**General**

1. **What features are provided in the Data Hub Azure?**

​See the [roadmap and releases](https://basproducts.atlassian.net/wiki/spaces/EDA/pages/3164307508/Roadmap+and+Releases) to see what features are available with each release. Release notes accompany each release to provide more details on updates and new features.  
Additionally, visit our [news page](https://bp365.sharepoint.com/sites/DataAnalyticsPlatforms/SitePages/Data%20Hub%20Azure%20Latest%20News.aspx) to get the latest news on Data Hub Azure.

1. **How can I learn to use the Data Hub?**

Please visit our [onboarding guide](https://bp365.sharepoint.com/sites/DataAnalyticsPlatforms/SitePages/Data%20Hub%20Azure%20Onboarding.aspx) to begin your learning journey on Data Hub Azure. The onboarding guide will introduce you to the Data Hub Azure and describe guidelines to follow when using the platform. The guide will also provide information on the sandbox offering which enable users to experiment with Data Hub Azure. For more information on the sandbox opportunity, see [sandbox section of the onboarding guide](https://bp365.sharepoint.com/sites/DataAnalyticsPlatforms/SitePages/Data%20Hub%20Azure%20Onboarding.aspx#%E2%80%8B%E2%80%8B%E2%80%8B%E2%80%8B%E2%80%8B%E2%80%8B%E2%80%8B%E2%80%8B%E2%80%8B%E2%80%8B%E2%80%8B%E2%80%8B%E2%80%8B%E2%80%8B%E2%80%8B%E2%80%8B%E2%80%8B%E2%80%8B%E2%80%8B%E2%80%8B%E2%80%8B5.-get-started).

1. **Where do I find technical information about nodes and domains available on Data Hub Azure?**

Visit our [​Node Information](https://basproducts.atlassian.net/wiki/spaces/EDA/pages/2789213089/Node+Information) Confluence page to find more information about deployed nodes and the target node design. On this confluence page you will find a list of deployed nodes and their relevant technical information.

1. **Where do I go to understand the terminology used in the Data Hub Azure?**

​Please visit our​ [glossary page](https://basproducts.atlassian.net/wiki/spaces/EDA/pages/3060334639/Glossary) in confluence to find definitions of common terminology used across Data Hub Azure.

There you can find the distinctions between some of the key concepts described in our [Operating Model](https://bp365.sharepoint.com/sites/DataAnalyticsPlatforms/SitePages/Data%20Hub%20Azure%20Platform%20OpModel.aspx):

1. Nodes vs Domains
2. Platform Ops vs Node Ops vs Domain Ops teams
3. Raw vs Enriched vs Curated data
4. **What is the Disaster Recovery plan for Data Hub Azure?**

Information on RPO (Recovery Point Objective) and RTO (Recovery Time Objective) for the hub and nodes will be maintained on our [​Recovery Objective - Data Platform](https://basproducts.atlassian.net/wiki/spaces/EDA/pages/3222044790/Recovery+Objective+-+Data+Platform) Confluence page. Further information on RPO and RTO will be released when available on this page.

**Access**

1. **How is access managed on the platform?**

The data access framework is owned by the Platform team. Data access and control is owned by the Data Domain Ops team. The Data Data Domain Ops team will have a data access operator who grants data access through the data access request process which includes approval from data owners. More details on this process, please visit our ​[Data Access Service](https://bp365.sharepoint.com/sites/DataAnalyticsPlatforms/Lists/Mesh%20FAQs/DispForm.aspx?ID=73&e=UeM0lm)page on confluence.

1. **How do I request access to Azure resources on a domain?**

**​**Users can request access to resources through the Azure Portal. The Data Domain Admins will process the request and add the user to a security group to access the resources. If you are unable to log in to the Azure Portal, contact your data domain administrator to request access.

You can find details about the node and data domains relevant to you, as well as your data domain custodian on the [Node Information](https://bp365.sharepoint.com/sites/DataIkuPlatformSharedService/_layouts/15/Doc.aspx?OR=teams&action=edit&sourcedoc=%7b43AAB39E-947D-44E2-8357-23F80FE1507F%7d) Confluence page.

1. **How can users access Data Hub Azure resources?**

Users can access Data Hub Azure resources directly via ZScaler (BP VPN Client) which enables secure authenticated connections to protected network resources over local and wide area networks. For more information on how to set this up, visit​ [tZScaler for the Azure Data Hub](https://basproducts.atlassian.net/wiki/spaces/EDA/pages/3227845189/ZScaler+for+the+Azure+Data+Hub" \o "https://basproducts.atlassian.net/wiki/spaces/EDA/pages/3227845189/ZScaler+for+the+Azure+Data+Hub)on confluence.

1. **How are access and resource groups controlled within each of the nodes?**

On the Data node, Platform Ops team access is applicable to those Data Node and underpinning resource groups which are created specifically for Core node services. Access to data node spoke is managed through an AAD account, wherein the Platform Ops would activate PIM (Privileged Identity Management) using a AAD group setup at the Management group level.

The resource groups defined in the node would fall into 3 categories:

MAPS Resource Groups - These are the RSGs that are created as part of Spoke provisioning. The access to MAPS RSGs are limited to Platform Ops and will be granted only on exceptional basis to other teams.Shared Core Node Services Resource Groups - These include resources such as the shared ingestion Data Factory, Databricks,  Hive & ADF Metastore SQL, SHIR, ADLS and Synapse. They are managed by the Platform Ops and consumed by Node Ops for the ingestion activities. Data Domain resources - These are the resource groups which are created and configured specifically for a single Data Domain. They are used by the Domain Ops groups and managed by the Domain Ops Admin groups.

For more details on the access and privileges within nodes, please visit the [Control Plane Resource Access](https://basproducts.atlassian.net/wiki/spaces/EDA/pages/3153397694/Control+Plane+Resource+Access)Confluence page.

**Roles & Responsibilities**

1. **What are the roles of the different teams working on the platform?**

To learn more about the roles of different teams working on the platform,  visit the Persona Category section of the Data Hub Azure [Personas page](https://bp365.sharepoint.com/sites/DataAnalyticsPlatforms/Lists/Mesh%20FAQs/DispForm.aspx?ID=66&e=j8JAbW)in Confluence

1. **Are the Node Ops team part of the dataWorx organisation?**

The ​Node Ops teams are responsible for non-production and production customer nodes and managing feature deployment on nodes. They enable the domain delivery groups and work closely with the central platform teams. ​In the current operating model, the Node Ops team is centrally owned by Azure Data Hub. As the platform matures, the Node Ops team will be owned by the dataWorx organisations.

1. **What can the Node Ops team do for dataWorx users?**

The Node Ops team supports dataWorx teams with a variety of capabilities including data node maintenance and capabilities, and feature deployment and release of blueprints. For more details on the specific activities Node Ops team would perform, please visit the [​Data Hub Azure Personas](https://bp365.sharepoint.com/sites/DataAnalyticsPlatforms/Lists/Mesh%20FAQs/DispForm.aspx?ID=68&e=6EGuiR)Confluence page.

1. **Who are the product and service owners of the platform?**

Product owners are accountable for Core Services and Feature Development. For example, they will define the features that enable users to deliver new functionality.

The service owner is responsible for management and maintenance of the platform service offerings, and ensuring it operates in line with bp's confidentiality, integrity and availability requirements. They are responsible for any products provided through the service.

1. **What are the responsibilities of the Node Ops team in ingestion pattern setup and configuration?**

The Node Ops team is responsible for configuring the ingestion patterns to deliver the data to the appropriate domain teams. They will monitor the delivery of the data. For more information on roles and responsibilities of teams involved in the ingestion setup and configuration, please visit the [​Ingestion Pattern Logical Workflow](https://basproducts.atlassian.net/wiki/spaces/EDA/pages/3050211159/Data+Registration+Ingestion+Service)in Confluence.

1. **Which users can create resources in a domain?**

End users cannot create new resources, but they can request resources from the Data Domain Admins.

**Data**

1. **How do I get access to data?**

Access to data in the Data Hub Azure is controlled through the use of Saviynt Enterprise Roles (SERs). Users who require access can request the SERs associated with their respective domains. Visit our [Accessing Data](https://basproducts.atlassian.net/wiki/spaces/EDA/pages/3153626398/Accessing+Data)page on Confluence to follow the instructions to request access to data.

1. **Can I use PII or secret data on the Data Hub Azure platform?**

Sensitive, PII, or secret data are not allowed on the Data Hub Azure platform. It is the responsibility of the users to ensure no sensitive, PII or secret data is ingested. PII data processing is expected to be available in Q2, for more details on this roadmap please visit our [Roadmap and Releases](https://basproducts.atlassian.net/wiki/spaces/EDA/pages/3164307508/Roadmap+and+Releases)Confluence Page.

1. **How do I ingest data?**

The process for data ingestion starts with registering data by providing all requested information on the data source, including approval from appropriate data owners. For detailed instructions on how to kick off this process visit our [Data Registration and Ingestion Service](https://basproducts.atlassian.net/wiki/spaces/EDA/pages/3050211159/Data+Registration+Ingestion+Service)Confluence page.

Please note that the Node Ops team is responsible for data ingestion into Data Hub Azure and will be supporting teams through this process.

1. **What is the best way to view the available ingestion patterns and the roadmap for those in development?**

Please visit the [Roadmap for Ingestion Patterns](https://basproducts.atlassian.net/wiki/spaces/EDA/pages/3218014772/Ingestion+Pattern+Roadmap)page in Confluence to view all available and planned ingestion patterns.

1. **Is there a data catalog I can use to search for data assets?**

This will be enabled in a later release (see [Data Catalog Service](https://basproducts.atlassian.net/wiki/spaces/EDA/pages/2807365835/Data+Catalog+Service)) when the Data Hub Azure is integrated with the bp Data Catalog. For an overview of the expected service, visit our [Data Catalog Service](https://basproducts.atlassian.net/wiki/spaces/EDA/pages/2807365835/Data+Catalog+Service)in Confluence.

1. **How do I prepare data on Data Hub Azure?**

Please visit the [Data Domain Development](https://basproducts.atlassian.net/wiki/spaces/EDA/pages/2746975511/Data+Domain+Development)guide where you can learn how to prepare data on the Data Hub Azure platform. ​The tools used to prepare data are Azure Data Factory, Azure Databricks and Azure Synapse SQL Pools. Usage of these tools is documented under the [Data Domain Development](https://basproducts.atlassian.net/wiki/spaces/EDA/pages/2746975511/Data+Domain+Development)guide.

1. **How do I combine data from different domains?**

This is not available in the current release. This capability will be available in release 1. For more information visit the ​[Roadmap and Release](https://basproducts.atlassian.net/wiki/spaces/EDA/pages/3164307508/Roadmap+and+Releases)Confluence page.

1. **How can one node access data from another node?**

This is not available in the current release. Intra-node consumption will be available in release 1 and inter-node consumption will be available in release 2. For more information visit the [​Roadmap and Release](https://basproducts.atlassian.net/wiki/spaces/EDA/pages/3164307508/Roadmap+and+Releases)Confluence page.

1. **Can data be moved between Azure and AWS data hubs?**

In this current release you cannot move data between Azure and AWS data hubs. This capability will be available in release 2. For more information visit the​ [Roadmap and Release](https://basproducts.atlassian.net/wiki/spaces/EDA/pages/3164307508/Roadmap+and+Releases)​ Confluence page.

1. **Which tools can I use to consume data from Data Hub Azure?**

Power BI can be used for data consumption in the Data Hub Azure. For more information on how to consume data using Power BI, visit the [Power BI Guide](https://basproducts.atlassian.net/wiki/spaces/EDA/pages/3216441975/Power+BI+Guide)in Confluence.

1. **How do I connect Power BI to a data source in Data Hub Azure?**

Prior to accessing Azure Data Hub sources with Power BI you’ll need the following:

Connection to BP network, ZScaler or native. Power BI Desktop application Your BP user account having access to a data domain (being in the Domain Ops group)  
Once you have those prerequisites in place, you can connect Power BI to a data source in Data Hub Azure. There are three sources to connect to via Power BI: Data Lake, Azure Synapse, and Databricks.

For details on how to connect to these sources, please visit the [Power BI Guide](https://basproducts.atlassian.net/wiki/spaces/EDA/pages/3216441975/Power+BI+Guide)in Confluence.

1. **Which Data Hub Azure data sources can I use with Power BI?**

The primary data sources are Synapse SQL Pools (similar to SQL Server) and Databricks. You can also use files uploaded to the Azure Data Lake Store workspace zone.

1. **Can I use Azure Data Studio to connect with Azure Synapse?**

You can use Azure Data Studio to connect with Azure Synapse. Azure Data Studio provides a friendlier UI to enable users to add data to Azure Synapse as well as create schemas and tables. For more information on how to use Azure Data Studio, visit our [Using Data in the Sandbox](https://basproducts.atlassian.net/wiki/spaces/EDA/pages/3073606074/Using+Data+in+the+Sandbox)Confluence page.

1. **How can data be exported outside of the Data Hub Azure?**

Depending on your use case and requirements, data can be exposed through Power BI and exported from there or downloaded through Storage Explorer.

Further technical egress is on the roadmap, more detail coming soon. To get the latest news on upcoming releases, visit our [news page](https://bp365.sharepoint.com/sites/DataAnalyticsPlatforms/Lists/Mesh%20FAQs/DispForm.aspx?ID=65&e=vaTNON).

**Technical**

1. **If I get a policy denied error in Azure Data Factory (ADF), what can I do?**

Please visit our support page [support page](https://bp365.sharepoint.com/sites/DataAnalyticsPlatforms/Lists/Mesh%20FAQs/DispForm.aspx?ID=76&e=8fXouu)to raise a ServiceNow ticket.

1. **I’m experiencing an issue connecting from Azure Data Factory (ADF) to Azure Data Lake Storage (ADLS) and Blob Storage.**

Please visit our support page [support page](https://bp365.sharepoint.com/sites/DataAnalyticsPlatforms/Lists/Mesh%20FAQs/DispForm.aspx?ID=76&e=8fXouu)to raise a ServiceNow ticket.

1. **I am experiencing issues with data type or file type when creating a schema on parquet files.**

If you specify that a field is a Date Time in a schema, and you're reading from a test-based format, the date needs to be a format that spark can parse. Here are a couple of options for handling that:

1) Pull the column in as a string and then add a subsequent transform that parses that string to a Date Time, passing it in the format of the Date Time (e.g. "yyyy-MM-dd", or "dd/MM/yyyy").

2) You can pass a datetimeformat option into a call to spark read (e.g. spark.read.csv.option("timestampFormat", "dd/MM/yyyy").

If you are still experiencing an issue, visit our [Support](https://bp365.sharepoint.com/sites/DataAnalyticsPlatforms/SitePages/Data%20Hub%20Azure%20Platform%20Support.aspx) page to raise a ServiceNow ticket.

1. **Is there an auto shutdown event scheduled for VM?**

Non-Prod Virtual Machines automatically shutdown at 7 pm local time of the VM region. To apply for an exception, you need to follow these steps found on our Auto shutdown Exception Request in SharePoint.

1. **I am able to establish the Azure Data Lake Storage (ADLS) connection from Data Factory (DF), how do I run the test connection?**

To test the connection, navigate to the Test Connection dialog. There is an option to test the connection to the service, and to a file path. It works in a similar way to the data set file browser, in that if you select file path and populate the container name the test will succeed.

**Sandbox Node**

1. **What is the Data Hub Azure Sandbox?**

​The Data Hub Azure Sandbox Node provides users with a controlled environment where they can explore the platform components and services. Users can expect a production-like environment to assess how the platform can support their specific use case.

The sandbox experience makes it easy for users to move from the sandbox to the platform by speeding up the on-boarding of their datasets.

It's important to note that the sandbox has limitations, including programmatic ingestion of data.

1. **How do I register my interest in using the Data Hub Azure Sandbox?**

​To register your interest in using the sandbox environment, please visit ServiceNow and raise a request against platform adoption configuration item. For more detail on how to do this, visit our [Support](https://bp365.sharepoint.com/sites/DataAnalyticsPlatforms/SitePages/Data%20Hub%20Azure%20Platform%20Support.aspx) Page.

1. **What can I do on the Sandbox Node?**

In the sandbox node, users have the opportunity to explore Data Hub Azure & node components in a controlled environment. While exploring, users can utilise learning pathways to help upskill on available services before onboarding onto the platform. For more information on the sandbox environment and what it can offer you please visit view our [Sandbox Onboarding Documentation](https://bp365.sharepoint.com/sites/DataAnalyticsPlatforms/SitePages/Data%20Hub%20Azure%20Onboarding.aspx#get-hands-on-by-experimenting-with-the-data-hub-azure-sandbox-environment)in SharePoint.

1. **Which resources are available on the sandbox?**

​The available resources include: Azure Data Factory, Azure Databricks workspace, Azure Key Vault, Jump VM, and Azure Synapse Analytics workspace with a Synapse SQL pool. In addition, users will have Azure Data Lake Storage accounts that are configured to mirror their setup on the Data Hub Azure platform. For more information on the sandbox environment and what the resources available please visit view our [Sandbox Onboarding Documentation](https://bp365.sharepoint.com/sites/DataAnalyticsPlatforms/SitePages/Data%20Hub%20Azure%20Onboarding.aspx#get-hands-on-by-experimenting-with-the-data-hub-azure-sandbox-environment)in SharePoint.

1. **Where is the Sandbox Node relative to other nodes?**

The sandbox is owned and managed by the central Data Ops Team. The sandbox node sits within a separate locked down subscription, with a resource group per sandbox domain (domains Databricks, Data Factory and Synapse).​ For more information on the sandbox environment and what it can offer you please visit view our [Sandbox Onboarding Documentation](https://bp365.sharepoint.com/sites/DataAnalyticsPlatforms/SitePages/Data%20Hub%20Azure%20Onboarding.aspx#get-hands-on-by-experimenting-with-the-data-hub-azure-sandbox-environment)​ in SharePoint.

1. **How do I get into a sandbox domain?**

To onboard onto the sandbox, first register your interest through a ServiceNow request (click our [​Support Page](https://bp365.sharepoint.com/sites/DataAnalyticsPlatforms/SitePages/Data%20Hub%20Azure%20Platform%20Support.aspx)for more guidance). Once you've raised interest, the Engagement and Adoption Team will work with you to completed a sandbox interest charter. The charter captures data requirements, number of users, reason for using the sandbox and how long they'll need the sandbox for. Your team is required to provide information on where you are working, your sandbox identifier, cost code, time of day usage in UTC and a list of users. This information is validated by the Engagement and Adoption team and Node Ops team to ensure all requirements are captured and the sandbox environment is configured the team's needs.

For more details on the full process please visit our [Sandbox Onboarding Process](https://pa-content.azureedge.net/:u:/rhttps:/bp365.sharepoint.com/sites/Data-Platform-Mesh/_layouts/15/Doc.aspx?sourcedoc=%7b9052C1B0-DBA3-4555-A4FD-F25D0717781D%7d&file=Sandbox%20Processes%20for%20Onboarding%20and%20Offboarding.vsdx&action=default&mobileredirect=true).

1. **How are new users added to the sandbox environment after the original setup?**

**​**Once the sandbox environment is setup with the original users requested, teams can request to add new users to the sandbox. The users of a sandbox environment are added to a security group per sandbox resource group with read/write permissions to Azure Data Lake Storage. ​Users that already exist in the group can request to add new members to the group in ServiceNow. Visit our [Support Page](https://bp365.sharepoint.com/sites/DataAnalyticsPlatforms/SitePages/Data%20Hub%20Azure%20Platform%20Support.aspx)to learn more about raising a ServiceNow ticket.

1. **How should data be added to the Azure Data Lake Storage (ADLS) in the Sandbox environment?**

There are several prerequisites that must be met before adding data into Azure Data Lake Storage in the Sandbox environment. Please review the​ [Data Usage Guidance](https://basproducts.atlassian.net/wiki/spaces/EDA/pages/3073606074/Using+Data+in+the+Sandbox#Data-Usage-Guide)to understand data usage planning, data uploading, data use and data clearance steps.

**##Content table**