Fabric CI/CD - What is there and what is coming

Mathias Thierbach & Gabi Muenster

SQL Day

16 edycja konferencji SQLDay

Data Community

13-15 maja 2024, WROCŁAW + ONLINE

partner platynowy

lingaro

partner złoty

DXC TECHNOLOGY

VOLVO

software one

partner srebrny

accenture

Nordcloud an IBM Company

Capgemini

TECHNOLOGY INNOVATION DATA KNOWLEDGE hdk

elitmind think bright

C&F

Gabi Muenster

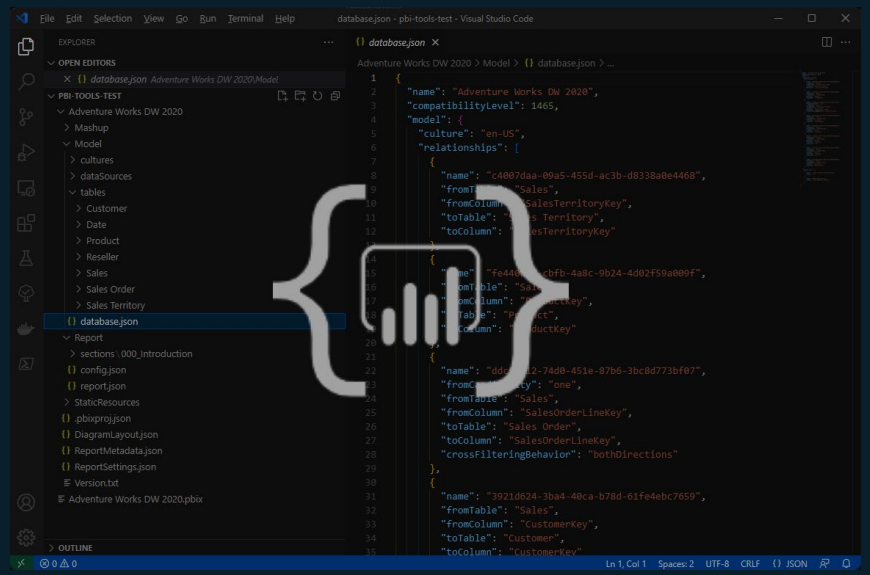
Senior Program Manager
Microsoft

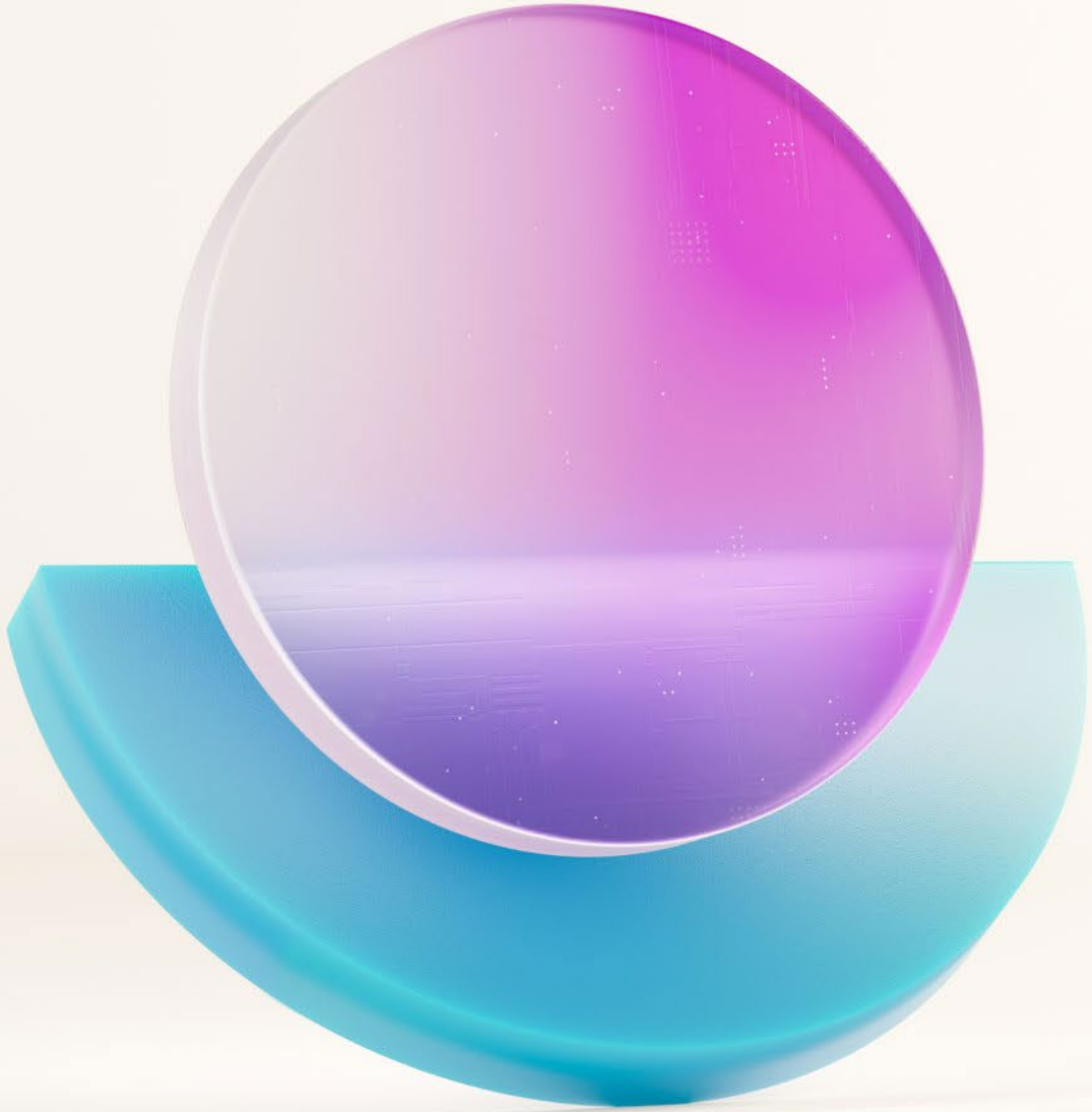
Linkedin.com/in/gabimuenster

Mathias Thierbach

Data Platform MVP
Owner pbi-tools Ltd
Developer of pbi-tools, #TMDL

linkedin.com/in/mthierba





Conceptual

CI / CD

Continuous Integration

- Small but frequent changes
- Working with multiple people on various aspects of the solution
- Merging back into main branch

Continuous Deployment

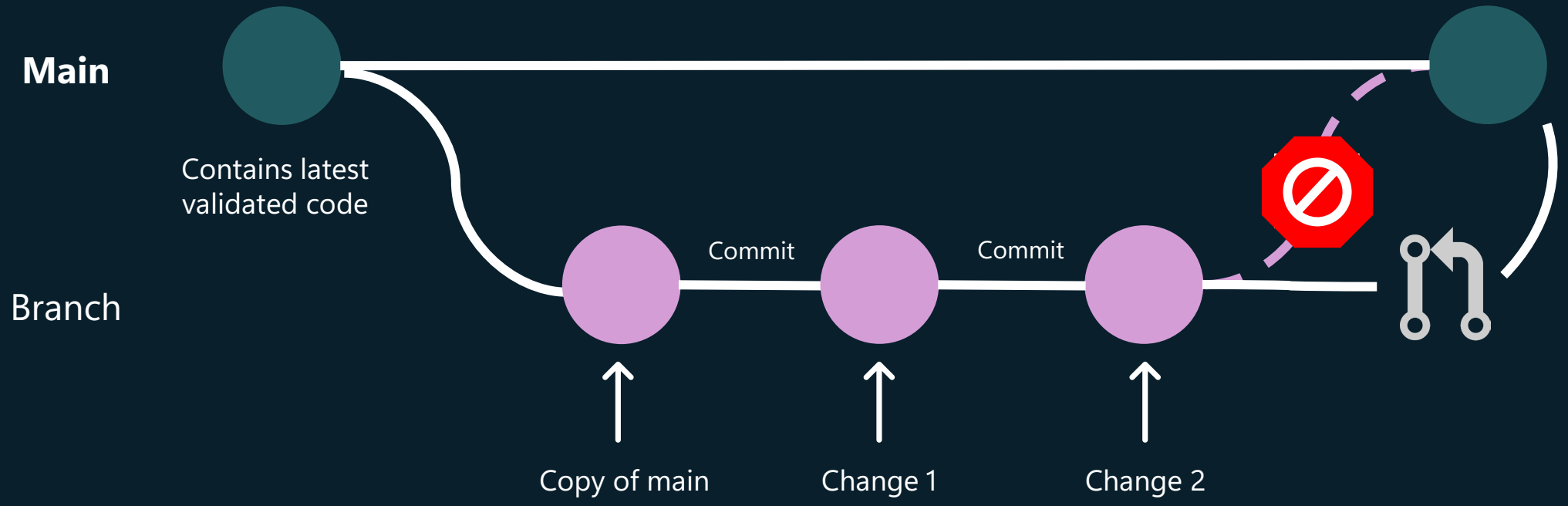
- Working in short cycles (often Agile)
- High frequency updates and releases
- Repeatable deployment process

DevOps vs DataOps vs GitOps

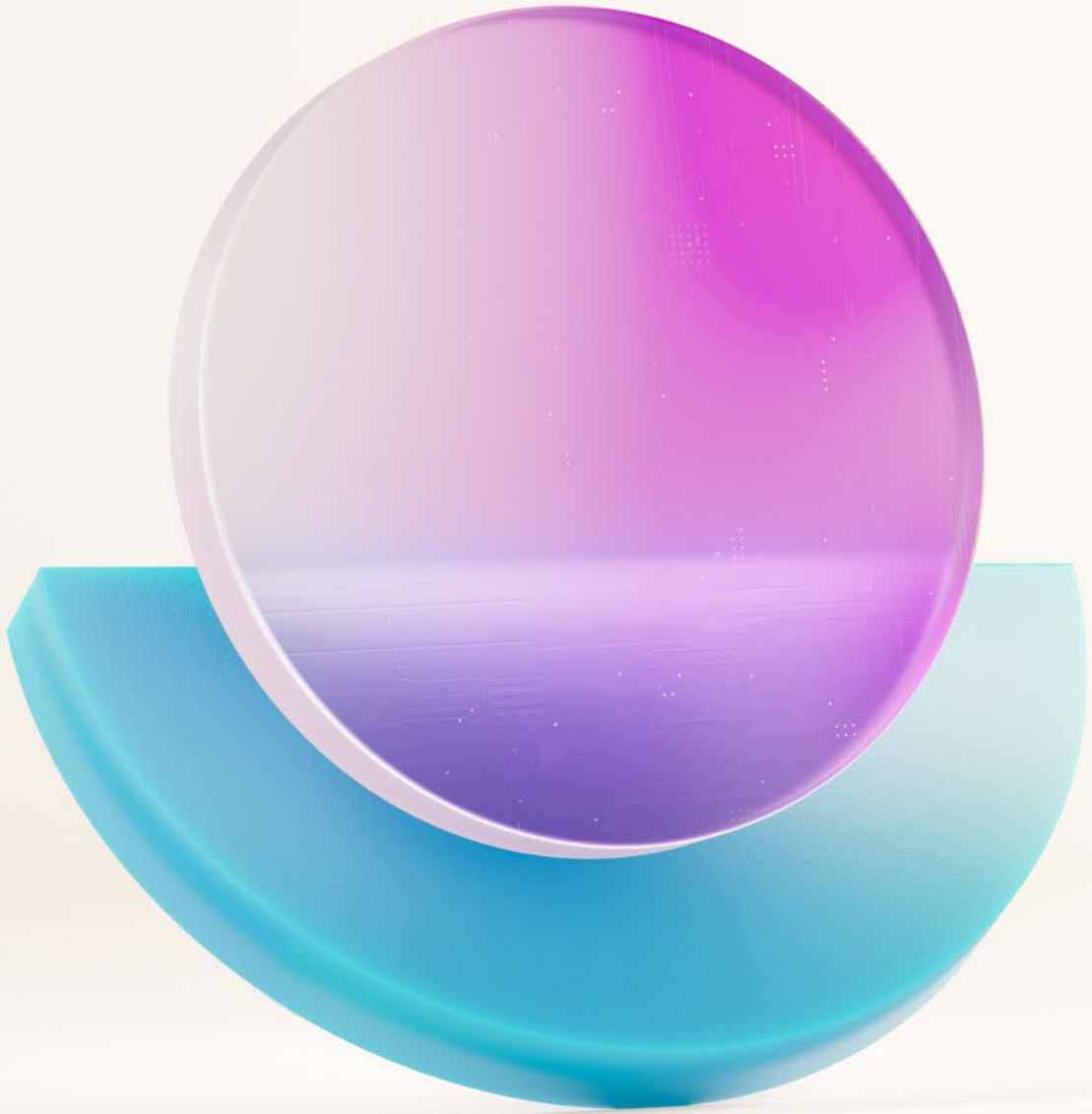
- Concepts are closely linked and somehow vague
- In common
- Use of a source control repository as the source-of-truth and for change management
- Use of staged deployments through multiple environments for QC purposes
- Reliance on automation (testing and deployment)
- We distinguish as follows:

	Applies to	Fabric Projects	Actions
DevOps	Source Code (applications)	Semantic Models (Metadata), Reports, Notebooks, Data Pipelines, Spark Jobs	Build, Package, Deploy
DataOps	Data assets	Lakehouse (Tables and FIles), Warehouse, KQL Database, Semantic Models (Refresh)	Create, Load, Delete
GitOps	Infrastructure	Workspaces, Users, Shared data sources, Gateways, Capacities, Azure resources	Create, Configure, Delete

Git concept


Main
Contains latest validated code
Branch
Commit
Commit
Copy of main
Change 1
Change 2

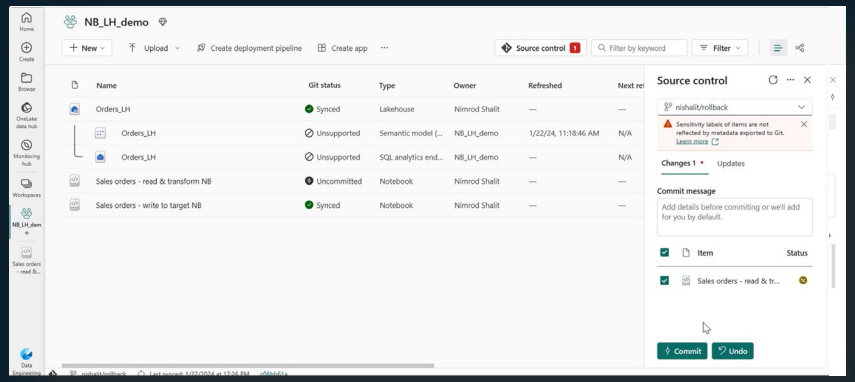
Staged approach



What we have today

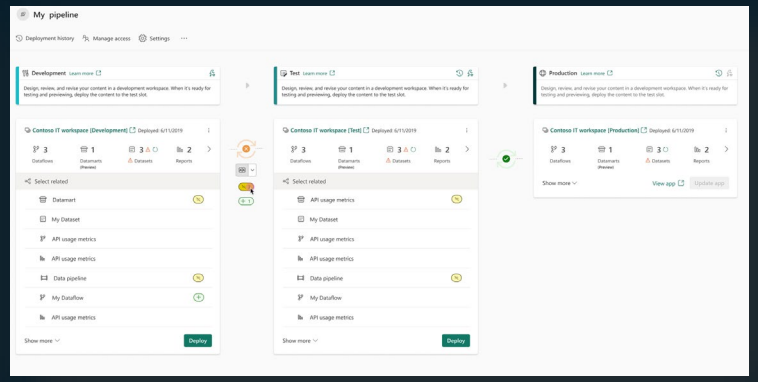
Fabric CI/CD platform

Built-in git integration



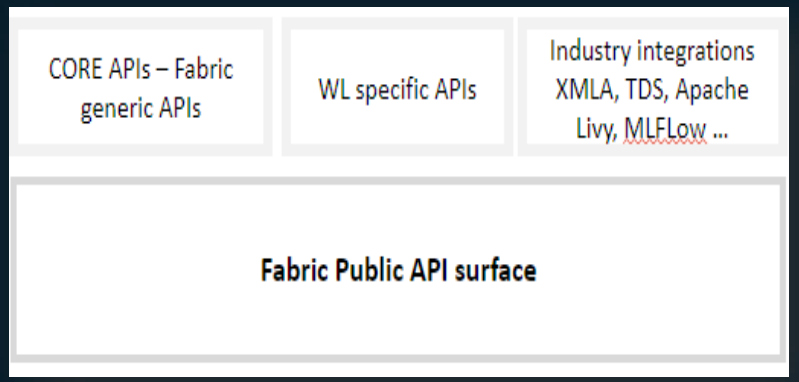
Public preview (with ADO)

Deployment pipelines



GA feature

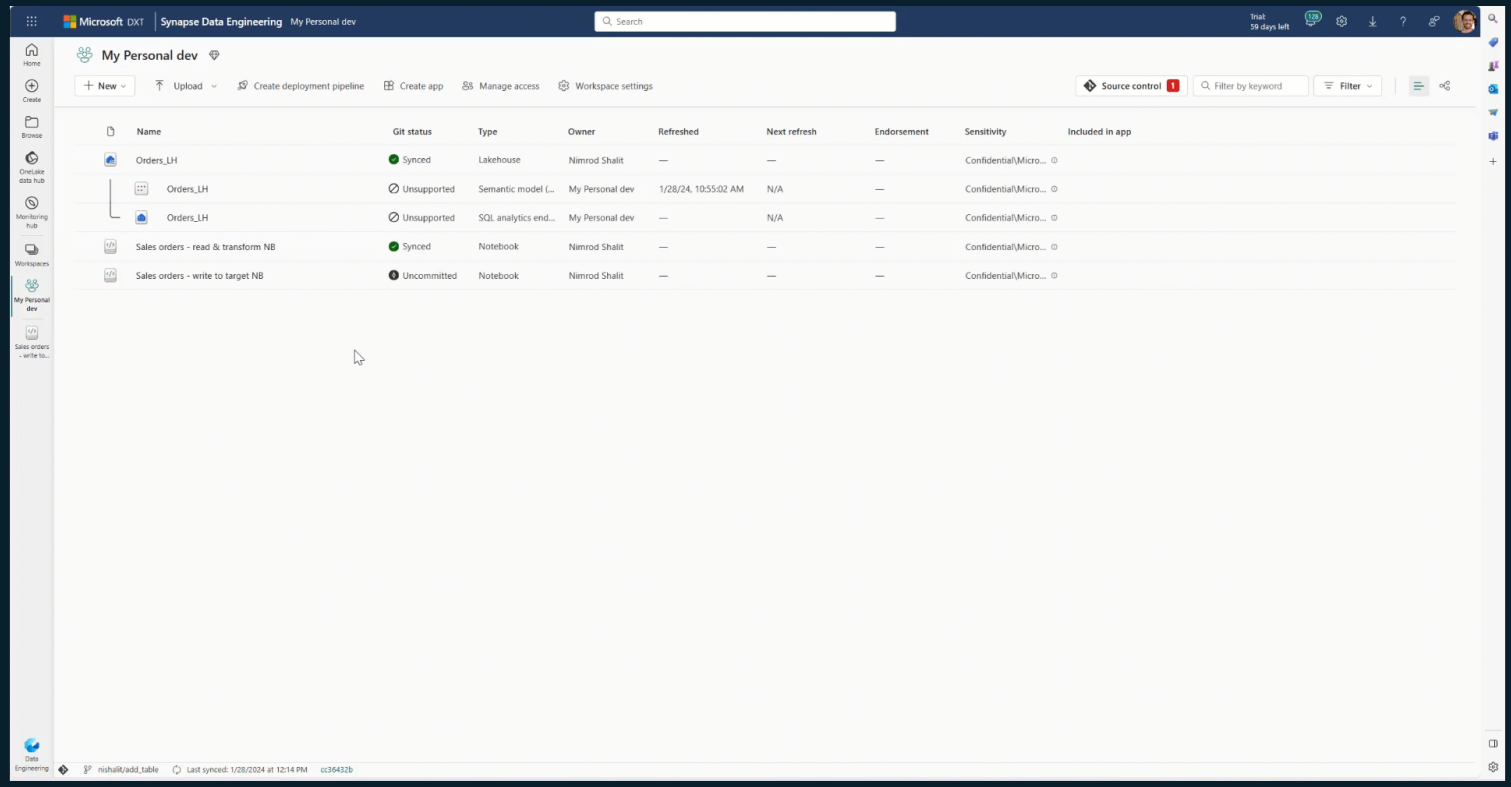
Fabric REST APIs



Public preview

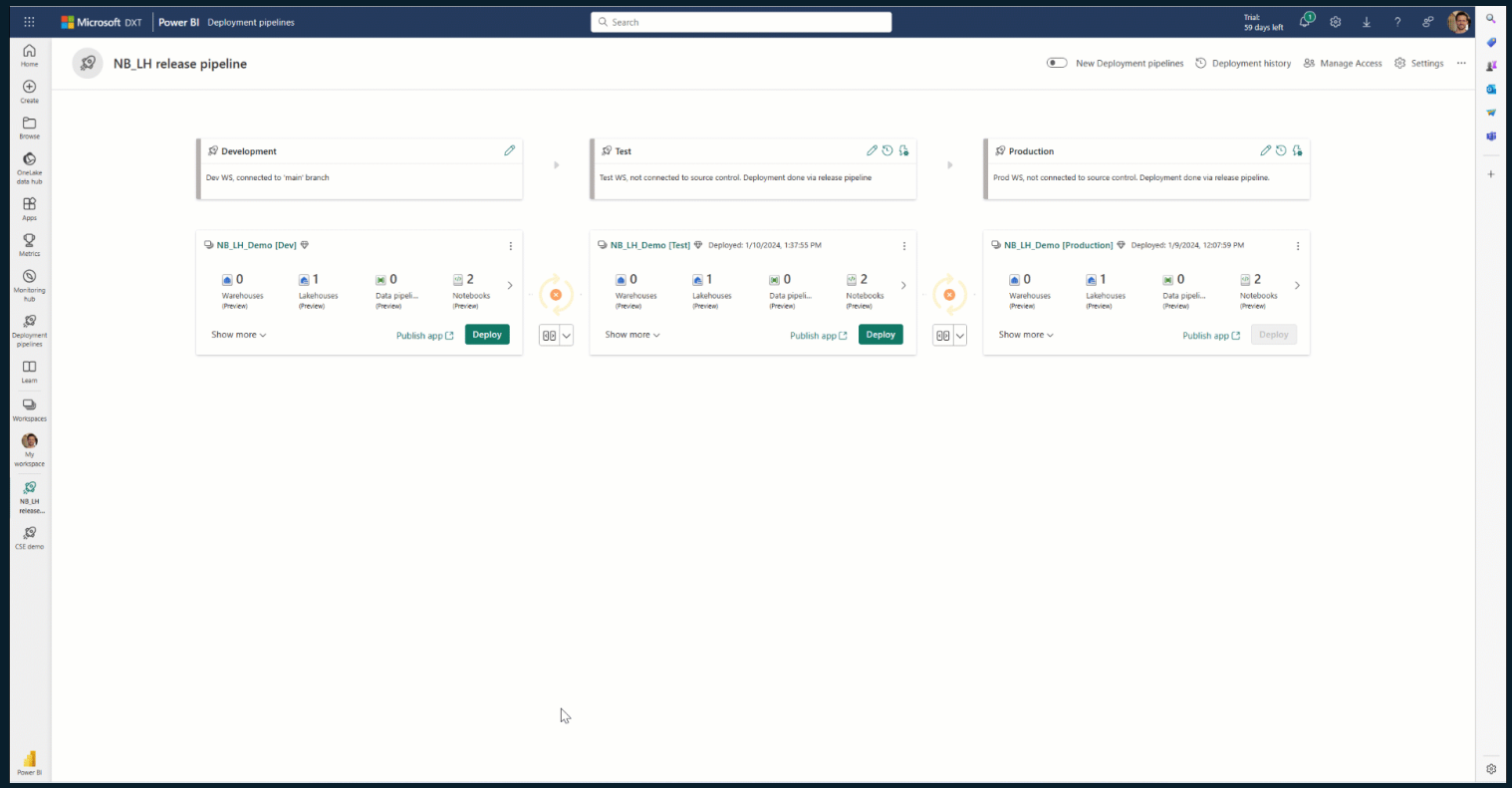
Git integration

- Sync a WS to a Git branch and back
- Git provider- Azure DevOps
- Supported items
  - Power BI- Report, Paginated report, Semantic models
  - Data Engineering- Notebook, Lakehouse
- Fabric git APIs – REST APIs & PS samples.
- Manage branches using ‘Checkout/Switch branch’
- Conflict resolution



Deployment Pipelines

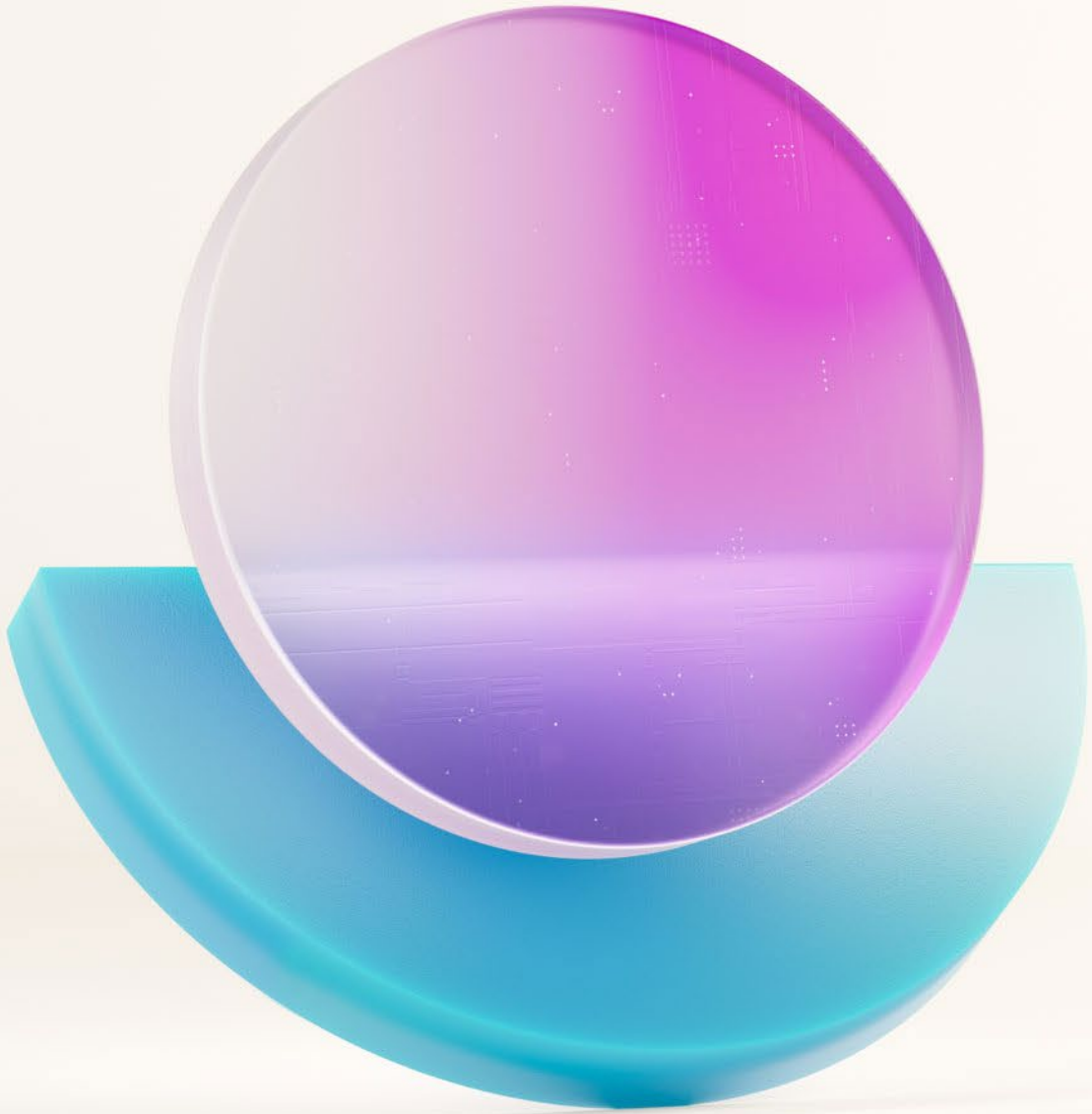
- Deploy items across WSs, and apply rules on configuration
- Supported items
  - Power BI- Report, Paginated Report, Semantic model, Dataflow Gen1, Dashboard, data mart
  - Data Engineering- Notebook, Lakehouse
- Compare changes on code-level
- Create a pipeline of 2-10 stages



Fabric User APIs

- Automated operations on behalf of Fabric users.
- Supported scenarios:
  - Item management *(see table)*
  - Item definition
  - Workspace management
  - Workspace access management
  - Execute item jobs

Item type	Create (without definition)	Get	Update	Delete	List
Dashboard	✗	✗	✗	✗	✓
DataPipeline	✗	✓	✓	✓	✓
Datamart	✗	✗	✗	✗	✓
Eventhouse	✓	✓	✓	✓	✓
Eventstream	✓	✓	✓	✓	✓
KQLDatabase	✗	✓	✓	✓	✓
KQLQueryset	✓	✓	✓	✓	✓
Lakehouse	✓	✓	✓	✓	✓
MLExperiment	✓	✓	✓	✓	✓
MLModel	✓	✓	✓	✓	✓
MirroredWarehouse	✗	✗	✗	✗	✓
Notebook	✓	✓	✓	✓	✓
PaginatedReport	✗	✗	✗	✗	✓
Report	✗	✓	✗	✓	✓
SemanticModel	✗	✓	✗	✓	✓
SparkJobDefinition	✓	✓	✓	✓	✓
SQLEndpoint	✗	✗	✗	✗	✓
Warehouse	✓	✓	✓	✓	✓



New announcements

New Fabric items supported in CI/CD

Item	Git integration	Deployment pipelines
Power BI reports	✓	✓
Power BI paginated reports	✓	✓
Power BI Semantic models*	✓	✓
Notebooks	✓	✓
Lakehouse	✓	✓
Data pipelines	✓	✓
Data Warehouse	✓	✓ (Soon)
Spark Job definition	✓	✓ (Soon)
Spark Environment	✓	✓ (Soon)

* No push datasets, DirectLake mode

E2E Automated CI/CD process

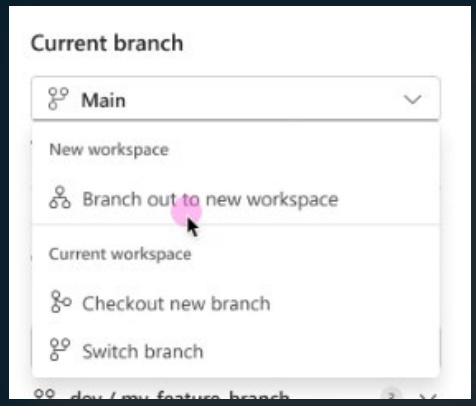
Enabling fully automated CI/CD processes using familiar tools

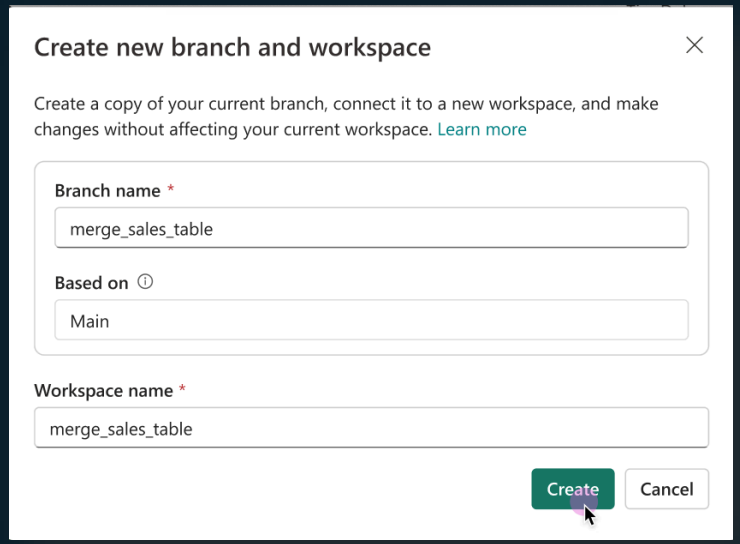
- Git APIs
  - Connect/Disconnect a workspace to a git branch
  - Commit to git
  - Update from git
  - Get item status

- Deployment pipelines APIs
  - Deploy all workspace content
  - Deploy selected items
  - Create a pipeline *(coming soon)*
  - Assign a workspace to a pipeline stage *(coming soon)*
  - Manage user access to a pipeline *(coming soon)*

Branch out to new workspace *(soon)*

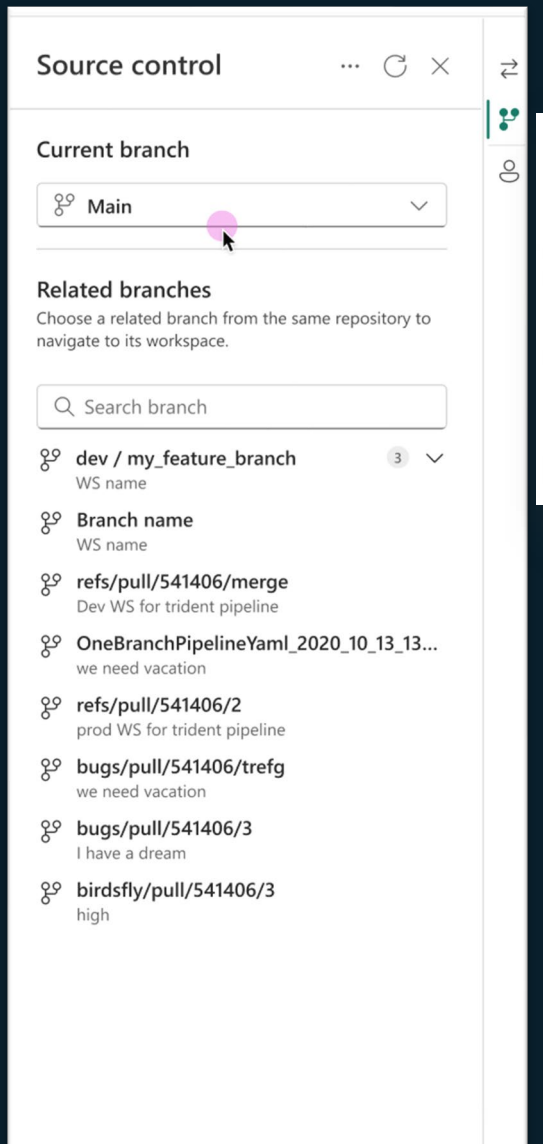
- **Branch out to new workspace –**
  - Create a new isolated work environment (workspace & branch) in a few clicks.
  - After clicking ‘Create’:
    - New workspace is created and assigned to a capacity
    - New branch is created, based on existing branch
    - New branch and workspace are connected
    - Branch content is synced into the workspace
  - Developer lands in new workspace, reducing ‘time-to-code’.

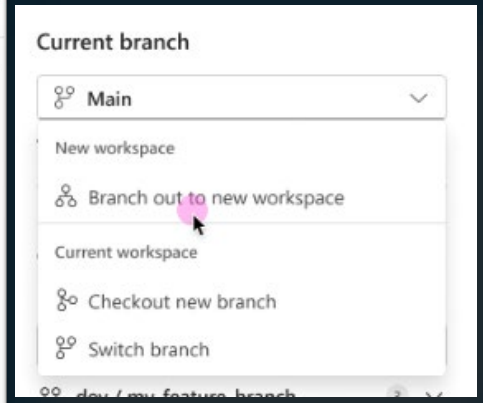


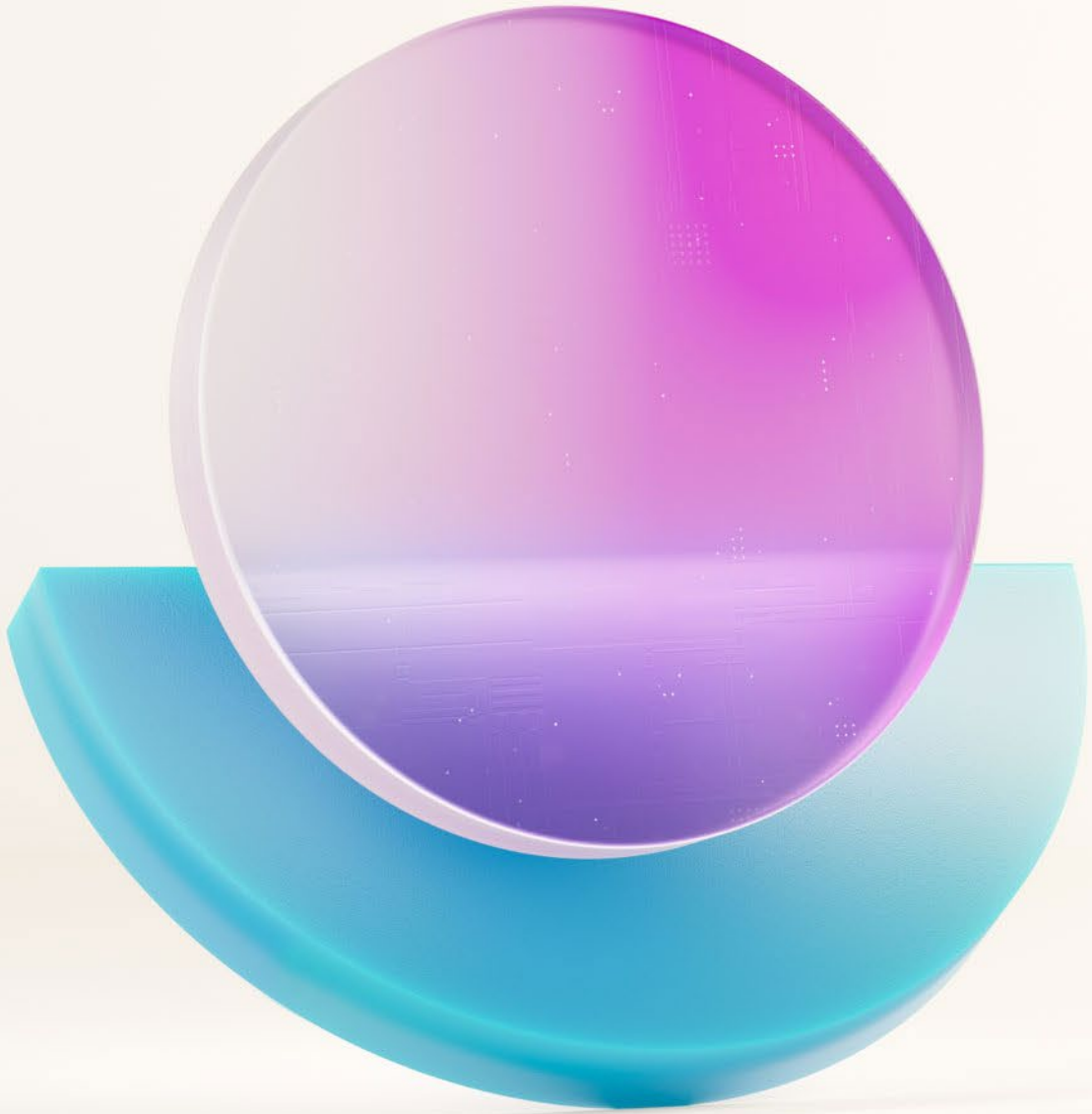


Contextual navigation *(soon)*

- **Branches tab**
  - Consolidate all branch related actions in a single place within the source control pane
- **Contextual navigation**
  - Find other branches and navigate to their respective workspaces.
  - Contextual list of the same ‘branch workspaces’.



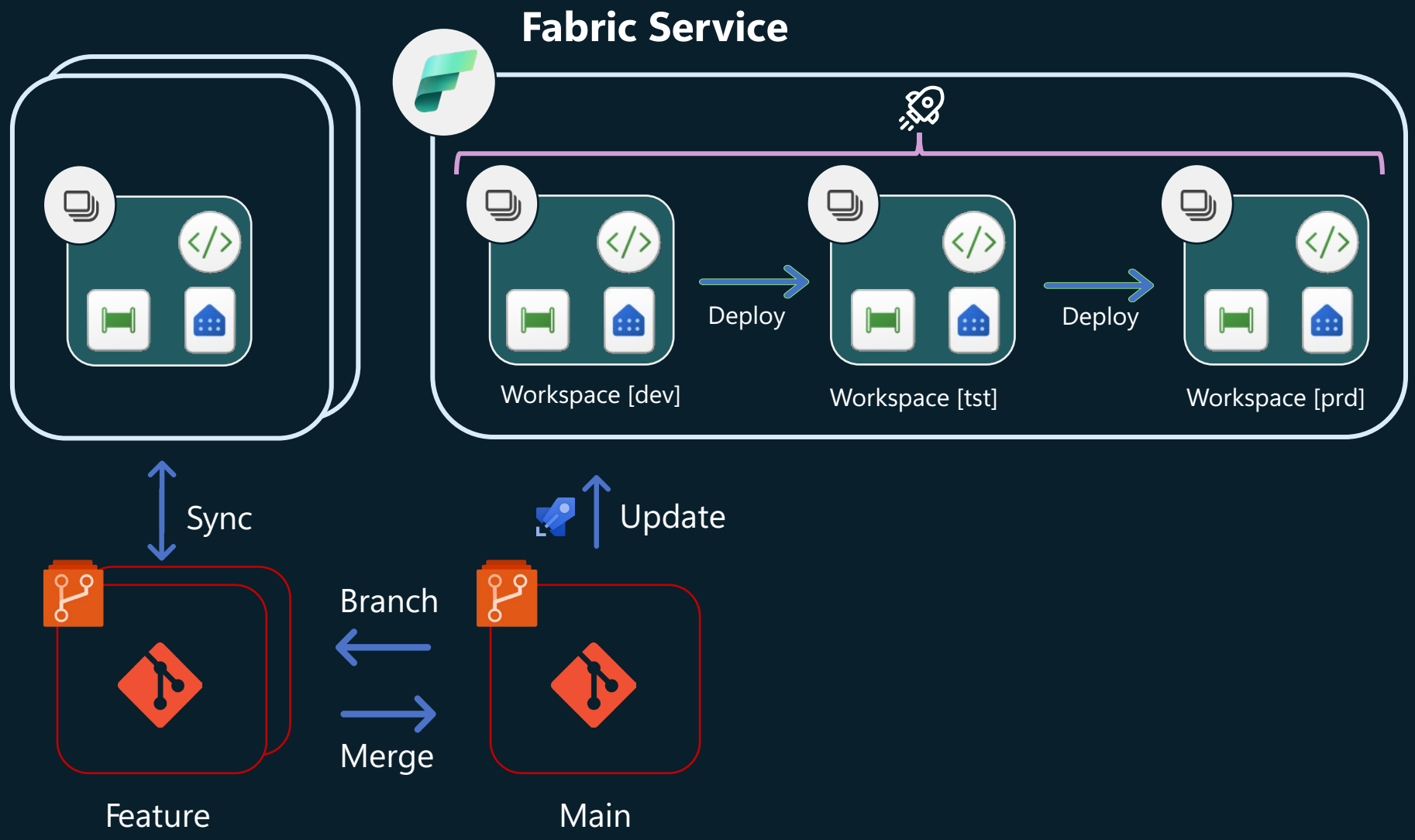




Deployment Strategies

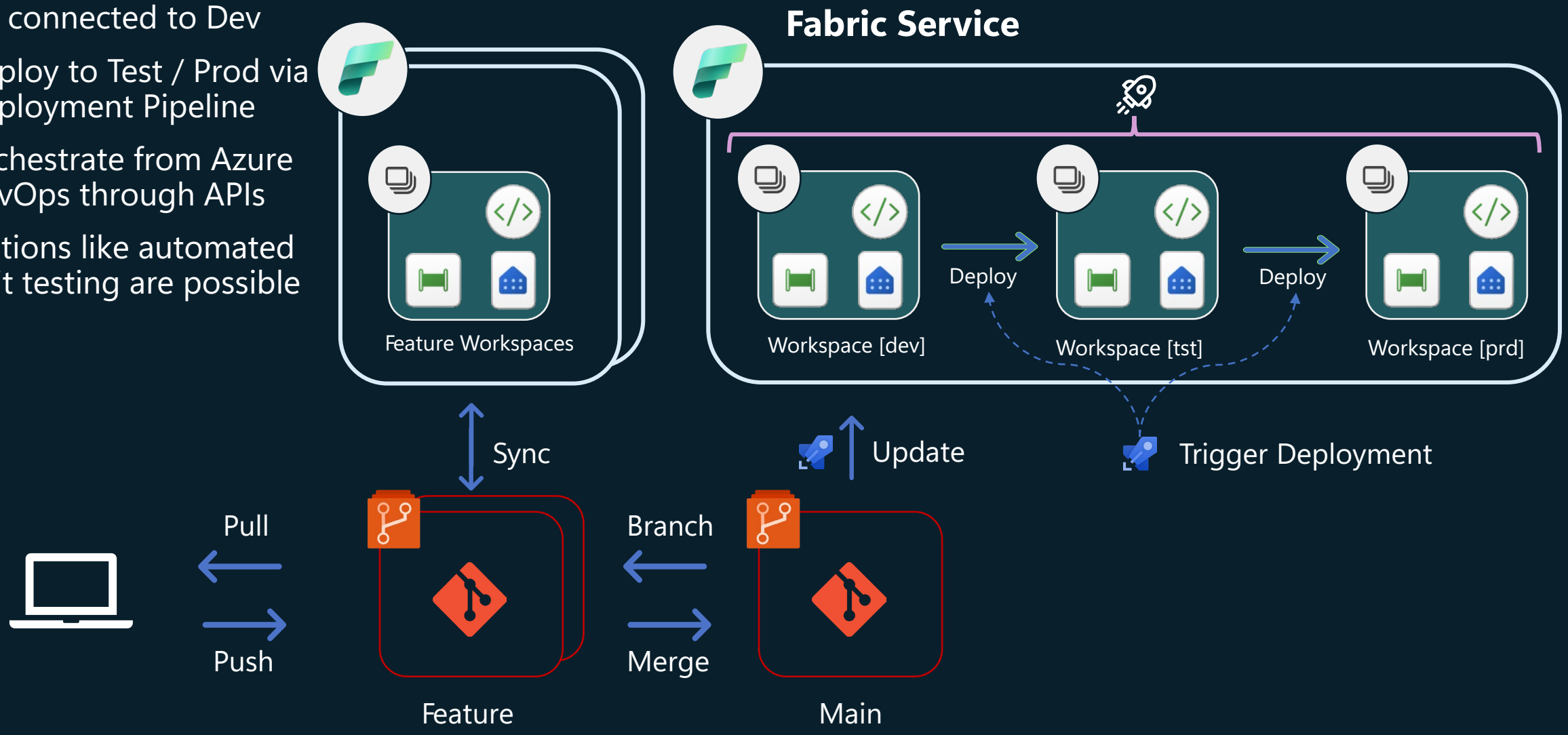
Scenario 1 (CD)

- Git connected to Dev
- Deploy to Test / Prod via Deployment Pipeline



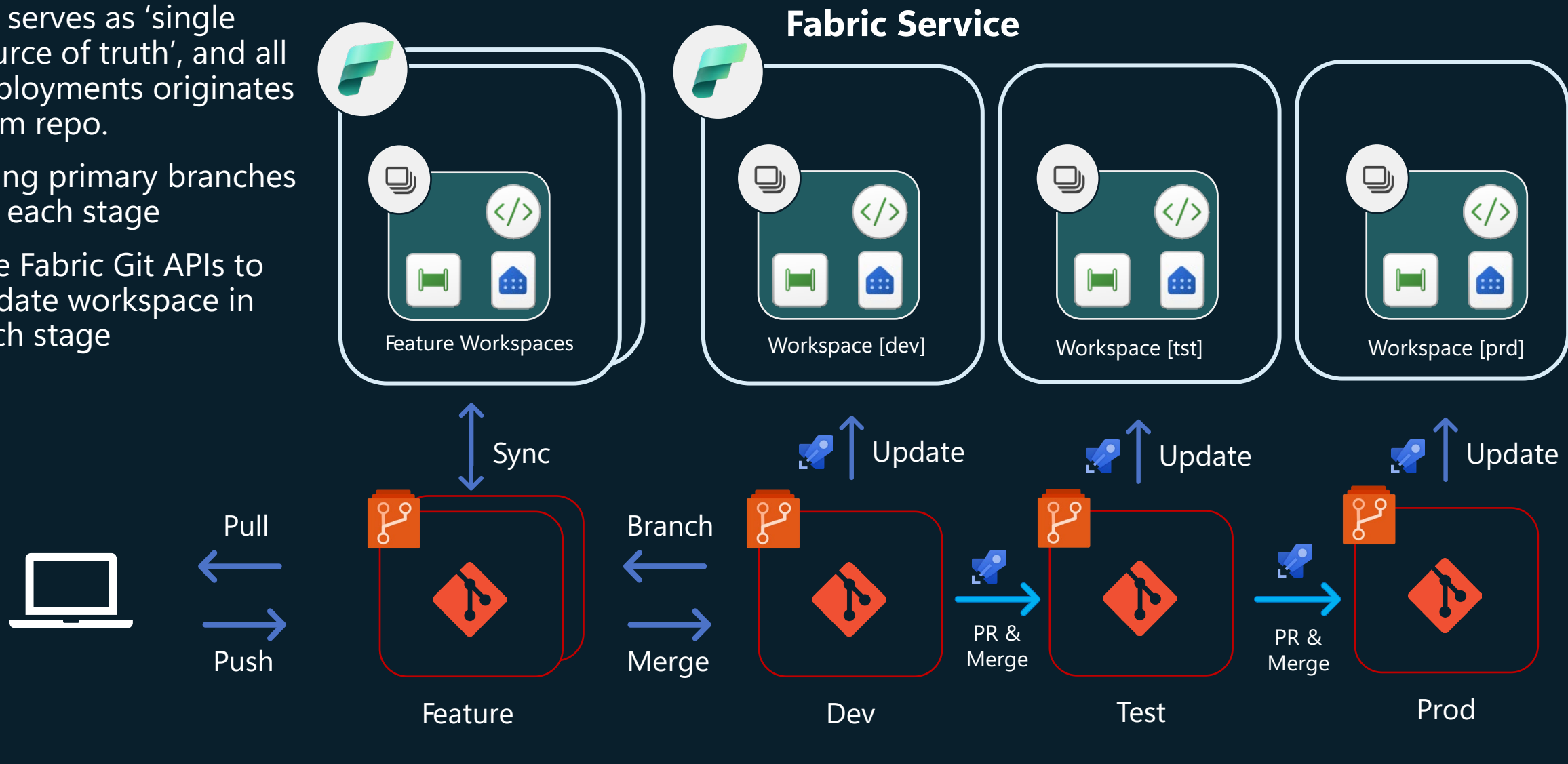
Scenario 2 – Git & Deployment pipelines

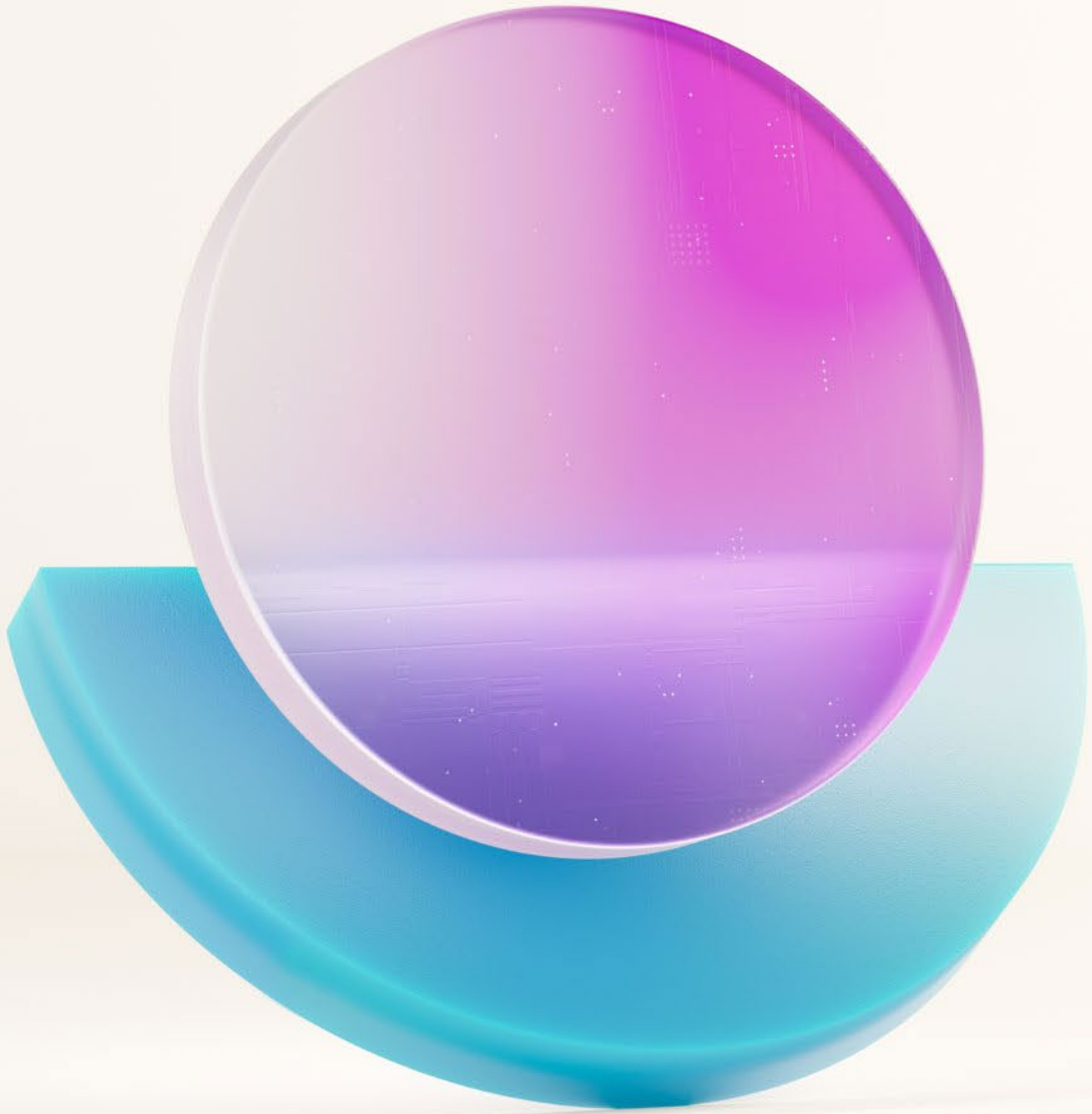
- Git connected to Dev
- Deploy to Test / Prod via Deployment Pipeline
- Orchestrate from Azure DevOps through APIs
- Options like automated unit testing are possible



Scenario 3 – Git- based deployments

- Git serves as ‘single source of truth’, and all deployments originates from repo.
- Using primary branches for each stage
- Use Fabric Git APIs to update workspace in each stage

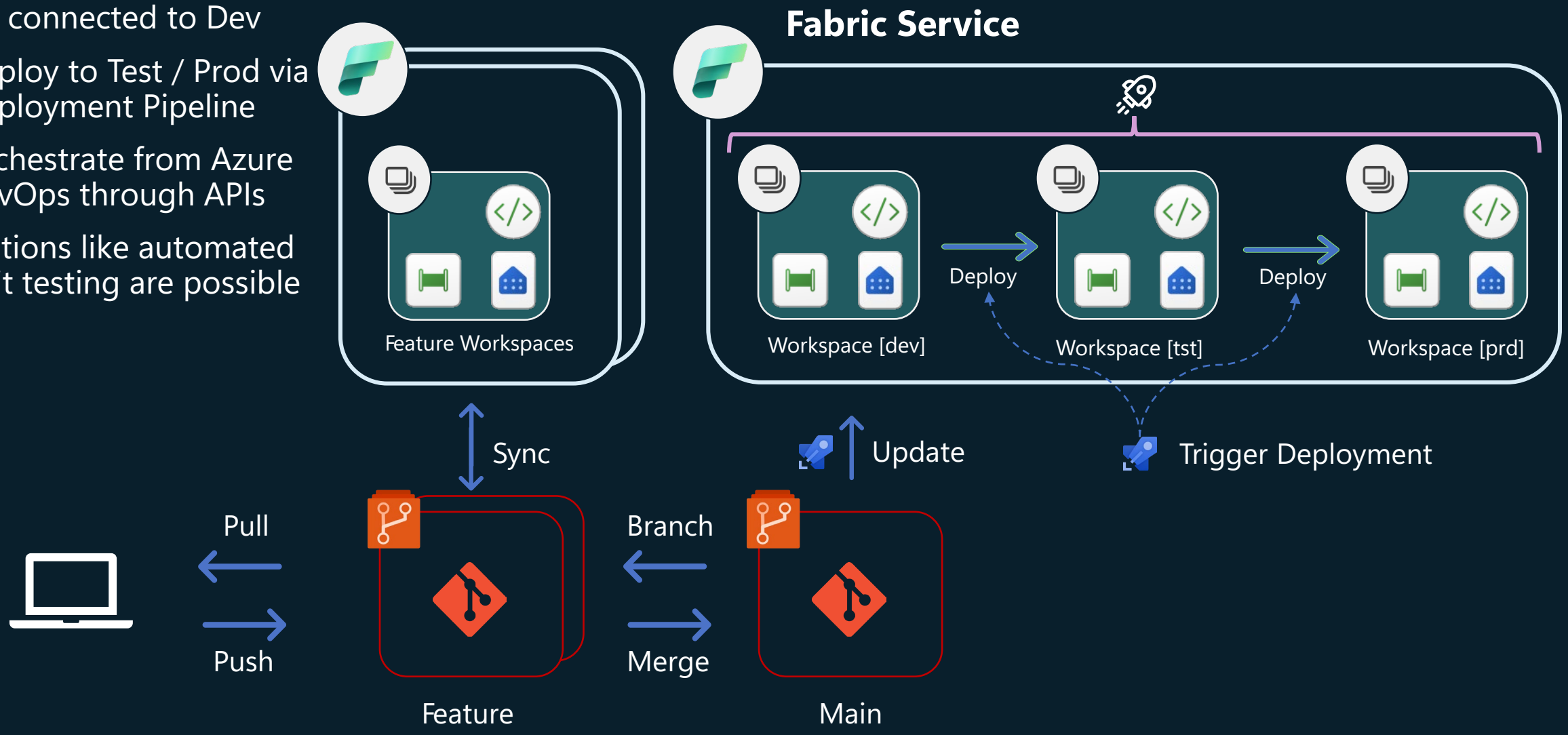


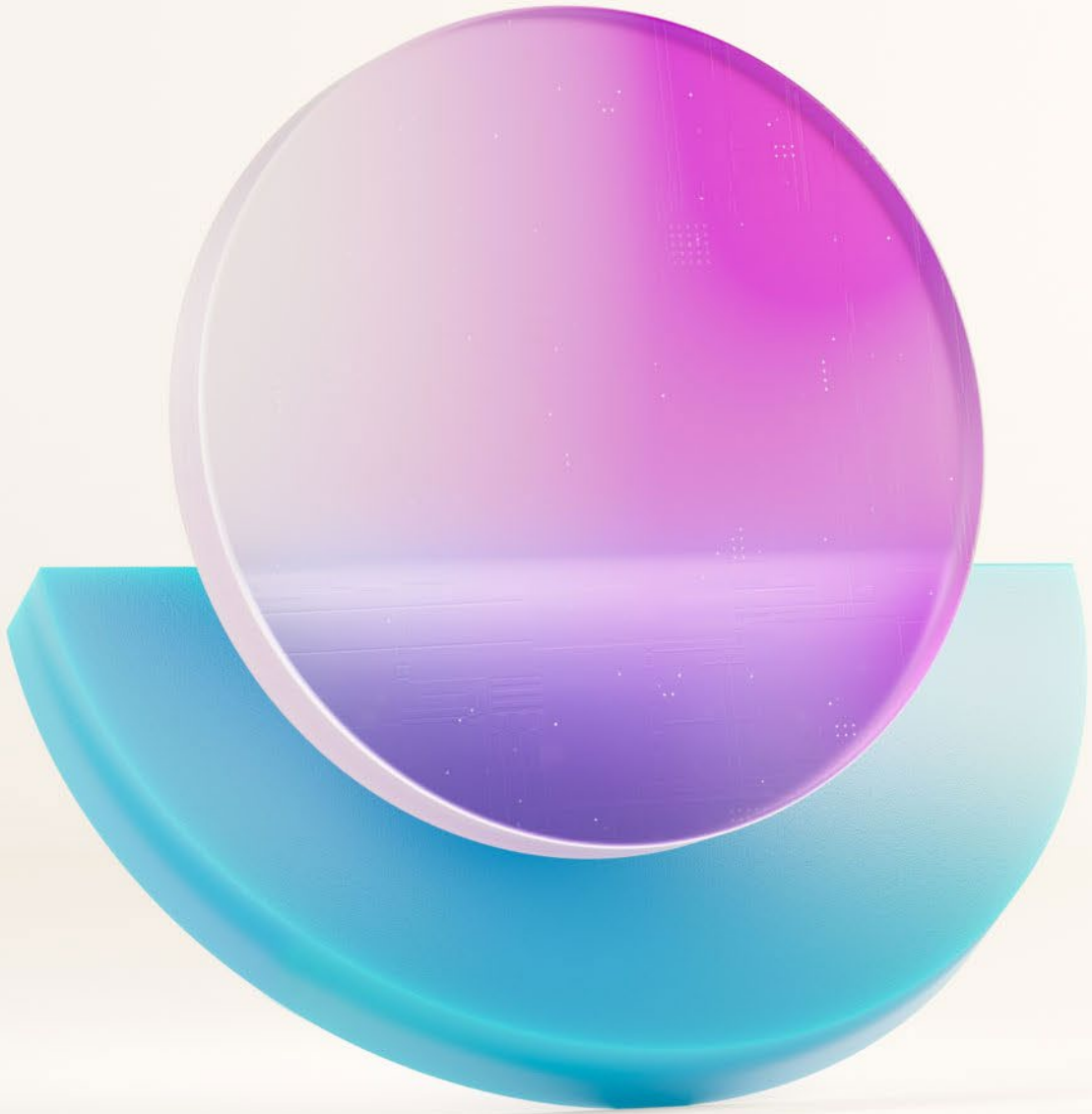


Demo

Scenario 2 – Git & Deployment pipelines

- Git connected to Dev
- Deploy to Test / Prod via Deployment Pipeline
- Orchestrate from Azure DevOps through APIs
- Options like automated unit testing are possible





Limitations & Roadmap

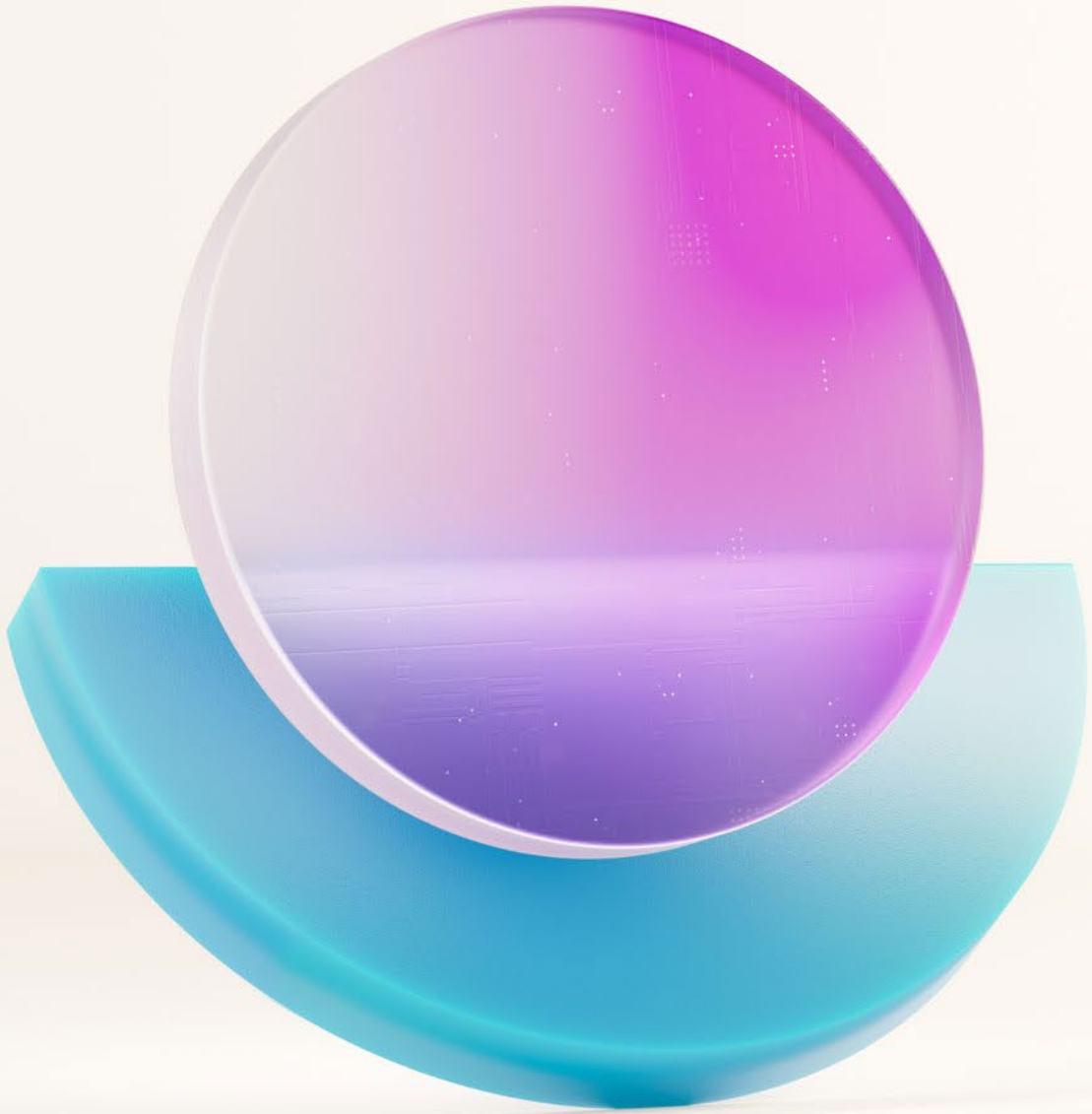
Limitations for automation (May 14th 2024)

Topic	Deployment Pipeline	GIT Integration	Fabric API / Custom Code
Shortcut creation	Not supported	Not supported	Fully supported
Service Principals	n/a	Not supported	Not supported
All item types	Not supported	Not supported	Not supported
Change connections	Deployment rules		
Change Report Connections	Only specific scenarios	Not supported	Report Rebind API
Data Load	Manual	Manual	Can be automated
Other Git Providers	n/a	Not (yet)	Fully supported
Automated Testing	Not supported	Not supported	Fully supported
Branching	n/a	Fixed names	Any
Folders	Not supported	Not supported	??

* List is not complete. Specific scenarios like Metadata references in Notebooks are also not yet supported.

Roadmap

- New git providers - GitHub and GitHub Enterprise
- Azure DevOps connections across different tenants
- Git General Availability (GA)
- Supported items:
  - Semantic models with DirectLake
  - Reflex
  - EventStream
  - Eventually, all items in fabric…
- Service Principal – gradual support
- Folders in Git integration
- Compare changes between workspace and Git branch



Closing

Resources

- Fabric Deployment Pipelines
  https://learn.microsoft.com/en-us/fabric/cicd/deployment-pipelines/intro-to-deployment-pipelines
- Fabric Git Integration
  https://learn.microsoft.com/en-us/fabric/cicd/git-integration/intro-to-git-integration
- API samples for CI/CD

  https://github.com/microsoft/fabric-samples/tree/main/e2e-samples
- Data Factory Testing Framework
  https://github.com/microsoft/data-factory-testing-framework/tree/main
- Microsoft Fabric Git REST APIs
  https://blog.fabric.microsoft.com/en-us/blog/automate-your-ci-cd-pipelines-with-microsoft-fabric-git-rest-apis