

Publishing and Managing Datasets



Agenda

- Comparing Imported Datasets to DirectQuery
 - Managing Dataset Refresh
 - Installing and Configuring On-premises Gateway
 - Designing Datasets using DirectQuery Mode



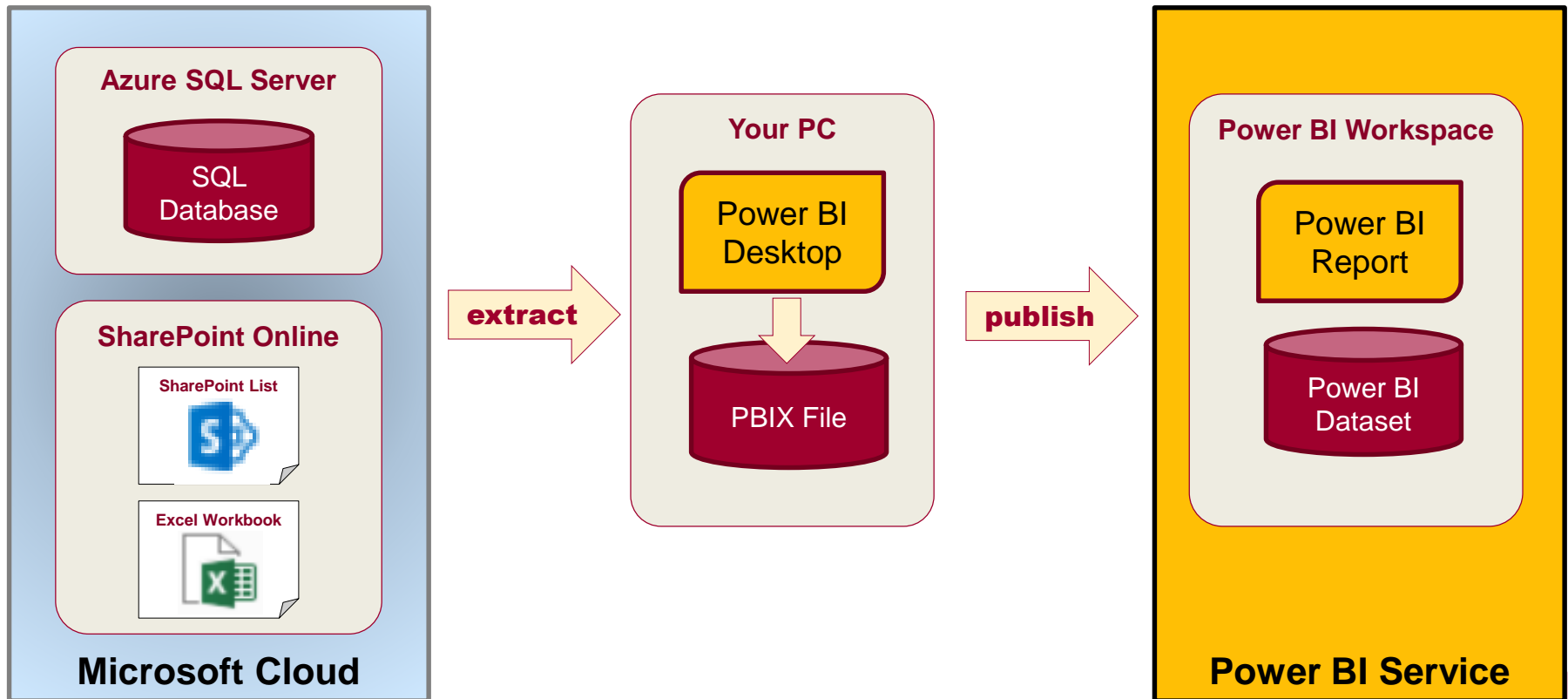
Imported Datasets Versus Live Connect

- Imported Dataset
 - Data is imported into storage within Power BI service
 - Entire dataset loaded into memory when in use
 - Reports and Datasets query data in imported dataset
 - Dataset size is maxed out at 1GB
 - Dataset must be refreshed when source data changes
- DirectQuery Dataset
 - No data is imported or cached in the Power BI service
 - Reports and Datasets query data with live connection
 - No need to refresh data
 - No need to worry about 1GB dataset size limitation
 - Limitation are placed on features for querying and data modeling



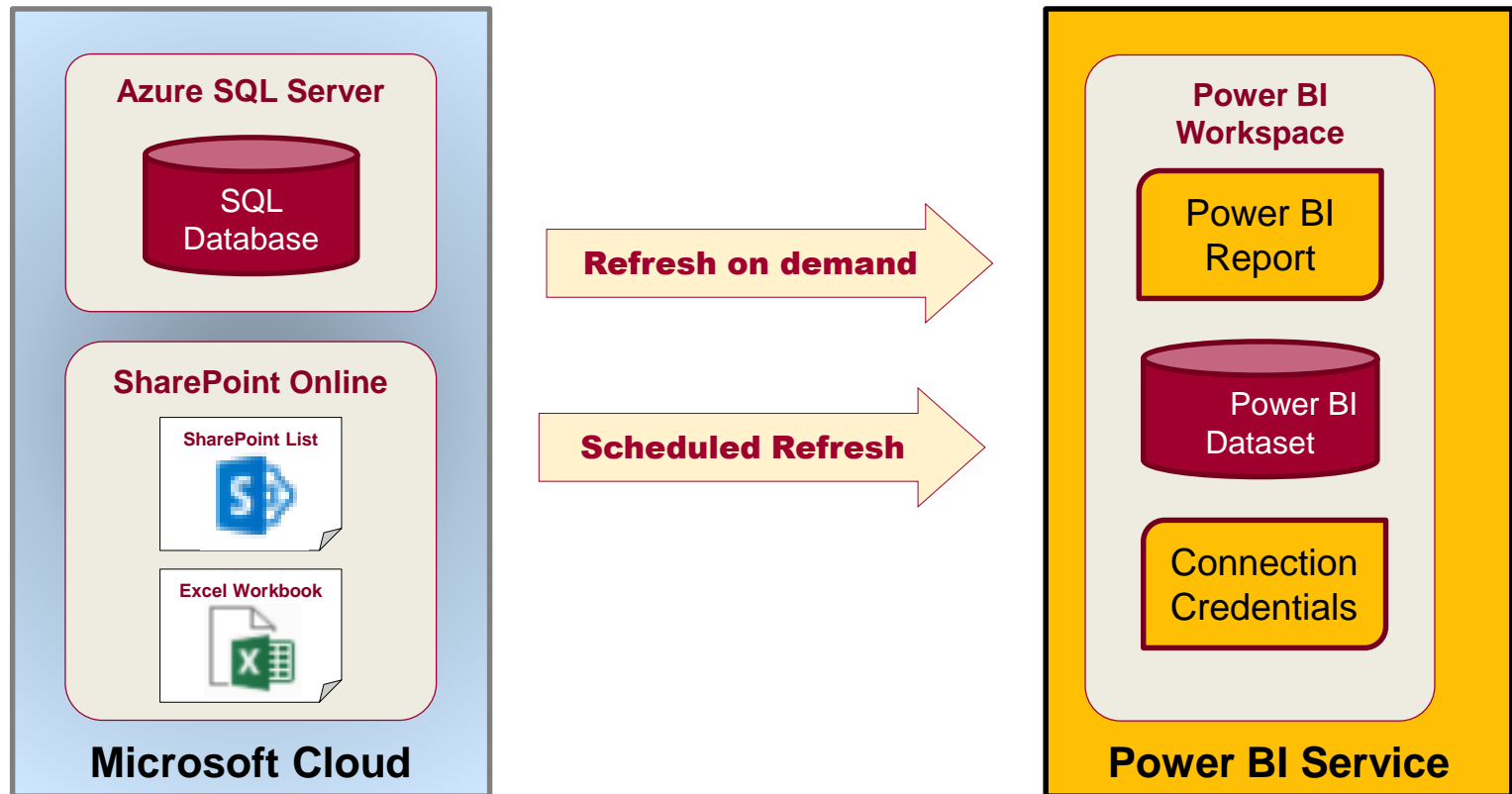
Importing Data from Cloud Data Sources

- Importing Datasets from Cloud Data Sources
 - Power BI Desktop saves data locally in PBIX file
 - PBIX file published to import dataset into Power BI Service



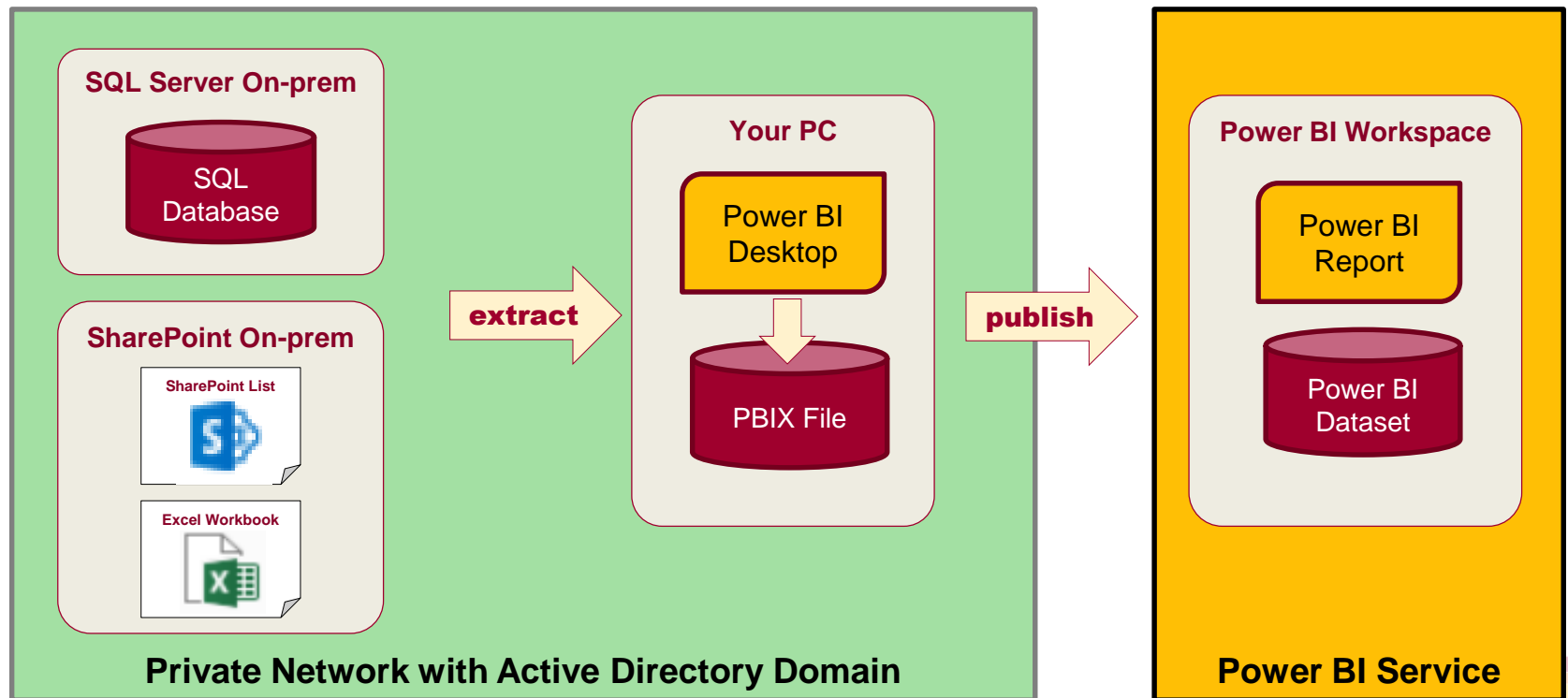
Refreshing Data from Cloud Data Sources

- Power BI Service Can Be Configured to Refresh Data
 - Power BI Desktop will never publish login credentials
 - You must configure data source credentials in Power BI service



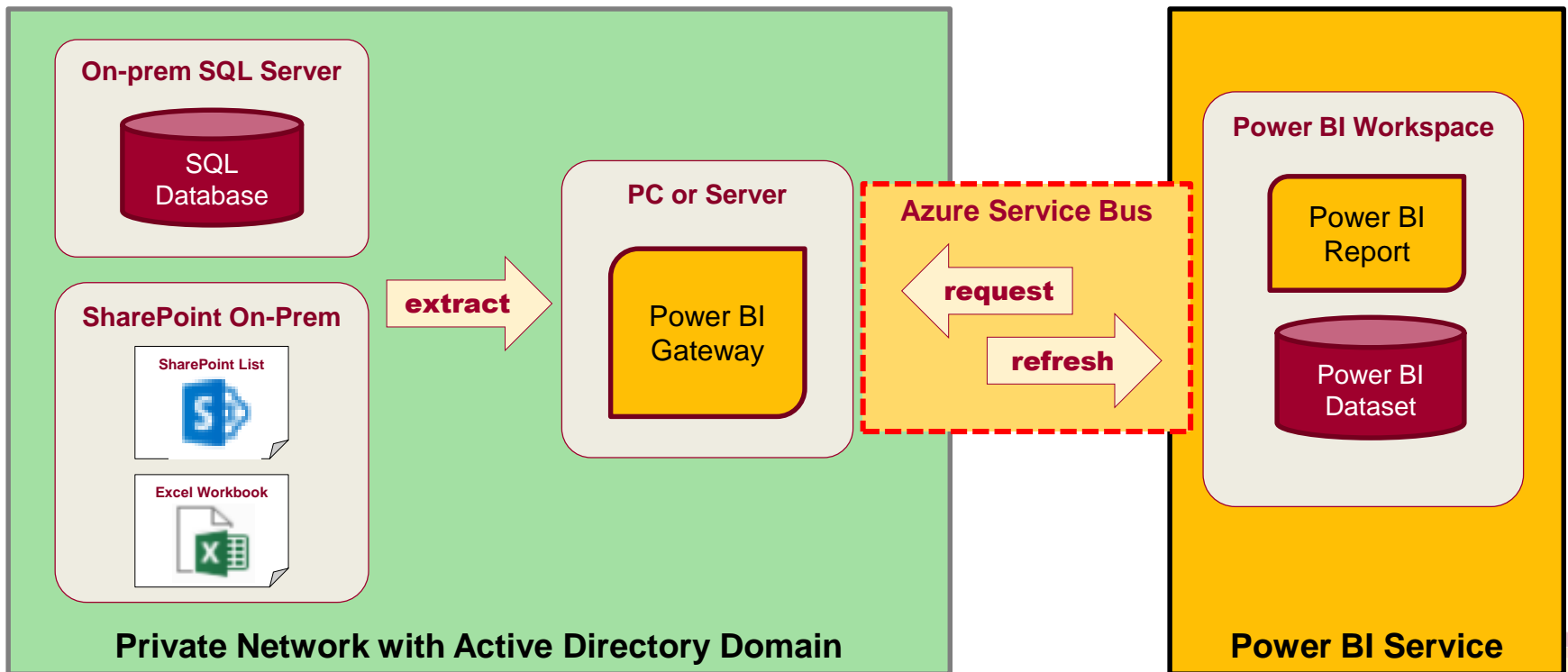
Importing Data from On-premises Data Sources

- What If Your Data Sources Live Inside an AD Domain?
 - Extracting and Publishing Data with PBIDT remains the same
 - Refreshing Data Becomes More Complicated
 - Power BI Service Cannot Directly Connect to Data Sources



What is a Power BI Gateway?

- Gateway connects Power BI to data source
 - Runs on PC or server joined to domain
 - Extracts data and returns it to Power BI service



Agenda

- ✓ Comparing Imported Datasets to DirectQuery
- Managing Dataset Refresh
 - Installing and Configuring On-premises Gateway
 - Designing Datasets using DirectQuery Mode



Configuring Credentials in the Cloud

- Power BI Desktop never Publishes Credentials
 - PBIDT caches credentials locally for its own use
 - Connection credentials must be configured in cloud
 - Power BI stores credentials in encrypted format

The image shows a screenshot of the Power BI Desktop interface. On the left, a panel titled 'Data source credentials' displays a yellow error message: 'Your data source can't be refreshed because the credentials are invalid. Please update your credentials and try again. [Show details](#)'. Below this, two data sources are listed: 'Web' and 'WingtipSalesDB-cpt.database.windows.net', each with an 'Edit credentials' link. On the right, a 'Configure Imported from Cloud' dialog box is open, showing fields for 'Server' (cpt.database.windows.net), 'Database' (WingtipSalesDB), 'Authentication method' (Basic), 'User name' (CptStudent), and 'Password' (masked with dots). 'Sign in' and 'Cancel' buttons are at the bottom right of the dialog.

▲ Data source credentials

⊗ Your data source can't be refreshed because the credentials are invalid. Please update your credentials and try again. [Show details](#)

Web [⊗ Edit credentials](#)

WingtipSalesDB-cpt.database.windows.net [⊗ Edit credentials](#)

Configure Imported from Cloud [×]

Server
cpt.database.windows.net

Database
WingtipSalesDB

Authentication method
Basic ▾

User name
CptStudent

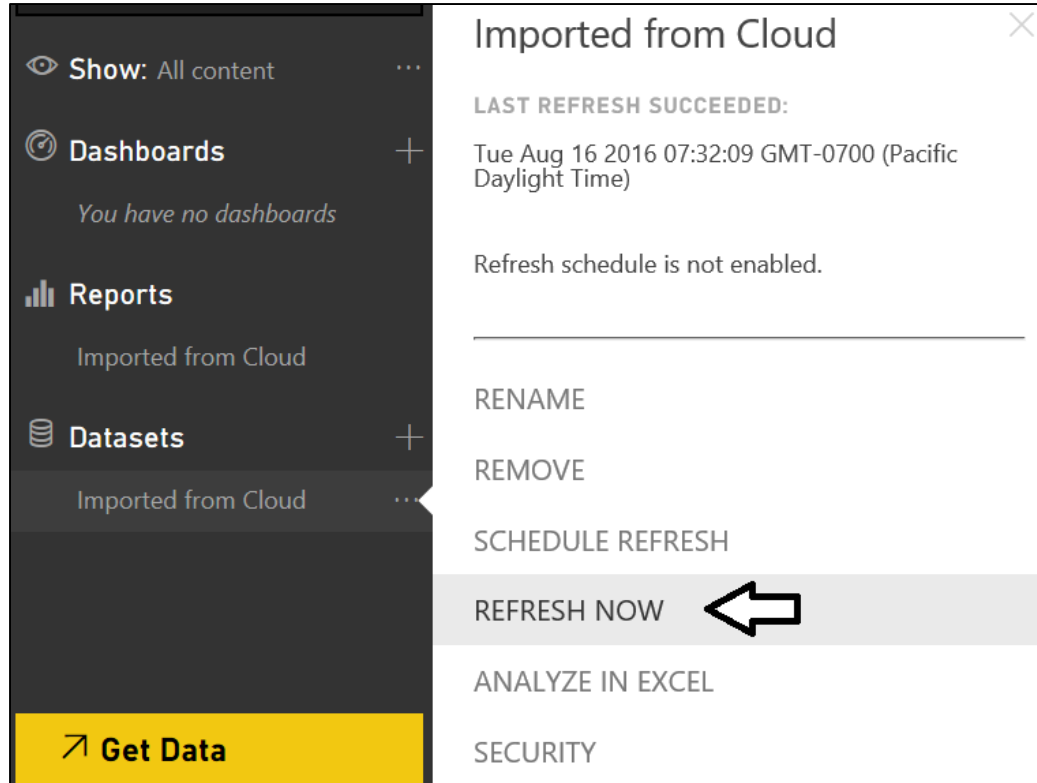
Password
●●●●●●●●

Sign in Cancel



Manual Refresh

- Once you have configured credentials...
 - You can refresh on demand from Dataset flyout menu



Configuring Scheduled Refresh

- Very easy to configure
 - Power BI Pro allows up to 8 refreshes per day
 - Power BI Standard allows one refresh per day

Schedule Refresh

Keep your data up to date

Yes ☒

Refresh frequency

Daily

Time zone

(UTC-05:00) Eastern Time (US and Canada)

Time

2 00 AM

[Add another time](#)

☒ Send refresh failure notification email to me

Apply Discard



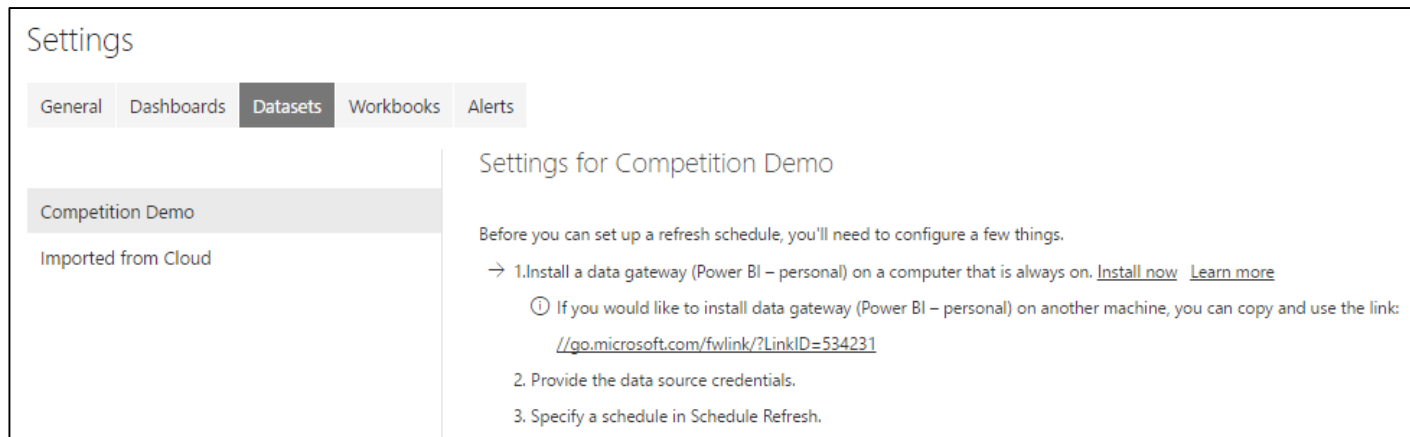
Agenda

- ✓ Comparing Imported Datasets to DirectQuery
- ✓ Managing Dataset Refresh
- Installing and Configuring On-premises Gateway
 - Designing Datasets using DirectQuery Mode



Data Refresh Might Require a Gateway

- Power Service cannot connect directly to...
 - Local files on your PC
 - Files in your local network
 - Data sources in an Active Directory domain



The screenshot shows the 'Settings' page in Power BI. The 'Datasets' tab is selected. On the left, a list of datasets includes 'Competition Demo' (Imported from Cloud). The main content area is titled 'Settings for Competition Demo' and contains instructions for setting up a refresh schedule. It states that a data gateway must be installed on a computer that is always on. It provides a link to 'Install now' and 'Learn more'. A note with an information icon explains that the gateway can be installed on another machine using a specific link: <https://go.microsoft.com/fwlink/?LinkID=534231>. The steps listed are: 1. Install a data gateway (Power BI – personal) on a computer that is always on. 2. Provide the data source credentials. 3. Specify a schedule in Schedule Refresh.

Settings

General Dashboards **Datasets** Workbooks Alerts

Competition Demo

Imported from Cloud

Settings for Competition Demo

Before you can set up a refresh schedule, you'll need to configure a few things.

→ 1. Install a data gateway (Power BI – personal) on a computer that is always on. [Install now](#) [Learn more](#)

① If you would like to install data gateway (Power BI – personal) on another machine, you can copy and use the link: <https://go.microsoft.com/fwlink/?LinkID=534231>

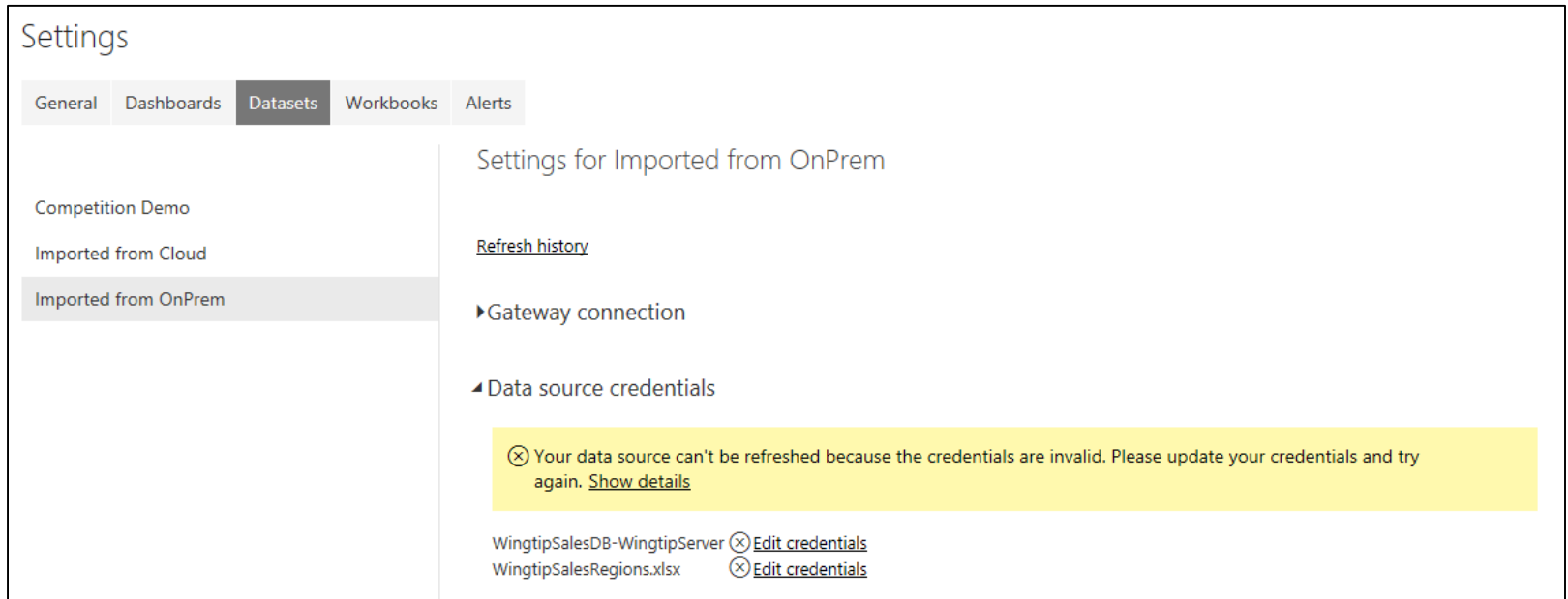
2. Provide the data source credentials.

3. Specify a schedule in Schedule Refresh.



When to Use On-Premises Gateway

- Common Motivation
 - Connecting to on-premises data sources for refresh



The screenshot shows the 'Settings' page in Power BI, specifically the 'Datasets' tab. The left sidebar lists 'Competition Demo', 'Imported from Cloud', and 'Imported from OnPrem' (which is selected). The main content area is titled 'Settings for Imported from OnPrem'. It includes a 'Refresh history' link, a 'Gateway connection' section with a right-pointing arrow, and a 'Data source credentials' section with a left-pointing arrow. A yellow warning box states: '⊗ Your data source can't be refreshed because the credentials are invalid. Please update your credentials and try again. [Show details](#)'. Below this, two data sources are listed: 'WingtipSalesDB-WingtipServer' and 'WingtipSalesRegions.xlsx', each with an '⊗ Edit credentials' link.

Settings

General Dashboards **Datasets** Workbooks Alerts

Competition Demo

Imported from Cloud

Imported from OnPrem

Settings for Imported from OnPrem

[Refresh history](#)

► Gateway connection

◄ Data source credentials

⊗ Your data source can't be refreshed because the credentials are invalid. Please update your credentials and try again. [Show details](#)

WingtipSalesDB-WingtipServer ⊗ [Edit credentials](#)

WingtipSalesRegions.xlsx ⊗ [Edit credentials](#)



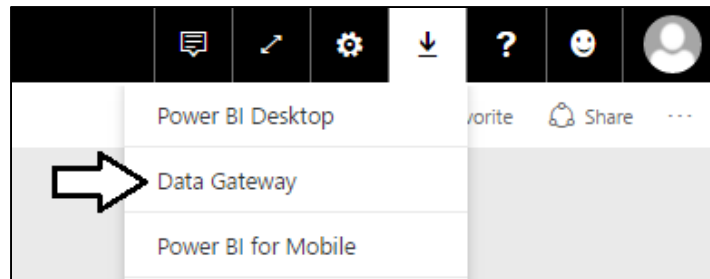
Power BI Gateway Types

- Power BI On-premises Gateway
 - Data transferred under identity of any user
 - Requires Power BI Pro and x64 bit Operating System
 - Adds support for DirectQuery mode
- Power BI Personal Gateway
 - Data transferred under identity of single user
 - Requires Power BI Pro and x64 bit Operating Systems
 - Can refresh data up to 8 times per day
 - Does not support DirectQuery mode

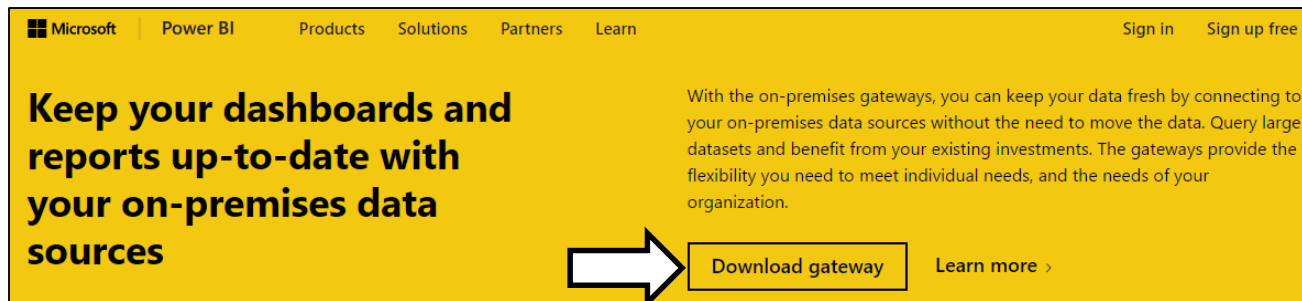


Downloading Power BI Gateway Install File

- Select **Data Gateway** from Download menu



- Click **Download** button on download page
<https://powerbi.microsoft.com/en-us/gateway/>



- Installation file for all gateways named **PowerBIGatewayInstall.exe**

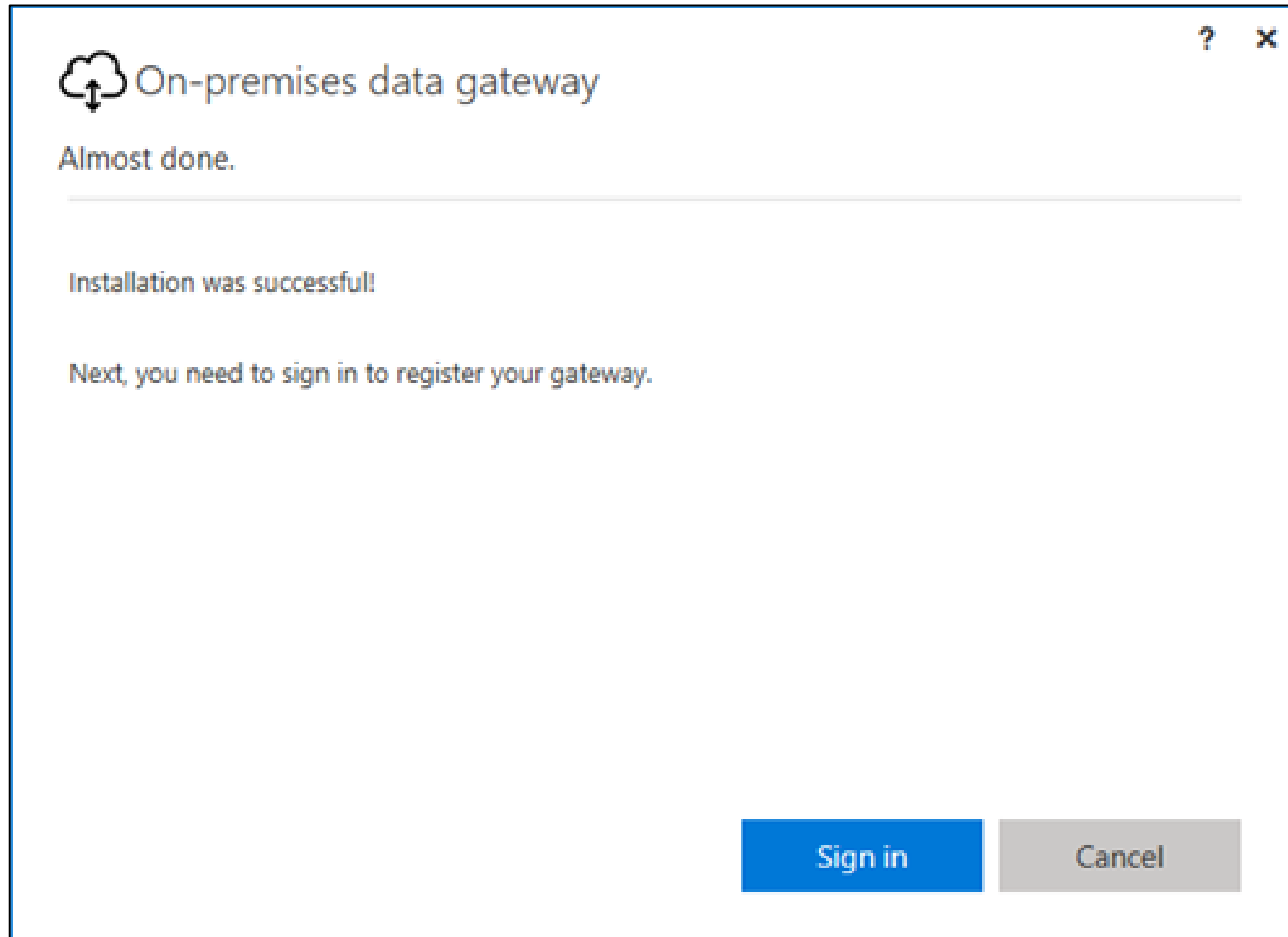


On-Premises Gateway


- Enables IT to deploy & manage shared gateways
 - Each gateway can service large number of users
 - IT can install and manage multiple gateways
 - Eliminates need to for individual users to install personal gateways
 - Replaces several older Power BI Gateways
- On-premises Gateway gives more control to IT
 - IT can centrally manage user access to underlying data sources
 - IT can monitor and gain visibility into gateway usage



On-Premises Gateway Installation




On-prem Gateway Names & Recovery Keys

 On-premises data gateway ? ✕

You are signed in as Student@PowerBiBootcamp.onmicrosoft.com and are ready to register the gateway.

New on-premises data gateway name

Recovery key (8 character minimum)

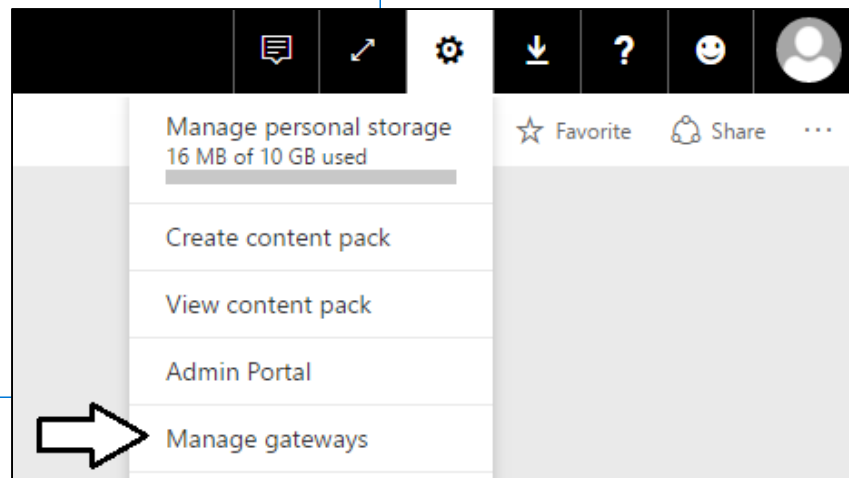
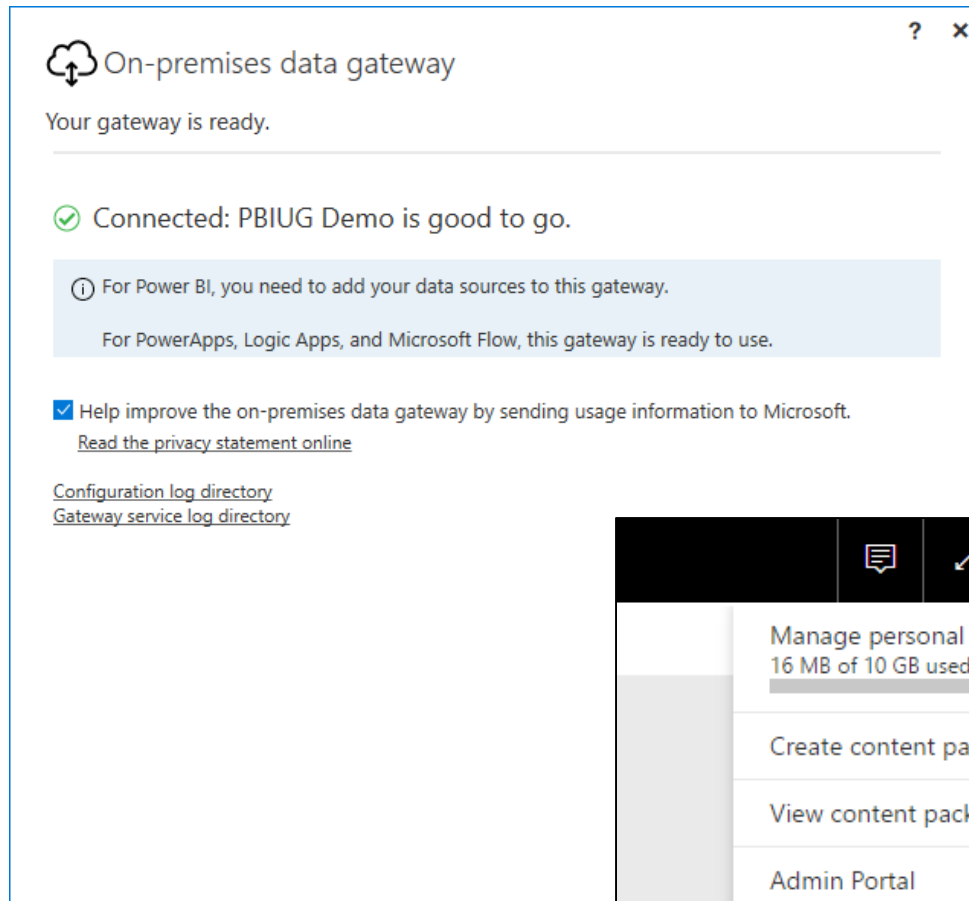
 This key is needed to restore the gateway and can't be changed. Record it in a safe place.

Confirm recovery key

< < Back Configure Cancel



Getting the Gateway Up and Running



Configuring Gateway in Power BI Service

Gateway Settings

Administrators

✓ Online: You are good to go.

① Add data sources to use the gateway

Gateway Name

PBIUG Demo

Department

Critical Path Trainig

Description

Contact Information

Student@PowerBiBootcamp.onmicrosoft.com

Apply

Discard

Gateway Settings

Administrators

People who can Administer this gateway

Enter email addresses

Add

☐ Ted Pattison

Remove



Configuring Data Sources in Gateway

Data Source Settings

Data Source Name

Data Source Type

Server

Database

Authentication Method

The credentials are encrypted using the key stored on-premises on the gateway server. [Learn more](#)

Username

Password

> Advanced settings

AddDiscard



Configuring Users

Data Source Settings Users

✓ Connection Successful

❗ Next Step: Go to the Users tab above and add users to this Data Source

Data Source Settings Users

People who can publish reports that use this data source

Enter email addresses

Add

☐ Ted Pattison



Outbound Ports – On-Premises Gateway

- Gateway communicates through outbound ports
 - Gateway does not require inbound ports

Domain names	Outbound ports	Description
*.powerbi.com	443	HTTPS
*.analysis.windows.net	443	HTTPS
*.login.windows.net	443	HTTPS
*.servicebus.windows.net	5671-5672	Advanced Message Queuing Protocol (AMQP)
*.servicebus.windows.net	443, 9350-9354	Listeners on Service Bus Relay over TCP (requires 443 for Access Control token acquisition)
*.frontend.clouddatahub.net	443	HTTPS
*.core.windows.net	443	HTTPS
login.microsoftonline.com	443	HTTPS
*.msftncsi.com	443	Used to test internet connectivity if the gateway is unreachable by the Power BI service.



Agenda

- ✓ Comparing Imported Datasets to DirectQuery
- ✓ Managing Dataset Refresh
- ✓ Installing and Configuring On-premises Gateway
- Designing Datasets using DirectQuery Mode





DEMO

Creating a Power BI Desktop Project using DirectQuery Mode

Supported Data Sources

Data source	Live/DirectQuery	User configured manual or scheduled refresh
Analysis Services Tabular	Yes	Yes
Analysis Services Multidimensional	Yes	Yes
SQL Server	Yes	Yes
SAP HANA	Yes	Yes
Oracle	Yes	Yes
Teradata	Yes	Yes
File	No	Yes
Folder	No	Yes
SharePoint list (on-premises)	No	Yes
Web	No	Yes
OData	No	Yes
IBM DB2	No	Yes
MySQL	No	Yes
Sybase	No	Yes



Summary

- ✓ Comparing Imported Datasets to DirectQuery
- ✓ Managing Dataset Refresh
- ✓ Installing and Configuring On-premises Gateway
- ✓ Designing Datasets using DirectQuery Mode

