quick Start reference deployment. There is no additional cost for using this Quick Start. See the pricing pages for each AWS service you will be using for full details.

2. On the **Select Template** page, keep the default setting for the template URL, and then choose **Next**.

3. On the **Specify Details** page, review the parameters for the template. Enter values for the parameters that require your input. For all other parameters, you can customize the default settings provided by the template.

In the following tables, parameters are listed and described separately for deploying PowerCenter into a new VPC or an existing VPC.

> **Note** The templates for the two scenarios share most, but not all, of the same parameters. For example, the template for an existing VPC prompts you for the VPC and subnet IDs in your existing VPC environment. You can also download the templates and edit them to create your own parameters based on your specific deployment scenario.

• Parameters for deployment into a new VPC:

View template

Network Configuration:

Parameter (name)	Default	Description
Availability Zones (AvailabilityZones)	*Requires input*	Select two Availability Zones that will be used to deploy Informatica PowerCenter. The Quick Start preserves the logical order you specify.
VPC CIDR (VPCCIDR)	10.0.0.0/16	CIDR block for the VPC. When you create a VPC, you must specify a range of IPv4 addresses for the VPC in the form of a Classless Inter-Domain Routing (CIDR) block. You can also optionally assign an IPv6 CIDR block to your VPC, and assign IPv6 CIDR blocks to your subnets. For more information about CIDR notation, see IETF RFC 4632.
Private Subnet 1 CIDR (PrivateSubnet1CIDR)	10.0.0.0/19	CIDR block for the private subnet located in Availability Zone 1.
Private Subnet 2 CIDR (PrivateSubnet2CIDR)	10.0.32.0/19	CIDR block for the private subnet located in Availability Zone 2.
Public Subnet 1 CIDR (PublicSubnet1CIDR)	10.0.128.0/20	CIDR block for the public (DMZ) subnet located in Availability Zone 1.
Public Subnet 2 CIDR (PublicSubnet2CIDR)	10.0.144.0/20	CIDR block for the public (DMZ) subnet located in Availability Zone 2.
Allowed Remote Access CIDR (RemoteAccessCIDR)	*Requires input*	The CIDR IP range that is permitted to access the Informatica domain. We recommend that you use a constrained CIDR range to reduce the potential of inbound attacks from unknown IP addresses. For example, if your IPv4 address is 203.0.113.25, specify 203.0.113.25/32 to list this single IPv4 address in CIDR notation. If your company allocates addresses from a range, specify the entire range, such as 203.0.113.0/24. For details, see VPCs and Subnets in the AWS documentation.

Amazon EC2 Configuration:

Parameter	Default	Description
Key Pair Name (KeyPairName)	*Requires input*	Public/private key pair, which allows you to connect securely to your instance after it launches. When you created an AWS account, this is the key pair you created in your preferred region.
Operating System (PowerCenterOS)	Windows Server	The operating system on which the Informatica domain will be deployed. The two options are Windows Server and Red Hat Enterprise Linux (RHEL).
Informatica Domain Instance Type (PowerCenterInstanceType)	c4.2xlarge	The EC2 instance type for the instance that hosts the Informatica domain.
Number of Instances (NumberOfPowerCenter Instances)	4	The number of PowerCenter nodes you want to provision. You can choose from 1, 2, or up to 16 nodes (in multiples of 2) to be part of the domain. Each node runs on a single EC2 instance.
Enable Elastic IP Addressing (EnableEIP)	Yes	Select **No** if you don't want to assign Elastic IP addresses to instances.
Create Elastic File System (CreateEFS)	Yes	(Linux only) Whether you want to create shared storage for your instances with Amazon EFS, if it's available in your region. For information about availability, see the AWS documentation.

Amazon RDS Configuration:

Parameter	Default	Description
Database Type (DatabaseType)	Oracle-Standard-Edition-One-11.2.0-License-Included	The type of database you'd like to use in Amazon RDS. Currently, the only choice is Oracle.
Database Name (RDSDatabaseName)	InfaDB	Database name for the RDS instance.
Database Username (DatabaseUsername)	*Requires input*	User name for the Amazon RDS database account (2-30 characters).
Informatica Database Password (DatabasePassword)	*Requires input*	Password for the Amazon RDS database account (8-30 characters). Only alphanumeric characters and underscore are allowed. This password should begin with an alphabetic character.
Confirm Password (DatabasePasswordConfirm)	*Requires input*	Reenter the password you specified for the **Informatica Database Password** parameter.

Informatica PowerCenter Configuration:

Parameter	Default	Description
Informatica Domain Name (InformaticaDomainName)	InfaDomain	The name for the Informatica domain.
Informatica Node Prefix (InformaticaNodeName)	Infa	The prefix for the nodes in the Informatica domain. The node number will be added to this string; e.g., Infa1, Infa2, etc.
Informatica Administrator Username (InformaticaDomain Username)	Administrator	The user name of the Administrator account for Informatica PowerCenter.
Informatica Administrator Password (InformaticaDomain Password)	*Requires input*	Administrator password for accessing PowerCenter. This string should be 8-30 characters long, with at least one special character, one number, one uppercase character, and one lowercase character. Double quotes ("), ampersands (&), and dollar signs ($) are not allowed.
Confirm Password (InformaticaDomain PasswordConfirm)	*Requires input*	Reenter the password you specified for the **Informatica Administrator Password** parameter.
Encryption Key Phrase (EncryptionKeyPhrase)	*Requires input*	The base word for generating an encryption key for the Informatica domain. This string should be 8-20 characters long, with at least one uppercase letter, one lowercase letter, and one number, and shouldn't contain any spaces.
Informatica License Key (LicenseKey)	*Optional*	If you have an existing PowerCenter license, specify the public URL to the S3 bucket where you've stored the license key file. Leave this parameter blank to deploy the Quick Start without a PowerCenter license; it can be provisioned later.

AWS Quick Start Configuration:

Parameter	Default	Description
Quick Start S3 Bucket Name (QSS3BucketName)	aws-quickstart	S3 bucket name for the Quick Start assets. This bucket name can include numbers, lowercase letters, uppercase letters, and hyphens (-), but should not start or end with a hyphen. You can specify your own bucket if you copy all the assets and submodules into it, if you want to override the Quick Start behavior for your specific implementation.
Quick Start S3 Key Prefix (QSS3KeyPrefix)	quickstart-informatica-powercenter/	S3 key prefix for the Quick Start assets. This prefix can include numbers, lowercase letters, uppercase letters, hyphens (-), and forward slashes (/), but should not start or end with a forward slash (which is automatically added). This parameter enables you to override the Quick Start behavior for your specific implementation.

aws

- **Parameters for deployment into an existing VPC:**

  View template for Windows   View template for RHEL

Network Configuration:

Parameter (name)	Default	Description
VPC ID (VPCID)	*Requires input*	ID of your existing VPC where you'd like to deploy PowerCenter (e.g., vpc-0343606e).
Public Subnet 1 ID (PublicSubnet1ID)	*Requires input*	Publicly accessible subnet ID located in Availability Zone 1 for the Informatica domain.
Public Subnet 2 ID (PublicSubnet1ID)	*Requires input*	Publicly accessible subnet ID located in Availability Zone 2 for the Informatica domain.
Informatica Database Subnets (DBSubnetIDs)	*Requires input*	IDs of two private subnets in the selected VPC. These must be in different Availability Zones in the selected VPC (e.g., us-west-1b, us-west-1c).
Allowed Remote Access CIDR (RemoteAccessCIDR)	*Requires input*	The CIDR IP range that is permitted to access the Informatica domain. We recommend that you use a constrained CIDR range to reduce the potential of inbound attacks from unknown IP addresses. For example, if your IPv4 address is 203.0.113.25, specify 203.0.113.25/32 to list this single IPv4 address in CIDR notation. If your company allocates addresses from a range, specify the entire range, such as 203.0.113.0/24. For details, see VPCs and Subnets in the AWS documentation.

Amazon EC2 Configuration:

Parameter	Default	Description
Key Pair Name (KeyPairName)	*Requires input*	Public/private key pair, which allows you to connect securely to your instance after it launches. When you created an AWS account, this is the key pair you created in your preferred region.
Operating System (PowerCenterOS)	Windows Server	The operating system on which the Informatica domain will be deployed. The two options are Windows Server and Red Hat Enterprise Linux (RHEL).
Informatica Domain Instance Type (PowerCenterInstanceType)	c4.2xlarge	The EC2 instance type for the instance that hosts the Informatica domain.
Number of Instances (NumberOfPowerCenter Instances)	4	The number of PowerCenter nodes you want to provision. You can choose from 1, 2, or up to 16 nodes (in multiples of 2) to be part of the domain. Each node runs on a single EC2 instance.

aws

Parameter	Default	Description
Enable Elastic IP Addressing (EnableEIP)	Yes	Select **No** if you don't want to assign Elastic IP addresses to instances.
Create Elastic File System (CreateEFS)	Yes	(Linux only) Whether you want to create shared storage for your instances with Amazon EFS, if it's available in your region. For information about availability, see the [AWS documentation](#).

Amazon RDS Configuration:

Parameter	Default	Description
Database Type (DatabaseType)	Oracle-Standard-Edition-One-11.2.0-License-Included	The type of database you'd like to use in Amazon RDS. Currently, the only choice is Oracle.
Database Name (RDSDatabaseName)	InfaDB	Database name for the RDS instance.
Database Username (DatabaseUsername)	*Requires input*	User name for the Amazon RDS database account (2-30 characters).
Informatica Database Password (DatabasePassword)	*Requires input*	Password for the Amazon RDS database account (8-30 characters). Only alphanumeric characters and underscore are allowed. This password should begin with an alphabetic character.
Confirm Password (DatabasePassword Confirm)	*Requires input*	Reenter the password you specified for the **Informatica Database Password** parameter.

Informatica PowerCenter Configuration:

Parameter	Default	Description
Informatica Domain Name (InformaticaDomainName)	InfaDomain	The name for the Informatica domain.
Informatica Node Prefix (InformaticaNodeName)	Infa	The prefix for the nodes in the Informatica domain. The node number will be added to this string; e.g., Infa1, Infa2, etc.
Informatica Administrator Username (InformaticaDomain Username)	Administrator	The user name of the Administrator account for Informatica PowerCenter.
Informatica Administrator Password (InformaticaDomain Password)	*Requires input*	Administrator password for accessing PowerCenter. This string should be 8-30 characters long, with at least one special character, one number, one uppercase character, and one lowercase character. Double quotes ("), ampersands (&), and dollar signs ($) are not allowed.

Parameter	Default	Description
Confirm Password (InformaticaDomain PasswordConfirm)	*Requires input*	Reenter the password you specified for the **Informatica Administrator Password** parameter.
Encryption Key Phrase (EncryptionKeyPhrase)	*Requires input*	The base word for generating an encryption key for the Informatica domain. This string should be 8-20 characters long, with at least one uppercase letter, one lowercase letter, and one number, and shouldn't contain any spaces.
Informatica License Key (LicenseKey)	*Optional*	If you have an existing PowerCenter license, specify the public URL to the S3 bucket where you've stored the license key file. Leave this parameter blank to deploy the Quick Start without a PowerCenter license; it can be provisioned later.

When you finish reviewing and customizing the parameters, choose **Next**.

4. On the **Options** page, you can [specify tags](#) (key-value pairs) for resources in your stack and [set advanced options](#). When you're done, choose **Next**.

5. On the **Review** page, review and confirm the template settings. Under **Capabilities**, select the check box to acknowledge that the template will create IAM resources.

6. Choose **Create** to deploy the stack.

7. Monitor the status of the stack. When the status is **CREATE_COMPLETE**, the deployment is complete.

8. You can use the URL displayed in the **Outputs** tab for the stack to view the resources that were created.

Step 3. Test the Deployment

1. Create folders in the PowerCenter repository.

2. Add users and assign permissions to folders.

3. Develop a mapping, session, and workflow, and then run the workflow.

   For more information about developing PowerCenter mappings, sessions, and workflows, see the PowerCenter documentation at http://network.informatica.com.

Troubleshooting

Q. I encountered a CREATE_FAILED error when I launched the Quick Start. What should I do?

A. If AWS CloudFormation fails to create the stack, we recommend that you relaunch the template with **Rollback on failure** set to **No**. (This setting is under **Advanced** in the AWS CloudFormation console, **Options** page.) With this setting, the stack's state will be retained and the instance will be left running, so you can troubleshoot the issue. (You'll want to look at the log files in `%ProgramFiles%\Amazon\EC2ConfigService` and `C:\cfn\log`.)

> **Important:** When you set **Rollback on failure** to **No**, you'll continue to incur AWS charges for this stack. Please make sure to delete the stack when you've finished troubleshooting.

For additional information, see [Troubleshooting AWS CloudFormation](#) on the AWS website or contact us on the [AWS Quick Start Discussion Forum](#).

Q. I encountered a size limitation error when I deployed the AWS CloudFormation templates.

A. We recommend that you launch the Quick Start templates from the location we've provided or from another S3 bucket. If you deploy the templates from a local copy on your computer or from a non-S3 location, you might encounter template size limitations when you create the stack. For more information about AWS CloudFormation limits, see the [AWS documentation](#).

Additional Resources

AWS services

- AWS CloudFormation
  https://aws.amazon.com/documentation/cloudformation/
- Amazon EBS
  https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/AmazonEBS.html
- Amazon EC2
  https://docs.aws.amazon.com/AWSEC2/latest/WindowsGuide/
- Amazon EFS
  https://aws.amazon.com/documentation/efs/

- Amazon VPC
  https://aws.amazon.com/documentation/vpc/

Informatica PowerCenter

- Datasheet
  https://www.informatica.com/content/dam/informatica-com/global/amer/us/collateral/brochure/powercenter_brochure_6659.pdf

- Architecture reference
  https://kb.informatica.com/proddocs/Product%20Documentation/5/PowerCenter_on_AWS_Architecture_en.pdf

- Solution brief
  https://www.informatica.com/content/dam/informatica-com/global/amer/us/collateral/solution-brief/informatica-for-aws_solution-brief_3180en.pdf

- PowerCenter on AWS Marketplace
  https://kb.informatica.com/whitepapers/4/Pages/1/504683.aspx

Quick Start reference deployments

- AWS Quick Start home page
  https://aws.amazon.com/quickstart/

Send Us Feedback

We welcome your questions and comments. Please post your feedback on the AWS Quick Start Discussion Forum.

You can visit our GitHub repository to download the templates and scripts for this Quick Start, and to share your customizations with others.

Document Revisions

Date	Change	In sections
April 2017	Added support for Elastic IP addressing and Amazon EFS	Changes to templates and throughout guide
November 2016	Initial publication	—

aws