

Extending SSIS with custom Data Flow components

Darren Green

Abstract

Get some real-world insights into developing data flow components for SSIS. This starts with an introduction to the data flow pipeline engine, and explains the real differences between adapters and the three sub-types of transformation. Understanding how the different types of component behave and manage data is key to writing components of your own, and probably should be required knowledge for anyone building packages at all. Using sample code throughout, I will show you how to write components, as well as highlighting best practice and lessons learned.

Terminology

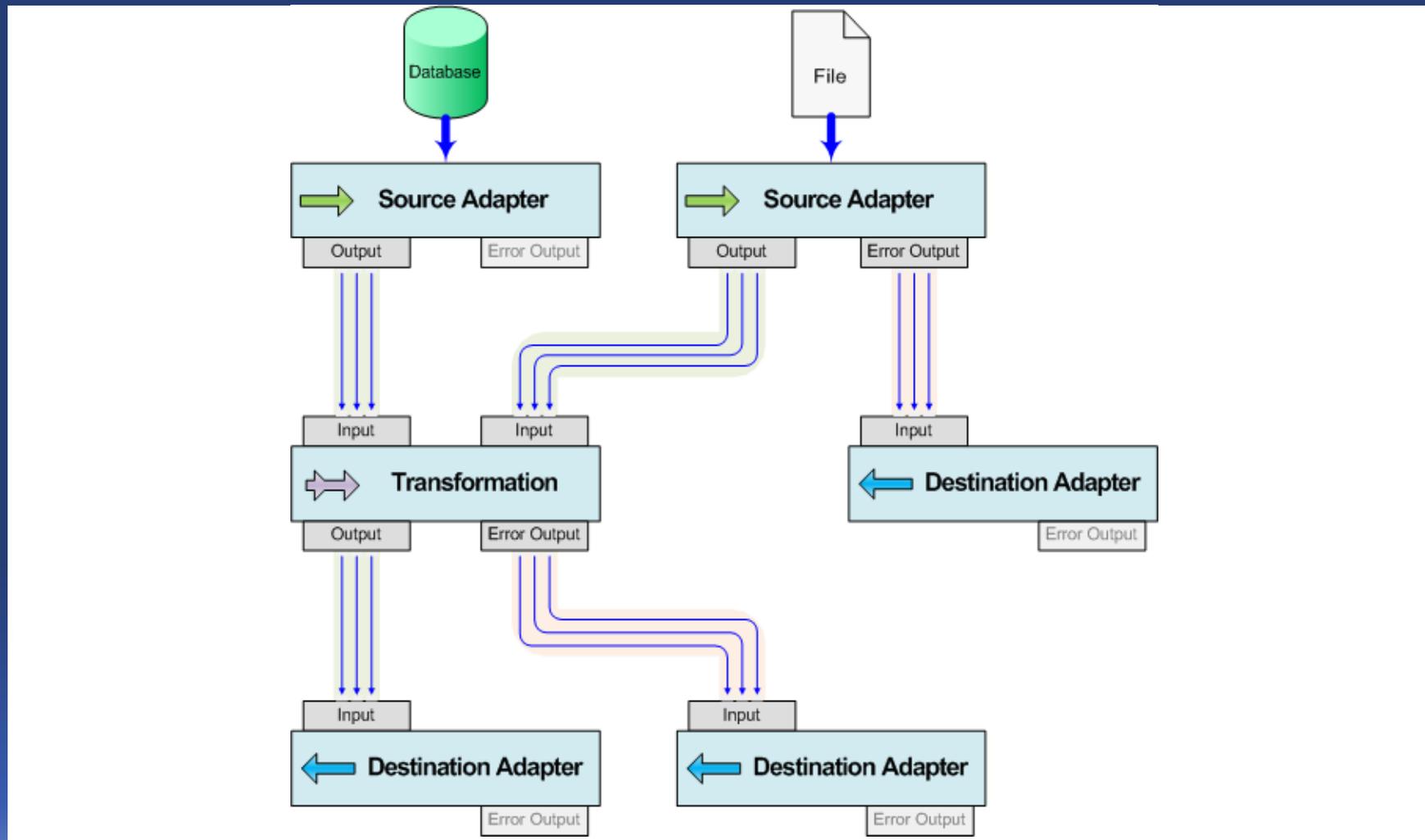
- Control Flow Task
- Data Flow Component
 - Source Adapter
 - Transformation
 - Destination Adapter

Component Types

- Source
 - One or more outputs, no inputs
- Destination
 - One or more inputs, no outputs
- Transformation
 - One or more inputs, one or more outputs

- *Error Outputs*

Data Flow Mechanics



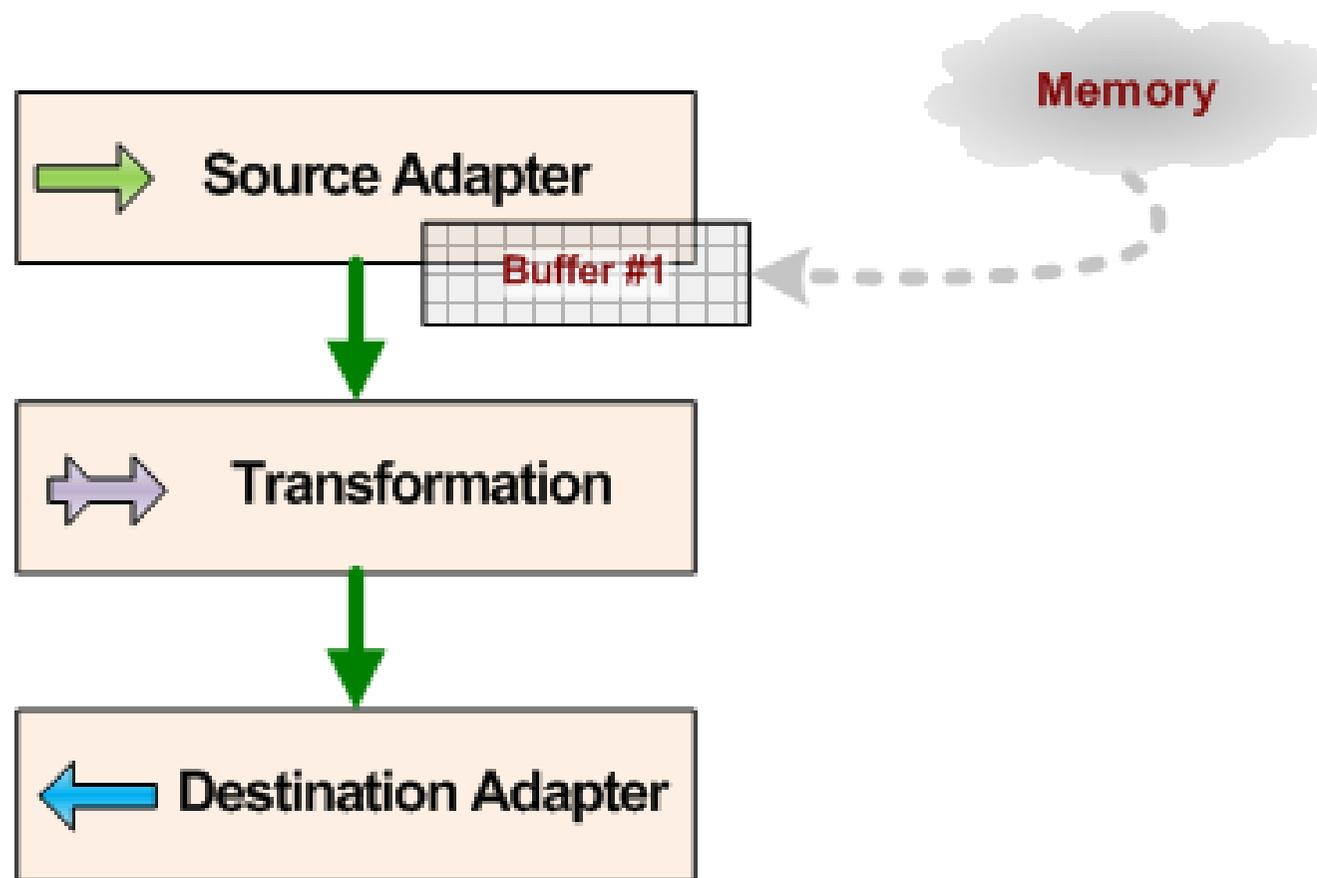


Components Demo

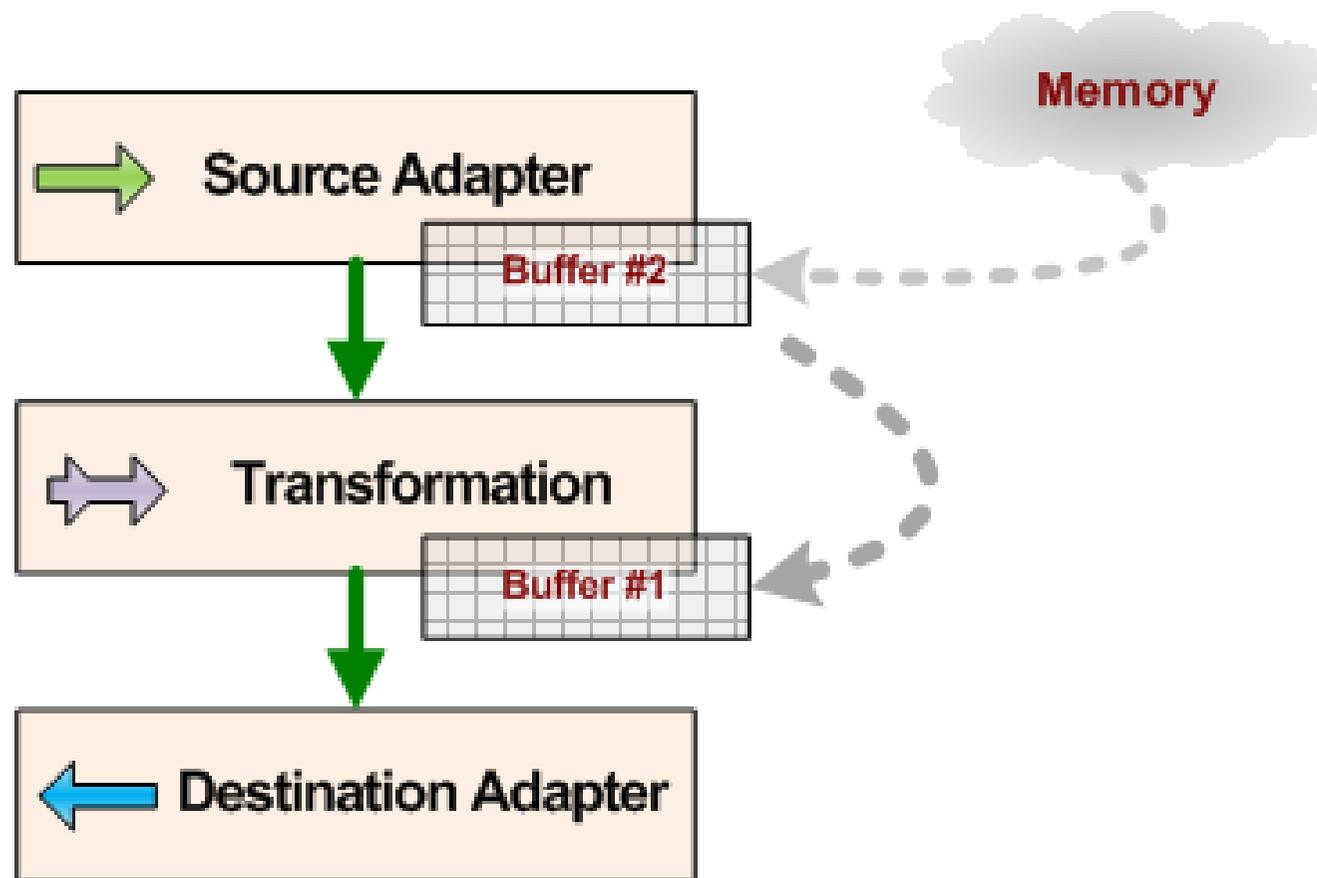
Buffers

- Memory structures for moving data in the pipeline
 - Working set of rows, passed between components
 - Multiple buffers in use at one time
- Fixed size determined by engine
 - number of rows = buffer size / row width
- Fixed internal structure
 - Cannot add a column, it already exists
 - Cannot change column, types or widths
 - Cannot remove a column, it is always there
 - Cannot add or remove a rows

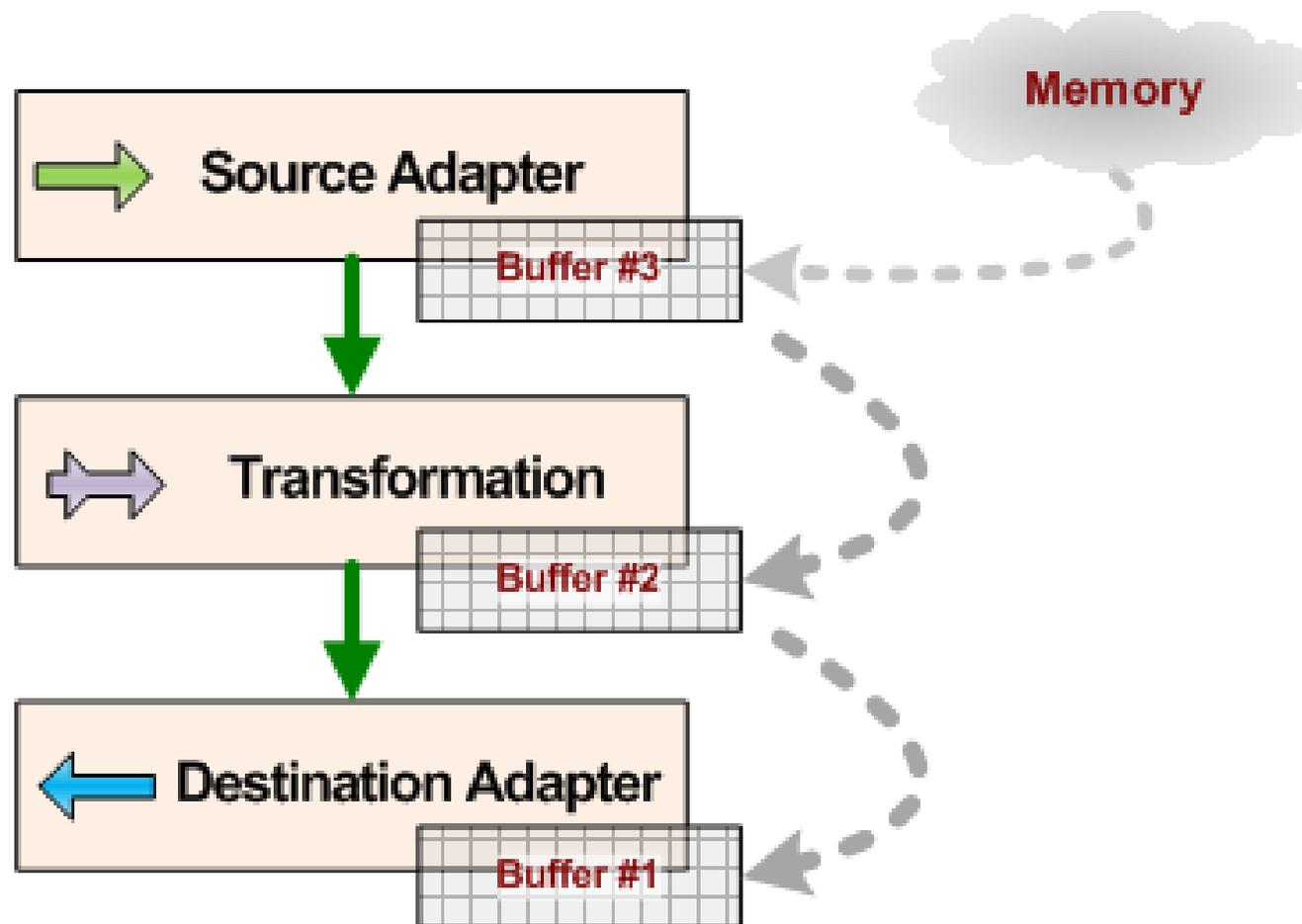
Buffers – Synchronous Data Flow



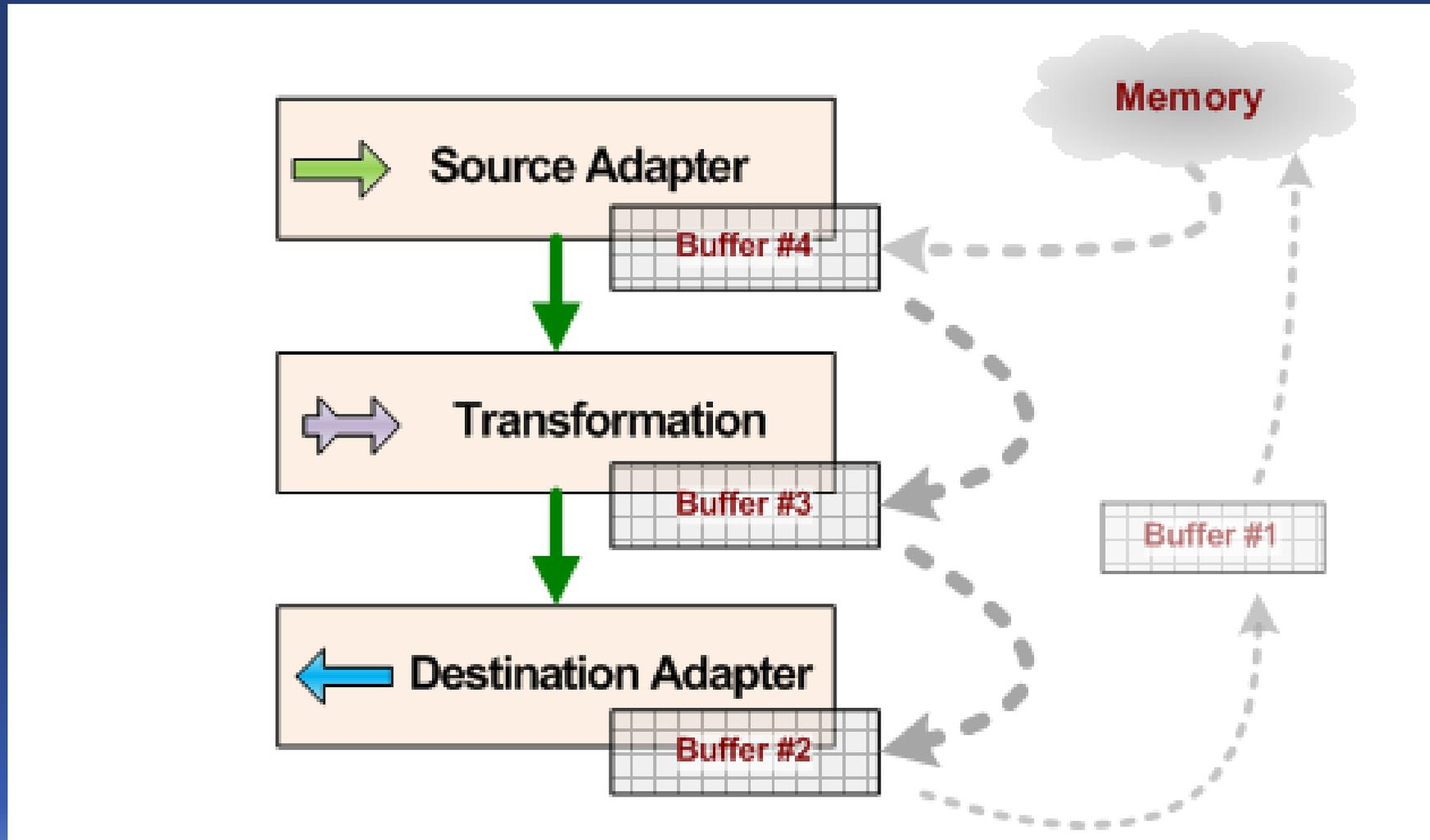
Buffers – Synchronous Data Flow



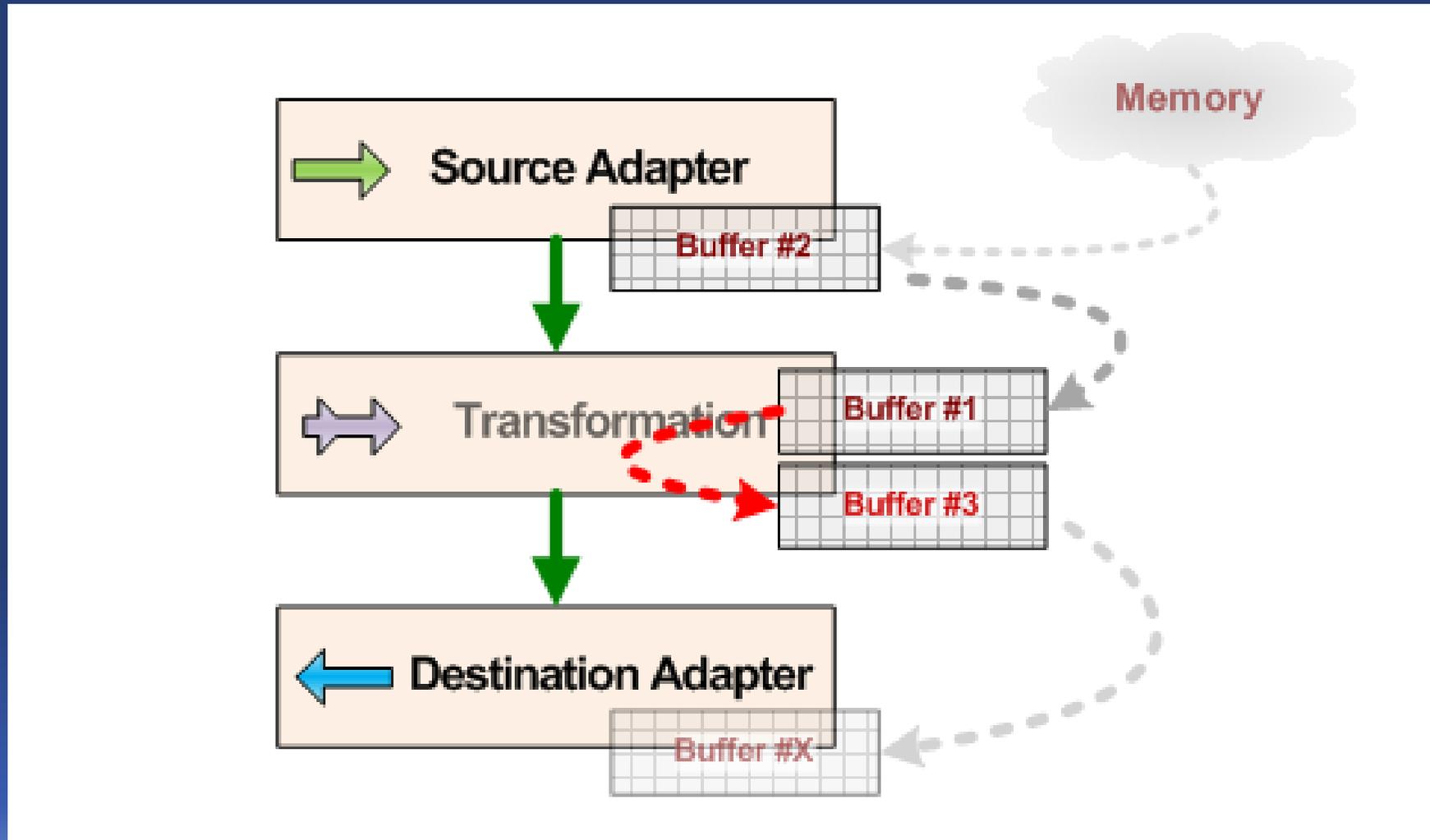
Buffers – Synchronous Data Flow



Buffers – Synchronous Data Flow



Buffers – Asynchronous Data Flow



Synchronous vs. Asynchronous

- Synchronous uses one buffer for the input and output
- Asynchronous uses separate input and output buffers
- *Expensive to create and populate a buffer!*



Components Demo

Design Tips

- Use custom properties
 - Label inputs and outputs with function
 - Relate input and output columns together
 - Relate source and target columns
 - *Validate properties, follow MS conventions*
- Use error outputs
 - Can be use for negative condition too
- Use design-time Component methods
 - Use to manage settings
 - Use to apply settings

Resources

- SQL Server Books Online (Get the Updates!)
- SQL Server Product Samples (Get the Updates!)
- <http://www.codeplex.com/MSFTISProdSamples>

- Additional Samples (Search for “SSIS”)
- <http://www.microsoft.com/downloads/>

- Darren Green
- darren.green@konesans.com

Feedback Forms

Thank you