

IBM InfoSphere Information Server on Cloud  
Version 1 Release 1

*User Guide*  
*(Last updated: 2017-06-20)*



**Note**

Before using this information and the product that it supports, read the information in [Notices and trademarks](#).

**Edition Notice**

This edition applies to version 1, release 1, modification 0 of IBM InfoSphere Information Server on Cloud and to all subsequent releases and modifications until otherwise indicated in new editions.

© **Copyright International Business Machines Corporation 2015, 2017.**

US Government Users Restricted Rights – Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

---

# Contents

- Chapter 1. Information Server on Cloud offerings..... 1**
- Chapter 2. Information Server on Cloud Enterprise Edition.....5**
  - Overview.....5
  - Available configurations.....7
  - Layout of IBM Information Server on Cloud Enterprise Edition server and client disks.....8
  - Information roadmap.....9
  - Product Overview.....9
  - Installing.....9
  - Administering.....10
  - Managing Metadata.....10
  - Getting Started with Consoles.....10
  - Connecting to data sources.....10
  - Getting started and using IBM Information Server on Cloud Enterprise Edition.....10
  - When you connect for the first time, follow these steps:.....11
  - After the initial connection, you can do any of the following tasks:.....12
- Chapter 3. Information Server on Cloud Data Quality..... 13**
  - Overview.....13
  - Available configurations.....15
  - Layout of IBM Information Server on Cloud Data Quality server and client disks.....16
  - Information roadmap.....17
  - IBM InfoSphere DataStage and QualityStage.....17
  - IBM InfoSphere Information Analyzer.....18
  - IBM InfoSphere Information Governance Catalog.....18
  - Getting started and using IBM Information Server on Cloud Data Quality.....19
  - When you connect for the first time, follow these steps:.....19
  - After the initial connection, you can do any of the following tasks:.....20
- Chapter 4. DataStage on Cloud..... 23**
  - Overview.....23
  - Available configurations.....24
  - Layout of IBM DataStage on Cloud server and client disks.....24
  - Information roadmap.....26
  - Product overview.....26
  - Getting started.....26
  - Using InfoSphere DataStage.....26
  - Troubleshooting and support.....27
  - Getting started and using IBM DataStage on Cloud.....27
  - To connect for the first time, follow these steps:.....27
  - After the initial connection, you can do any of the following tasks:.....28
- Chapter 5. DataStage on Cloud Designer Client..... 29**
  - Overview.....29
  - Available configuration.....30
  - Layout of IBM® DataStage® on Cloud Designer Client disks.....30
  - Information roadmap.....30
  - Product overview.....31
  - Getting started.....31
  - Using InfoSphere DataStage.....31

Troubleshooting and support.....	31
Getting started and using IBM DataStage on Cloud Designer Client.....	31
When you connect for the first time, follow these steps:.....	32
After the initial connection, you can do any of the following tasks:.....	32
Adding and connecting extra DataStage on Cloud Designer Client machines.....	33
<b>Chapter 6. Information Governance Catalog on Cloud.....</b>	<b>35</b>
Overview.....	35
Available configurations.....	37
Layout of IBM Information Governance Catalog on Cloud server and client disks.....	37
Information roadmap.....	38
Getting started and using IBM Information Governance Catalog on Cloud.....	39
When you connect for the first time, follow these steps:.....	40
After the initial connection, you can do any of the following tasks:.....	40
<b>Chapter 7. Connecting to other systems.....</b>	<b>41</b>
Connecting to the IBM InfoSphere DataStage and QualityStage Designer client.....	41
Connecting to an on-premises computer.....	42
Connecting to IBM Db2 on Cloud.....	42
Example.....	43
Connecting on an on-premises Db2 database instance.....	43
Example.....	44
<b>Chapter 8. Administering cloud offerings on the server.....</b>	<b>47</b>
Open ports on server and client machines.....	47
Open ports for incoming traffic on the server machine.....	47
Open ports for incoming traffic on client machines.....	48
Starting services after the Information Server on Cloud server machine reboots.....	49
Setting up firewall security.....	49
About these tasks.....	49
To change the security level or to manage rules for the Information Server on Cloud server firewall.....	49
To show the Microsoft Windows firewall profiles on the Information Server on Cloud client.....	51
To show a list of all open ports on the Information Server on Cloud client.....	51
To block an open port on the Information Server on Cloud client.....	51
Scripts to change the iptables firewall settings.....	51
Enhancing security of Information Server on Cloud computers.....	53
Setting up SSH keys.....	53
Allowing SSH access to specific IP addresses.....	53
Managing LUKS keys on the IBM Information Server on Cloud server.....	54
<b>Chapter 9. Troubleshooting cloud offerings.....</b>	<b>55</b>
Cannot install a patch on services tier machine.....	55
Symptoms.....	55
Resolving the problem.....	55
<b>Notices.....</b>	<b>57</b>

# Chapter 1. Information Server on Cloud offerings

IBM® Information Server on Cloud offerings provide data integration and governance products as a service on the global cloud infrastructure of IBM SoftLayer®. The rich features of on-premises IBM InfoSphere® Information Server components are provided without the cost and complexity of deploying the infrastructure. Optional add-on services can also be added to maintain and manage the infrastructure.

Information Server on Cloud helps to reduce the time that is required to provision and deploy data integration and governance. As a result, your IT resources are free to innovate and develop new solutions.

Information Server on Cloud includes the following offerings:

<b>IBM Information Server on Cloud Enterprise Edition</b>	Information Server on Cloud Enterprise Edition provides end-to-end information integration capabilities to help you understand, govern, create, maintain, transform, and deliver quality data.
<b>IBM Information Server on Cloud Data Quality</b>	Information Server on Cloud Data Quality cleanses data and monitors data quality.
<b>IBM DataStage® on Cloud</b>	DataStage on Cloud provides a framework to design and run jobs that transform and cleanse your data.
<b>IBM® DataStage® on Cloud Designer Client</b>	DataStage on Cloud Designer Client provides a framework to create, design, and develop DataStage jobs. This offering includes all IBM InfoSphere Information Server clients. As a result, you can connect to the Information Server on Cloud server and work with the clients.
<b>IBM Information Governance Catalog on Cloud</b>	Information Governance Catalog on Cloud provides a standardized approach to discover and govern your business assets, and to define a common business language.

In the following table, a check mark (✓) indicates that the component is available with the listed offering. An asterisk (\*) next to a check mark indicates limits on the product availability. Check the license agreement for details.

*Table 1: Available components per offering*

InfoSphere Information Server component	Information Server on Cloud Enterprise Edition	Information Server on Cloud Data Quality	DataStage on Cloud	DataStage on Cloud Designer Client	Information Governance Catalog on Cloud
InfoSphere DataStage	✓		✓		✓*
InfoSphere QualityStage®	✓	✓			✓
InfoSphere DataStage and QualityStage Designer	✓	✓	✓	✓	✓*
InfoSphere DataStage and QualityStage Director	✓	✓	✓	✓	✓
InfoSphere Data Click	✓			InfoSphere Data Click client only	

Table 1: Available components per offering (continued)

<b>InfoSphere Information Server component</b>	<b>Information Server on Cloud Enterprise Edition</b>	<b>Information Server on Cloud Data Quality</b>	<b>DataStage on Cloud</b>	<b>DataStage on Cloud Designer Client</b>	<b>Information Governance Catalog on Cloud</b>
InfoSphere FastTrack	✓			InfoSphere FastTrack client only	
InfoSphere Information Analyzer	InfoSphere Information Analyzer server	InfoSphere Information Analyzer server		InfoSphere Information Analyzer client only	
InfoSphere Information Governance Catalog	✓	✓*		InfoSphere Information Governance Catalog client only	✓
IBM Glossary Anywhere	✓	✓	✓	Glossary Anywhere installer only	✓
InfoSphere Information Governance Dashboard	✓	✓	✓	✓	✓
InfoSphere Information Services Director	✓	✓	✓	InfoSphere Information Services Director client only	
InfoSphere Metadata Integration Bridges and the metadata interchange agent	✓	✓	✓	✓	✓
InfoSphere Metadata Asset Manager	✓	✓	✓	✓	✓
InfoSphere Information Server istool command-line utility	✓	✓	✓	✓	✓
InfoSphere Information Server Manager client, Multi-Client Manager	✓	✓	✓	✓	✓
IBM Business Process Manager Standard	✓				

*Table 1: Available components per offering (continued)*

<b>InfoSphere Information Server component</b>	<b>Information Server on Cloud Enterprise Edition</b>	<b>Information Server on Cloud Data Quality</b>	<b>DataStage on Cloud</b>	<b>DataStage on Cloud Designer Client</b>	<b>Information Governance Catalog on Cloud</b>
IBM Cognos® Business Intelligence	✓	✓			
IBM InfoSphere Data Architect	✓				✓





---

## Chapter 2. Information Server on Cloud Enterprise Edition

IBM® Information Server on Cloud Enterprise Edition provides a hosted environment that you configure and control. You can use Information Server on Cloud Enterprise Edition to extend the reach of your business by leveraging cloud offerings, while you reduce the costs that are associated with providing these services. Different plans are available so that any size business can access the powerful and scalable platform by IBM.

---

### Overview

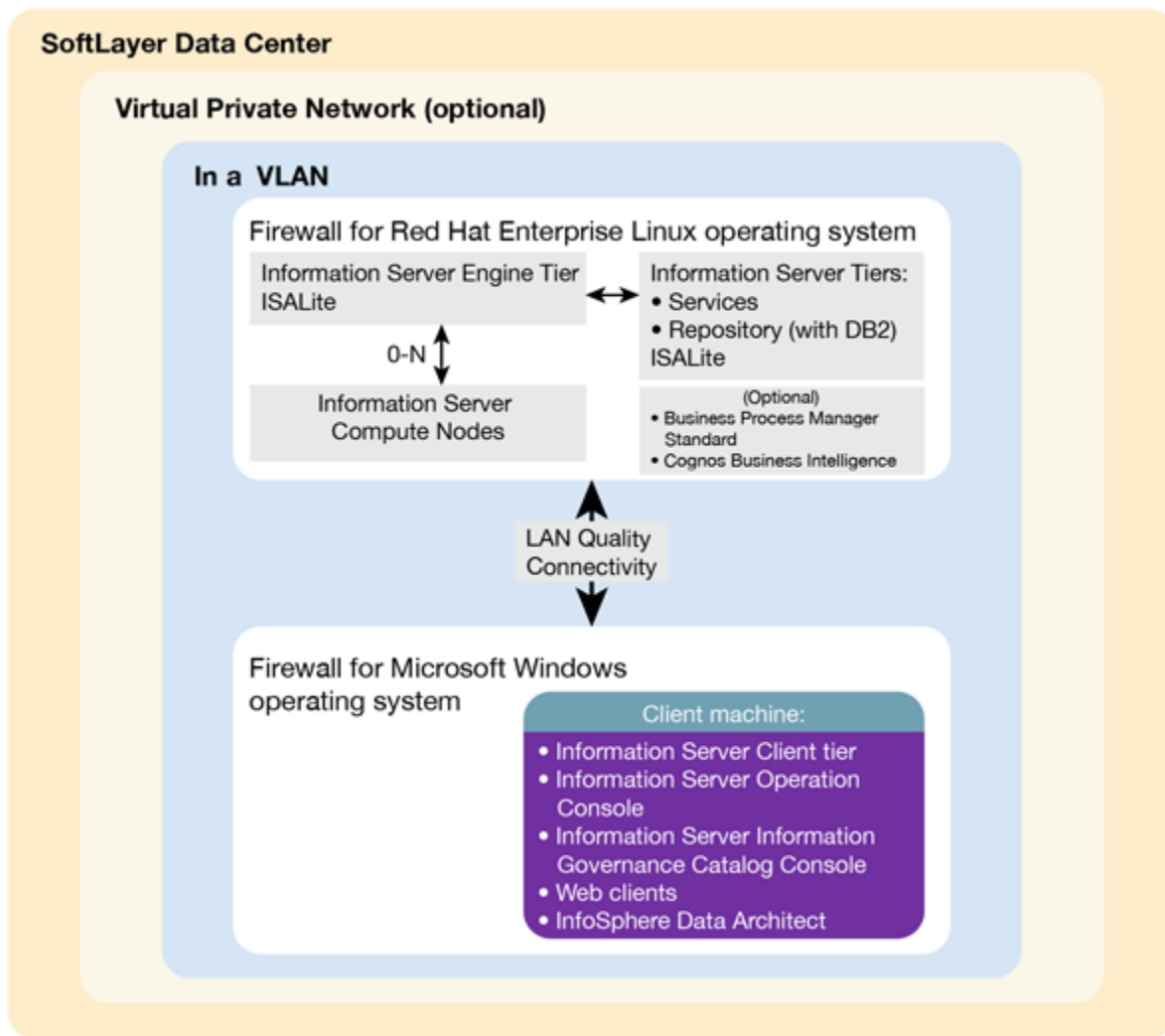
IBM® Information Server on Cloud Enterprise Edition is a data integration software platform that helps organizations derive more value from the complex, heterogeneous information spread across their systems. Information Server on Cloud Enterprise Edition provides all of the functions of its on-premises counterpart, IBM InfoSphere® Information Server.

Information Server on Cloud Enterprise Edition includes the following key features:

- Removes the burden of deployment so that you can focus on core development of integration and governance assets
- Integrates and governs products without the need for an on-premises infrastructure
- Reduces time to value with enterprise-grade features
- Optimizes performance and reliability by using a virtual machine or a dedicated bare metal server
- Keeps infrastructure costs in line with the changing needs of the business
- Provides flexibility and agility to meet dynamic business needs
- Works with cloud and on-premises data sources
- Scales according to your business demands

Information Server on Cloud Enterprise Edition uses the characteristics of software-as-a-service (SaaS). You select the plan size based on your needs. IBM provisions the machines and deploys the Information Server on Cloud Enterprise Edition software.

The following figure shows the topology of the server and client machines in a typical deployment.



As a hosted offering, you have the same control over your data in the cloud as in the on-premises system:

- Actively monitor and report any issues that you encounter with IBM Software as a Service (SaaS).
- Maintain the software platform of your cloud offering and the operating system to meet your security standards.
- Maintain software firewalls on servers that face the internet in a manner to provide required protection.
- Develop parallel jobs to transform and cleanse data, and develop server jobs to transform data. Establish connectivity between data sources and applications. Develop your own workload, business rules, monitoring, and scheduling for all jobs. You are responsible for the quality and performance of programs, applications, and jobs that you develop.
- Ensure the continuity, compatibility, and performance of the IBM SaaS platform by installing only permissible software, including any open source packages.
- Regularly upgrade the environment and operating system of your cloud offering.
- Create and maintain regular backups of data.
- Create and maintain high availability configurations.

The following managed add-on services are available to maintain and manage the infrastructure:

**Jump start** This setup service provides up to 50 hours of remote consulting time for startup activities.

**Accelerator** This setup service provides up to 50 hours of remote consulting time to perform various scoped activities.

**Silver** This service provides monthly remote consulting time for operations and maintenance activities.

**Gold** This service provides monthly remote consulting time for operations and maintenance activities. The service includes everything that is provided by the Silver service and delivers extra activities.

Note the following limitations and restrictions of Information Server on Cloud Enterprise Edition:

- If your offering is designated as “Non-Production“, Information Server on Cloud Enterprise Edition can be deployed only as part of your development and test environments for internal non-production activities. These activities include, but are not limited to: testing, performance tuning, fault diagnosis, internal benchmarking, staging, quality assurance activity, developing internally used additions or extensions to the offering by using published application programming interfaces.
- Users must not modify the configuration file that is needed to run the job for parallel processing.

## Available configurations

IBM® Information Server on Cloud Enterprise Edition servers for the small and medium plans are virtual servers with dedicated CPUs. The servers in the large plan are in a bare metal environment.

Select the offering plan that fits your usage and environment needs.

*Table 2: Offering sizes: small production and non-production*

Tier	Memory (GB)	Number of cores	Network speed and bandwidth	First disk	Second disk
Engine	16	4	1 Gbps with 250 GB bandwidth	100 GB storage area network (SAN)	500 GB SAN
Service metadata	16	4	1 Gbps with 250 GB bandwidth	100 GB storage area network (SAN)	500 GB SAN
(Optional) BPM, Cognos®	16	4	1 Gbps with 1000 GB bandwidth	100 GB storage area network (SAN)	500 GB SAN

*Table 3: Offering sizes: medium production and non-production*

Tier	Memory (GB)	Number of cores	Network speed and bandwidth	First disk	Second disk
Engine	32	8	1 Gbps with 1000 GB bandwidth	100 GB SAN	1 TB SAN
Service metadata	32	8	1 Gbps with 1000 GB bandwidth	100 GB SAN	1 TB SAN
(Optional) BPM, Cognos	32	8	1 Gbps with 1000 GB bandwidth	100 GB SAN	1 TB SAN

*Table 4: Offering sizes: large production*

Tier	Memory (GB)	Number of cores	Network speed and bandwidth	First disk	Second disk
Engine	64	12	1 Gbps with 5000 GB bandwidth	1.7 TB SSD	1.7 TB SSD
Service metadata	32	8	1 Gbps with 5000 GB bandwidth	960 GB SSD	960 GB SSD
(Optional) BPM, Cognos	32	8	1 Gbps with 5000 GB bandwidth	960 GB SSD	960 GB SSD

Small and medium offerings include the following configuration:

- One virtual server machine with services and repository tiers

- One virtual server machine with the engine tier
- Optional: One virtual server machine with IBM Business Process Manager Standard, and IBM Cognos Business Intelligence.
- One client machine for small offerings and three client machines for medium offerings, including IBM InfoSphere® Data Architect.

The number of client machines is based on the number of concurrent users. For small offerings, two concurrent users are allowed. For medium offerings, five concurrent users are allowed. By default, Microsoft Windows operating system allows two concurrent users to access the machine by using Remote Desktop Connection.

Large-size offering includes the following configuration:

- Two bare metal machines with IBM Information Server on Cloud Enterprise Edition. One machine has the services and repository tiers while the other machine has the engine tier.
- Optional: One bare metal machine with IBM Business Process Manager Standard, and IBM Cognos Business Intelligence.
- Five client machines based on 10 concurrent users, including IBM InfoSphere Data Architect.

## Layout of IBM Information Server on Cloud Enterprise Edition server and client disks

The layout of the Information Server on Cloud Enterprise Edition server and client disks depends on the plan size of your system.

### Virtual servers for small and medium plans

Information Server on Cloud Enterprise Edition comes with two virtual servers. One server has the services and repository tiers and the other server has the engine tier. Optionally, you have a third virtual server with IBM® Business Process Manager Standard, and IBM Cognos® Business Intelligence.

The small and medium plans come with two Storage Area Network (SAN) disks. The Red Hat Enterprise Linux operating system is on the first SAN disk in both the small and medium plans. The second SAN disk is encrypted by using Linux Unified Key Setup (LUKS).

The encryption key details are provided in the Welcome letter from the IBM Sales Representative. It is recommended that you add your own key and remove the supplied key before you use the system.

The product is installed on the /opt directory. User data can be stored on /data directory. Both directories are on the partition /dev/xvdc1 that is encrypted.

Mount point	Encryption status	Partition name	Partition size	Disk	Used for
/boot	None	/dev/xvda1	256 MB	Primary disk is /dev/xvda	Kernel boot area and operating system data
/	None	/dev/xvda2	99.8 GB	Primary disk is /dev/xvda	Kernel boot area and operating system data
/disk1	LUKS	/dev/xvdc1	500 GB for small plan. 1000 GB for medium plan	Primary disk is /dev/xvdc	Product installation. Directories /data and /opt are created on this disk.
SWAP	None	/dev/xvdb1	2 GB	/dev/xvdb	Swap space for paging

## Bare metal server for large plan

The large plan comes with two bare metal machines, and optionally a third one with IBM Business Process Manager Standard and IBM Cognos Business Intelligence. RAID level 1 implementation makes them appear as a single disk.

The disk is divided into four partitions. The Red Hat Enterprise Linux operating system is on a 10 GB partition. The boot data is on a 256 MB partition. The swap space is on a 2 GB partition. The remaining space is on another partition that is encrypted by using LUKS.

Mount point	Encryption status	Partition name	Partition size	Disk	Used for
/boot	None	/dev/sda1	256 MB	/dev/sda	Kernel boot data
SWAP	None	/dev/sda2	2 GB	/dev/sda	Swap space for paging
/	None	/dev/sda3	10 GB	/dev/sda	Operating system data
/disk1	LUKS	/dev/sda5	About 900 GB for the services tier, about 1.6 TB for the engine tier, and optionally about 900 GB for BPM and Cognos	/dev/sda	Product installation. Directories / data, /installables, and /opt are created on this disk.

## Client for all plans

The Information Server on Cloud Enterprise Edition client machine configuration is the same for all plans sizes. The client machine has two Storage Area Network (SAN) disks that are 100 GB each. One disk is drive C for the Microsoft Windows operating system. The other disk is drive F, and it is an empty disk.

## Information roadmap

This roadmap lists information resources that are available for users who are new to the IBM® Information Server on Cloud Enterprise Edition products. These resources provide information about various subject areas.

## Product Overview

- **Introduction to InfoSphere® Information Server** InfoSphere Information Server features an integrated set of product modules, or suite components, including the following: IBM InfoSphere Information Governance Catalog, IBM InfoSphere DataStage®, IBM InfoSphere FastTrack, IBM InfoSphere Information Analyzer, IBM InfoSphere Information Services Director, and IBM InfoSphere QualityStage®.
- **Integration scenarios** Information integration is a complex activity that affects every part of an organization. To address the most common integration business problems, these integration scenarios show how you can deploy and use InfoSphere Information Server and the InfoSphere Foundation Tools components together in an integrated fashion. The integration scenarios focus on data quality within a data warehouse implementation.
- **Release notes** The release notes contain information that is important for the successful installation and use of the product.

## Installing

- **Installing InfoSphere Information Servers software.** This section includes information about planning, preparing the target computers, installing, and configuring the software. It also explains how to create highly available, scalable configurations.
- **Troubleshooting installations.** This section provide descriptions of possible installation problems and the steps to correct them.

- **Migrating to InfoSphere Information Server, Version 11.5** This section describes how to migrate from existing installations to new InfoSphere Information Server, Version 11.5, installations.

## Administering

- **Managing Logs** You can configure the log messages that are generated when activities run in the suite.
- **Managing Schedules** In the IBM InfoSphere Information Server Web console, you can query all of the schedules that are defined across all of the suite components, check their status, history, and forecast, perform maintenance tasks such as purging the schedule execution history, and stop or start existing schedules to prevent system overload.
- **Temporary file cleanup and database optimization** If you are concerned about disk space, you can occasionally clear some files and directories to lower your disk usage. You can also run database tools to ensure good repository database performance.

## Managing Metadata

- **Importing and managing assets by using IBM InfoSphere Metadata Asset Manager** You can use bridges and connectors to import metadata into a staging area, where you can analyze and preview the contents of the import before you share it to the metadata repository. You can browse, search, and manage assets that are in the metadata repository. You can export database assets.
- **Exchanging metadata by using IBM InfoSphere Metadata Interchange Bridges** By using bridges you can import metadata from a wide range of sources, including IBM Cognos, SAP BusinessObjects, CA ERwin, and other tools. The prerequisites and import parameters for bridges are documented in detail in this section.
- **Common metadata assets** Common metadata assets are stored in the metadata repository and shared between tools in the InfoSphere™ Information Server suite. Common metadata assets include logical and physical data model assets, implemented data resources, and business intelligence (BI) assets.
- **Managing assets by using the command line** You can use the command line to move assets between different metadata repositories for environments such as development, test, and production. You can also query and delete common metadata assets and generate glossary assets from BI assets and logical data models.

## Getting Started with Consoles

- **Opening consoles and clients by using the Launchpad** The Launchpad is a standard, single web interface for opening the various clients or consoles for IBM InfoSphere Information Server.
- **Opening consoles and clients without using the Launchpad** You can open the various clients or consoles for InfoSphere Information Server by using their URL.
- **Automatic login from web-based applications** The login page of the IBM InfoSphere Information Server console and web-based applications contain an option that enables you to log in automatically in your subsequent attempts.
- **IBM InfoSphere Information Server console overview** The IBM InfoSphere Information Server console is a rich-client-based interface for activities such as creating and managing projects, setting project-level security, analyzing data with IBM InfoSphere Information Analyzer, enabling information services with IBM InfoSphere Information Services Director, and running reports.

## Connecting to data sources

- **Connecting to data sources** IBM InfoSphere Information Server connectivity options enable jobs to transfer data between InfoSphere Information Server and data sources.

## Getting started and using IBM Information Server on Cloud Enterprise Edition

---

You must set up your connection to Information Server on Cloud Enterprise Edition. Information Server on Cloud Enterprise Edition provides all of the functions of its on-premises counterpart, IBM® InfoSphere® Information Server. It is in an IBM SoftLayer® hosted environment.

**Prerequisite:** You must know the IP address and the credentials of an account on the Information Server on Cloud Enterprise Edition server and client computers. This information is in the “Access Credentials and Initial Setup” section of the Welcome letter that you received from your IBM Sales Representative.

The Information Server on Cloud Enterprise Edition client is on a Microsoft Windows virtual machine that is hosted on SoftLayer. When you connect to the client, you can access the client user interfaces. McAfee anti-virus software is installed on the client machine.

The Information Server on Cloud Enterprise Edition servers are on Red Hat Enterprise Linux virtual or bare metal computers that are hosted on SoftLayer. The services and repository tiers are on one computer. The engine tier is on the other computer. Optionally, IBM Business Process Manager Standard and IBM Cognos® Business Intelligence are on the separate Red Hat Enterprise Linux virtual or bare metal computers. When you connect to the servers, you can access the IBM InfoSphere Information Server engine, services, and repository tiers. You can restart Information Server on Cloud Enterprise Edition and do administrative tasks.

The default firewall configuration of server machines allows SSH connections only from client machines. You must first connect to a client machine by using a remote desktop connection, and then from the client machine you can connect to server machines by using SSH. After you log in to a server machine, you can change the firewall configurations to allow SSH connections from other machines. Communication between the server and client systems happens through a private IP. If you want to access the server from an on-premises client machine, you must modify the iptable rules.

**Note:** The Add Subscription window in Subscription Manager is disabled when you use a host name in the URL. To enable the window, use an IP address in the Subscription Manager URL.

## When you connect for the first time, follow these steps:

1. From your local computer, connect to a client machine by using Remote Desktop Connection. See [step 1.e](#) for details.
2. To modify the firewall to allow SSH access to the servers from your local computer, [set up firewall security](#).
3. On your local computer, use an SSH client to connect to the server machines. For example, you can use PuTTY to connect to the servers. The user ID, password, and IP address of the servers are in your Welcome letter.
4. On the Information Server on Cloud Enterprise Edition servers, [run the ISALite tool](#) on all tiers. The *IS\_install\_path* for the server computers is `/opt/IBM/InformationServer`.
5. Connect the Information Server on Cloud Enterprise Edition client to the Information Server on Cloud Enterprise Edition servers by following these steps:
  - **If your local computer is in a Microsoft Windows environment**
    - a. On your local computer, go to the **Start** menu. Click **Accessories > Remote Desktop Connection**.
    - b. Enter the IP address of the Microsoft Windows computer that hosts your Information Server on Cloud Enterprise Edition client. Click **Connect**.
    - c. In the Windows Security window, enter the user name and password for the Information Server on Cloud Enterprise Edition client. The user ID, password, and IP address of the client are in your Welcome letter.  
**Important:** Do not include the domain name with the user name.
    - d. In the Information Server on Cloud Enterprise Edition client, open the file `C:\Windows\System32\drivers\etc\hosts`. Make sure that entries with the private IP exist for the Information Server on Cloud Enterprise Edition servers that you are connecting to. **Important:** The server IP must be a private IP. You cannot connect to Information Server on Cloud Enterprise Edition servers when you use a public IP.
  - **If your local computer is in an Apple Mac environment**
    - a. On your local computer, install Microsoft Remote Desktop from the Apple App Store.
    - b. Click the Microsoft Remote Desktop icon, and then click **Open**.
    - c. In the Microsoft Remote Desktop window, click **New**.
    - d. In the Edit Remote Desktops window, supply the following information:
      - In the PC name field, type in the IP address of the cloud client machine.
      - In the User name and Password fields, type in the Windows user name and password that are in the Welcome letter.
6. Verify the connection and installation on the Information Server on Cloud Enterprise Edition client by following these steps:
  - [Run the ISALite tool](#). The *IS\_install\_path* for the client computer is `C:\IBM\InformationServer`.
  - [Test the installation](#) of the Information Server on Cloud Enterprise Edition client.

7. Optional: Enable multiple users to open remote sessions to the Information Server on Cloud Enterprise Edition client by following these steps on the client computer:
  - a. [Create user accounts](#).
  - b. [Give users permission to do a remote desktop connection](#). Each Windows machine allows two concurrent user sessions. The number of Windows client machines that come with different sizes of your Information Server on Cloud Enterprise Edition offering are based on the number of concurrent users that are allowed in that offering. The small offering size has a maximum of two concurrent sessions. The medium offering size has a maximum of five concurrent sessions. The large offering size has a maximum of 10 concurrent sessions.
8. Reset the password for users and administrators on the Information Server on Cloud Enterprise Edition servers.
9. Optional: If you choose to include IBM Business Process Manager Standard, or IBM Cognos Business Intelligence in your environment, you must configure them. For details, see:
  - [IBM Business Process Manager Standard: Installing and configuring IBM BPM Standard](#)
  - [IBM Cognos Business Intelligence](#):
    - [IBM InfoSphere Information Governance Dashboard](#)
    - [Configure the IBM InfoSphere Business Glossary URI](#)

**Note:** IBM Business Process Manager Standard Process Center and Process Server are installed and running. If you do not plan to use both of them, stop the service.

### After the initial connection, you can do any of the following tasks:

- [Open the InfoSphere DataStage® and QualityStage® Designer client](#)
- [Connect to an on-premises computer](#)
- [Connect to an IBM Db2 on Cloud database](#)
- [Connect to an on-premises Db2® database instance](#)
- [Perform general administration and security tasks](#)
- Open Information Server on Cloud Enterprise Edition clients from the client machine by following either method.
  - For browser clients, open Microsoft Internet Explorer. In the Favorites Bar, click the **IIS\_launchpad** bookmark. From the Launchpad window, click the icon of the client of InfoSphere Information Server.
  - For thick clients, on your desktop click **Start > Programs > IBM Infosphere Information Server > client-name**.

You can connect to any of the following InfoSphere Information Server clients that were provisioned with their server component:

- IBM InfoSphere Information Governance Catalog
- IBM InfoSphere Information Analyzer
- IBM InfoSphere Metadata Integration Bridges and the metadata interchange agent
- IBM InfoSphere Metadata Asset Manager
- IBM InfoSphere Information Server istool command-line utility
- IBM InfoSphere Information Server Manager client
- Multi-Client Manager
- IBM InfoSphere DataStage and QualityStage Administrator
- IBM InfoSphere DataStage and QualityStage Designer
- IBM InfoSphere DataStage and QualityStage Director.

### Related information

- [Enhancing security of Information Server on Cloud computers](#)



---

## Chapter 3. Information Server on Cloud Data Quality

IBM® Information Server on Cloud Data Quality is included in Information Server on Cloud Enterprise Edition, but it can be used as an independent service. Use Information Server on Cloud Data Quality to cleanse data and monitor data quality.

### Overview

---

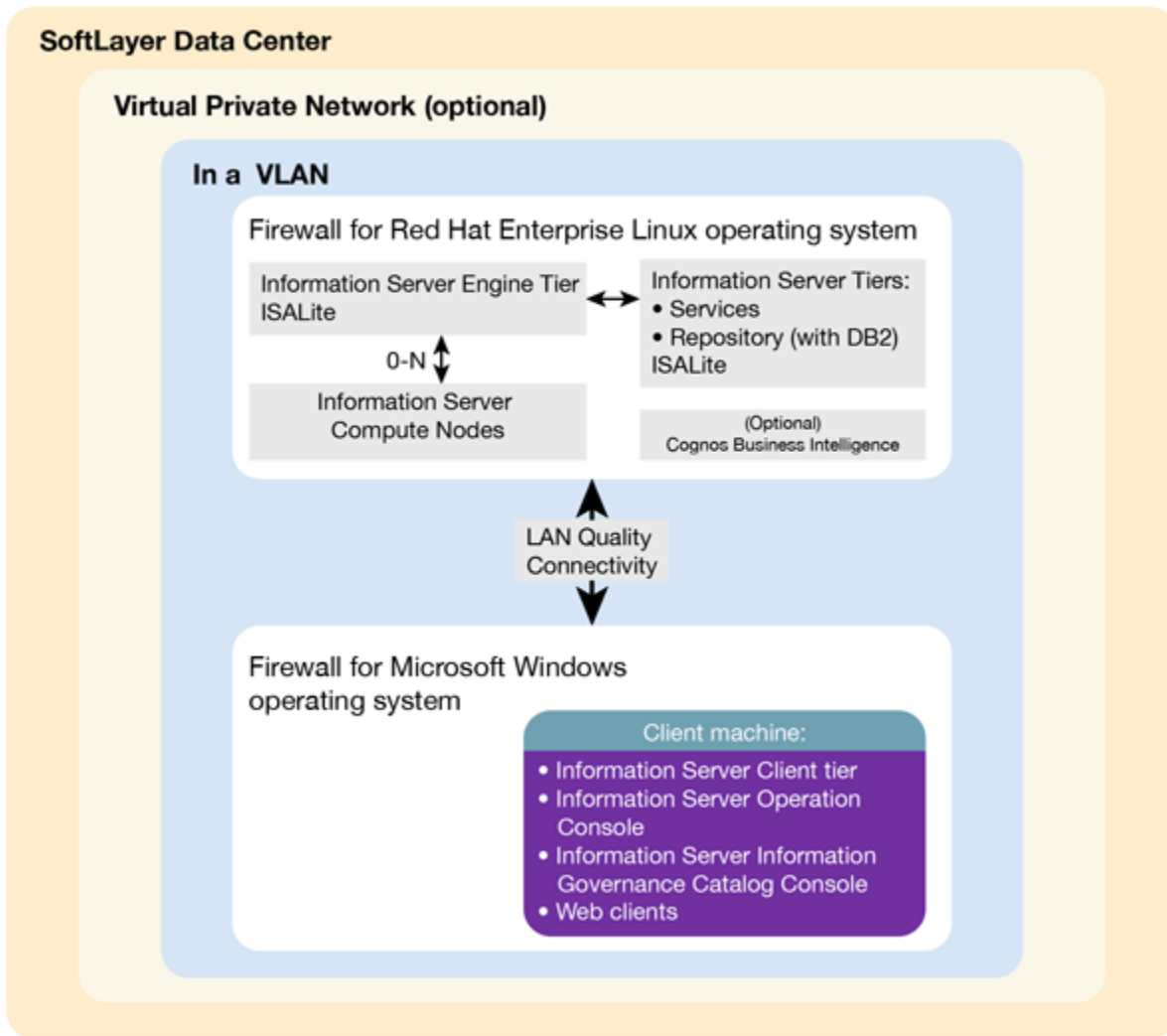
IBM® Information Server on Cloud Data Quality is a data integration software platform that helps organizations derive more value from the complex, heterogeneous information spread across their systems. You can establish and manage high-quality data by using Information Server on Cloud Data Quality. You can cleanse data and monitor data quality on an ongoing basis to turn your data into trusted information.

Information Server on Cloud Data Quality helps you to do the following tasks:

- Deliver customizable data-cleansing functions in batch and in near real time
- Monitor and maintain data quality
- Provide a unified environment and support for advanced data validation rules

Information Server on Cloud Data Quality uses the characteristics of software-as-a-service (SaaS). You select the plan size based on your needs. IBM provisions the machines and deploys the Information Server on Cloud Data Quality software.

The following figure shows the topology of the server and client machines in a typical deployment.



As a hosted offering, you have the same control over your data in the cloud as in the on-premises system:

- Actively monitor and report any issues that you encounter with IBM Software as a Service (SaaS).
- Maintain the software platform of your cloud offering and the operating system to meet your security standards.
- Maintain software firewalls on servers that face the internet in a manner to provide required protection.
- Develop parallel jobs to transform and cleanse data, and develop server jobs to transform data. Establish connectivity between data sources and applications. Develop your own workload, business rules, monitoring, and scheduling for all jobs. You are responsible for the quality and performance of programs, applications, and jobs that you develop.
- Ensure the continuity, compatibility, and performance of the IBM SaaS platform by installing only permissible software, including any open source packages.
- Regularly upgrade the environment and operating system of your cloud offering.
- Create and maintain regular backups of data.
- Create and maintain high availability configurations.

The following managed add-on services are available to maintain and manage the infrastructure:

**Jump start** This setup service provides up to 50 hours of remote consulting time for startup activities.

**Accelerator** This setup service provides up to 50 hours of remote consulting time to perform various scoped activities.

**Silver** This service provides monthly remote consulting time for operations and maintenance activities.

**Gold** This service provides monthly remote consulting time for operations and maintenance activities. The service includes everything that is provided by the Silver service and delivers extra activities.

Note the following limitations and restrictions of Information Server on Cloud Data Quality:

- If your offering is designated as “Non-Production”, Information Server on Cloud Data Quality can be deployed only as part of your development and test environments for internal non-production activities. These activities include, but are not limited to: testing, performance tuning, fault diagnosis, internal benchmarking, staging, quality assurance activity, developing internally used additions or extensions to the offering by using published application programming interfaces.
- Users must not modify the configuration file that is needed to run the job for parallel processing.

## Available configurations

IBM® Information Server on Cloud Data Quality servers for the small and medium plans are virtual servers with dedicated CPUs. The servers in the large plan are in a bare metal environment.

Select the offering plan that fits your usage and environment needs.

*Table 7: Offering sizes: small production and non-production*

Tier	Memory (GB)	Number of cores	Network speed and bandwidth	First disk	Second disk
Engine	16	4	1 Gbps with 250 GB bandwidth	100 GB storage area network (SAN)	500 GB SAN
Service metadata	16	4	1 Gbps with 250 GB bandwidth	100 GB storage area network (SAN)	500 GB SAN
(Optional) Cognos	16	4	1 Gbps with 1000 GB bandwidth	100 GB storage area network (SAN)	500 GB SAN

*Table 8: Offering sizes: medium production and non-production*

Tier	Memory (GB)	Number of cores	Network speed and bandwidth	First disk	Second disk
Engine	32	8	1 Gbps with 1000 GB bandwidth	100 GB SAN	1 TB SAN
Service metadata	32	8	1 Gbps with 1000 GB bandwidth	100 GB SAN	1 TB SAN
(Optional) Cognos	32	8	1 Gbps with 1000 GB bandwidth	100 GB SAN	1 TB SAN

*Table 9: Offering sizes: large production*

Tier	Memory (GB)	Number of cores	Network speed and bandwidth	First disk	Second disk
Engine	64	12	1 Gbps with 5000 GB bandwidth	1.7 TB SSD	1.7 TB SSD
Service metadata	32	8	1 Gbps with 5000 GB bandwidth	960 GB SSD	960 GB SSD
(Optional) Cognos	32	8	1 Gbps with 5000 GB bandwidth	960 GB SSD	960 GB SSD

Small and medium offerings include the following configuration:

- One virtual server machine with services and repository tiers.

- One virtual server machine with the engine tier.
- Optional: One virtual machine with IBM Cognos® Business Intelligence.
- One client machine for small offerings and three client machines for medium offerings.

The number of client machines is based on the number of concurrent users. For small offerings, two concurrent users are allowed. For medium offerings, five concurrent users are allowed. By default, Microsoft Windows operating system allows two concurrent users to access the machine by using Remote Desktop Connection.

Large-size offering includes the following configuration:

- Two bare metal machines with Information Server on Cloud Data Quality. One machine has the services and repository tiers while the other machine has the engine tier.
- Optional: One bare metal machine with IBM Cognos Business Intelligence.
- Five client machines based on 10 concurrent users.

## Layout of IBM Information Server on Cloud Data Quality server and client disks

The layout of the Information Server on Cloud Data Quality server and client disks depends on the plan size of your system.

### Virtual servers for small and medium plans

Information Server on Cloud Data Quality comes with two virtual servers. One server has the services and repository tiers and the other server has the engine tier. Optionally, you have a third virtual server with IBM® Cognos® Business Intelligence.

The small and medium plans come with two Storage Area Network (SAN) disks. The Red Hat Enterprise Linux operating system is on the first SAN disk in both the small and medium plans. The second SAN disk is encrypted by using Linux Unified Key Setup (LUKS).

The encryption key details are provided in the Welcome letter from the IBM Sales Representative. It is recommended that you add your own key and remove the supplied key before you use the system.

The product is installed on the /opt directory. User data can be stored on /data directory. Both directories are on the partition /dev/xvdc1 that is encrypted.

*Table 10: Small and medium disk layout*

Mount point	Encryption status	Partition name	Partition size	Disk	Used for
/boot	None	/dev/xvda1	256 MB	Primary disk is /dev/xvda	Kernel boot area and operating system data
/	None	/dev/xvda2	99.8 GB	Primary disk is /dev/xvda	Kernel boot area and operating system data
/disk1	LUKS	/dev/xvdc1	500 GB for small plan. 1000 GB for medium plan	Secondary disk is /dev/xvdc	Product installation. Directories /data and /opt are created on this disk.
SWAP	None	/dev/xvdb1	2 GB	/dev/xvdb	Swap space for paging

### Bare metal server for large plan

The large plan also comes with two servers. One server has a Solid State Drive (SSD) disk of about 960 GB for the services tier. The other server has an SSD disk of about 1.6 TB for the engine tier. Optionally, you can have additional

server with IBM Cognos Business Intelligence with Solid State Drive (SSD) disk of about 960 GB. RAID level 1 implementation makes them appear as a single disk.

The disk is divided into four partitions. The Red Hat Enterprise Linux operating system is on a 10 GB partition. The boot data is on a 256 MB partition. The swap space is on a 2 GB partition. The remaining space is on another partition that is encrypted by using LUKS.

*Table 11: Bare metal disk layout*

Mount point	Encryption status	Partition name	Partition size	Disk	Used for
/boot	None	/dev/sda1	256 MB	/dev/sda	Kernel boot data
SWAP	None	/dev/sda2	2 GB	/dev/sda	Swap space for paging
/	None	/dev/sda3	10 GB	/dev/sda	Operating system data
/disk1	LUKS	/dev/sda5	About 900 GB for the services tier, about 1.6 TB for the engine tier, and optionally about 900 GB for Cognos	/dev/sda	Product installation. Directories / data, / installables, and /opt are created on this disk.

#### Client for all plans

The Information Server on Cloud Data Quality client machine configuration is the same for all plans. The client machine has two SAN disks that are 100 GB each. One disk is drive C for the Microsoft Windows operating system. The other disk is drive F, and it is an empty disk.

## Information roadmap

This roadmap lists information resources that are available for users who are new to the data quality products from IBM®. These products include IBM InfoSphere® DataStage® and QualityStage®, IBM InfoSphere Information Analyzer, and IBM InfoSphere Information Governance Catalog. These resources provide information about various subject areas, such as learning basic skills and troubleshooting.

### IBM InfoSphere DataStage and QualityStage

- **Getting started with InfoSphere DataStage and InfoSphere QualityStage** Use tutorials to learn the skills that you need to develop parallel jobs that transform data and parallel jobs that cleanse data.
- **Designing DataStage and QualityStage jobs** You design InfoSphere DataStage and QualityStage jobs using the IBM InfoSphere DataStage and QualityStage Designer client.
- **Cleansing data with InfoSphere QualityStage jobs** The cleansing process can include, but is not limited to, eliminating redundant, obsolete, or inaccurate data. Clean data is a critical component for accurate information, reports, and analyses. Throughout your organization, people make business decisions based on data that is provided to them. By cleansing data, you provide high-quality data.
- **Standardizing address data with InfoSphere QualityStage** IBM InfoSphere QualityStage add-on modules standardize address records and match those records to postal validation reference files. These modules let you concentrate on the quality of your address data so that you do not have to become an expert in international postal standards.
- **Running InfoSphere DataStage and QualityStage jobs** You run your InfoSphere DataStage and QualityStage jobs from the InfoSphere DataStage and QualityStage Director client.

- **Monitoring jobs** You can use the IBM InfoSphere DataStage and QualityStage Operations Console to monitor the job runs, services, system resources, and workload management queues on several IBM InfoSphere Information Server engines.
- **Administering workload management** You can use the workload management queues to control the starting of parallel and server jobs.
- **Administering projects** IBM InfoSphere DataStage and QualityStage jobs are organized in projects, along with associated design items.

## IBM InfoSphere Information Analyzer

- **Methodology and best practices** Use IBM InfoSphere Information Analyzer to understand the content, structure, and overall quality of your data at a given point in time.
- **Configuring InfoSphere Information Analyzer** After InfoSphere Information Analyzer is installed, you must complete additional configuration steps before you can use it. You should also validate any settings created during installation.
- **Managing metadata** After you import metadata into the metadata repository, you can add or modify information about the metadata such as the description of a table. You can also add information such as contacts, policies, and terms in IBM InfoSphere Information Governance Catalog to the imported schemas, directories, tables, files, and data fields.
- **Analyzing data with the on-premise client** Before you can begin analyzing data, you must understand the project environment, open a project, review the dashboard, and decide which analysis jobs or analysis reviews you want to conduct.
- **Analyzing data with InfoSphere Information Analyzer thin client** You can analyze data sets, and view and edit analysis results in a browser. You can also view data quality scores for data sets and columns, view data rules and quality rules associated with data sets, and create or delete quality rules.
- **Validating data by using data rules** You can define and run data rules, which evaluate or validate specific conditions associated with your data sources. Data rules can be used to extend your data profiling analysis, to test and evaluate data quality, or to improve your understanding of data integration requirements.
- **Publishing and transferring analysis results** You can view analysis results and publish it to the metadata repository. The analysis results can also be transferred into InfoSphere Optim™ and InfoSphere Guardium®.
- **Importing and exporting projects from the client** You can import and export projects and project assets and move them between instances of the IBM InfoSphere Information Server metadata repository by using the InfoSphere Information Analyzer import and export wizards.
- **Running and scheduling InfoSphere Information Analyzer jobs** You can run and schedule IBM InfoSphere Information Analyzer jobs outside of the InfoSphere Information Analyzer client.
- **Managing tables** You can manage reference tables and tables that are generated during analysis. You can open a table, delete a table, export a table to a file, add a table to the project, and view the metadata in a table.
- **Developing InfoSphere Information Analyzer applications with the HTTP API** A Hypertext Transfer Protocol (HTTP) application programming interface (API) is provided with IBM InfoSphere Information Analyzer. You can develop applications with this API to access and analyze InfoSphere Information Analyzer content.
- **Reports for information analysis** You can create reports that summarize analysis results and show details about your project. Reports are saved in the metadata repository and can be accessed by any user who is authorized to view them.

## IBM InfoSphere Information Governance Catalog

- **Overview** Learn about the features and benefits of InfoSphere Information Governance Catalog.
- **Designing the catalog** A catalog is an authoritative dictionary of the assets that are used throughout the organization. One of the main benefits of a well-designed catalog is increased trust and confidence in organization information. Planning, designing, and publishing a catalog involves several tasks.
- **Administering the catalog** The catalog is composed of glossary assets (terms, categories, information governance policies, and information governance rules) and information assets. You can assign security roles and permissions to users to control access to the catalog. You can also assign workflow roles, assign users as stewards, define custom attributes, and define external asset types. In addition, you can configure the display of glossary assets.
- **Governing your data** You can create, edit, and delete catalog assets. Some assets are created when they are imported into the catalog. You can assign stewards and assets, define and assign custom attributes, and extend data flows for lineage reports by importing assets. In addition, you can configure assets for lineage analysis reports. Assets of different types that have a common business purpose can be grouped in a collection.

- **Viewing catalog content** You can view, search, browse, and query the catalog to find catalog assets.
- **Expanding catalog capabilities**
- Look up terms while you work in other applications from the Microsoft Windows desktop by using Glossary Anywhere.
- Develop and extend applications by using InfoSphere Information Governance Catalog for Eclipse. You can access your glossary content from within your Eclipse-based development tool.
- Write client applications to access and author catalog content by using InfoSphere Information Governance Catalog REST API. You can integrate your catalog content into other software tools and portals.

## Getting started and using IBM Information Server on Cloud Data Quality

You must set up your connection to Information Server on Cloud Data Quality. Information Server on Cloud Data Quality provides all of the functions of its on-premises counterparts, IBM® InfoSphere® Information Analyzer, IBM InfoSphere Information Governance Catalog, and IBM InfoSphere DataStage® and QualityStage®. It is in an IBM SoftLayer® hosted environment.

**Prerequisite:** You must know the IP address and the credentials of an account on the Information Server on Cloud Data Quality server and client computers. This information is in the “Access Credentials and Initial Setup” section of the Welcome letter that you received from your IBM Sales Representative.

The Information Server on Cloud Data Quality client is on a Microsoft Windows virtual machine that is hosted on SoftLayer. When you connect to the client, you can access the client user interfaces. McAfee anti-virus software is installed on the client machine.

The Information Server on Cloud Data Quality servers are on Red Hat Enterprise Linux virtual or bare metal computers that are hosted on SoftLayer. The services and repository tiers are on one computer. The engine tier is on the other computer. Optionally, IBM Cognos® Business Intelligence is on a separate Red Hat Enterprise Linux virtual or bare metal computer. When you connect to the servers, you can access the IBM InfoSphere Information Server engine, services, and repository tiers. You can restart Information Server on Cloud Data Quality and do administrative tasks.

The default firewall configuration of server machines allows SSH connections only from client machines. You must first connect to a client machine by using a remote desktop connection, and then from the client machine you can connect to server machines by using SSH. After you log in to a server machine, you can change the firewall configurations to allow SSH connections from other machines. Communication between the server and client systems happens through a private IP. If you want to access the server from an on-premises client machine, you must modify the iptable rules.

**Note:** The Add Subscription window in Subscription Manager is disabled when you use a host name in the URL. To enable the window, use an IP address in the Subscription Manager URL.

### When you connect for the first time, follow these steps:

1. From your local computer, connect to a client machine by using Remote Desktop Connection. For details, see [step 1.e](#).
2. To modify the firewall to allow SSH access to the servers from your local computer, [set up firewall security](#).
3. On your local computer, use an SSH client to connect to the server machines. For example, you can use PuTTY to connect to the servers. The user ID, password, and IP address of the servers are in your Welcome letter.
4. On the Information Server on Cloud Data Quality servers, [run the ISALite tool](#). The `IS_install_path` for the server computers is `/opt/IBM/InformationServer`
5. Connect the Information Server on Cloud Data Quality client to the Information Server on Cloud Data Quality servers by following these steps:
  - **If your local computer is in a Microsoft Windows environment**
    - a. On your local computer, go to the **Start** menu. Click **Accessories > Remote Desktop Connection**.
    - b. Enter the IP address of the Microsoft Windows computer that hosts your Information Server on Cloud Data Quality client. Click **Connect**.
    - c. In the Windows Security window, enter the user name and password for the Information Server on Cloud Data Quality client. The user ID, password, and IP address of the client are in your Welcome letter. **Important:** Do not include the domain name with the user name.

- d. In the Information Server on Cloud Data Quality client, open the file C:\Windows\System32\drivers\etc\hosts. Make sure that an entry with the private IP exists in the file for the Information Server on Cloud Data Quality servers that you are connecting to. **Important** The server IP must be a private IP. You cannot open other Information Server on Cloud Data Quality clients when you use a public IP.
- **If your local computer is in an Apple Mac environment**
  - a. On your local computer, install Microsoft Remote Desktop from the Apple App Store.
  - b. Click the Microsoft Remote Desktop icon, and then click **Open**.
  - c. In the Microsoft Remote Desktop window, click **New**.
  - d. In the Edit Remote Desktops window, supply the following information:
    - In the PC name field, type in the IP address of the cloud client machine.
    - In the User name and Password fields, type in the Windows user name and password that are in the Welcome letter.
6. Verify the connection and installation on the Information Server on Cloud Data Quality client by following these steps:
  - Run the ISALite tool. The *IS\_install\_path* for the client computer is C:\IBM\InformationServer.
  - Test the installation of the Information Server on Cloud Data Quality client.
7. Optional: Enable multiple users to open remote sessions to the Information Server on Cloud Data Quality client by following these steps on the client computer:
  - a. Create user accounts.
  - b. Give users permission to do a remote desktop connection. The number of concurrent users in a remote session to the client is based on the offering size of your Information Server on Cloud Data Quality. The small offering size has a maximum of two concurrent sessions. The medium offering size has a maximum of five concurrent sessions. The large offering size has a maximum of 10 concurrent sessions.
8. Reset the password for users and administrators on the Information Server on Cloud Data Quality servers.
9. Optional: If you choose to include IBM Cognos Business Intelligence in your environment, you must configure it. For details, see:
  - IBM InfoSphere Information Governance Dashboard
  - Configure the IBM InfoSphere Business Glossary URI

### After the initial connection, you can do any of the following tasks:

- Open the InfoSphere DataStage® and QualityStage® Designer client
- Connect to an on-premises computer
- Connect to an IBM Db2 on Cloud database
- Connect to an on-premises Db2® database instance
- Perform general administration and security tasks
- Open Information Server on Cloud Enterprise Edition clients from the client machine by following either method.
  - For browser clients, open Microsoft Internet Explorer. In the Favorites Bar, click the **IIS\_launchpad** bookmark. From the Launchpad window, click the icon of the client of InfoSphere Information Server.
  - For thick clients, on your desktop click **Start > Programs > IBM Infosphere Information Server > client-name**.

You can connect to any of the following InfoSphere Information Server clients that were provisioned with their server component:

- IBM InfoSphere Information Governance Catalog
- IBM InfoSphere Information Analyzer
- IBM InfoSphere Metadata Integration Bridges and the metadata interchange agent
- IBM InfoSphere Metadata Asset Manager
- IBM InfoSphere Information Server istool command-line utility
- IBM InfoSphere Information Server Manager client
- Multi-Client Manager
- IBM InfoSphere DataStage and QualityStage Administrator
- IBM InfoSphere DataStage and QualityStage Designer



- IBM InfoSphere DataStage and QualityStage Director.

**Related information**

- [Enhancing security of Information Server on Cloud computers](#)



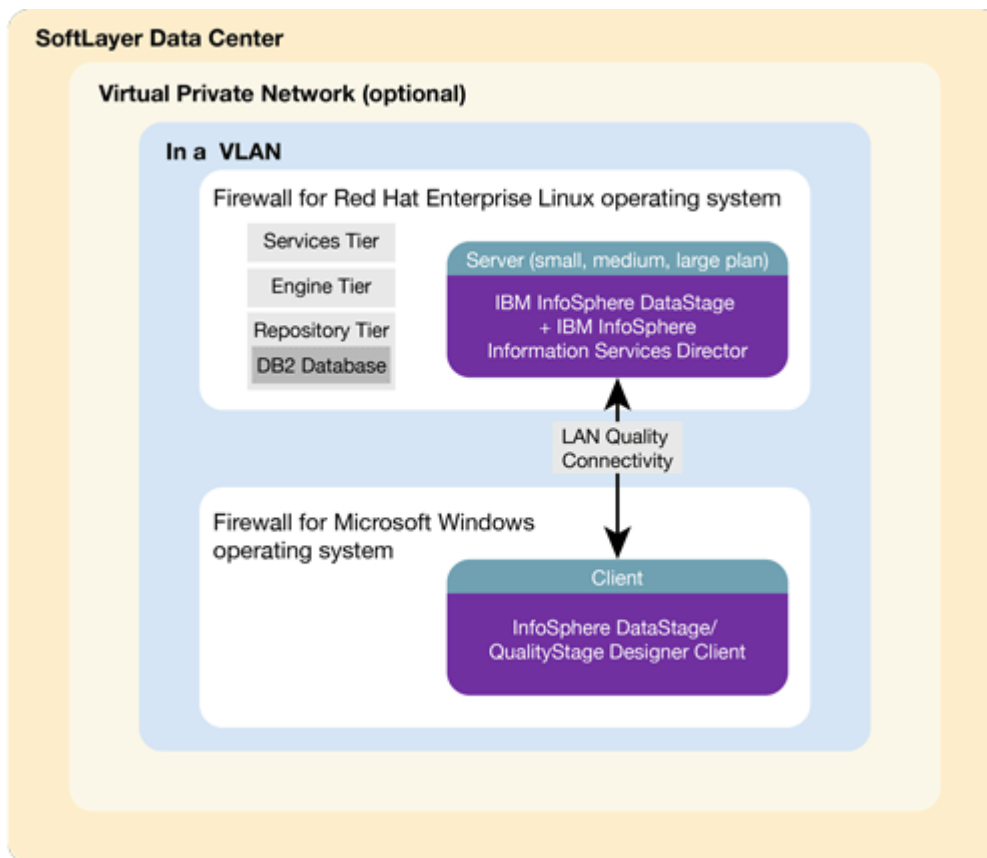
# Chapter 4. DataStage on Cloud

IBM® DataStage® on Cloud provides a hosted environment that you configure and control. You can use DataStage on Cloud to extend the reach of your business by leveraging cloud offerings, while you reduce the costs that are associated with providing these services. DataStage on Cloud provides several different plans so that any size business can access the powerful and scalable ETL platform by IBM.

## Overview

IBM® DataStage® on Cloud is a data integration tool for designing, developing, and running jobs that move and transform data. DataStage on Cloud provides all of the functions of its on-premises counterpart, IBM InfoSphere® DataStage.

DataStage on Cloud combines some of the characteristics of infrastructure-as-a-service (IaaS) and platform-as-a-service (PaaS). You select the plan size based on your needs. IBM provisions the machine and deploys the DataStage on Cloud software.



As a hosted offering, you have the same control over your data in the cloud as in the on-premises system:

- Actively monitor and report any issues that you encounter with IBM Software as a Service (SaaS).
- Maintain the software platform of your cloud offering and the operating system to meet your security standards.
- Maintain software firewalls on servers that face the internet in a manner to provide required protection.
- Develop integration, transformation, and other jobs. Establish connectivity between data sources and applications. You can also develop your own workload, business rules, monitoring, and scheduling for all jobs. You are responsible for the quality and performance of programs, applications, and jobs that you develop.
- Ensure the continuity, compatibility, and performance of the IBM SaaS platform by installing only permissible software, including any open source packages.

- Regularly upgrade the environment and operating system of your cloud offering.
- Create and maintain regular backups of data.
- Create and maintain high availability configurations.

The following managed add-on services are available to maintain and manage the infrastructure:

**Jump start** This setup service provides up to 50 hours of remote consulting time for startup activities.

**Accelerator** This setup service provides up to 50 hours of remote consulting time to perform various scoped activities.

**Silver** This service provides monthly remote consulting time for operations and maintenance activities.

**Gold** This service provides monthly remote consulting time for operations and maintenance activities. The service includes everything that is provided by the Silver service and delivers extra activities.

Note the following limitations and restrictions of DataStage on Cloud:

- If your offering is designated as “Non-Production,” DataStage on Cloud can be deployed only as part of your development and test environments for internal non-production activities. These activities include, but are not limited to: testing, performance tuning, fault diagnosis, internal benchmarking, staging, quality assurance activity, developing internally used additions or extensions to the offering by using published application programming interfaces.
- Users must not modify the configuration file that is needed to run the job for parallel processing.
- Usage of IBM InfoSphere Information Services Director with DataStage on Cloud has the following restrictions:
  - Batch jobs are restricted to Topology I batch jobs. For details, see [Designing IBM InfoSphere DataStage and QualityStage® jobs as services](#).
  - Restricted to SOAP over HTTP as service binding.

## Available configurations

IBM® DataStage® on Cloud servers for the small and medium plans are virtual servers with dedicated CPUs. The servers in the large plan are in a bare metal environment.

Select the offering plan that fits your usage and environment needs.

*Table 12: DataStage on Cloud available configurations*

Offering	Environment	Memory (GB)	Number of Cores	Network speed (Gbps)	First disk	Second disk
Small	Production Non-production	16	4	1	100 GB Storage area network (SAN)	500 GB SAN
Medium	Production Non-production	32	8	1	100 GB SAN	1 TB SAN
Large	Production	64	12	1	2-960 GB SSD	

Small and medium offerings include the following configuration:

- One virtual machine with DataStage on Cloud, IBM InfoSphere® Information Services Director
- One virtual machine with a Microsoft Windows operating system and all client utilities installed

Large-size offering includes the following configuration:

- One bare metal machine with DataStage on Cloud, IBM InfoSphere Information Services Director
- Two virtual machines with a Microsoft Windows operating system and all client utilities installed

## Layout of IBM DataStage on Cloud server and client disks

The layout of the DataStage® on Cloud server and client disks depends on the plan size of your system.

### Virtual server for small and medium plans

The small and medium plans come with two Storage Area Network (SAN) disks. The Red Hat Enterprise Linux operating system is on the first SAN disk in both the small and medium plans. The second SAN disk is encrypted by using Linux Unified Key Setup (LUKS) for production offerings of both these sizes.

The encryption key details are provided in the Welcome letter from the IBM® Sales Representative. It is recommended that you add your own key and remove the supplied key before you use the system.

The product is installed on the /opt directory. User data can be stored on /data directory. Both directories are on the partition /dev/xvdc1 that is encrypted for production grade parts.

*Table 13: Small and medium server disk layout*

Mount point	Encryption status	Partition name	Partition size	Disk	Used for
/boot	None	/dev/xvda1	256 MB	Primary disk is /dev/xvda	Kernel boot area and operating system data
/	None	/dev/xvda2	99.8 GB	Primary disk is /dev/xvda	Kernel boot area and operating system data
/disk1	LUKS	/dev/xvdc1	500 GB for small plan. 1000 GB for medium plan	Secondary disk is /dev/xvdc	Product installation. Directories /data and /opt are created on this disk.
SWAP	None	/dev/xvdb1	2 GB	/dev/xvdb	Swap space for paging

### Bare metal server for large plan

The large plan comes with two Solid State Drive (SSD) disks that are 960 GB each. RAID level 1 implementation makes them appear as a single disk.

The disk is divided into four partitions. The Red Hat Enterprise Linux operating system is on a 10 GB partition. The boot data is on a 256 MB partition. The swap space is on a 2 GB partition. The remaining space is on another partition that is encrypted by using LUKS.

*Table 14: Bare metal server disk layout*

Mount point	Encryption status	Partition name	Partition size	Disk	Used for
/boot	None	/dev/sda1	256 MB	/dev/sda	Kernel boot data
SWAP	None	/dev/sda2	2 GB	/dev/sda	Swap space for paging
/	None	/dev/sda3	10 GB	/dev/sda	Operating system data

Table 14: Bare metal server disk layout (continued)

Mount point	Encryption status	Partition name	Partition size	Disk	Used for
/disk1	LUKS	/dev/sda5	About 900 GB	/dev/sda	Product installation. Directories /data, /installables, and /opt are created on this disk.

### Client for all plans

The DataStage on Cloud client machine configuration is the same for all plans sizes. The client machine has two SAN disks that are 100 GB each. One disk is C: for the Microsoft Windows operating system. The second disk is F:, and it is an empty disk.

## Information roadmap

This roadmap lists information resources that are available for users who are new to the DataStage® products from IBM®. These resources provide information about various subject areas, such as learning basic skills and troubleshooting.

### Product overview

- **InfoSphere® DataStage features** Learn about the features and benefits of InfoSphere DataStage.
- **InfoSphere DataStage product documentation** The InfoSphere Information Server Knowledge Center provides you with InfoSphere DataStage concepts and usage information to help new users prepare for using your new system.
- **InfoSphere DataStage developerWorks® forum** Use this forum to interact with other InfoSphere DataStage users to better understand how to design, build, debug and deploy jobs for information collection, integration and transformation.

### Getting started

- **Overview of InfoSphere DataStage** This overview covers InfoSphere DataStage job life cycles, job designs, and how the product integrates with the rest of the InfoSphere Information Server suite.
- **Tutorial: Creating parallel jobs** This tutorial demonstrates how you can use InfoSphere DataStage to develop jobs that extract, transform, and load data. By transforming and cleansing the source data and applying consistent formatting, you enhance the quality of the data.

### Using InfoSphere DataStage

- **Designing InfoSphere DataStage jobs** This topic explains how to design InfoSphere DataStage jobs using the IBM InfoSphere DataStage and QualityStage® Designer client. The Designer client gives you the tools that you need to create jobs that extract, transform, load, and check the quality of data.
- **Developing InfoSphere DataStage parallel jobs** This topic covers how to design parallel jobs to transform and to cleanse data. Parallel jobs are compiled and run on the InfoSphere Information Server engine.
- **Developing server jobs** Learn how server jobs are compiled and run on the InfoSphere Information Server engine. Such jobs connect to a data source, extract and transform data, and write data to a target database or file, such as a data warehouse.
- **Deploying jobs** Use the InfoSphere Information Server Manager to move InfoSphere DataStage assets between projects on the same engine, or on different engines. You can also use the InfoSphere Information Server Manager to move assets from one domain to another.
- **Running InfoSphere DataStage jobs** This topic covers how to run InfoSphere DataStage from the IBM InfoSphere DataStage and QualityStage Director client.

- **Administering workload management** Use the workload management queues to control the starting of parallel and server jobs.
- **Monitoring jobs** Use IBM InfoSphere DataStage and QualityStage Operations Console to monitor the job runs, services, system resources, and workload management queues on several InfoSphere Information Server engines.
- **Parallel job reference** Use this reference material to perform more advanced operations with parallel jobs.

## Troubleshooting and support

- **Troubleshooting InfoSphere DataStage** Use the Troubleshooting page to find common troubleshooting topics.
- **IBM Support Portal** Use the Support Portal for InfoSphere DataStage to search for known problems and APARs.

## Getting started and using IBM DataStage on Cloud

---

You must set up your connection to DataStage® on Cloud to design, develop, and run jobs that move and transform data. DataStage on Cloud provides all of the functions of its on-premises counterpart, IBM® InfoSphere® DataStage. DataStage on Cloud is in an IBM SoftLayer® cloud hosted environment that offers deployment without the cost, complexity, and risk of managing your own infrastructure.

**Prerequisite:** You must know the IP address and the credentials of an account on the DataStage on Cloud server and client computer. This information is in the “Access Credentials and Initial Setup” section of the Welcome letter that you received from your IBM Sales Representative.

The DataStage on Cloud client is on a Microsoft Windows virtual machine that is hosted on SoftLayer. When you connect to the client, you can access the client user interfaces, create jobs, and run them. McAfee anti-virus software is installed on the client machine.

The DataStage on Cloud server is on a Red Hat Enterprise Linux virtual or bare metal computer that is hosted on SoftLayer. When you connect to the server, you can access the IBM InfoSphere Information Server engine, services, and repository tiers. You can restart InfoSphere DataStage, edit the dsenv file, and do administrative tasks.

The default firewall configuration of server machines allows SSH connections only from client machines. You must first connect to a client machine by using a remote desktop connection, and then from the client machine you can connect to server machines by using SSH. After you log in to a server machine, you can change the firewall configurations to allow SSH connections from other machines. Communication between the server and client systems happens through a private IP. If you want to access the server from an on-premises client machine, you must modify the iptable rules.

### To connect for the first time, follow these steps:

1. From your local computer, connect to a client machine by using Remote Desktop Connection. For details, see [step 5](#).
2. To modify the firewall to allow SSH access to the servers from your local computer, [set up firewall security](#).
3. On your local computer, use an SSH client to connect to the server machines. For example, you can use PuTTY to connect to the servers. The user ID, password, and IP address of the servers are in your Welcome letter.
4. On the DataStage on Cloud server, [run the ISALite tool](#). The *IS\_install\_path* for the server computers is /opt/IBM/InformationServer.
5. {#connecte} Connect to the DataStage on Cloud client to the DataStage on Cloud server by following these steps
  - **If your local computer is in a Microsoft Windows environment**
    - a. On your local computer, go to the **Start** menu. Click **Accessories > Remote Desktop Connection**.
    - b. Enter the IP address of the Microsoft Windows computer that hosts your DataStage on Cloud client. Click **Connect**.
    - c. In the Windows Security window, enter the user name and password for the DataStage on Cloud client. The user ID, password, and IP address of the client are in your Welcome letter. **Important:** Do not include the domain name with the user name.
    - d. In the DataStage on Cloud client, open the file C:\Windows\System32\drivers\etc\hosts. Make sure that an entry with the private IP exists in the file for the DataStage on Cloud server that you are connecting to. **Important:** The server IP must be a private IP. You cannot open other client machines when you use a public IP.

- **If your local computer is in an Apple Mac environment**

- a. On your local computer, install Microsoft Remote Desktop from the Apple App Store.
  - b. Click the Microsoft Remote Desktop icon, and then click **Open**.
  - c. In the Microsoft Remote Desktop window, click **New**.
  - d. In the Edit Remote Desktops window, supply the following information:
    - 1) In the PC name field, type in the IP address of the cloud client machine.
    - 2) In the User name and Password fields, type in the Windows user name and password that are in the Welcome letter.
6. Verify the connection and installation on the DataStage on Cloud client by following these steps:
    - Run the ISALite tool. The *IS\_install\_path* for the client computer is C:\IBM\InformationServer.
    - Test the installation of the IBM InfoSphere DataStage and QualityStage® Administrator and InfoSphere DataStage and QualityStage Designer clients.
  7. Optional: Enable multiple users to open remote sessions to the DataStage on Cloud client by following these steps on the client computer:
    - a. Create user accounts.
    - b. Give users permission to do a remote desktop connection. The number of concurrent users in a remote session to the client is based on the offering size of your DataStage on Cloud. The small offering size has a maximum of two concurrent sessions. The medium offering size has a maximum of five concurrent sessions. The large offering size has a maximum of 10 concurrent sessions.
  8. Install Adobe Flash Player on the client computer to enable use of the Hierarchical Data stage in parallel job designs. The version of Adobe Flash Player to use is listed in the Web Browser Plug-Ins table of the System Requirements page.
  9. Reset the password for users on the DataStage on Cloud client. Likewise, reset the password for administrators on the DataStage on Cloud server.

**After the initial connection, you can do any of the following tasks:**

- Open the InfoSphere DataStage® and QualityStage® Designer client
- Connect to an on-premises computer
- Connect to an IBM Db2 on Cloud database
- Connect to an on-premises Db2® database instance
- Perform general administration and security tasks

**Related information**

- Getting started and using IBM Information Governance Catalog on Cloud



## Chapter 5. DataStage on Cloud Designer Client

IBM® DataStage® on Cloud Designer Client provides a hosted environment that you configure and control. You can use DataStage on Cloud Designer Client is included in IBM Information Server on Cloud Enterprise Edition, but it can be used as an independent service. Use DataStage on Cloud Designer Client to create, design, and develop DataStage jobs.

### Overview

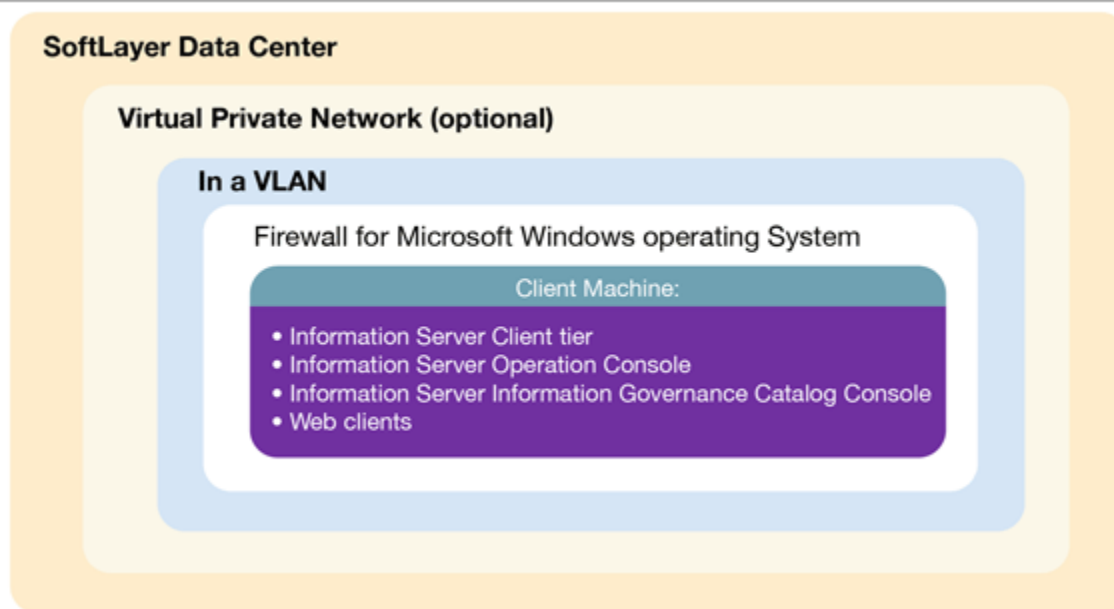
IBM® DataStage® on Cloud Designer Client is a data integration software platform that helps organizations derive more value from the complex, heterogeneous information spread across their systems. DataStage on Cloud Designer Client provides a single unified platform to understand, cleanse, transform, and deliver trustworthy and context-rich information.

DataStage on Cloud Designer Client includes the following IBM Information Server on Cloud clients:

- IBM InfoSphere® Information Governance Catalog
- IBM InfoSphere Information Analyzer
- IBM InfoSphere Metadata Integration Bridges and the metadata interchange agent
- IBM InfoSphere Metadata Asset Manager
- IBM InfoSphere Information Server istool command-line utility
- IBM InfoSphere Information Server Manager client, Multi-Client Manager
- IBM InfoSphere DataStage and QualityStage® Administrator
- IBM InfoSphere DataStage and QualityStage Designer
- IBM InfoSphere DataStage and QualityStage Director

DataStage on Cloud Designer Client uses the characteristics of software-as-a-service (SaaS). You select the plan size based on your needs. IBM provisions the machine and deploys the DataStage on Cloud Designer Client software.

The following figure shows the topology of the client machine in a typical deployment.



As a hosted offering, you have the same control over your data in the cloud as in the on-premises system:

- Actively monitor and report any issues that you encounter with IBM Software as a Service (SaaS).
- Maintain the software platform of your cloud offering and the operating system to meet your security standards.
- Maintain software firewalls on servers that face the internet in a manner to provide required protection.

- Develop integration, transformation, and other jobs. Establish connectivity between data sources and applications. You can also develop your own workload, business rules, monitoring, and scheduling for all jobs. You are responsible for the quality and performance of programs, applications, and jobs that you develop.
- Ensure the continuity, compatibility, and performance of the IBM SaaS platform by installing only permissible software, including any open source packages.
- Regularly upgrade the environment and operating system of your cloud offering.
- Create and maintain regular backups of data.
- Create and maintain high availability configurations.

The following managed add-on services are available to maintain and manage the infrastructure:

**Jump start** This setup service provides up to 50 hours of remote consulting time for startup activities.

**Accelerator** This setup service provides up to 50 hours of remote consulting time to perform various scoped activities.

**Silver** This service provides monthly remote consulting time for operations and maintenance activities.

**Gold** This service provides monthly remote consulting time for operations and maintenance activities. The service includes everything that is provided by the Silver service and delivers extra activities.

**Restriction:** All Information Server on Cloud clients are installed, but you can use only the clients of components that are installed at the server side.

## Available configuration

IBM® DataStage® on Cloud Designer Client is on virtual machine with dedicated CPUs.

The following table lists the configuration of the client machine.

Configuration	Size
Memory (GB)	4
Number of cores	4
Network speed (Gbps)	1 Gbps with 1000 GB bandwidth
First disk	100 GB Storage Area Network (SAN)
Second disk	100 GB SAN

The virtual machine has a Microsoft Windows operating system environment. All IBM Information Server on Cloud client utilities are installed on DataStage on Cloud Designer Client.

## Layout of IBM® DataStage® on Cloud Designer Client disks

The DataStage® on Cloud Designer Client disk comes in one size for all plans.

DataStage on Cloud Designer Client is a client virtual machine with two Storage Area Network (SAN) disks that are 100 GB each. One disk is drive C for the Microsoft Windows operating system. The other disk is drive F, and it is an empty disk.

DataStage on Cloud Designer Client can be used in combination with any of the servers that are offered with IBM® Information Server on Cloud and IBM DataStage on Cloud offerings.

## Information roadmap

This roadmap lists information resources that are available for users who are new to the DataStage® products from IBM®. These resources provide information about various subject areas, such as learning basic skills and troubleshooting.

## Product overview

- **InfoSphere® DataStage features** Learn about the features and benefits of InfoSphere DataStage.
- **InfoSphere DataStage product documentation** The InfoSphere Information Server Knowledge Center provides you with InfoSphere DataStage concepts and usage information to help new users prepare for using your new system.
- **InfoSphere DataStage developerWorks® forum** Use this forum to interact with other InfoSphere DataStage users to better understand how to design, build, debug and deploy jobs for information collection, integration and transformation.

## Getting started

- **Overview of InfoSphere DataStage** This overview covers InfoSphere DataStage job life cycles, job designs, and how the product integrates with the rest of the InfoSphere Information Server suite.
- **Tutorial: Creating parallel jobs** This tutorial demonstrates how you can use InfoSphere DataStage to develop jobs that extract, transform, and load data. By transforming and cleansing the source data and applying consistent formatting, you enhance the quality of the data.

## Using InfoSphere DataStage

- **Designing InfoSphere DataStage jobs** This topic explains how to design InfoSphere DataStage jobs using the IBM InfoSphere DataStage and QualityStage® Designer client. The Designer client gives you the tools that you need to create jobs that extract, transform, load, and check the quality of data.
- **Developing InfoSphere DataStage parallel jobs** This topic covers how to design parallel jobs to transform and to cleanse data. Parallel jobs are compiled and run on the InfoSphere Information Server engine.
- **Developing server jobs** Learn how server jobs are compiled and run on the InfoSphere Information Server engine. Such jobs connect to a data source, extract and transform data, and write data to a target database or file, such as a data warehouse.
- **Deploying jobs** Use the InfoSphere Information Server Manager to move InfoSphere DataStage assets between projects on the same engine, or on different engines. You can also use the InfoSphere Information Server Manager to move assets from one domain to another.
- **Running InfoSphere DataStage jobs** This topic covers how to run InfoSphere DataStage from the IBM InfoSphere DataStage and QualityStage Director client.
- **Administering workload management** Use the workload management queues to control the starting of parallel and server jobs.
- **Monitoring jobs** Use IBM InfoSphere DataStage and QualityStage Operations Console to monitor the job runs, services, system resources, and workload management queues on several InfoSphere Information Server engines.
- **Parallel job reference** Use this reference material to perform more advanced operations with parallel jobs.

## Troubleshooting and support

- **Troubleshooting InfoSphere DataStage** Use the Troubleshooting page to find common troubleshooting topics.
- **IBM Support Portal** Use the Support Portal for InfoSphere DataStage to search for known problems and APARs.

## Getting started and using IBM DataStage on Cloud Designer Client

---

You must set up your connection to DataStage® on Cloud Designer Client to design, develop, and run jobs that move and transform data. DataStage on Cloud Designer Client includes all clients of IBM® InfoSphere® Information Server. It is in an IBM SoftLayer® cloud hosted environment.

**Prerequisite:** You must know the IP address and the credentials of an account on DataStage on Cloud Designer Client computer. This information is in the “Access Credentials and Initial Setup” section of the Welcome letter that you received from your IBM Sales Representative.

DataStage on Cloud Designer Client is on a Microsoft Windows virtual machine. When you connect to DataStage on Cloud Designer Client, you can access client user interfaces, create jobs, and run them. McAfee anti-virus software is installed on the client machine.

Before using DataStage on Cloud Designer Client with any of the IBM Information Server on Cloud, IBM Information Server on Cloud Data Quality, and IBM DataStage on Cloud offerings, required connections must be set up between the client and servers.

## When you connect for the first time, follow these steps:

1. Connect DataStage on Cloud Designer Client to the Information Server on Cloud server by following these steps:
  - **If your local computer is in a Microsoft Windows environment**
    - a. On your local computer, go to the **Start** menu. Click **Accessories > Remote Desktop Connection**.
    - b. Enter the IP address of the Microsoft Windows computer that hosts your DataStage on Cloud Designer client. Click **Connect**.
    - c. In the Windows Security window, enter the user name and password for the DataStage on Cloud Designer client. The user ID, password, and IP address of the client are in your Welcome letter. **Important:** Do not include the domain name with the user name.
    - d. In the DataStage on Cloud Designer client, open the file C:\Windows\System32\drivers\etc\hosts. Make sure that an entry with the private IP exists in the file for the Information Server on Cloud server that you are connecting to. **Important:** The server IP must be a private IP. The connection from DataStage on Cloud Designer Client fails when you use a public IP.
  - **If your local computer is in an Apple Mac environment**
    - a. On your local computer, install Microsoft Remote Desktop from the Apple App Store.
    - b. Click the Microsoft Remote Desktop icon, and then click **Open**.
    - c. In the Microsoft Remote Desktop window, click **New**.
    - d. In the Edit Remote Desktops window, supply the following information:
      - In the PC name field, type in the IP address of the cloud client machine.
      - In the User name and Password fields, type in the Windows user name and password that are in the Welcome letter.
2. Verify the connection and installation on DataStage on Cloud Designer Client. On the client computer, follow these steps:
  - Run the ISALite tool. The *IS\_install\_path* for the client computer is C:\IBM\InformationServer.
  - Test the installation of the IBM InfoSphere DataStage and QualityStage® Administrator and InfoSphere DataStage and QualityStage Designer clients.
3. Optional: On the client computer, enable multiple users to open remote sessions to DataStage on Cloud Designer Client by following these steps.
  - a. Create user accounts.
  - b. Give users permission to do a remote desktop connection.
4. Reset the password for users on DataStage on Cloud Designer Client.

## After the initial connection, you can do any of the following tasks:

- You can add additional DataStage on Cloud Designer Client machines to your environment.
- Depending on the clients that you purchased with the offering, you can create jobs, create data rules, and other tasks. DataStage on Cloud Designer Client includes the following InfoSphere Information Server clients:
  - IBM InfoSphere Information Governance Catalog
  - IBM InfoSphere Information Analyzer
  - IBM InfoSphere Information Governance Dashboard
  - IBM InfoSphere Information Services Director
  - IBM InfoSphere Data Click
  - IBM InfoSphere FastTrack
  - IBM InfoSphere Metadata Integration Bridges and the metadata interchange agent
  - IBM InfoSphere Information Server istool command-line utility
  - IBM InfoSphere Information Server Manager client
  - Multi-Client Manager

- IBM InfoSphere DataStage and QualityStage Administrator client
- IBM InfoSphere DataStage and QualityStage Designer client
- IBM InfoSphere DataStage and QualityStage Director client
- Open the InfoSphere DataStage® and QualityStage® Designer client
- Connect to an on-premises computer
- Connect to an IBM Db2 on Cloud database
- Connect to an on-premises Db2® database instance
- Perform general administration and security tasks

As your computing needs grow, you can add and connect more DataStage on Cloud Designer Client machines in your Information Server on Cloud environment.

### Related tasks

- Enhancing security of Information Server on Cloud computers

## Adding and connecting extra DataStage on Cloud Designer Client machines

You can add extra IBM® DataStage® on Cloud Designer Client machines in your IBM Information Server on Cloud environment as your computing needs increase.

**Prerequisite:** You must be logged in to an administrative account on the Information Server on Cloud server.

You use scripts to add more DataStage on Cloud Designer Client machines to your Information Server on Cloud environment.

The scripts are in directory `/opt/IBM/scripts` on the Information Server on Cloud server.

Information Server on Cloud servers come with iptables firewall setup. You change the iptables rules to allow traffic between the new DataStage on Cloud Designer Client client machine and the servers.

1. To use the new DataStage on Cloud Designer Client machine with IBM Information Server on Cloud Data Quality and IBM Information Server on Cloud Enterprise Edition, follow these steps:
  - a. In the services tier machine, do the following steps in a command-line window:
    - 1) Open the file `/etc/hosts`, and then add a host entry for the new DataStage on Cloud Designer Client machine.
    - 2) Run the script `allow_services_tier_from_ip.sh` with the private IP address of the new DataStage on Cloud Designer Client machine. For example, the command might be: `allow_services_tier_from_ip.sh 10.xxx.xxx.xxx`.
  - b. In the engine tier machine, run the script `allow_engine_tier_from_ip.sh` with the private IP address of the new DataStage on Cloud Designer Client machine. This script permits traffic from the DataStage on Cloud Designer Client machine to the engine tier machine.
  - c. In the new DataStage on Cloud Designer Client machine, open the Microsoft Windows PowerShell command-line tool. Run the following command: `netsh advfirewall firewall add rule name="Open HTTPS port for IS" protocol=tcp localport=19443 dir=in remoteip=<Private_ip_of_the_services_tier> profile=any action=allow enable=Yes`. The `<Private_ip_of_the_services_tier>` value is the IP address of the services tier of the Information Server on Cloud machine. This command adds a rule to the iptables firewall to permits traffic from the DataStage on Cloud Designer Client machine to the services tier machine.
2. To use the new DataStage on Cloud Designer Client machine with IBM DataStage on Cloud server, follow these steps:
  - a. In the new DataStage on Cloud Designer Client machine, open the `C:\Windows\System32\drivers\etc\hosts` file. Add a host entry for the DataStage on Cloud server.
  - b. In the DataStage on Cloud server, open the file `/etc/hosts`. Add a host entry for the DataStage on Cloud Designer Client machine.
  - c. Update the iptables firewall rules by following these steps:
    - 1) On the engine tier machine of the DataStage on Cloud server, run the script `allow_datastage_from_ip.sh` with the private IP address of the new DataStage on Cloud Designer Client machine. For example, the command might be: `allow_datastage_from_ip.sh 10.xxx.xxx.xxx`. **Tip:** Check the list of open ports that are opened by

using this script. Depending on the list of services that you are using from the DataStage on Cloud Designer Client machine, you might want to close some ports.

- 2) In the new DataStage on Cloud Designer Client machine, open the Windows PowerShell command-line tool. Run the command `netsh advfirewall firewall add rule name="Open HTTPS port for IS" protocol=tcp localport=19443 dir=in remoteip=Private_ip_of_datastage_server profile=any action=allow enable=Yes`. This command opens HTTPS port 19443 on the DataStage on Cloud Designer Client machine to the DataStage on Cloud server.

---

## Chapter 6. Information Governance Catalog on Cloud

IBM® Information Governance Catalog on Cloud provides a hosted environment that you configure and control. You can use Information Governance Catalog on Cloud to extend the reach of your business by leveraging cloud offerings, while you reduce the costs that are associated with providing these services. Information Governance Catalog on Cloud provides several different plans so that any size business can access the powerful and scalable ETL platform by IBM.

### Overview

---

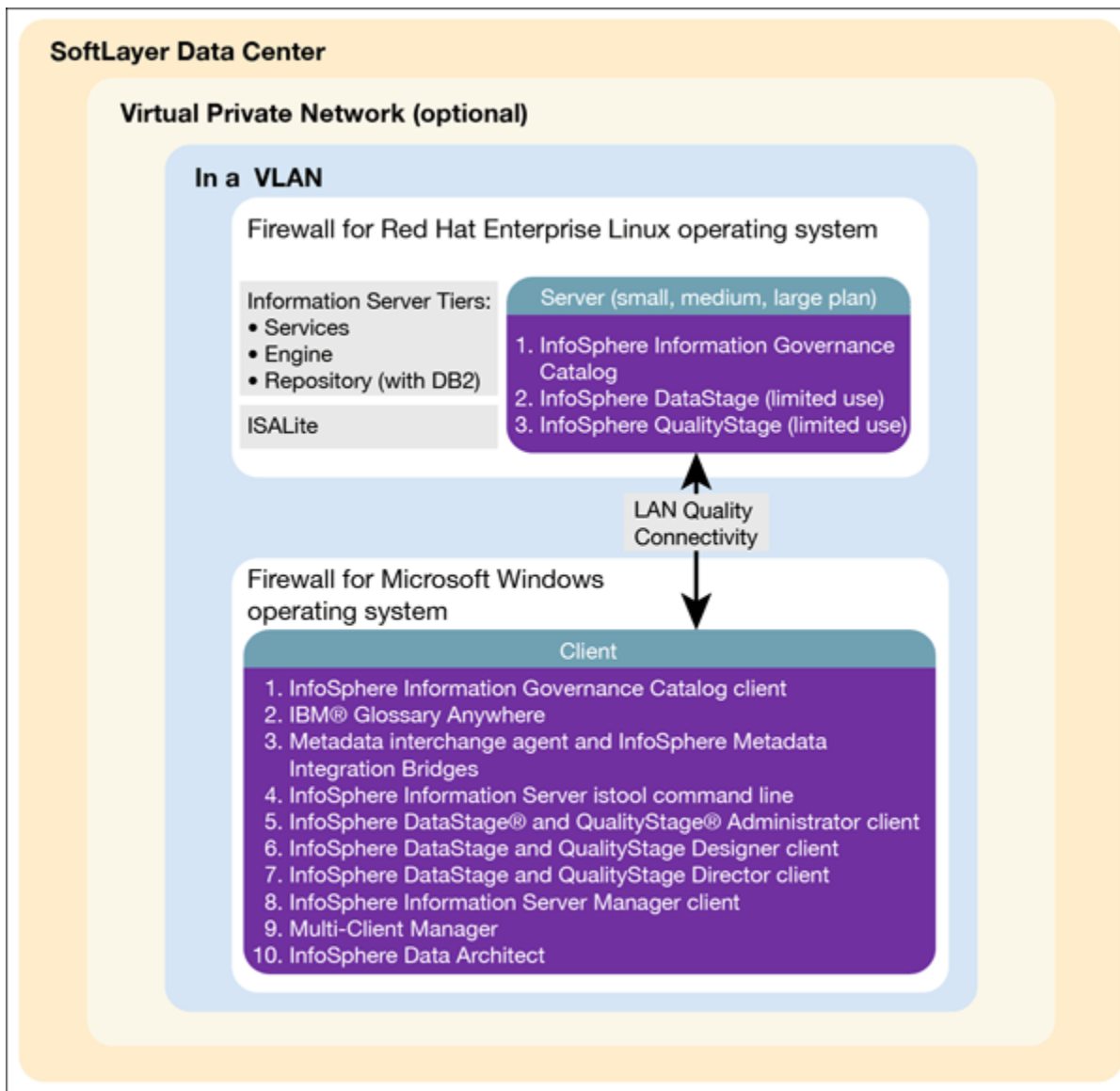
IBM® Information Governance Catalog on Cloud is an interactive, web-based tool in the cloud that enables users to create, manage, and share an enterprise vocabulary and classification system in a central catalog. Information Governance Catalog on Cloud provides all of the functions of its on-premises counterpart, IBM InfoSphere® Information Governance Catalog.

You can do the following tasks by using Information Governance Catalog on Cloud:

- Create a well-documented, end-to-end information blueprint to align your business requirements with your enterprise and reference architectures before you start a strategic project
- Establish a common business language and manage business perspectives about information to align those views with the IT perspective
- Manage and explore data lineage to create trusted information that supports data governance and compliance efforts
- Provide a solid foundation for different types of information integration and governance projects, including information integration, lifecycle management, and security initiatives

Information Governance Catalog on Cloud uses the characteristics of software-as-a-service (SaaS). You select the plan size based on your needs. IBM provisions the machine and deploys the Information Governance Catalog on Cloud software.

The following figure shows the topology of the server and client machines in a typical deployment.



As a hosted offering, you have the same control over your data in the cloud as in the on-premises system:

- Actively monitor and report any issues that you encounter with IBM Software as a Service (SaaS).
- Maintain the software platform of your cloud offering and the operating system to meet your security standards.
- Maintain software firewalls on servers that face the internet in a manner to provide required protection.
- Develop integration, transformation, and other jobs. Establish connectivity between data sources and applications. You can also develop your own workload, business rules, monitoring, and scheduling for all jobs. You are responsible for the quality and performance of programs, applications, and jobs that you develop.
- Ensure the continuity, compatibility, and performance of the IBM SaaS platform by installing only permissible software, including any open source packages.
- Regularly upgrade the environment and operating system of your cloud offering.
- Create and maintain regular backups of data.
- Create and maintain high availability configurations.

The following managed add-on services are available to maintain and manage the infrastructure:

**Jump start** This setup service provides up to 50 hours of remote consulting time for startup activities.

**Accelerator** This setup service provides up to 50 hours of remote consulting time to perform various scoped activities.

**Silver** This service provides monthly remote consulting time for operations and maintenance activities.



**Gold** This service provides monthly remote consulting time for operations and maintenance activities. The service includes everything that is provided by the Silver service and delivers extra activities.

**Restriction:** With Information Governance Catalog on Cloud, users are not authorized to use any of the following components or functions:

- Run a job function of IBM InfoSphere DataStage®
- Design a new asset of IBM InfoSphere DataStage and QualityStage® Designer

## Available configurations

IBM® Information Governance Catalog on Cloud servers for the small and medium plans are virtual servers with public CPUs. The servers in the large plan are in a bare metal environment.

Select the offering plan that fits your usage and environment needs.

Offering	Memory (GB)	Number of Cores	Network speed and bandwidth	First disk	Second disk
Small	12	4	1 Gbps with 250 GB bandwidth	100 GB Storage area network (SAN)	100 GB SAN
Medium	16	8	1 Gbps with 250 GB bandwidth	100 GB SAN	500 GB SAN
Large	64	12	1 Gbps with 1000 GB bandwidth	960 GB SSD	960 GB SSD

A client machine is provided with all offering sizes. The client is in a Microsoft Windows environment with 6 GB of memory, two cores, and two 100 GB SAN storage. The network speed is 1 Gbps, with 250 GB bandwidth.

The following IBM Information Server on Cloud client tier components are installed with Information Governance Catalog on Cloud on client machine:

- IBM InfoSphere® Information Governance Catalog client
- IBM Glossary Anywhere
- IBM InfoSphere Metadata Integration Bridges and the metadata interchange agent
- IBM InfoSphere Metadata Asset Manager
- IBM InfoSphere Information Server istool command-line utility
- IBM InfoSphere Information Server Manager client
- Multi-Client Manager
- IBM InfoSphere DataStage® and QualityStage® Administrator client
- IBM InfoSphere DataStage and QualityStage Designer client
- IBM InfoSphere DataStage and QualityStage Director client
- IBM InfoSphere Data Architect client

These client tier components import or create assets and their metadata in the metadata repository. The assets and metadata can then be accessed by Information Governance Catalog on Cloud.

## Layout of IBM Information Governance Catalog on Cloud server and client disks

The layout of the Information Governance Catalog on Cloud server and client disks depends on the plan size of your system.

### Virtual server for small and medium plans

The small and medium plans come with two Storage Area Network (SAN) disks. The Red Hat Enterprise Linux operating system is on the first SAN disk in both the small and medium plans. The second SAN disk is encrypted by using Linux Unified Key Setup (LUKS) for both these sizes.

The encryption key details are provided in the Welcome letter from the IBM® Sales Representative. It is recommended that you add your own key and remove the supplied key before you use the system.

The product is installed on the /disk1 directory. User data can be stored on /data directory. Both directories are on the partition /dev/xvdc1 that is encrypted.

*Table 17: Small and medium disk layout*

Mount point	Encryption status	Partition name	Partition size	Disk	Used for
/boot	None	/dev/xvda1	256 MB	Primary disk is /dev/xvda	Kernel boot area and operating system data
/	None	/dev/xvda2	99.8 GB	Primary disk is /dev/xvda	Kernel boot area and operating system data
/disk1	LUKS	/dev/xvdc1	100 GB for small plan. 500 GB for medium plan	Secondary disk is /dev/xvdc	Installation files and creation of /data and /opt directories.
SWAP	None	/dev/xvdb1	2 GB	/dev/xvdb	Swap space for paging

### Bare metal server for large plan

The large plan comes with two Solid State Drive (SSD) disks that are 960 GB each. RAID level 1 implementation makes them appear as a single disk.

The disk is divided into four partitions. The Red Hat Enterprise Linux operating system is on a 10 GB partition. The boot data is on a 248 MB partition. The swap space is on a 2 GB partition. The remaining space is on another partition that is encrypted by using LUKS.

*Table 18: Bare metal disk layout*

Mount point	Encryption status	Partition name	Partition size	Disk	Used for
/boot	None	/dev/sda1	248 MB	/dev/sda	Kernel boot data
SWAP	None	/dev/sda3	2 GB	/dev/sda	Swap space for paging
/	None	/dev/sda2	10 GB	/dev/sda	Operating system data
/opt	LUKS	/dev/mapper/cryptData	About 900 GB	/dev/sda	Installation files and creation of /opt directory.

### Client for all plans

The Information Governance Catalog on Cloud client machine configuration is the same for all plans sizes. The client machine has two SAN disks that are 100 GB each. One disk is C: for the Microsoft Windows operating system. The second disk is F:, and it is an empty disk.

## Information roadmap

This roadmap lists information resources that are available for users who are new to the IBM® InfoSphere® Information Governance Catalog products. These resources provide information about various subject areas.

- **Overview** Learn about the features and benefits of InfoSphere Information Governance Catalog.
- **Designing the catalog** A catalog is an authoritative dictionary of the assets that are used throughout the organization. One of the main benefits of a well-designed catalog is increased trust and confidence in organization information. Planning, designing, and publishing a catalog involves several tasks.
- **Administering the catalog** The catalog is composed of glossary assets (terms, categories, information governance policies, and information governance rules) and information assets. You can assign security roles and permissions to users to control access to the catalog. You can also assign workflow roles, assign users as stewards, define custom attributes, and define external asset types. In addition, you can configure the display of glossary assets.
- **Governing your data** You can create, edit, and delete catalog assets. Some assets are created when they are imported into the catalog. You can assign stewards and assets, define and assign custom attributes, and extend data flows for lineage reports by importing assets. In addition, you can configure assets for lineage analysis reports. Assets of different types that have a common business purpose can be grouped in a collection.
- **Viewing catalog content** You can view, search, browse, and query the catalog to find catalog assets.
- **Expanding catalog capabilities**
  - **Look up terms** while you work in other applications from the Microsoft Windows desktop by using Glossary Anywhere.
  - **Develop and extend applications** by using InfoSphere Information Governance Catalog for Eclipse. You can access your glossary content from within your Eclipse-based development tool.
  - Write client applications to access and author catalog content by using InfoSphere Information Governance Catalog REST API. You can integrate your catalog content into other software tools and portals.
- **Redbooks publication and Support**
  - **Information Governance Principles and Practices for a Big Data Landscape** Published in 2014, this publication describes how the IBM Big Data Platform provides the integrated capabilities that are required for the adoption of Information Governance in the big data landscape. The IBM Big Data Platform, which is coupled with a framework for Information Governance, provides an approach to build, manage, and gain significant value from the big data landscape.
  - **IBM Support Portal** Use the Support Portal for IBM InfoSphere Information Server to search for known problems and APARs.

## Getting started and using IBM Information Governance Catalog on Cloud

---

You must set up your connection to Information Governance Catalog on Cloud. Information Governance Catalog on Cloud provides all of the functions of its on-premises counterpart, IBM® InfoSphere® Information Governance Catalog. It is deployed in an IBM SoftLayer® hosted environment.

**Prerequisite:** You must know the IP address and the credentials of an account on the Information Governance Catalog on Cloud server and client computers. This information is in the “Access Credentials and Initial Setup“ section of the Welcome letter that you received from your IBM Sales Representative.

**Prerequisite:** If your DataStage® jobs use the Hierarchical Data stage in parallel job designs, you must install Adobe Flash Player on the client computer. The version of Adobe Flash Player to use is listed in the Web Browser Plug-Ins table of the System Requirements page.

The Information Governance Catalog on Cloud client is on a Microsoft Windows virtual machine that is hosted on SoftLayer. When you connect to the client, you can access the client user interfaces. McAfee anti-virus software is installed on the client machine.

The Information Governance Catalog on Cloud server is on a Red Hat Enterprise Linux virtual or bare metal computer that is hosted on SoftLayer. When you connect to the server, you can access the IBM InfoSphere Information Server engine, services, and repository tiers. You can restart Information Governance Catalog on Cloud and do administrative tasks.

The Information Governance Catalog on Cloud license does not permit you to run DataStage jobs, but you can import those jobs by using DataStage on Cloud Designer Client. Information Governance Catalog on Cloud uses the job definitions to draw the lineage graph.

The default firewall configuration of server machines allows SSH connections only from client machines. You must first connect to a client machine by using a remote desktop connection, and then from the client machine you can connect to server machines by using SSH. After you log in to a server machine, you can change the firewall configurations to

allow SSH connections from other machines. Communication between the server and client systems happens through a private IP. If you want to access the server from an on-premises client machine, you must modify the iptable rules.

### When you connect for the first time, follow these steps:

1. From your local computer, connect to a client machine by using Remote Desktop Connection. For details, see [step 1.e](#).
2. To modify the firewall to allow SSH access to the servers from your local computer, [set up firewall security](#).
3. On your local computer, use an SSH client to connect to the server machines. For example, you can use PuTTY to connect to the servers. The user ID, password, and IP address of the servers are in your Welcome letter.
4. On the Information Governance Catalog on Cloud server, [run the ISALite tool](#). The *IS\_install\_path* for the server computers is /opt/IBM/InformationServer
5. Connect to the Information Governance Catalog on Cloud client to the Information Governance Catalog on Cloud server by following these steps:
  - **If your local computer is in a Microsoft Windows environment**
    - a. On your local computer, go to the **Start** menu. Click **Accessories > Remote Desktop Connection**.
    - b. Enter the IP address of the Microsoft Windows computer that hosts your Information Governance Catalog on Cloud client. Click **Connect**.
    - c. In the Windows Security window, enter the user name and password for the Information Governance Catalog on Cloud client. The user ID, password, and IP address of the client are in your Welcome letter. **Important:** Do not include the domain name with the user name.
    - d. In the Information Governance Catalog on Cloud client, open the file C:\Windows\System32\drivers\etc\hosts. Make sure that an entry with the private IP exists in the file for the Information Governance Catalog on Cloud server that you are connecting to. **Important:** The server IP must be a private IP. You cannot open other client machines when you use a public IP.
  - **If your local computer is in an Apple Mac environment**
    - a. On your local computer, install Microsoft Remote Desktop from the Apple App Store.
    - b. Click the Microsoft Remote Desktop icon, and then click **Open**.
    - c. In the Microsoft Remote Desktop window, click **New**.
    - d. In the Edit Remote Desktops window, supply the following information:
      - In the PC name field, type in the IP address of the cloud client machine.
      - In the User name and Password fields, type in the Windows user name and password that are in the Welcome letter.
6. Verify the connection and installation on the Information Governance Catalog on Cloud client by following these steps:
  - Run the ISALite tool. The *IS\_install\_path* for the client computer is C:\IBM\InformationServer.
  - Test the installation of the Information Governance Catalog on Cloud client.
7. Optional: Enable multiple users to open remote sessions to the Information Governance Catalog on Cloud client by following these steps on the client computer:
  - a. [Create user accounts](#).
  - b. [Give users permission](#) to do a remote desktop connection.
8. [Reset the password](#) for users and administrators on the Information Governance Catalog on Cloud server.

### After the initial connection, you can do any of the following tasks:

- [Open the InfoSphere DataStage® and QualityStage® Designer client](#)
- [Connect to an on-premises computer](#)
- [Perform general administration and security tasks](#)

#### Related information

- [Getting started and using IBM DataStage on Cloud](#)
- [Enhancing security of Information Server on Cloud computers](#)

## Chapter 7. Connecting to other systems

You can connect your IBM Information Server cloud offerings to on-premises systems and to cloud systems. Not all cloud offerings can connect to every system.

The following table lists a sample of which system can connect to Information Server cloud offerings. For information about how to connect, click the links in the **Connect to these systems** column.

Note that with network connection available, cloud offerings can connect to the same data sources as their on-premise equivalents. For details, see [https://www.ibm.com/support/knowledgecenter/SSZJPZ\\_11.5.0/com.ibm.swg.im.iis.connect.nav.doc/containers/cont\\_iisinfo\\_connect.html](https://www.ibm.com/support/knowledgecenter/SSZJPZ_11.5.0/com.ibm.swg.im.iis.connect.nav.doc/containers/cont_iisinfo_connect.html)

*Table 19: Sample of systems that can connect to cloud offerings*

Connect to these systems	IBM® Information Server on Cloud Enterprise Edition	IBM DataStage® on Cloud	IBM Information Governance Catalog on Cloud	IBM Information Server Data Quality on Cloud
<a href="#">IBM InfoSphere DataStage and QualityStage® Designer client</a>	Yes	Yes	Yes	Yes
<a href="#">On-premises computer</a>	Yes	Yes	Yes	Yes
<a href="#">IBM Db2 on Cloud database</a>	Yes	Yes	No	Yes
<a href="#">On-premises Db2® database instance</a>	Yes	Yes	No	Yes

### Connecting to the IBM InfoSphere DataStage and QualityStage Designer client

You connect to the InfoSphere® DataStage® and QualityStage® Designer client of IBM® DataStage on Cloud from IBM Information Server on Cloud offerings.

**Prerequisite:** You must know the IP address, host name, and credentials of the DataStage on Cloud server. This information is in the Welcome letter from your IBM Sales Representative.

You can connect from these cloud offerings:

- IBM Information Server on Cloud Enterprise Edition
- IBM DataStage on Cloud
- IBM Information Governance Catalog on Cloud
- IBM Information Server on Cloud Data Quality

1. On the Microsoft Windows machine where the DataStage on Cloud client is installed, click **Start > Programs > IBM InfoSphere DataStage and QualityStage Designer client**.

2. Enter the following information:

- **Host name of the services tier:** The host name or IP address of the services tier on the DataStage on Cloud server in the following format: <IP\_hostname\_services\_tier>:<port\_number>. For example, 169.XX.XX.XX:9446.
- **User name:** The user name of the IBM InfoSphere Information Server administrator user on the DataStage on Cloud server. For example, isadmin.
- **Password:** The password of the InfoSphere Information Server administrator user.
- **Project:** The name of the default project is dstage1. For example, <host\_name>/dstage1.

3. Click **Login**, and then accept and install the certificate that is generated. You can now use the InfoSphere DataStage and QualityStage Designer client on the cloud.

## Connecting to an on-premises computer

---

You can connect to an on-premises computer from IBM® Information Server on Cloud cloud offerings.

**Prerequisite:** You must know the IP address and credentials of an administrator user of the on-premises computer.

**Prerequisite:** You must know the VPN client that you want to use in your DataStage® on Cloud server or client machines.

You can connect to an on-premises computer from these cloud offerings:

- IBM Information Server on Cloud Enterprise Edition
- IBM DataStage on Cloud
- IBM Information Governance Catalog on Cloud
- IBM Information Server on Cloud Data Quality

A VPN client must be installed on your DataStage on Cloud server or client machine only if your on-premise machine is behind a firewall. Then, after the VPN client is installed, you can access the on-premises system.

You can set up the VPN client to always run on the DataStage on Cloud server. Alternatively, you can run the VPN client only when you run a DataStage job that accesses computers behind a firewall.

Your on-premise system might be a server that is running in a UNIX or Microsoft Windows operating system environment. The VPN client enables a job that is running in the DataStage on Cloud to read from and write to an on-premise system.

1. If the on-premise machine that you need to connect to is behind a firewall, do these steps before you connect for the first time:
  - a. Log in to the DataStage on Cloud server machine that you want to connect from.
  - b. Install a VPN client that is appropriate for the operating system of the cloud machine. The DataStage on Cloud server is on a Linux operating system.
  - c. On your server machine, ping the on-premises computer. This step verifies that the connection is valid. After the connection is available, you can connect to the on-premise machines by using DataStage jobs to read and write data.
2. To run the VPN client only during job execution, design the job to open up the firewall at the beginning of the job execution and to close the firewall when the job execution is done. Use a command-line or API of your VPN client in an OSH script or in a stage to open and close the firewall.

## Connecting to IBM Db2 on Cloud

---

You can connect some IBM® Information Server cloud offerings to your IBM Db2 on Cloud instance.

**Prerequisite:** You must know the IP address or host name, port number, and credentials for a user to connect to IBM InfoSphere® DataStage® and QualityStage® Designer on the client machine.

**Prerequisite:** You must know the JDBC URL and credentials to connect to the Db2 on Cloud instance.

You can connect to your IBM Db2 on Cloud instance from these cloud offerings:

- IBM Information Server on Cloud Enterprise Edition
- IBM DataStage on Cloud
- IBM Information Server on Cloud Data Quality

Then, you can read data from Db2 on Cloud or write to Db2 on Cloud from other sources, such as a flat file or RDBMS stages.

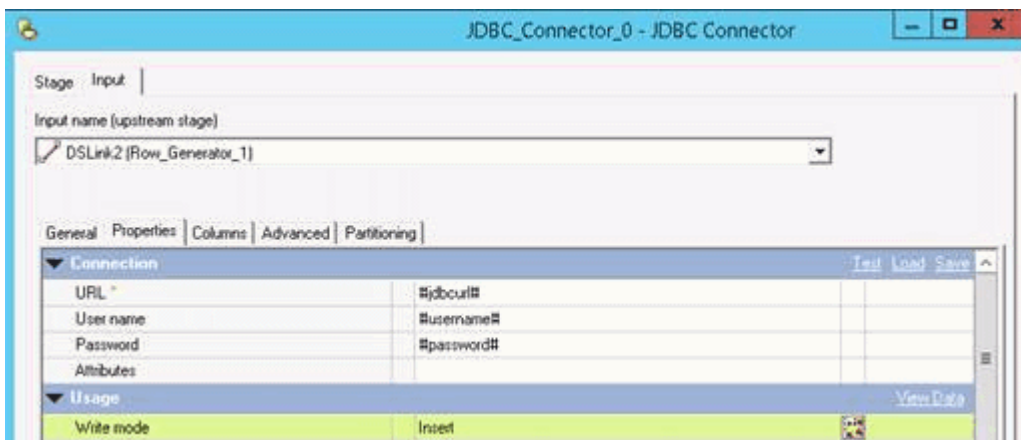
You can use the JDBC Connector stage to connect to Db2 on Cloud. For more information about how to use this stage, see [JDBC data sources](#).

For example, suppose that you have a Db2 on Cloud service that runs on IBM Bluemix®. You can connect to that Db2 on Cloud instance by using the JDBC Connector stage when you run a job on your DataStage on Cloud server. You use the credentials that are required to connect to the Db2 on Cloud service.

1. On your DataStage on Cloud client machine, log in with the credentials of a user in the InfoSphere DataStage and QualityStage Designer client.
2. In the InfoSphere DataStage and QualityStage Designer client, design a job that includes a JDBC Connector stage by following these steps:
  - a. Click **File > New** from the menu.
  - b. In the New window, select the Parallel Job icon, and then click **OK**.
  - c. Add the JDBC Connector stage to the job by following these steps:
    - 1) In the palette, select the Database category.
    - 2) Locate JDBC Connector in the list of available stage types.
    - 3) Drag the JDBC Connector stage icon to the canvas.
    - 4) Rename the JDBC Connector stage. Choose a name that indicates the role of the stage in the job.
  - d. Create the necessary links and more stages for the job:
    - For a job that reads data from a JDBC data source, create the next stage in the job, and then create an output link from the JDBC Connector stage to the next stage.
    - For a job that writes data to a JDBC data source, create one or more links from other stages in the job to the JDBC Connector stage.
  - e. Save the job.
3. Get the JDBC URL, user name, and password of your Db2 on Cloud instance by following these steps:
  - a. Go to your dashboard in Bluemix.
  - b. From the list of services, select Db2 on Cloud. The connection parameters are available in JSON format under the topic “Service Credentials”.
4. Enter the connection parameters of your Db2 on Cloud instance in the JDBC connector stage.
5. Compile the job, and then run it.

## Example

You design a job that generates data by using the Row Generator stage. This stage writes to your Db2 on Cloud instance by using a JDBC connector. You provide the connection information to the JDBC Connector stage, compile the job, and then run the job.



## Connecting on an on-premises Db2 database instance

You can connect some IBM® Information Server cloud offerings to your Db2® database instance on an on-premises computer.

**Prerequisite:** You must know the IP address or host name, port number, and credentials for a Db2 user to connect to the Db2 database on the on-premises computer.

**Prerequisite:** You must know the IP address or host name, port number, and credentials of the DataStage® on Cloud server to connect from the IBM InfoSphere® DataStage and QualityStage® Designer client.

**Prerequisite:** If the on-premises Db2 database is installed on a system that is behind a firewall, follow the steps to [connect to an on-premises computer](#).

You can connect to your on-premises Db2 database instance from these cloud offerings:

- IBM Information Server on Cloud Enterprise Edition
- IBM DataStage on Cloud
- IBM Information Server Data Quality on Cloud

Then, you can move data from your on-premises database to the cloud, and from the cloud to your on-premises database.

You can use the Db2 Connector stage from the DataStage on Cloud server to read data from the on-premises Db2 database. For more information about how to use this stage, see [Db2 connector](#).

You catalog the on-premises database to the DataStage on Cloud to connect them with each other.

1. Verify the connection between the on-premises computer and the DataStage on Cloud server by using the **ping** command.
2. On the DataStage on Cloud server, configure the Db2 database connector. The Db2 connector requires that you catalog each on-premises database to Db2 on the DataStage on Cloud server.
3. On the DataStage on Cloud server, log in as a Db2 instance user.
4. Catalog the on-premises database on the DataStage on Cloud server by doing these commands:

```
db2 catalog tcpip node node_name remote remote_ip_address db2_port_number
db2 catalog database database_name at node node_name
db2 terminate
```

If you do the commands inside the Db2 prompt, do not include “db2” in the command.

5. At the Db2 prompt, confirm that the on-premises Db2 database is cataloged:

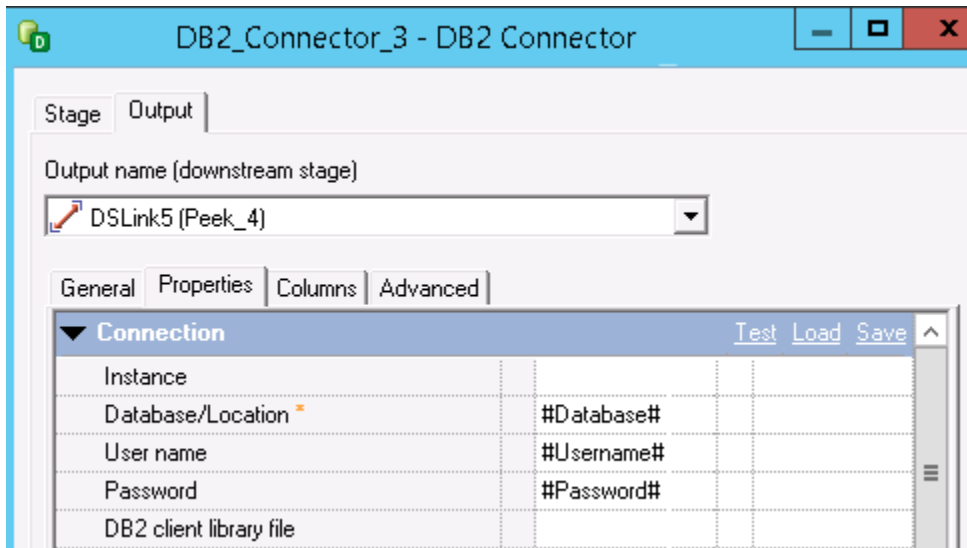
```
list database directory
```

## Example

You create a job to transfer data from the on-premises Db2 database to a sequential file on your DataStage on Cloud server. You design a job by using the Db2 connector and Sequential File stages.

In Db2 Connector stage properties, you provide the database name, user name, and password details of your on-premises Db2 database.





Complete the job design with other details, compile the job, and then run the job.



# Chapter 8. Administering cloud offerings on the server

You administer your IBM® Information Server on Cloud Enterprise Edition offerings on the server with the same utilities and tools that you use for your on-premises systems.

Information Server on Cloud Enterprise Edition offerings are based on IBM InfoSphere® Information Server. As a result, administrative tasks for your Information Server on Cloud server are managed by the same utilities and tools as for your on-premises systems. See [Administering](#) for detailed information about administering your system.

You can do the following administrative tasks:

- **Start services and application server services** For information, see [Starting services \(Linux, UNIX\)](#).
- **Shut down services and application server services** For more information, see [Shutting down services \(Linux, UNIX\)](#).
- **Change passwords** For more information about changing passwords, see [Changing the password for the administrator and database accounts](#).
- **Encrypt Db2® database with native encryption** This encryption is done on all machines. For more information, see [Db2 native encryption](#).
- **Back up and restore the services tier, engine tier, and repository tier** For more information, see [Backing up IBM InfoSphere Information Server components](#) and [Performing a manual migration](#).

## Open ports on server and client machines

At the operating system level, the firewall is implemented on Linux-based server machines by using iptables. On the Windows client machines, the firewall is implemented by using netsh. Although most ports are blocked, some ports are open by default on the server and client machines.

### Open ports for incoming traffic on the server machine

Ports that are required for Information Server on Cloud to work are opened on private IP addresses. For details about these ports, see [Configuring your network](#).

Note the following changes in port configuration for Information Server on Cloud:

- The default ORB Listener port number is changed from 9100 to 9108
- ASB Agent's object port is fixed to 33311
- The default Information Server console port is changed from 9443 to 9446
- The default WebSphere® Application Server Bootstrap port is changed from 2809 to 2825

To connect to Information Server on Cloud from any client machine other than the one that is bundled with your part number, you must also open ports from that machine.

Category	Protocol	Port/Type	Reason that the port is open	From
HTTP+SSL	TCP	443	To ensure that HTTP communication works	All
SSH	TCP	Given in the Welcome letter	To allow SSH connections	Client Machines
Bootstrap	UDP	68	To receive messages from BOOTP configuration server at boot time	All

Table 20: Open ports on other Linux server machines in the offering (continued)

Category	Protocol	Port/Type	Reason that the port is open	From
PING	ICMP	Echo-request	To allow ping	All

## Open ports for incoming traffic on client machines

Table 21: Open ports on Information Server on Cloud clients

Port	Protocol	Profile	Reason that the port is open	Open from which machine
5986	TCP	Domain	Install the product by using Microsoft Windows Remote Management	All machines
3389	TCP	Domain	Connect to the client machine by using Microsoft Remote Desktop Connection	All machines
19443	TCP	Any	HTTPS port for service tier	All
Echo-request	ICMP	Any	Allow ping	All

(Optional) Open ports for incoming traffic on IBM Business Process Manager on Cloud

Table 22: Open ports if you decide to use IBM® Business Process Manager Standard

Component	Port	From which machine
BPM Process Center	WAS_Secure_port: 9043. WC_defaulthost_secure: 9443. DCS_UNICAST_ADDRESS: 9354. SIB_ENDPOINT_SECURE_ADDRESS: 7286. SIP_DEFAULTHOST_SECURE: 5061. IPC_CONNECTOR_ADDRESS: 9633.	Client machines. Information Server Services tier machine
BPM Process Server	WAS_Secure_port: 9045. WC_defaulthost_secure: 9444. DCS_UNICAST_ADDRESS: 9357. SIB_ENDPOINT_SECURE_ADDRESS: 7287. SIP_DEFAULTHOST_SECURE: 5063. IPC_CONNECTOR_ADDRESS: 9635.	Client machines. Information Server Services tier machine
SSH	Given in the Welcome letter	Client machines
Bootstrap	68	All
PING	Echo-request	All

(Optional) Open ports for incoming traffic on IBM Cognos Business Intelligence on Cloud

<i>Table 23: Open ports if you decide to use IBM Cognos® Business Intelligence</i>		
<b>Component</b>	<b>Port</b>	<b>From which machine</b>
IBM Cognos Business Intelligence	Dispatcher port: 9300. Log Server port: 9362	Client machines; Information Server Services tier machine
SSH	Given in the Welcome letter	Client machines
Bootstrap	68	All
PING	Echo-request	All

(Optional) Open ports on Information Server Services tier machine for incoming traffic from BPM and Cognos machine

<i>Table 24: Open ports if you decide to use IBM Business Process Manager Standard and IBM Cognos Business Intelligence</i>		
<b>Component</b>	<b>Port</b>	<b>From which machine</b>
Information Server Services tier	IIS Console port: 9446. Db2 port: 50000	BPM. Cognos

## Starting services after the Information Server on Cloud server machine reboots

When the Information Server on Cloud server machine restarts, services are started automatically. In some cases, you might need to start the services manually.

It takes several minutes for all services on the server machine to complete their startup. If you try to log in to any Information Server on Cloud client tools as soon as the server machine restarts, the services on the server machine might not yet be started. As a result, you cannot log in.

1. Wait a few minutes for all services to start, and then log in to either the IBM® InfoSphere® Information Server Web console.
2. If you still cannot log in, follow the steps to start the services manually.

## Setting up firewall security

You can change the security level of your IBM® Information Server on Cloud server by configuring the iptables firewall. On the client, you can use the netsh command from Microsoft Windows to show the firewall profile and to list open ports.

**Prerequisite:** You must know the credentials of a user and the IP address of the Information Server on Cloud client and servers. This information is in the “Access Credentials and Initial Setup” section of the Welcome letter that you received from your IBM® Sales Representative.

### About these tasks

When the Information Server on Cloud server and client are delivered, the network access to the Linux server and the Microsoft Windows client is secured.

If you need to access ports of the Information Server on Cloud server from your on-premises application, you must modify the iptables rules.

On the Information Server on Cloud client, you can show the firewall profile, list the open ports, and close an open port.

### To change the security level or to manage rules for the Information Server on Cloud server firewall

1. Log in to the root account of the Information Server on Cloud server.
2. Open a command-line window, and then type the command sudo.
3. Do any of following commands from the prompt:

### **List the rules in iptables**

You can print the rules with the line numbers. This command is useful when you need to insert a rule at a specific location in iptables.

```
iptables -L --line-numbers
```

### **Restrict access to the Information Server on Cloud server from a range of IP addresses**

You can restrict access to a subnet so that only those IP addresses can use ssh to access your Information Server on Cloud server.

```
iptables -A INPUT -i eth0 -p tcp -s <ip_subnet_address> --dport 22 -m state --state NEW,ESTABLISHED -j ACCEPT
```

where <ip\_subnet\_address> is the range of IP addresses in the subnet. For example, the following rule allows incoming ssh connections from only the 192.xxx.xxx.xxx subnet:

```
iptables -A INPUT -i eth0 -p tcp -s 192.xxx.xxx.xxx/24 --dport 22 -m state --state NEW,ESTABLISHED -j ACCEPT
```

### **Restrict access to the Information Server on Cloud server from a single IP address**

This restriction is the most stringent.

```
iptables -A INPUT -i eth0 -p tcp -s <ip_address> --dport 22 -m state --state NEW,ESTABLISHED -j ACCEPT
```

where <ip\_address> is the single IP address that is permitted to access the server. For example, the following rule allows incoming ssh connections from a single port on 192.xxx.xxx.xxx:

```
iptables -A INPUT -i eth0 -p tcp -s 192.xxx.xxx.xxx --dport 22 -m state --state NEW,ESTABLISHED -j ACCEPT
```

### **Open a port from another machine**

To connect to the Information Server on Cloud server from any other machine, you must open the IBM InfoSphere® Information Server ports from that machine. For a list of ports that are open by default for incoming traffic on the server machine, see the topic [Open ports on server and client machines](#).

```
iptables -A INPUT -p tcp -s <ip_address> --dport 9446 -j ACCEPT
```

where <ip\_address> is the IP address of the machine for which you want to open port 9446 on the Information Server on Cloud server machine. For example, the following rule opens port 9446 on the server for incoming traffic from port 9446 on 192.xxx.xxx.xxx:

```
iptables -A INPUT -p tcp -s 192.xxx.xxx.xxx --dport 9446 -j ACCEPT
```

### **Open the port for a fix pack or patch installation**

Fix packs and patches can be installed to the existing Information Server on Cloud server environment. By default, port 8446 is used for the installation. To do the installation in GUI mode from a remote machine, 8446 port must be opened.

```
iptables -A INPUT -p tcp -s <ip_address> --dport 8446 -j ACCEPT
```

where <ip\_address> is the IP address for which you want to open port 8446 on the Information Server on Cloud server machine. For example, the following rule opens port 8446 on the server for incoming traffic from port 8446 on 192.xxx.xxx.xxx:

```
iptables -A INPUT -p tcp -s 192.xxx.xxx.xxx --dport 8446 -j ACCEPT
```

If you use a port other than 8446 for installation, use the other port in the iptables command.

### **Add a rule at a specific line in the iptables**

The order of the rules is important. Rules are appended to the end of existing rules. For example, if the rule that you added follows a DROP rule that matches incoming traffic, the DROP rule is considered before the newly added rule. You can add a rule at a specific line in the iptables to push subsequent rules down in the list.

```
iptables -I INPUT line_number -p tcp -s <ip_address> --dport 22 -j ACCEPT
```

For example, the following rule adds a rule at line 6 to accept incoming traffic on port 22 from 192.xxx.xxx.xxx:

```
iptables -I INPUT 6 -p tcp -s 192.xxx.xxx.xxx --dport 22 -j ACCEPT
```

### **Delete all rules in iptables**

**Important:** Before you remove all rules, be sure that the default INPUT policy is ACCEPT. Otherwise, you cannot connect to the Information Server on Cloud server from anywhere.</div>

```
iptables -F
```

### **Save rules in iptables permanently**

To persist the modifications to the iptables rules after the system restarts, you must update the file `/etc/sysconfig/iptables`.

```
iptables-save > /etc/sysconfig/iptables
```

## **To show the Microsoft Windows firewall profiles on the Information Server on Cloud client**

1. Access the Information Server on Cloud client by using Microsoft Remote Desktop Connection.
2. Open a command-line window on the client.
3. Run the command `netsh advfirewall show allprofiles`.

## **To show a list of all open ports on the Information Server on Cloud client**

1. Access the Information Server on Cloud client by using Microsoft Remote Desktop Connection.
2. On the Information Server on Cloud client, click **Start > Administrative Tools > Server Manager**.
3. In the upper-right corner of the Server Manager window, click **Tools**. From the list, select **Windows Firewall with Advanced Security**.
4. In the left pane, click **Inbound Rules** to see which ports are open for incoming network traffic.

## **To block an open port on the Information Server on Cloud client**

Do the previous task. Right-click the port number, and then select the appropriate action.

## **Scripts to change the iptables firewall settings**

You can change the settings of the iptables firewall to suit the needs of your organization.

Use scripts to help manage the iptables firewall settings on the cloud server. The following scripts are located in directory `/opt/IBM/scripts/` on the server. To use these scripts, you must be logged in to an administrative account on the server.

Refer to [Adding and connecting extra DataStage on Cloud Designer Client machines](#) to modify the iptable rules to access the Information Server on Cloud Enterprise Edition server and Information server on Cloud Data Quality server from an IBM® InfoSphere® Information Server client other than the Information Server on Cloud client that is provided as part of the offering.

### **allow\_datastage\_from\_ip.sh**

Use this script to access the DataStage® on Cloud server from an IBM® InfoSphere® DataStage client other than the Information Server on Cloud client that is provided as part of the offering.

The command line parameter `IP_address` is the IP address of the computer that you want to allow access to the DataStage on Cloud server.

For example, the following command enables `192.xxx.xxx.xxx` to access the DataStage on Cloud server.

```
allow_datastage_from_ip.sh 192.xxx.xxx.xxx
```

### **allow\_ping\_from\_ip.sh**

Use this script to be able to ping the Information Server on Cloud server from a specific computer.

The command line parameter `IP_address` is the IP address of the computer that you want to allow to ping the Information Server on Cloud server.

For example, the following command enables `192.xxx.xxx.xxx` to ping the server.

```
allow_ping_from_ip.sh 192.xxx.xxx.xxx
```

### **allow\_port\_from\_ip.sh**

Use this script to allow access to a specific port on the Information Server on Cloud server from a specific computer.

The first command line parameter `IP_address` is the IP address of the specific computer. The second parameter `Port_number` is the port on the server to allow the connection from `IP_address`.

For example, the following command 8446 enables `192.xxx.xxx.xxx` to connect to the server at port 8446.

```
allow_port_from_ip.sh 192.xxx.xxx.xxx
```

### **delete\_rule\_with\_linenumber.sh**

Use this script to delete a rule at a specific line of the iptables firewall on the Information Server on Cloud server.

The command line parameter `line_number_of_rule` is the line number of the rule that you want to delete.

**Tip:** To print the line numbers of iptables rules, use the command: `iptables -L --line-numbers`.

For example, the following command deletes the rule at line 20.

```
delete_rule_with_linenumber.sh 20
```

### **save\_iptables\_changes\_redhat.sh**

Use this script to persist the changes to iptables rules that you made on the Information Server on Cloud server. If the changes are not persisted, the changes might be lost when you restart the server.

### **allow\_services\_tier\_from\_bpm\_cognos\_ip.sh**

Use this script to access the Information Server on Cloud server services tier for Information Server on Cloud Data Quality and Information Server on Cloud Enterprise Edition from the machine where IBM Business Process Manager and IBM Cognos® Business Intelligence are installed.

The command line parameter `IP_address` is the IP address of the computer that you want to allow access to the Information Server on Cloud services tier machine.

For example, the following command enables `192.xxx.xxx.xxx` to access the Information Server on Cloud services tier server.

```
allow_services_tier_from_bpm_cognos_ip.sh 192.xxx.xxx.xxx
```

### **Related information:**

- [Getting started and using DataStage on Cloud](#)
- [Getting started and using IBM Information Governance Catalog on Cloud](#)



## Enhancing security of Information Server on Cloud computers

---

You can improve the security of your Information Server on Cloud computers by using SSH key pairs, changing the default SSH port, and restricting SSH access to specific ports.

### Related tasks

- [Getting started and using IBM Information Server on Cloud Enterprise Edition](#)
- [Getting started and using IBM Information Server on Cloud Data Quality](#)
- [Getting started and using IBM DataStage on Cloud Designer Client](#)
- [Getting started and using IBM Information Governance Catalog on Cloud](#)

## Setting up SSH keys

You can use SSH keys to restrict access to your Information Server on Cloud server from specific computers.

**Prerequisite:** You must be logged in to the Information Server on Cloud server in the root account.

You can restrict access to the Information Server on Cloud server by using SSH keys. You create the key pairs on the local computer or virtual machine from which to connect to the Information Server on Cloud server, and then copy the keys to the server.

Next, you disable SSH login password authentication. Because this step disables connections to the cloud server that use login and password authentication, you cannot connect to the Information Server on Cloud server if you cannot connect to your local computer that generated the SSH keys.

1. On your local computer, type the command **ssh-keygen** to create SSH key pairs. Store the key pair in the default location. By default, private keys are stored in your home directory of `/home/username/.ssh/id_rsa`. Public keys are stored in `id_rsa.pub`.
2. Copy the public SSH key to the Information Server on Cloud server that you want to connect to from your local computer. In the Information Server on Cloud server, append the public key to the file `$HOME/.ssh/authorized_keys`. The directory `$HOME` might be `/home/<user_name>`, where `<user_name>` is your login name. For the root user, `$HOME` might be `/root`.
3. Disable SSH logon by following these steps:
  - a. At the command-line prompt on your local computer, open the file `/etc/ssh/sshd_config`.
  - b. Add the line `PasswordAuthentication no`. If this line exists but is commented out, uncomment the line.
  - c. Run the command `service sshd restart` to restart the SSHD service. **Important:** After you restart the SSHD service, you cannot connect to the Information Server on Cloud server by using login and password. To avoid being locked out of the Information Server on Cloud server, test your connection to the server by using SSH keys before you restart the SSHD service.

## Allowing SSH access to specific IP addresses

You can allow SSH access to the IBM® Information Server on Cloud server from specific IP addresses to prevent intrusion and hacking.

**Prerequisite:** You must be logged in to the Information Server on Cloud server in the root account.

**Prerequisite:** You must know the port that is used for SSH. The port number is in the “Access Credentials and Initial Setup” section of the Welcome letter that you received from your IBM Sales Representative. It is also in the file `/etc/ssh/sshd_config` on the Information Server on Cloud server.

The SSH port of the Information Server on Cloud server is open only to client machines that are provisioned with the Information Server on Cloud Enterprise Edition offering. This procedure opens the Information Server on Cloud server to other machines.

Run the `allow_port_from_ip.sh` script. The first argument is the IP of the computer you want to give access to. The second argument is the SSH port number.

For example, the following command allows port 1234 from the IP address 192.xx.xx.xx to access the Information Server on Cloud server.

```
allow_port_from_ip.sh 192.xx.xx.xx 1234
```

## Managing LUKS keys on the IBM Information Server on Cloud server

Any partition that is encrypted by using Linux Unified Key Setup (LUKS) can have eight different keys. You can use any of the keys to open the encrypted partition. In addition, you can add new keys or remove existing ones according to your needs.

**Prerequisite:** You must know the credentials of the root user and the IP address of the Information Server on Cloud server. This information is in the “Access Credentials and Initial Setup” section of the Welcome letter that you received from your IBM® Sales Representative.

One LUKS key is specified in the “Encryption Details” section of the Welcome letter that you received from your IBM Sales Representative. This key file is used from /etc/crypttab file to unlock the partition when the cloud server reboots. For details about encrypted partitions on the Information Server on Cloud server, see the following topics:

- [Layout of Information Server Enterprise Edition on Cloud server and client disks](#)
- [Layout of Information Server Data Quality on Cloud server and client disks](#)
- [Layout of DataStage on Cloud server and client disks](#)
- [Layout of DataStage Designer Client on Cloud server and client disks](#)
- [Layout of Information Governance Catalog on Cloud server and client disks](#)

1. On the Information Server on Cloud server, open a command-line window.
2. Do any of the following tasks to manage your keys. The commands can be run from any directory.

### **Add new LUKS key**

```
cryptsetup luksAddKey <partition_name>
```

The parameter <partition\_name> is the partition that is encrypted.

For example, the partition for virtual servers is /dev/xvdc1. For bare metal servers, the partition is /dev/sda5.

Type in the existing key that is given in the Welcome letter. When prompted, enter the name of the new key.

### **Remove existing LUKS keys**

```
cryptsetup luksRemoveKey <partition_name>
```

When prompted, enter the name of the specific key that you want to remove.

### **Add a key from a file**

```
cryptsetup luksRemoveKey <partition_name> <keyfile_name>
```

The parameter <keyfile\_name> must include the full file path of the file.

The command asks for an existing key and then reads the new key from the file.

---

# Chapter 9. Troubleshooting cloud offerings

These topics contain troubleshooting information for Information Server on Cloud.

## Cannot install a patch on services tier machine

---

When you try to install a patch for Information Server on Cloud Enterprise Edition and Information Server on Cloud Data Quality on services tier machine, you are unable to open the installation wizard.

### Symptoms

During the installation, you receive a URL, which you can open on a Windows machine and run the installation by using the wizard (graphical mode). When you open the URL, you are unable to proceed.

The issue occurs on the Internet Explorer browser.

### Resolving the problem

To resolve the issue, complete the following steps:

1. In the URL, replace the host name with the private IP address of service tier.
2. Configure the Compatibility View settings of Internet Explorer browser:
  - a. Click **Tools > Compatibility View settings**.
  - b. Specify the URL with private IP address in the Add to this website field and then click Add.
  - c. Click **Save**.



# Notices

---

This information was developed for products and services offered in the U.S.A. This material may be available from IBM in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

## Notices

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing  
IBM Corporation  
North Castle Drive  
Armonk, NY 10504-1785 U.S.A.

For license inquiries regarding double-byte character set (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

Intellectual Property Licensing  
Legal and Intellectual Property Law  
IBM Japan Ltd.  
19-21, Nihonbashi-Hakozakicho, Chuo-ku  
Tokyo 103-8510, Japan

**The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law:** INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Corporation  
J46A/G4  
555 Bailey Avenue  
San Jose, CA 95141-1003 U.S.A.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

All statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

This information is for planning purposes only. The information herein is subject to change before the products described become available.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

#### COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

Each copy or any portion of these sample programs or any derivative work, must include a copyright notice as follows:

© (your company name) (year). Portions of this code are derived from IBM Corp. Sample Programs. © Copyright IBM Corp. \_enter the year or years\_. All rights reserved.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

#### **Privacy policy considerations**

IBM Software products, including software as a service solutions, ("Software Offerings") may use cookies or other technologies to collect product usage information, to help improve the end user experience, to tailor interactions with the end user or for other purposes. In many cases no personally identifiable information is collected by the Software Offerings. Some of our Software Offerings can help enable you to collect personally identifiable information. If this Software Offering uses cookies to collect personally identifiable information, specific information about this offering's use of cookies is set forth below.

Depending upon the configurations deployed, this Software Offering may use session or persistent cookies. If a product or component is not listed, that product or component does not use cookies.

The table shows how each InfoSphere® Information Server product and component uses cookies.

Table 25: Use of cookies by InfoSphere Information Server products and components

Product module	Component or feature	Type of cookie that is used	Collect this data	Purpose of data	Disabling the cookies
Any (part of InfoSphere Information Server installation)	InfoSphere Information Server web console	<ul style="list-style-type: none"> <li>• Session</li> <li>• Persistent</li> </ul>	User name	<ul style="list-style-type: none"> <li>• Session management</li> <li>• Authentication</li> </ul>	Cannot be disabled
Any (part of InfoSphere Information Server installation)	InfoSphere Metadata Asset Manager	<ul style="list-style-type: none"> <li>• Session</li> <li>• Persistent</li> </ul>	No personally identifiable information	<ul style="list-style-type: none"> <li>• Session management</li> <li>• Authentication</li> <li>• Enhanced user usability</li> <li>• Single sign-on configuration</li> </ul>	Cannot be disabled
InfoSphere DataStage®	Big Data File stage	<ul style="list-style-type: none"> <li>• Session</li> <li>• Persistent</li> </ul>	<ul style="list-style-type: none"> <li>• User name</li> <li>• Digital signature</li> <li>• Session ID</li> </ul>	<ul style="list-style-type: none"> <li>• Session management</li> <li>• Authentication</li> <li>• Single sign-on configuration</li> </ul>	Cannot be disabled
InfoSphere DataStage	XML stage	Session	Internal identifiers	<ul style="list-style-type: none"> <li>• Session management</li> <li>• Authentication</li> </ul>	Cannot be disabled
InfoSphere DataStage	IBM InfoSphere DataStage and QualityStage® Operations Console	Session	No personally identifiable information	<ul style="list-style-type: none"> <li>• Session management</li> <li>• Authentication</li> </ul>	Cannot be disabled
InfoSphere Data Click	InfoSphere Information Server web console	<ul style="list-style-type: none"> <li>• Session</li> <li>• Persistent</li> </ul>	User name	<ul style="list-style-type: none"> <li>• Session management</li> <li>• Authentication</li> </ul>	Cannot be disabled
InfoSphere QualityStage Standardization Rules Designer	InfoSphere Information Server web console	<ul style="list-style-type: none"> <li>• Session</li> <li>• Persistent</li> </ul>	User name	<ul style="list-style-type: none"> <li>• Session management</li> <li>• Authentication</li> </ul>	Cannot be disabled
InfoSphere Information Governance Catalog		<ul style="list-style-type: none"> <li>• Session</li> <li>• Persistent</li> </ul>	<ul style="list-style-type: none"> <li>• User name</li> <li>• Internal identifiers</li> <li>• State of the tree</li> </ul>	<ul style="list-style-type: none"> <li>• Session management</li> <li>• Authentication</li> <li>• Single sign-on configuration</li> </ul>	Cannot be disabled
InfoSphere Information Analyzer	Data Rules stage in the InfoSphere DataStage and QualityStage Designer client	Session	Session ID	Session management	Cannot be disabled

If the configurations deployed for this Software Offering provide you as customer the ability to collect personally identifiable information from end users via cookies and other technologies, you should seek your own legal advice about any laws applicable to such data collection, including any requirements for notice and consent.

For more information about the use of various technologies, including cookies, for these purposes, see IBM's Privacy Policy at <http://www.ibm.com/privacy> and IBM's Online Privacy Statement at <http://www.ibm.com/privacy/details> the section entitled "Cookies, Web Beacons and Other Technologies" and the "IBM Software Products and Software-as-a-Service Privacy Statement" at <http://www.ibm.com/software/info/product-privacy>.

## Terms and conditions for product documentation

Permissions for the use of these publications are granted subject to the following terms and conditions.

- Applicability** These terms and conditions are in addition to any terms of use for the IBM website.
- Personal use** You may reproduce these publications for your personal, noncommercial use provided that all proprietary notices are preserved. You may not distribute, display or make derivative work of these publications, or any portion thereof, without the express consent of IBM.
- Commercial use** You may reproduce, distribute and display these publications solely within your enterprise provided that all proprietary notices are preserved. You may not make derivative works of these publications, or reproduce, distribute or display these publications or any portion thereof outside your enterprise, without the express consent of IBM.
- Rights** Except as expressly granted in this permission, no other permissions, licenses or rights are granted, either express or implied, to the publications or any information, data, software or other intellectual property contained therein.
- IBM reserves the right to withdraw the permissions granted herein whenever, in its discretion, the use of the publications is detrimental to its interest or, as determined by IBM, the above instructions are not being properly followed.
- You may not download, export or re-export this information except in full compliance with all applicable laws and regulations, including all United States export laws and regulations.
- IBM MAKES NO GUARANTEE ABOUT THE CONTENT OF THESE PUBLICATIONS. THE PUBLICATIONS ARE PROVIDED "AS-IS" AND WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, NON-INFRINGEMENT, AND FITNESS FOR A PARTICULAR PURPOSE.

## Trademarks

IBM, the IBM logo, and [ibm.com](http://www.ibm.com)<sup>®</sup> are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml).

The following terms are trademarks or registered trademarks of other companies:

Adobe is a registered trademark of Adobe Systems Incorporated in the United States, and/or other countries.

Intel and Itanium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows and Windows NT are trademarks of Microsoft Corporation in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Java<sup>™</sup> and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

The United States Postal Service owns the following trademarks: CASS, CASS Certified, DPV, LACS<sup>Link</sup>, ZIP, ZIP + 4, ZIP Code, Post Office, Postal Service, USPS and United States Postal Service. IBM Corporation is a non-exclusive DPV and LACS<sup>Link</sup> licensee of the United States Postal Service.

Other company, product or service names may be trademarks or service marks of others.





