

GCN Tutorial

EEE 525

Agenda

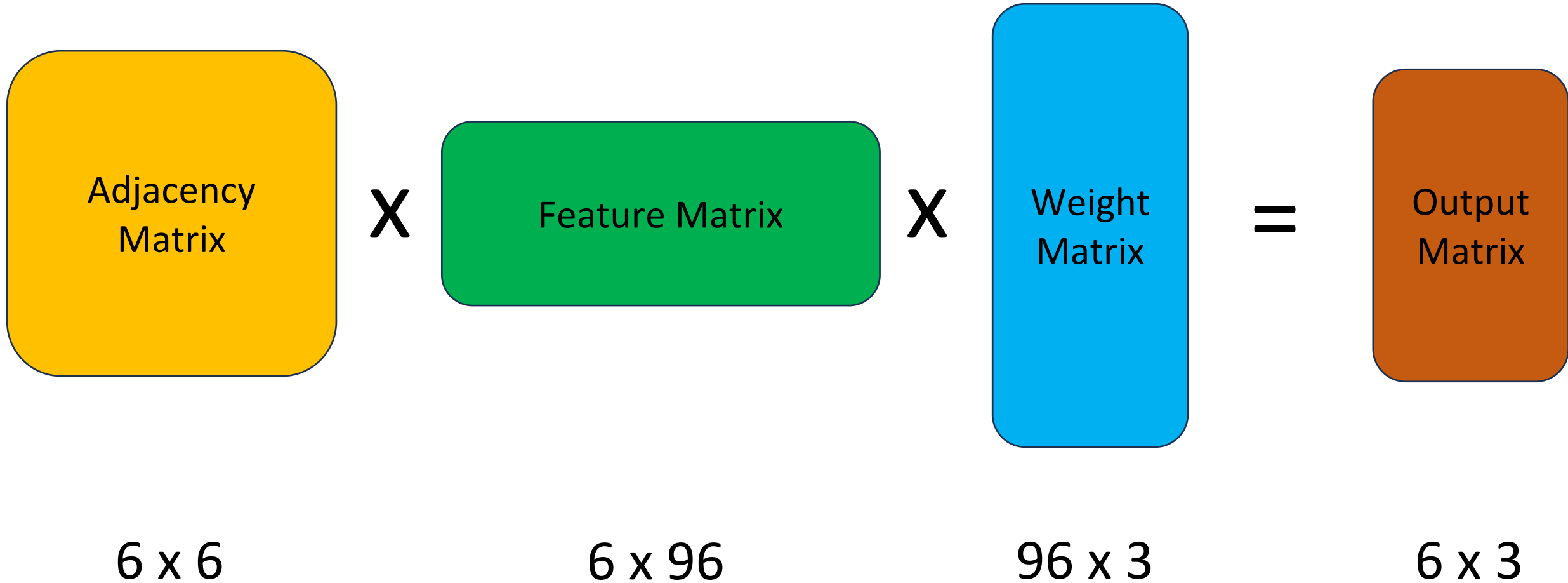
- Overview and understanding GCN
- Data flow and steps to complete GCN operation
- Basic architecture for GCN design
- Tips and tricks

Overview and understanding GCN

GCN Overview

- Graph Convolutional Neural Networks (GCNs)
- GCNs are a type of neural network designed for analyzing and processing data represented as graphs
- Graphs consist of nodes and edges, making them suitable for modeling relationships and connections

Very Basic Overview (Matrix Multiplication)



The 3 Matrices

- Feature Matrix
- Weight Matrix
- Adjacency Matrix

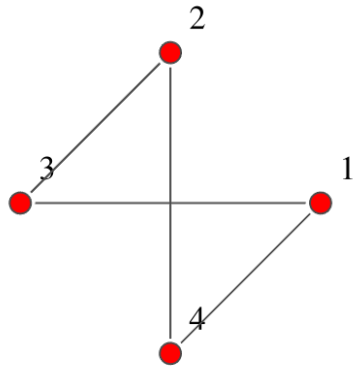
Feature Matrix

- The feature matrix represents the features associated with each node in the graph
- Data inputs to any other machine learning model
- If there are N nodes in the graph, the feature matrix will have N rows

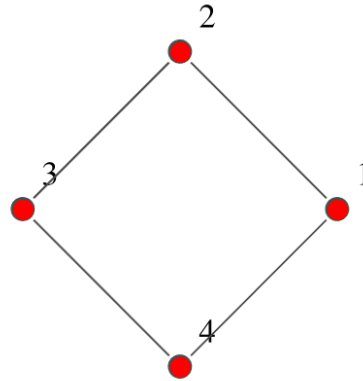
Weight Matrix

- The weight matrix contains parameters that are learned during the training process
- These parameters are often initialized randomly and updated using optimization techniques like gradient descent
- The weights determine how much influence each neighbor has on the central node's feature update

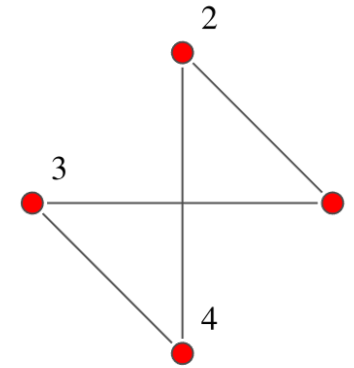
Adjacency Matrix



$$\begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$$



$$\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$



$$\begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

Example (Social Networks)

- Features Matrix: A list of users and their interests
- Weight Matrix: Tuned parameters from training the model
- Adjacency Matrix: The users that are connected to each other
- $FM \times WM$: Users and how much influence they might have on each other
- $ADJ \times FM \times WM$: Leverage connection data to build communities

Extra Resources

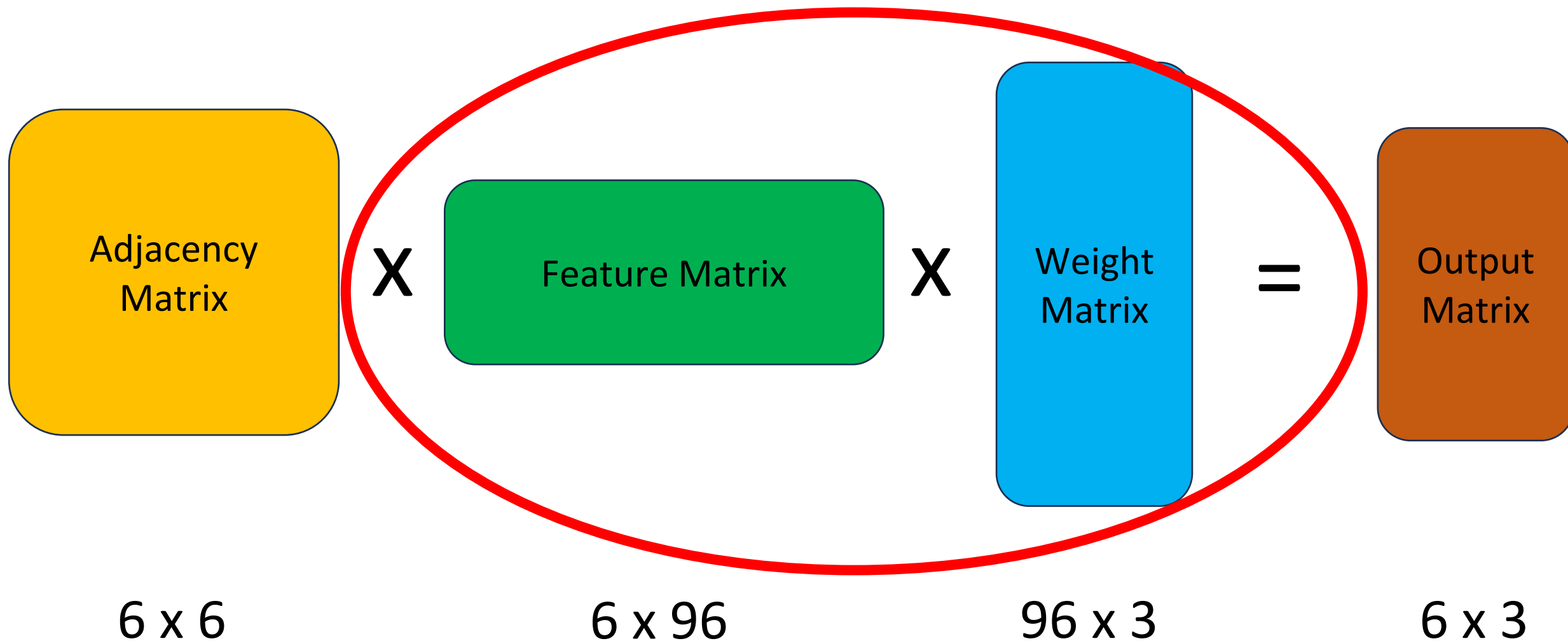
- <https://towardsdatascience.com/graph-convolutional-networks-explained-d88566682b8f>
- <https://distill.pub/2021/gnn-intro/>
- <https://www.youtube.com/watch?v=F3PgltDzllc>

Data flow and steps to complete GCN operation

Steps Overview

- Transformation: Multiplication of feature and weight matrix
- Combination: Multiplication of COO and FM_WM (streaming)
- Argmax Function: Node classification

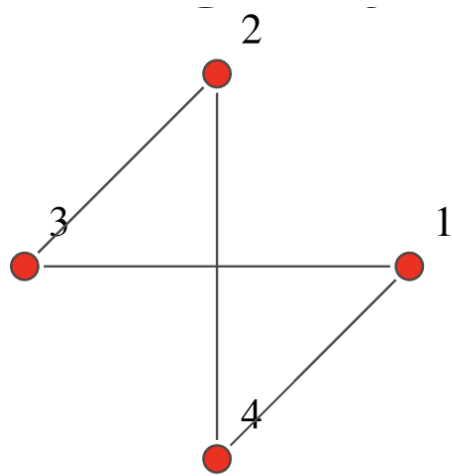
???Why do this multiplication first???



Transformation Steps

1. Load a row of data from the feature matrix and store locally
2. Loop through all the columns of the weight matrix
3. Perform vector multiplication
4. Store vector multiplication output
5. Load next row of data from the feature matrix and store locally
6. Go to step 2

Understanding Adjacency vs Coordinate Format (COO)



$$\begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$$

Edge 1	Edge 2	Edge 3	Edge 4
001	001	010	011
100	011	100	010

Combination Step

1. After completing the the transformation step
2. Start at the first edge of the COO matrix
- 3. Translate the COO data to matrix multiplication of ADJ x (FM_WM)**
4. Go to the next edge in the COO matrix
5. Repeat step 3

Argmax Function: Node Classification

1. After completing the the transformation and combination step
2. ADJ_FM_WM Matrix should be a 6 x 3 matrix
3. Loop through the rows in memory block that stores ADJ_FM_WM
4. Find which of the 3 columns contain the largest value per row
5. Store and output to the TB

See next slide for example

Argmax function example

Column 1	Column 2	Column 3
121	543	374
45	643	304
145	130	18
547	485	766
35	567	43
234	654	789

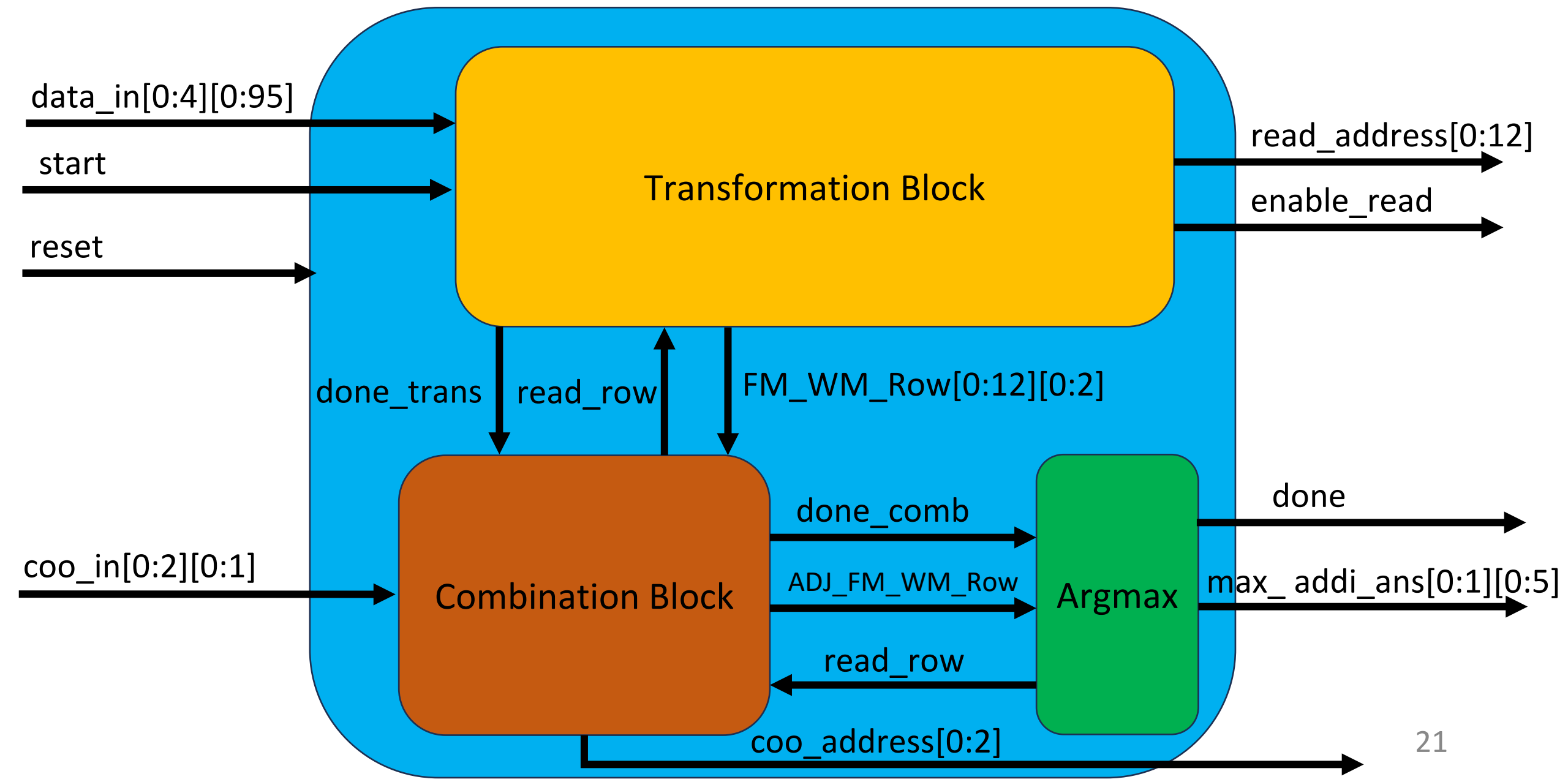
ADJ_FM_WM

Classification
2
2
1
3
2
3

Answer

Basic architecture for GCN design

GCN Block Overview



GCN inputs

data_in: rows/cols from both Weight and Feature matrix depending on the address

start: TB telling the DUT to start performing operations

coo_in: COO data that will be streamed in

clk: Clock

reset: Reset

GCN Outputs

read_address: The address of both rows/cols of the Weight and Feature matrices
(See Later slide about the address Scheme)

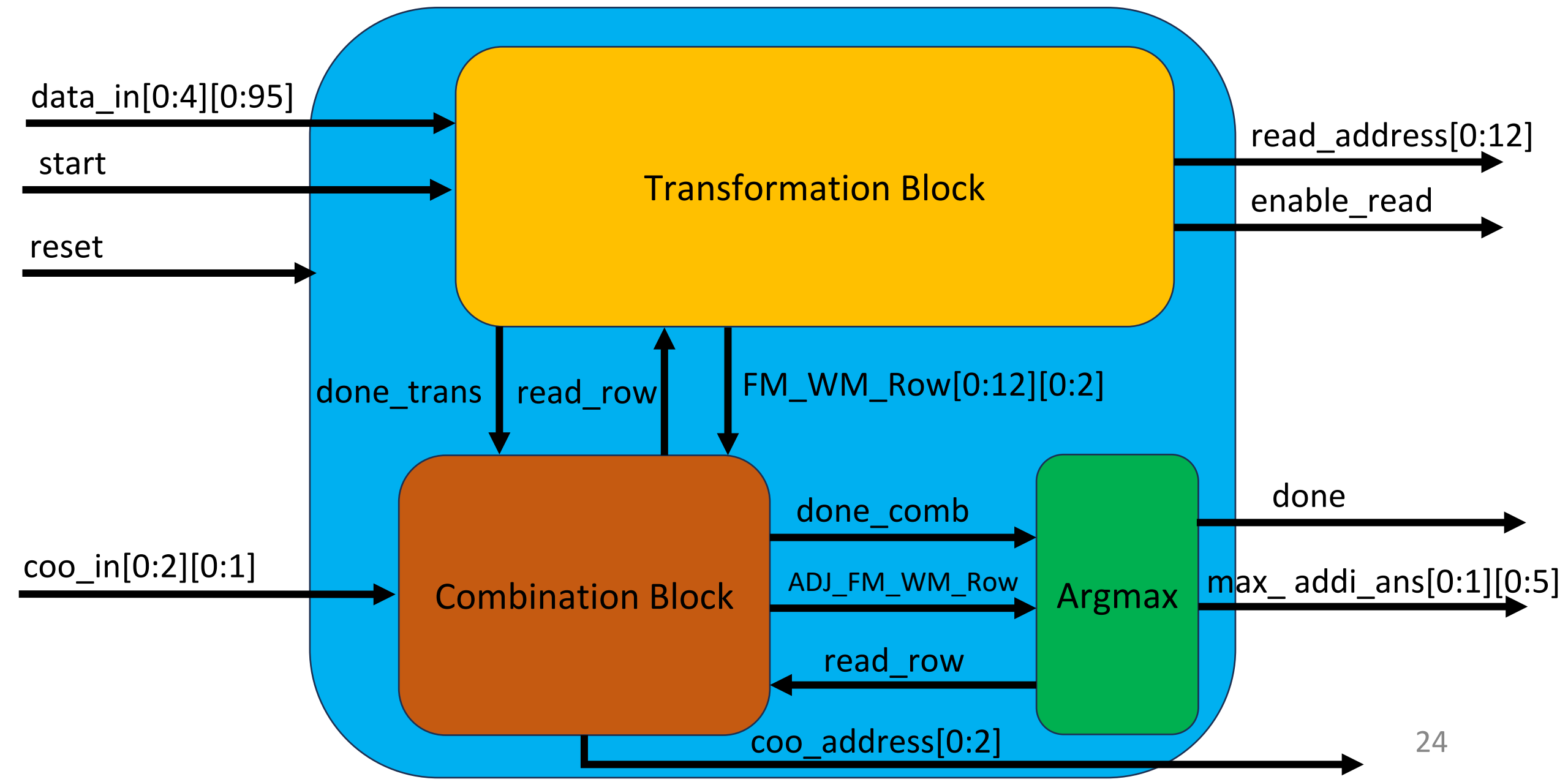
enable_read: Signal enabling the DUT to read rows/cols of the WM & FM from the TB

done: The GCN calculation is complete

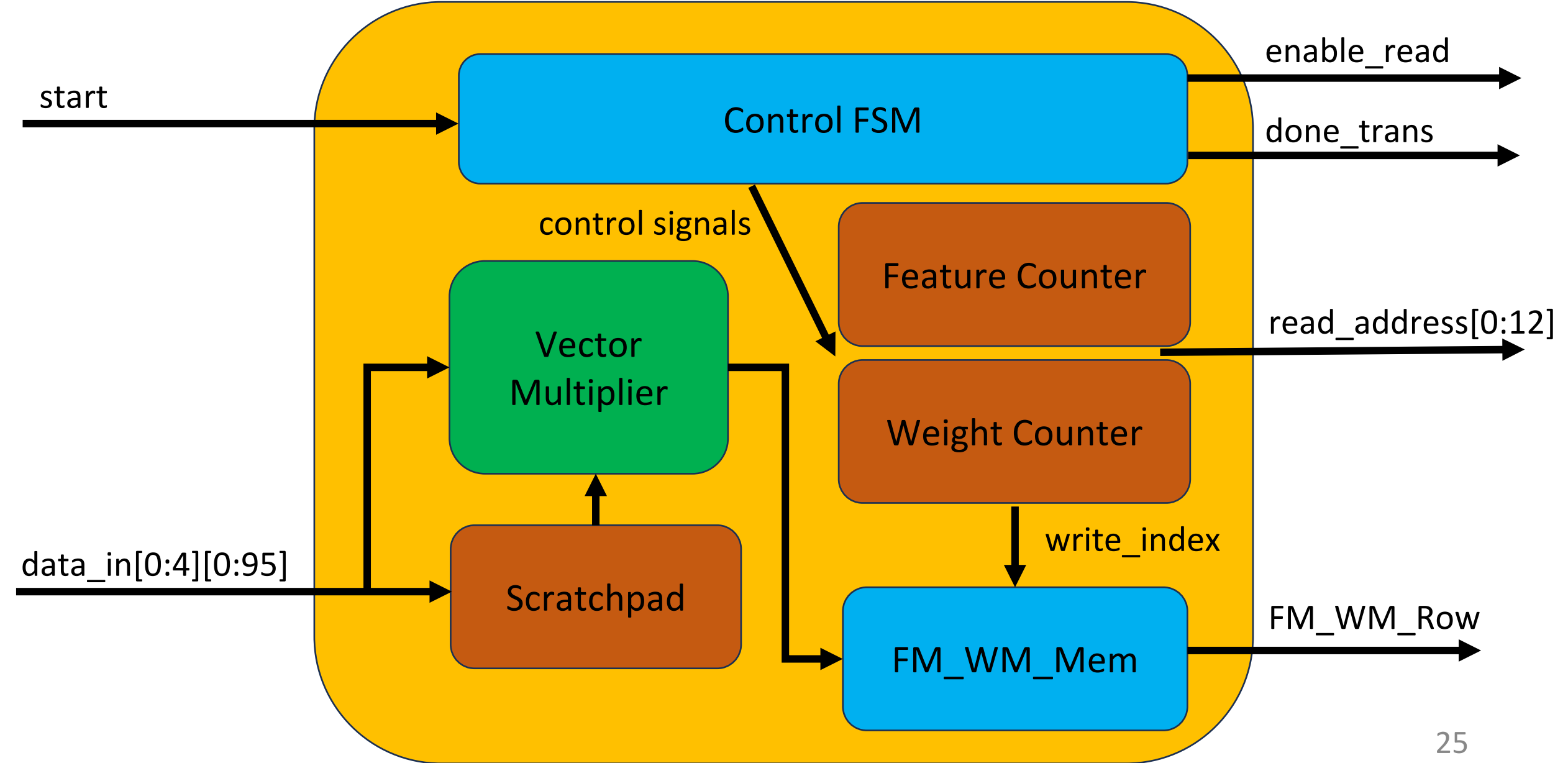
max_addi_answer: The classification for all the nodes in the graph

coo_address: Address of the column in the coo_matrix

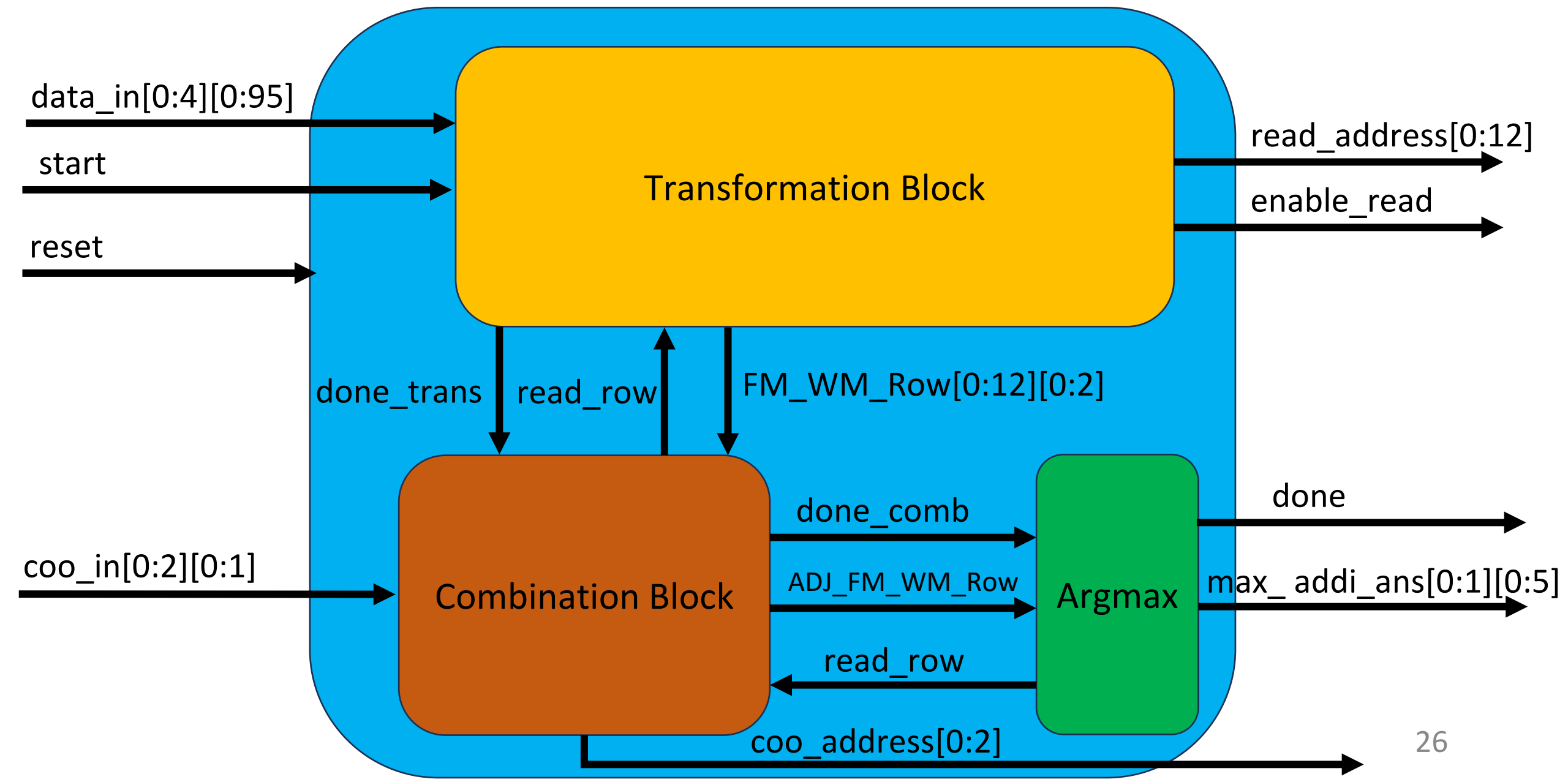
GCN Block Overview



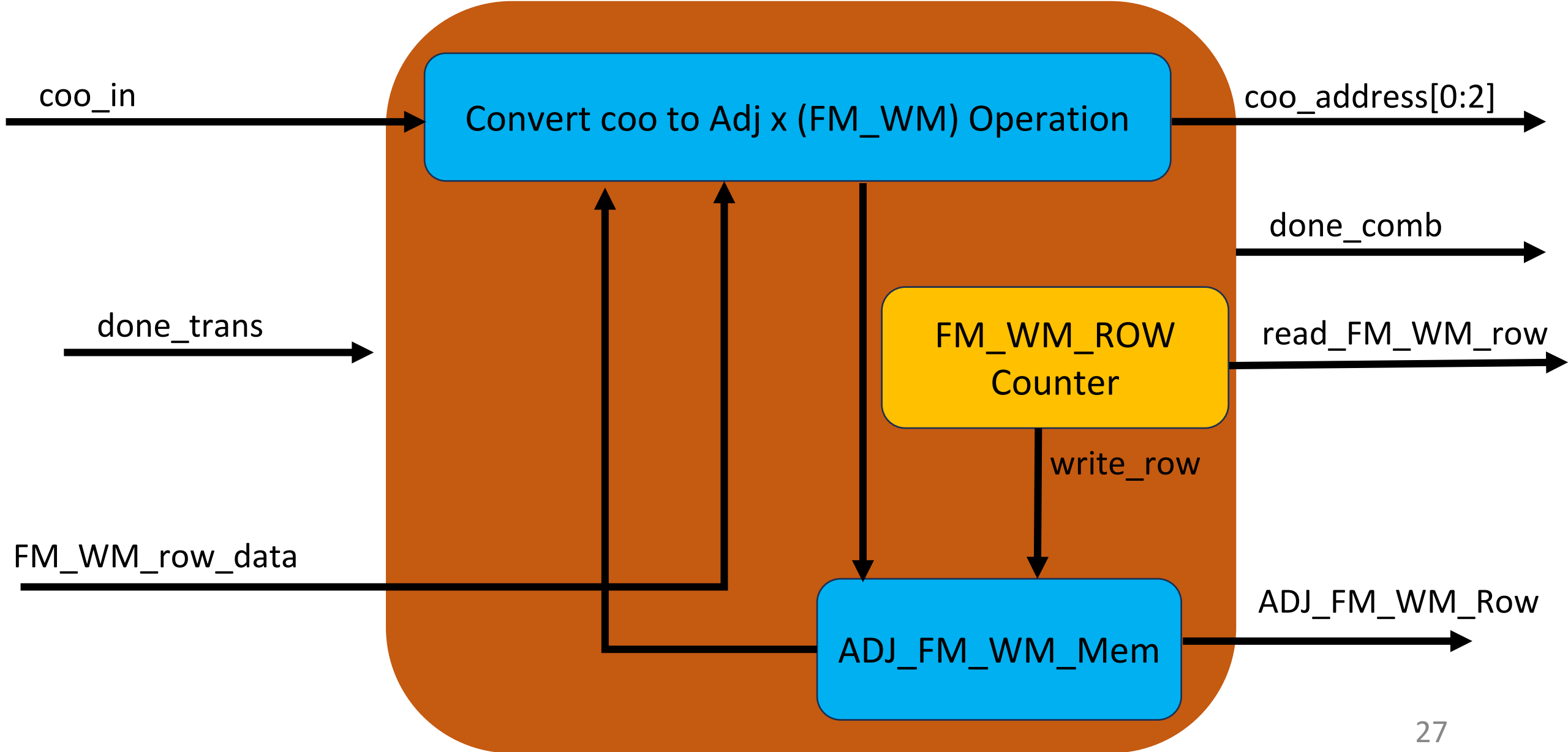
Transform Block Overview



GCN Block Overview



Combination Block Overview

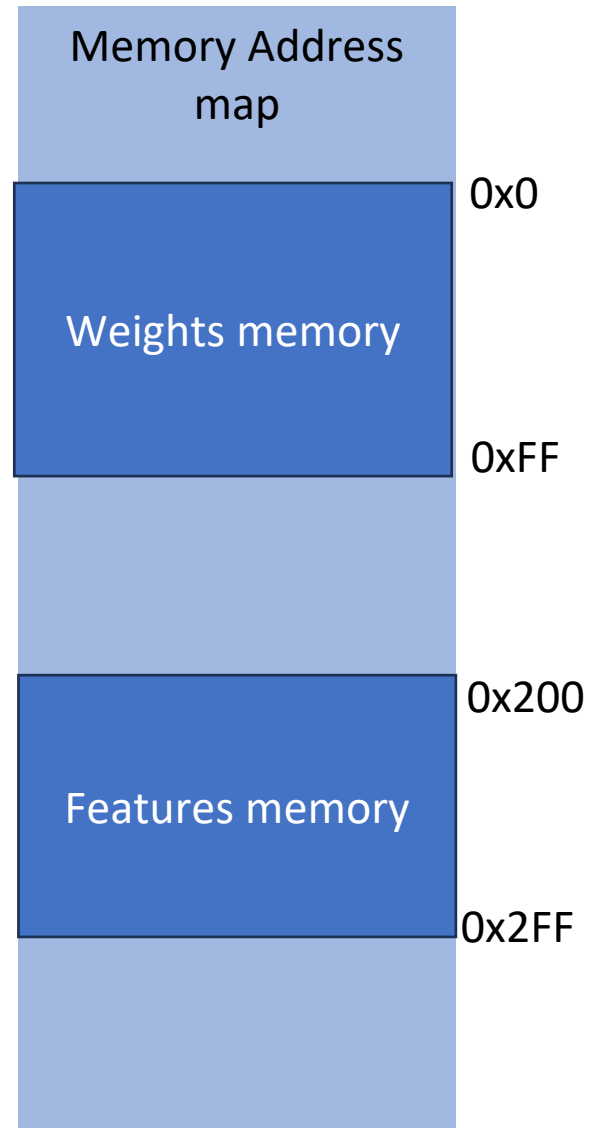


Argmax Block Overview



?

Virtual Memory Scheme Overview



Blocks you will be given

- FSM in transform block
- FM_WM_Mem in transform block
- Scratchpad in transform block
- ADJ_FM_WM_Mem in combination block

Tips and Tricks

Tips and Tricks

- Draw a picture
- START EARLY
- Come to office hours for questions
- Test incrementally not just at the very end of writing RTL
- ChatGPT will not always write synthesizable code (!!!Be Careful!!!)

Tips and Tricks

- Luke: RTL & TB
- START EARLY
- Come to office hours for questions
- Test incrementally not just at the very end of writing RTL
- ChatGPT will not always write synthesizable code (!!!Be Careful!!!)