



Kicking off compute on Gadi with Globus flows

KISARU LIYANAGE, NCI

MATTHEW DOWNTON, NCI

Globus for Bioinformatics Workflows

- Bioinformatics workflows often require large data transfers followed by workflow execution
- Traditional approach:
 - Manually initiate and monitor data transfers
 - Manually launch workflows after data arrival
 - Prone to delays and human error
- Improved approach with Globus Flows:
 - Automate reliable data transfers using **Globus Transfer**
 - Automatically trigger/monitor downstream workflows using **Globus Compute**
 - Enables faster, reproducible, and fully automated workflows

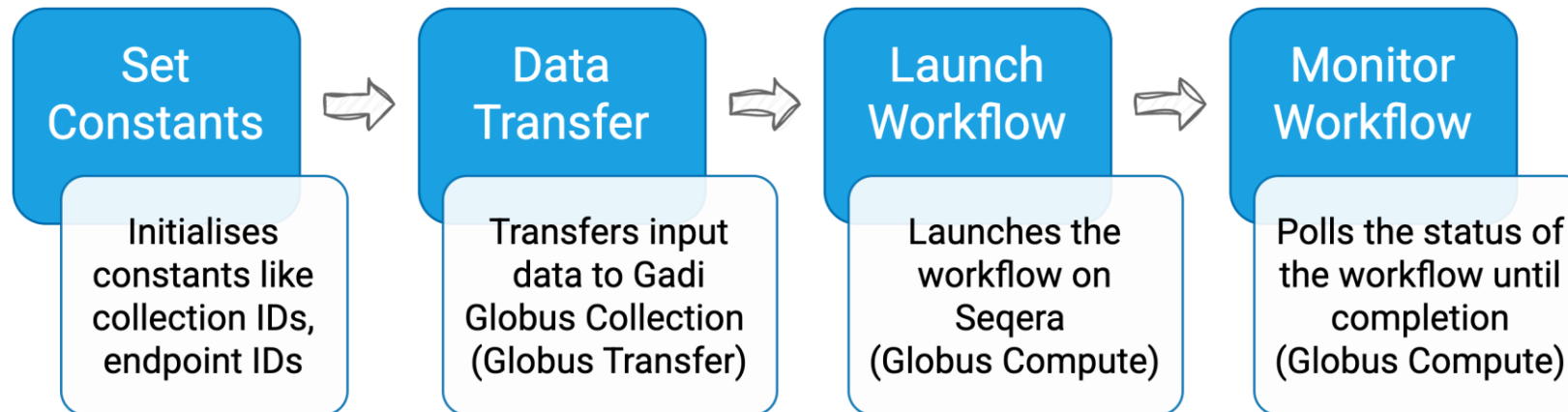


Globus Flow Design

Two approaches

1. Data transfer → Launch workflow within a persistent session
2. Data transfer → Launch workflow on Seqera platform

[GitHub repository: <https://github.com/kisarur/gadi-genomics-globus-flow>]





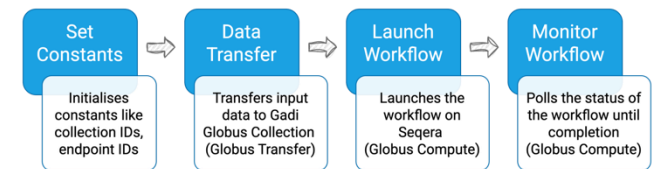
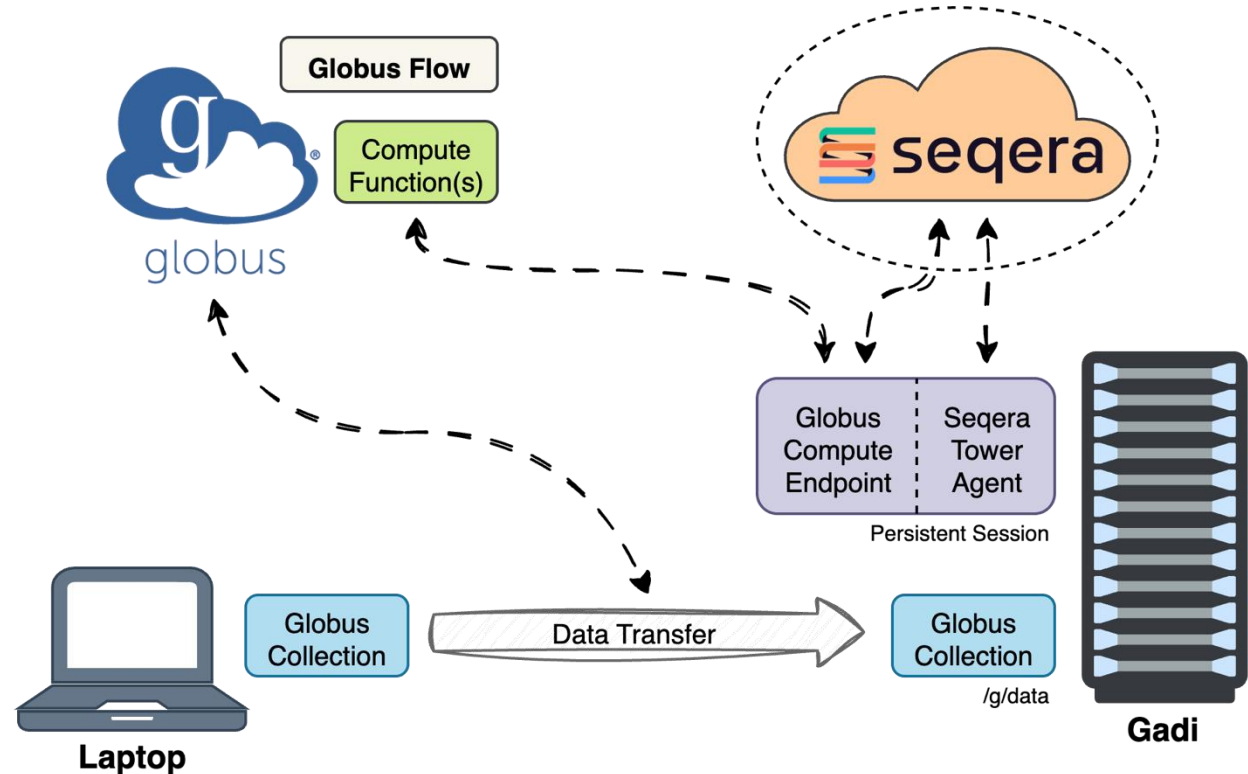
System Architecture Overview

Components

- Globus Flow
 - Flow Definition
 - Input Schema
- Globus Collections
- Globus Compute Endpoint
- Compute Function(s)
- Seqera Tower Agent (Optional)

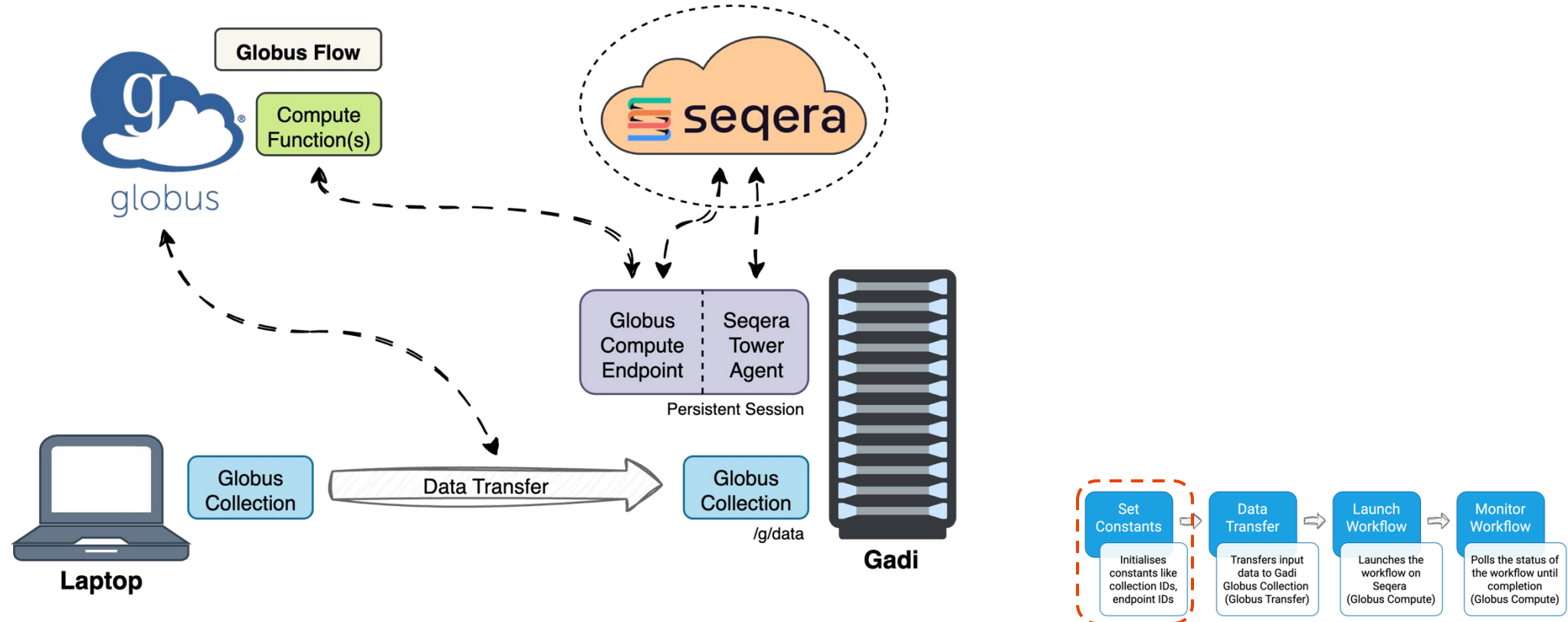
System Architecture Overview

Interconnection



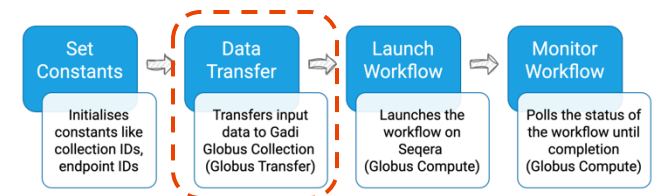
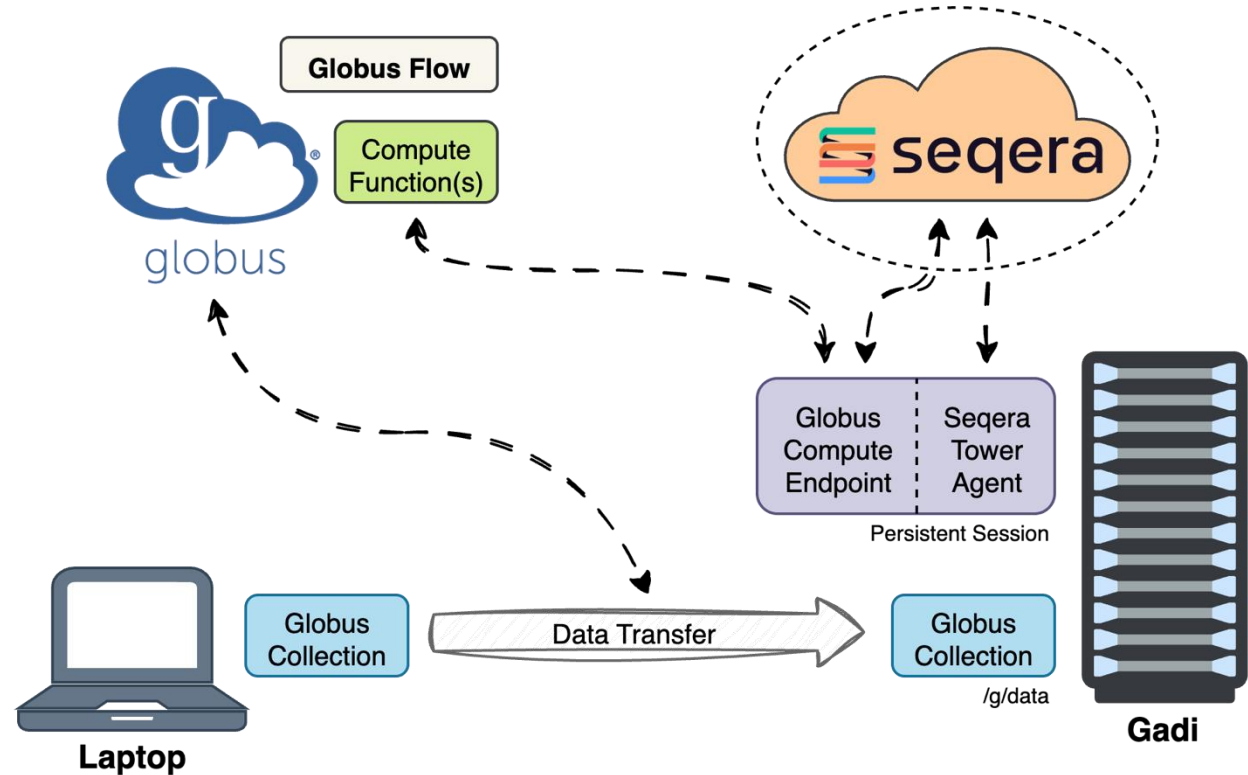
System Architecture Overview

Interconnection



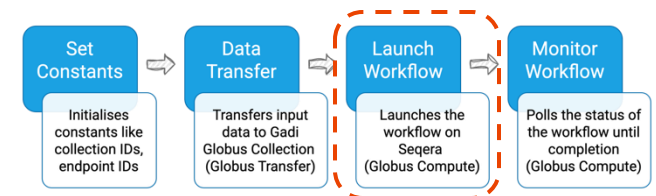
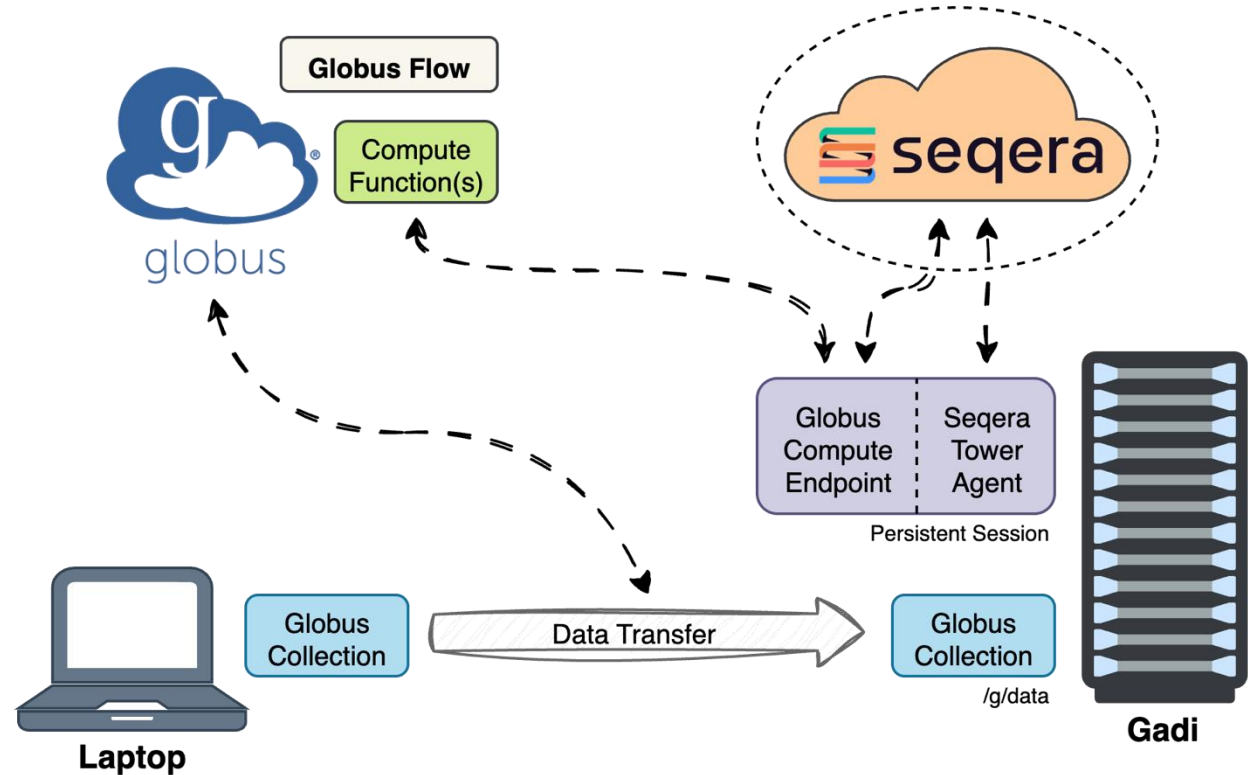
System Architecture Overview

Interconnection



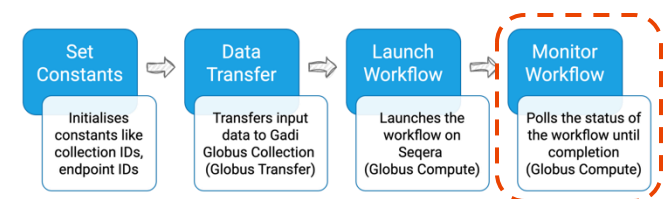
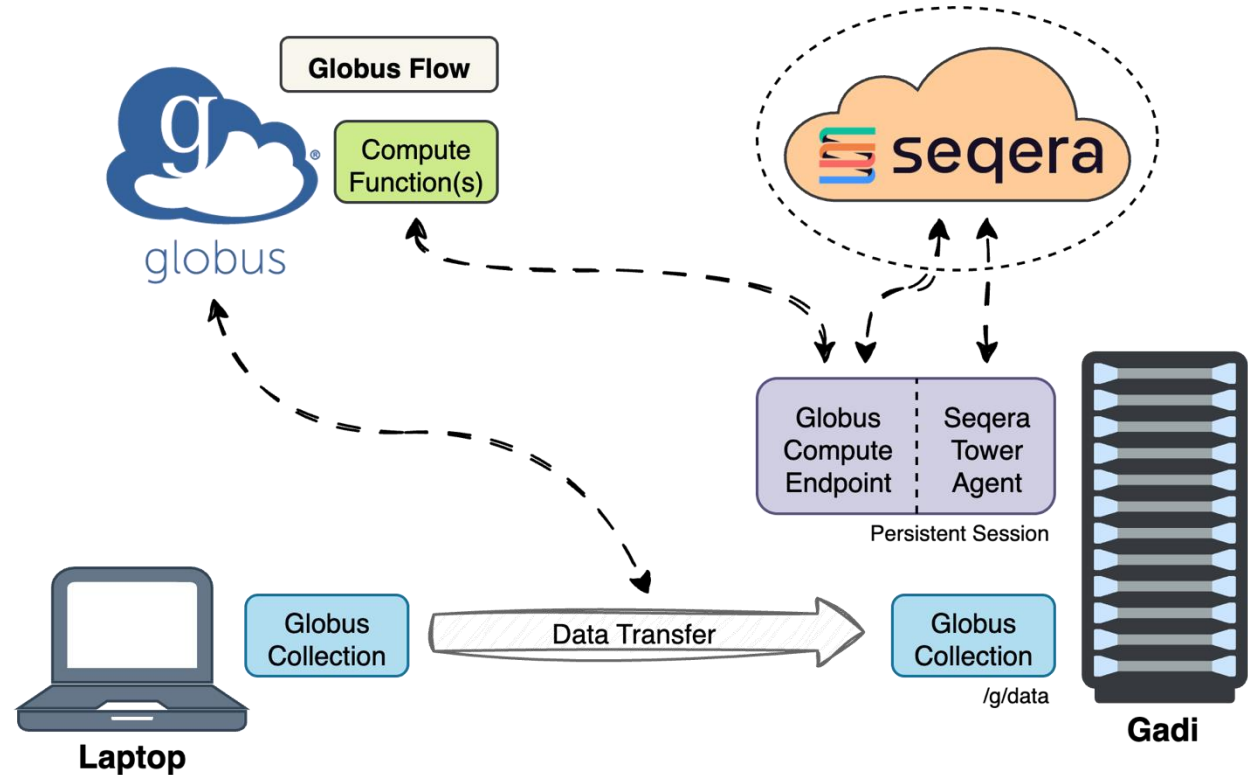
System Architecture Overview

Interconnection



System Architecture Overview

Interconnection



Demo

