Progress

Progress DataDirect Hybrid Data Pipeline

Security Statement

Release 4.6.1

Progress DataDirect Hybrid Data Pipeline

Copyright

Preface

This document describes risk and vulnerability management strategies that are in place for Progress DataDirect Hybrid Data Pipeline security and governance.

Hybrid Data Pipeline leverages robust security mechanisms. Since it is typically deployed on-premises or through a managed private cloud, it is important for customers to ensure that end-to-end security is appropriately configured and managed. This includes leveraging your own defense in depth strategy, network security, encryption, access control, and other safeguards. The defense in depth security principle is a layering of security technologies and process of safeguarding the environment against known and emerging threats. Each ingress and egress point to must be managed, inspected, and validated through routine self and third-party assessment to prevent, identify, and correct vulnerabilities to protect against compromise. Coupled with appropriate on-premises IT security controls, Hybrid Data Pipeline can be an invaluable strategy for secure data access.

1

Hybrid Data Pipeline Architecture

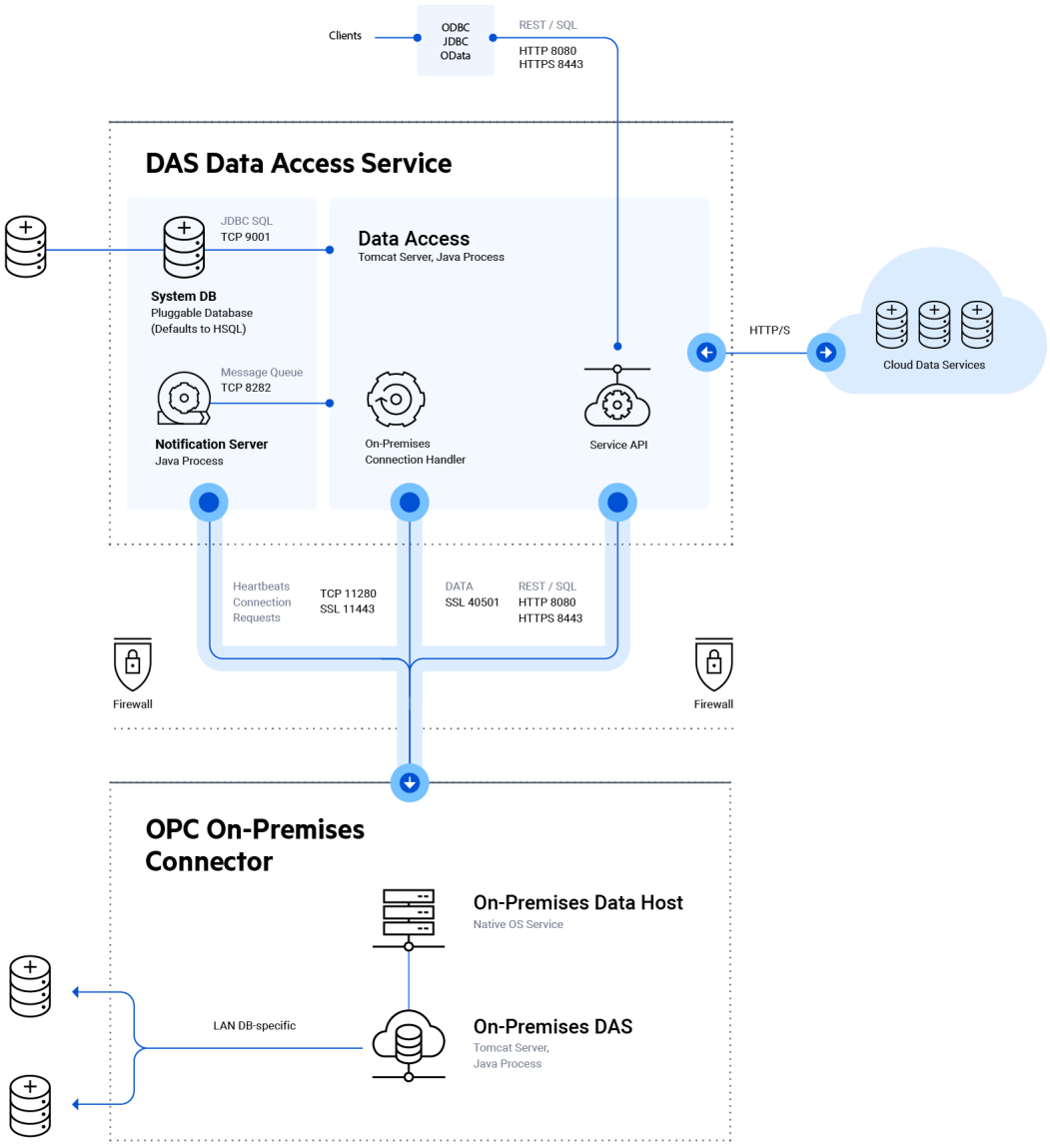
For details, see the following topics:

Overview of Operation

Understanding the security architecture of Hybrid Data Pipeline is critical to developing your own secure data access service. Figure 1 shows a high-level view of the various components and processes within Hybrid Data Pipeline. The key components include the clients, the Hybrid Data Pipeline server (which contains the DAS or Data Access Service), and the On-Premises Connector (OPC).

Note: The configuration of ports may differ in a cluster deployment or deployment using a load balancer.

Figure 1: Security Architecture



Clients

Hybrid Data Pipeline interfaces with applications (such as BI analytics or data management tools) using an industry standard data access interface for SQL and REST, typically ODBC, JDBC, or OData. When using ODBC or JDBC, Hybrid Data Pipeline must be deployed with a corresponding ODBC or JDBC driver. OData connectivity is handled by an OData layer within Hybrid Data Pipeline, and therefore does not require the use of a driver or separate component. Clients communicate with the main Hybrid Data Pipeline server over HTTPS to read or write data to a downstream application or database that may reside in the cloud or behind the firewall. Figure 1 illustrates secure communication over port 8080 or 8443 (configurable). HTTP is not recommended but usable for testing purposes.

Hybrid Data Pipeline Server

The Hybrid Data Pipeline server contains the DAS or Data Access Service as shown in Figure 1. The DAS is responsible for managing user and service configuration, brokering the flow of data between clients and databases, managing state, and communicating with on-premises connectors and more.

Storage of user and service configuration is handled by the System DB. HSQL is the embedded option, but the database is pluggable, allowing administrators to use their preferred relational database (such as Oracle, SQL Server, or MySQL). All communication with the pluggable database can be configured to occur over an encrypted channel, and the sensitive data at rest will always be encrypted. See Encryption for specific details.

When accessing data from a cloud data service (such as Salesforce) or a cloud platform (such as Microsoft Azure), the HDPS should be configured to use HTTPS. The exact configuration support varies by the endpoint being accessed.

Hybrid Data Pipeline may be deployed on one or more nodes behind a load balancer to scale out a solution. In this scenario, Hybrid Data Pipeline may be configured to use HTTPS for communication between individual nodes and between the load balancer and the nodes themselves.

On-Premises Connector (OPC)

The On-Premises Connector or OPC typically resides behind a firewall and is responsible for brokering the data access between the DAS and a database endpoint that is accessible via a LAN. A single Hybrid Data Pipeline instance can connect to one or more OPCs.

An OPC establishes a connection with the DAS using an outbound request to the Service API using HTTPS, as shown in Figure 1. The server will create a new data endpoint and return that address to the OPC. The OPC will then establish a mutually authenticated, SSL-secured session with the DAS to handle the flow of data from the LAN-accessible database to the application connected to the client.

OPC Authentication with the DAS

A unique random AES256 encryption key is generated for each Hybrid Data Pipeline instance. During the creation of installer redistribution files, this key with an added salt value is encrypted using a master AES256 encryption key and placed in the `OnPremise.properties` file as the AuthKey value. For example:

```
AuthKey=E2C884E892C59AA9D5A67CC3D045E28B20C5B7DF337BFD72972E768
   64581FC3E6A2456BAA37B80B58500B83A34D643BED491383FA75F2708A3684
   8A5789EBFEA28B2C25317C7DB85DA76D4DE4CFAE6E
```

During installation of the OPC, the user password provided to the OPC installer is encrypted with salt using the unique Hybrid Data Pipeline encryption key from the AuthKey value. The encrypted password is then placed in the `OnPremise.properties` file as the Auth value. For example:

```
Auth=0389310B6641AB1AFF844D79256402C92CC92B8F89986F5CB60B9BAE95
   F823B165CDB34653BD1CD202B6A373368FE996EB85ECEA71D1CE56A3941FCA8
   B2CA959
```

The OPC uses the master AES256 encryption key to decode the unique Hybrid Data Pipeline AES256 encryption key stored in AuthKey. This key is then used to decrypt the user password stored in Auth. The user ID and password are used to authenticate the OPC with the DAS over an HTTPS connection.

2

Deployment Patterns

For details, see the following topics:

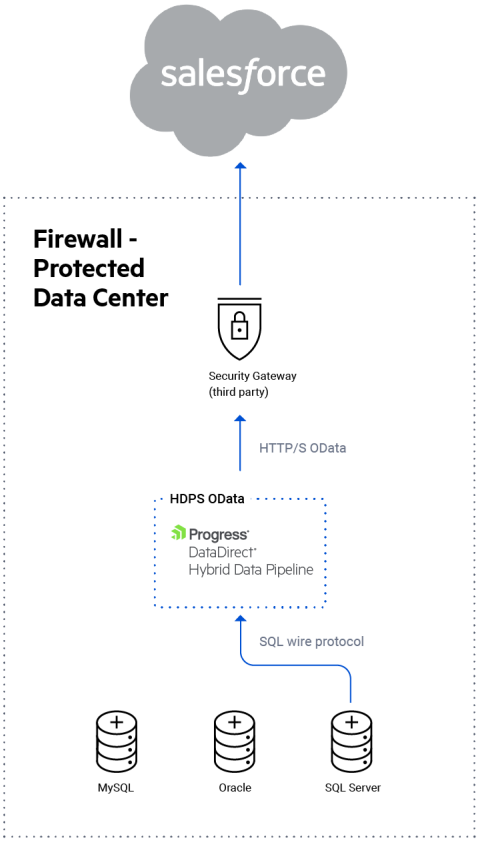
- Server-only Deployment
- OPC Deployment
- Cloud ISV Deployment: Data Integration
- Cloud ISV Deployment: Expose Data for Business Intelligence

Server-only Deployment

Many organizations can achieve cloud to on-premises data access using a server-only deployment of Hybrid Data Pipeline. This deployment pattern relies on a security gateway appliance to broker a pass-through communication channel between the cloud application and on-premises database. Figure 2 contains an example use case to highlight this deployment. In this example, Salesforce is accessing External Data Objects using a secure HTTPS OData connection. The organization's security gateway appliance is configured to accept this connection from Salesforce only and redirect it to Hybrid Data Pipeline. The security gateway and Hybrid Data Pipeline would be configured using mutual authenticated HTTP(S).

The Hybrid Data Pipeline server itself would access a local, LAN-accessible database (in this case, SQL Server) using an SSL-enabled or non-SSL connection, communicating using the database wire protocol.

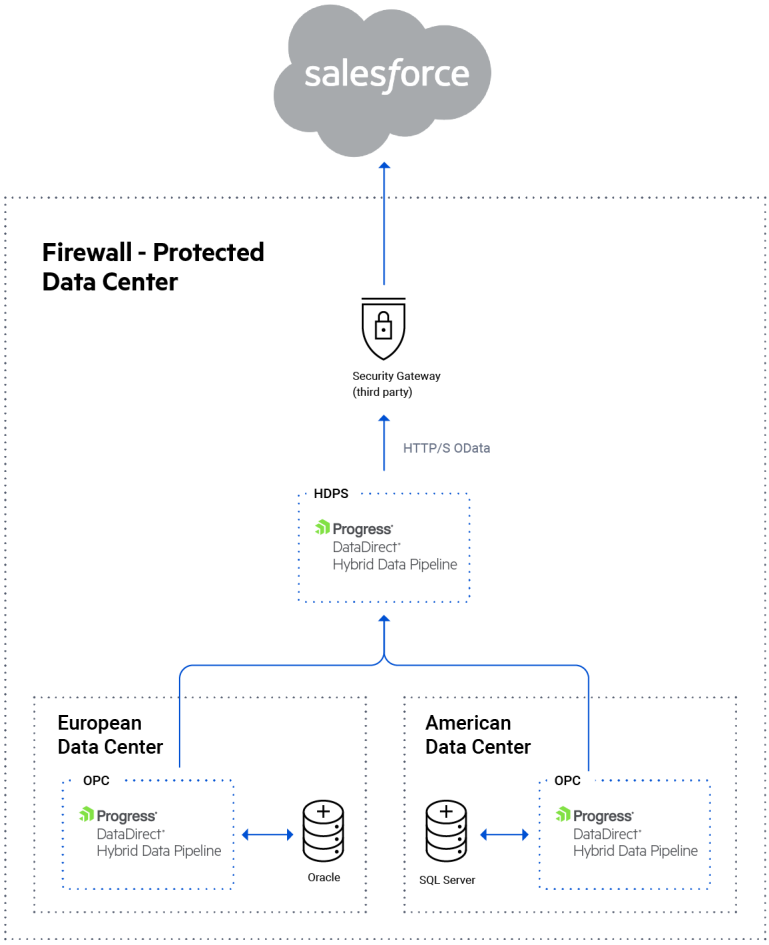
Figure 2: Server-only Deployment



OPC Deployment

Large organizations with multiple data centers or a highly complex network security strategy can use the OPC to assist with configuring access across multiple layers of security. Figure 3 highlights a deployment pattern for an organization with different data centers to service geographically organized divisions of their business. Here, the organization again wants to surface on-premises data to their Salesforce instance, but in this case, the data is further segmented by regional data centers.

Figure 3: OPC Configuration.

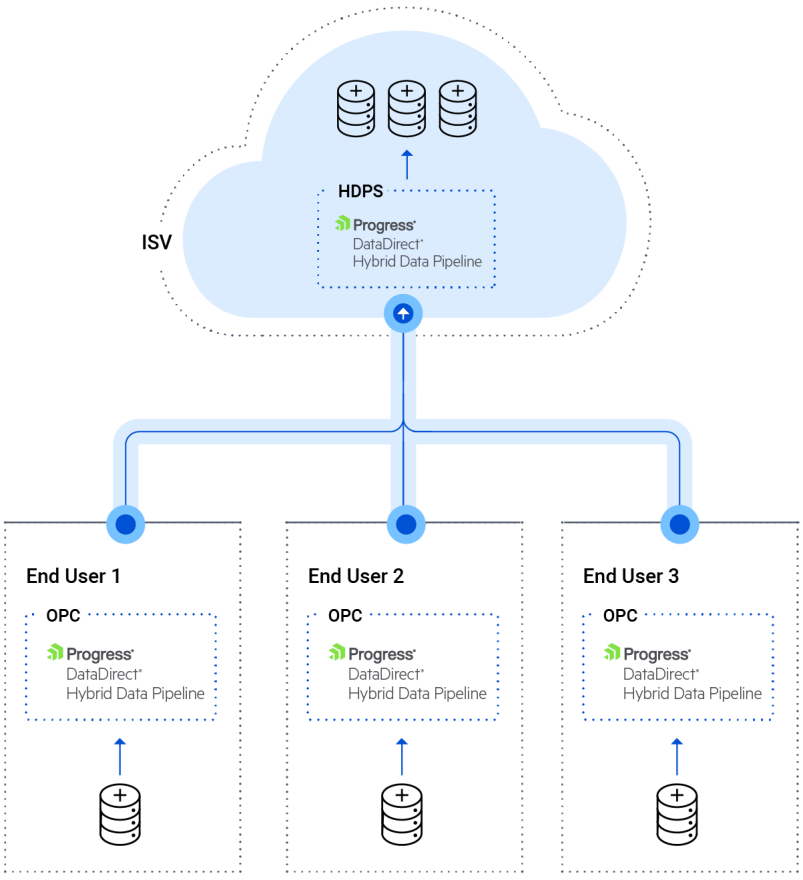


Cloud ISV Deployment: Data Integration

Unlike the last two enterprise deployment patterns, cloud software vendors will typically use an inverted pattern designed to streamline the creation of a data pipeline to their end-users' data. It is inverted in the sense that Hybrid Data Pipeline is hosted in the cloud, rather than on-premises or in a DMZ.

Figure 4 represents a typical deployment that a cloud ISV might use when integrating with legacy customer data residing behind a firewall. In this diagram, an OPC is white labeled and deployed at each end-user site, allowing the ISV to access this data from the cloud. Typically, the ISV will use this setup to provide real-time access to on-premises data, as external data objects without duplicating the data. Or, the ISV may want to start ingesting this data for analytics or data management, or to facilitate migration from on-premises systems to a hybrid or pure cloud deployment.

Figure 4: Cloud ISV Deployment for Data Integration

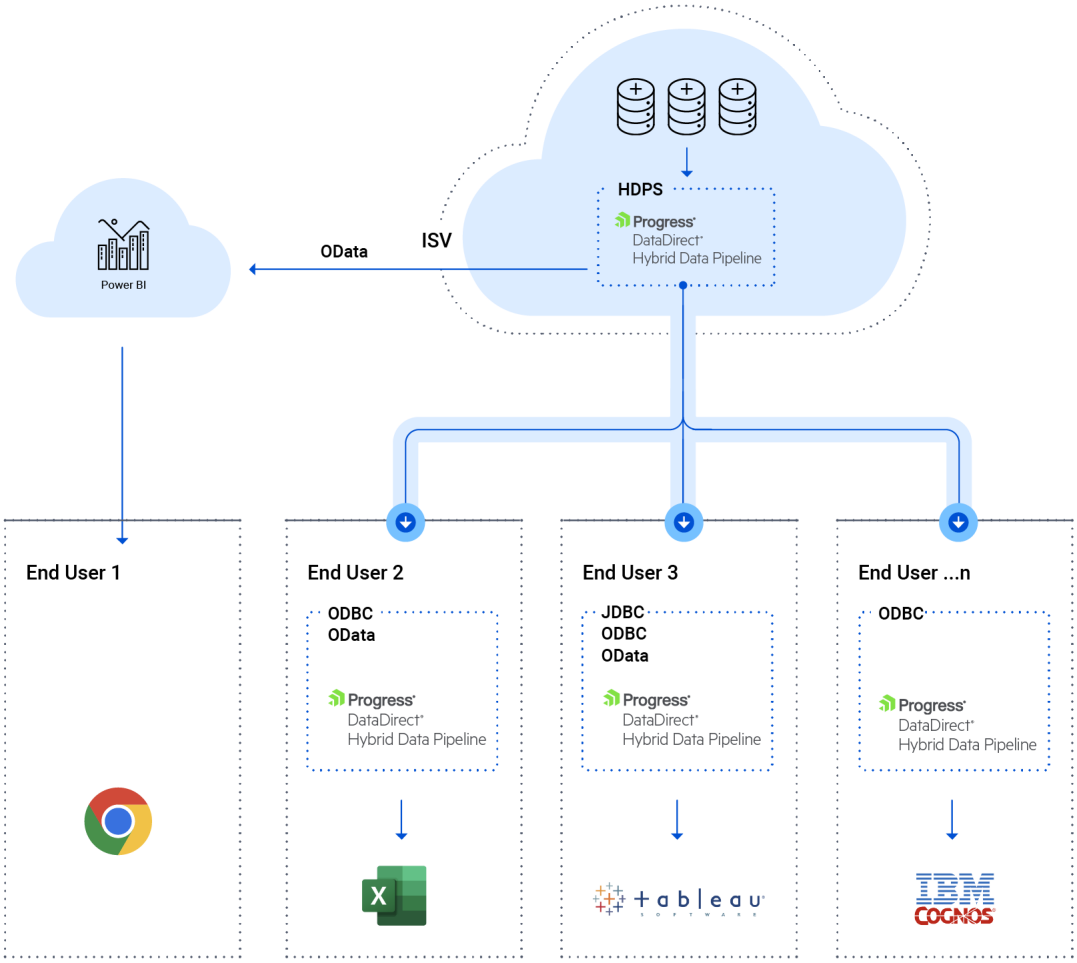


Cloud ISV Deployment: Expose Data for Business Intelligence

Another common scenario for an ISV is to provide end user site access to their cloud-hosted database for business intelligence (BI). In this deployment, end-users want the flexibility to use their favorite BI or reporting tool (such as Excel, Cognos, Business Objects, Crystal Reports, Tableau, Qlik, or Microstrategy).

Figure 5 represents an architecture that would allow an ISV to expose cloud-hosted data to end-users, without exposing the underlying database directly to the internet. Hybrid Data Pipeline makes it easy to configure a secure and scalable, HTTPS-based communication channel to the database via standards-based SQL and REST interfaces compatible with virtually any end-user BI tool.

Figure 5: ISV Deployment for Exposing Data for Business Intelligence



3

Proactive Mitigation and Remediation

For details, see the following topics:

- Software Development Life Cycle

Software Development Life Cycle

Progress operates an advanced development organization that prides itself on the security of our products. Security is pervasive across the Software Development Life Cycle, from tooling strategies to process to testing to staff culture (see DataDirect Product Life Cycle for more details). Secure coding, application testing, and continual developer security training are pillars of our program.

Hybrid Data Pipeline code is built, reviewed and validated by developers using Open Web Application Security Project (OWASP) guidance to minimize potential for vulnerabilities. Development and technical support teams undergo secure coding training and security awareness education.

4

Strategies for Reducing Public Cloud Risk

For details, see the following topics:

- Identification and Monitoring
- Compliance
- Encryption
- Penetration Testing
- DataDirect Security Guidelines

Identification and Monitoring

When choosing your hosting solution, it is important to ensure that end-to-end security is appropriately configured and managed. For example, many cloud hosting services handle load-balancing and firewall containments, real-time application query filtering, and 24x7 monitoring. Application servers should be configured with malware and anti-virus protections, patched and hardened against known flaws and vulnerabilities, and subject to periodic third-party assessments including penetration testing.

Compliance

General Data Protection Regulation

On May 25, 2018, the General Data Protection Regulation (GDPR) took effect in the European Union (EU). The GDPR expands the rights granted to EU individuals and places many obligations on organizations that market to, track or handle EU personal data. The GDPR must be adhered to by organizations that are located in the EU or do business in the EU that collect, store, transfer or use personal data about EU individuals.

Personal data is any information that relates to an identified or identifiable living individual. Different pieces of information, which collected together can lead to the identification of a particular person, also constitutes personal data. Examples of personal data may include:

- A name and surname
- A home address
- An email address such as name.surname@company.com
- An identification card number
- Location data
- An Internet Protocol (IP) address

Storage of Personal Data using Hybrid Data Pipeline Server

Hybrid Data Pipeline is responsible for managing user and service configuration, brokering the flow of data between clients and databases, and communicating with on-premises connectors (OPC). Data (personal or non-personal) that is brokered between the clients and databases via DAS and/or OPCs is not permanently persisted during transmission.

Storage of user and service configuration information is handled by the System DB. The user information stored within the System DB may contain personal data specific to user names, credentials, and IP addresses.

Progress operates its IT and development infrastructure in general alignment with SOC2, NIST 800-53, and ISO2700X best practices.

As with identification and monitoring, it is important to ensure that compliance is appropriately configured and managed by your hosting solution.

Encryption

Hybrid Data Pipeline can be configured to accept only authorized (cryptographic) communications from known administrative endpoints. All customer-sensitive data elements (including remote credential/database pairings stored) are protected by encryption, both at rest (AES-256) and in transit. Hybrid Data Pipeline runs on an Apache Tomcat web server. TLS 1.2 is the minimum supported version of TLS. Client components such as the ODBC and JDBC driver default to TLS 1.2. All user passwords are encrypted using SHA-256-bit one-way hash with per-user salt. The following table summarizes supported encryption protocols for Hybrid Data Pipeline components.

Component	Protocols Accepted
Hybrid Data Pipeline Server	TLS 1.2 and 1.3

Component	Protocols Accepted
ODBC and JDBC drivers	TLS 1.2 and 1.3
On-Premises Connector (OPC)	TLS 1.2 and 1.3

Penetration Testing

Secure code reviews and third-party penetration testing are performed quarterly as a validation of our thorough monthly internal, redundant testing and evaluation methodologies by our own team of certified ethical hackers and security experts.

DataDirect Security Guidelines

The Progress DataDirect Security Guidelines outline the general principles under which Progress manages the reporting, management, discussion, and disclosure of Security Vulnerabilities discovered in DataDirect software, including Hybrid Data Pipeline, and related components. Please refer to Progress DataDirect Security Guidelines for more details.