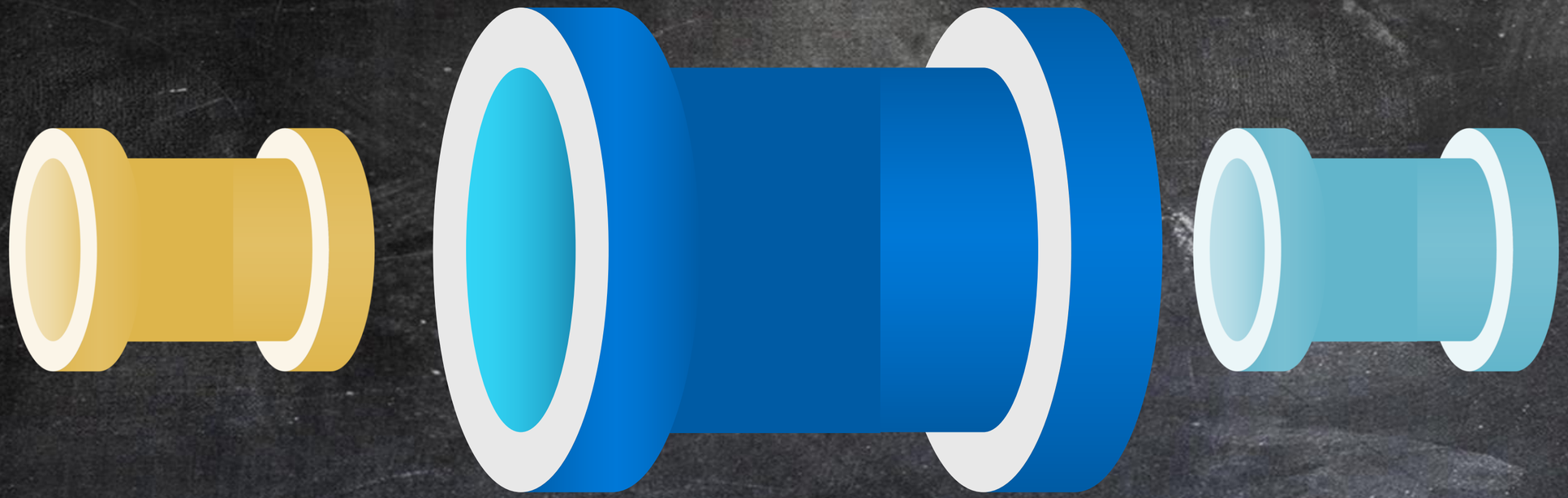
Cloud Formations

Proudly presents...

Integration Pipelines



Paul Andrew

Co-Founder & Director

Chief Technology Officer

/mrpaulandrew

@mrpaulandrew

In/mrpaulandrew

- Mentor | Author
- Speaker | Podcast Host
- Event Organiser

SQL Server 2000

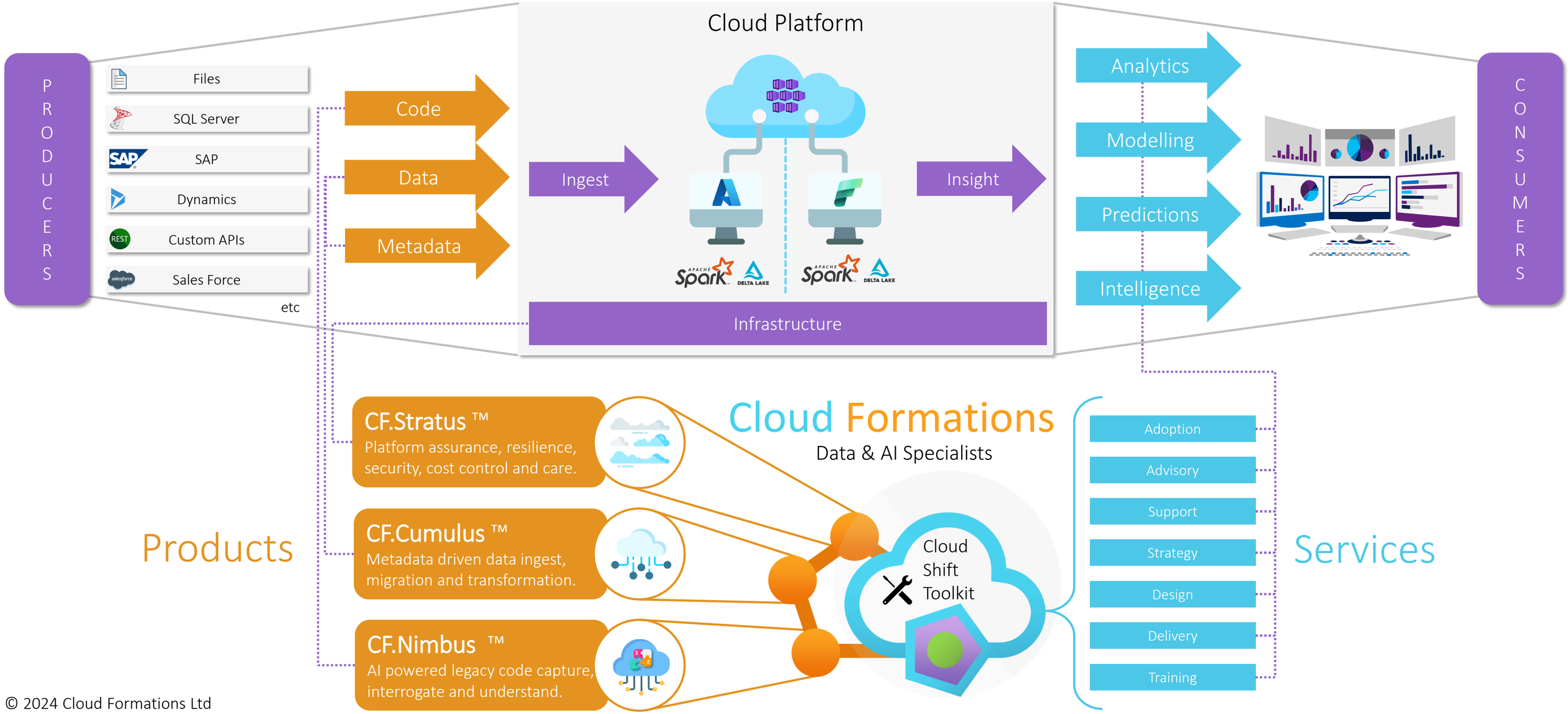
Cloud Formations

How Many Monitors Do You Have?

Our Cloud Shift Toolkit – A Typical Data Journey

Cloud Formations

Couple our AI driven innovative **Products** with our industry leading **Services** to get your data moving, delivering use cases and unlocking business value as part of our **Toolkit**.



© 2024 Cloud Formations Ltd

About You

With regards to data technologies:

- Role
- Strengths
- Weaknesses
- Aspirations

Cloud Formations - Knowledge Transfer & Training

© 2024 Cloud Formations Ltd

Agenda: Data Integration Pipelines

Fundamentals to Level 300

- **Module 1:** Pipeline Fundamentals
  - An Evolution of Orchestration Services
  - Core Components
  - Common Activities
  - Execution Dependencies
- **Module 2:** Integration Runtimes & Gateways
  - Orchestration Compute
    - Azure
    - Hosted
    - SSIS
    - Airflow
  - Using Gateways vs IRs
- **Module 3:** Data Transformation
  - Data Flows
  - Power Query Injection
  - Spark Configuration
  - Use Cases
- **Module 4:** Dynamic Pipelines
  - Expressions & Interpolation
  - Simple Metadata Driven Execution
  - Dynamic Content Chains
  - Reference Names

<< BREAK

???

- **Module 5:** Pipeline Extensibility
  - Azure Batch Service
  - Pipeline Custom Activities
  - Azure Management API
  - Functions
- **Labs**
  - Create Azure resources
  - Build a copy pipeline
  - Create a reusable pipeline
  - Author a data flow
  - Monitor factory activity
  - Explore Synapse pipelines
  - Explore Fabric pipelines
  - Mini-project

<< LUNCH

???

- **Module 6:** Execution Parallelism
  - Control Flow Scale Out
  - Concurrency Limits
  - Internal vs External Activities
  - Metadata Driven Frameworks
- **Module 7:** VNet Integration
  - Private Endpoints
  - Managed VNet's
  - Firewall Bypass
- **Module 8:** Security
  - Service Principals
  - Managed Identities
  - Key Vault Integration & Return Values
  - Customer Managed Keys
  - Pipeline Access & Permissions

<< BREAK

???

- **Module 9:** Monitoring & Alerting
  - Studio Monitoring
  - Log Analytics & Kusto Queries
  - Operational Dashboards
  - Alerting Options
- **Module 10:** Solution Testing
  - Development Time Validation
  - Test Coverage
  - NUnit Tests
- **Module 11:** CI/CD
  - Source Control vs Developer UI
  - Basic ARM Template Deployments
  - Advanced Deployment Patterns
- **Module 12:** Final Thoughts
  - Costs & Conclusions
  - Best Practices

Cloud Formations - Knowledge Transfer & Training

© 2024 Cloud Formations Ltd

Agenda: Data Integration Pipelines

Fundamentals to Level 300

- Module 1: Pipeline Fundamentals
  - The History of Azure Orchestration
  - Synapse Analytics vs Data Factory vs Microsoft Fabric
  - Integration Components
  - Common Activities
  - Execution Dependencies
- Module 2: Integration Runtime Design Patterns
  - Compute Types
    - Azure
    - Hosted
    - SSIS
  - Patterns & Configuration
- Module 3: Data Transformation
  - Data Flows
  - Power Query Injection
  - Spark Configuration
  - Use Cases

<< BREAK

- Module 4: Dynamic Pipelines
  - Expressions & Interpolation
  - Simple Metadata Driven Execution
  - Dynamic Content Chains
  - Reference Names
- Module 5: Pipeline Extensibility
  - Azure Batch Service
  - Pipeline Custom Activities
  - Azure Manag
  - Azure Functi
- Labs
  - Create Azure resources
  - Build a copy pipeline
  - Create a reusable pipeline
  - Author a data flow
- Module 6: Execution Parallelism
  - Control Flow Scale Out
  - Concurrency Limitations
  - Internal vs External Activities
  - Orchestration Framework - procfwk.com
- Module 7: VNet Integration
  - Private Endpoints
  - Managed VNet's
  - Firewall Bypass

Breadth

Depth

- Monitor factory activity
- Explore Synapse pipelines
- Explore Fabric pipelines
- Mini-project

<< LUNCH

- Module 8: Security
  - Service Principals
  - Managed Identities
  - Azure Key Vault Integration
  - Customer Managed Keys
  - Pipeline Access & Permissions

<< BREAK

- Module 9: Monitoring & Alerting
  - Studio Monitoring
  - Log Analytics & Kusto Queries
  - Operational Dashboards
  - Advanced Alerting
- Module 10: Solution Testing
  - Development Time Validation
  - Test Coverage
  - NUnit Tests
- Module 11: CI/CD
  - Source Control vs Developer UI
  - Basic ARM Template Deployments
  - Advanced Deployment Patterns
- Module 12: Final Thoughts
  - Costs & Conclusions
  - Best Practices

Cloud Formations - Knowledge Transfer & Training

© 2024 Cloud Formations Ltd

Agenda: Data Integration Pipelines

Fundamentals to Level 300

- Module 1: Pipeline Fundamentals
  - The History of Azure Orchestration
  - Synapse Analytics vs Data Factory vs Microsoft Fabric
  - Integration Components
  - Common Activities
  - Execution Dependencies
- Module 2: Integration Runtime Design Patterns
  - Compute Types
    - Azure
    - Hosted
    - SSIS
  - Patterns & Configuration
- Module 3: Data Transformation
  - Data Flows
  - Power Query Injection
  - Spark Configuration
  - Use Cases

<< BREAK

- Module 4: Dynamic Pipelines
  - Expressions & Interpolation
  - Simple Metadata Driven Execution
  - Dynamic Content Chains
  - Reference Names
- Module 5: Pipeline Extensibility
  - Azure Batch Service
  - Pipeline Custom Activities
  - Azure Management API
  - Azure Functions
- Labs
  - Create Azure resources
  - Build a copy pipeline
  - Create a reusable pipeline
  - Author a data flow
- Module 6: Execution Parallelism
  - Control Flow Scale Out
  - Concurrency Limitations
  - Internal vs External Activities
  - Orchestration Framework
- Module 7: VNet Integration
  - Private Endpoints
  - Managed VNet's
  - Firewall Bypass

Development

<< LUNCH

- Monitor factory activity
- Explore Synapse pipelines
- Explore Fabric pipelines
- Mini-project

Production

- Module 8: Security
  - Service Principals
  - Managed Identities
  - Azure Key Vault Integration
  - Customer Managed Keys
  - Pipeline Access & Permissions

<< BREAK

- Module 9: Monitoring & Alerting
  - Studio Monitoring
  - Log Analytics & Kusto Queries
  - Operational Dashboards
  - Advanced Alerting
- Module 10: Solution Testing
  - Development Time Validation
  - Test Coverage
  - NUnit Tests
- Module 11: CI/CD
  - Source Control vs Developer UI
  - Basic ARM Template Deployments
  - Advanced Deployment Patterns
- Module 12: Final Thoughts
  - Costs & Conclusions
  - Best Practices

Cloud Formations - Knowledge Transfer & Training

© 2024 Cloud Formations Ltd

Agenda: Data Integration Pipelines

Fundamentals to Level 300

- **Module 1:** Pipeline Fundamentals
  - An Evolution of Orchestration Services
  - Core Components
  - Common Activities
  - Execution Dependencies
- **Module 2:** Integration Runtimes & Gateways
  - Orchestration Compute
    - Azure
    - Hosted
    - SSIS
    - Airflow
  - Using Gateways vs IRs
- **Module 3:** Data Transformation
  - Data Flows
  - Power Query Injection
  - Spark Configuration
  - Use Cases
- **Module 4:** Dynamic Pipelines
  - Expressions & Interpolation
  - Simple Metadata Driven Execution
  - Dynamic Content Chains
  - Reference Names

<< BREAK

11:15 to 11:40

- **Module 5:** Pipeline Extensibility
  - Azure Batch Service
  - Pipeline Custom Activities
  - Azure Management API
  - Functions
- **Labs**
  - Create Azure resources
  - Build a copy pipeline
  - Create a reusable pipeline
  - Author a data flow
  - Monitor factory activity
  - Explore Synapse pipelines
  - Explore Fabric pipelines
  - Mini-project

<< LUNCH

13:30 to 14:15

- **Module 6:** Execution Parallelism
  - Control Flow Scale Out
  - Concurrency Limits
  - Internal vs External Activities
  - Metadata Driven Frameworks
- **Module 7:** VNet Integration
  - Private Endpoints
  - Managed VNet's
  - Firewall Bypass
- **Module 8:** Security
  - Service Principals
  - Managed Identities
  - Key Vault Integration & Return Values
  - Customer Managed Keys
  - Pipeline Access & Permissions
- **Module 9:** Monitoring & Alerting
  - Studio Monitoring
  - Log Analytics & Kusto Queries
  - Operational Dashboards
  - Alerting Options

<< BREAK

15:35 to 16:00

- **Module 10:** Solution Testing
  - Development Time Validation
  - Test Coverage
  - NUnit Tests
- **Module 11:** CI/CD
  - Source Control vs Developer UI
  - Basic ARM Template Deployments
  - Advanced Deployment Patterns
- **Module 12:** Final Thoughts
  - Costs & Conclusions
  - Best Practices

Cloud Formations - Knowledge Transfer & Training

© 2024 Cloud Formations Ltd