Pagerank Algorithm

AHDL project

Nikunj Rajput
Electrical and Computer Engineering
NYU Tandon School of Engineering
New York, USA
Nr1536@nyu.edu

Vaishnavi Nandedkar
Electrical and Computer Engineering
NYU Tandon School of Engineering
New York, USA
Vn549@nyu.edu

Abstract—This document here presents a report on the project PageRank that was implemented using Verilog language in VSIM.

Index Terms— Introduction, Architecture, Design Choices, Results, Future scope.

I. INTRODUCTION

The Pagerank algorithm is the cornerstone of web search. The Page Rank algorithm is used in the web search algorithms. When we search for a particular key word the search engine gives us several websites with the first having highest importance.

The importance of the websites depends on the number of citations its is having. The code which we have designed is used to calculate a score. The website which has maximum value of score has the highest importance.

II. (PROJECT ARCHICHECTURE)

This document outlines the methodology adopted to create a verilog program for Pagerank Algorithm which is the cornerstone for web search. This covers the results obtained after NOC logic implementation for 64 websites and gives out the Top 10 websites with the cumulative score for the respective websites.

Following figure (2) shows the basic architecture of our module which includes 4 structures which has 16 websites each. These 16 websites communicate with each other using point to point communication. For the communication between

2 different structures we are using Network on Chip. For this NOC we are having 8 modules, 2 modules for each structure.

Communication Methodology :

Following figure (2) shows the communication mechanism of our module. As shown in the figure whenever a website needs score of another website which is stored in another structure it sends out a signal to the request module. The request module then generates a request packet which is shown in figure (3). For our program the the request packet for first structure is stored in dataInE1 variable. This request packet is given as the input to request router. The request router based on its Round-Robin policy produces an output packet which, for structure 1, is stored in dataOutE1. This output packet is given as the input to the responder module, which generates a data packet as shown in
figure (4). The data packet for structure 1 in our program is stored in dataInE_resp, which is then given as the input to responder router. Responder router then based on its Round-Robin policy provides an output which is stored in dataOutE_resp. The value stored in dataOutE_resp is used in
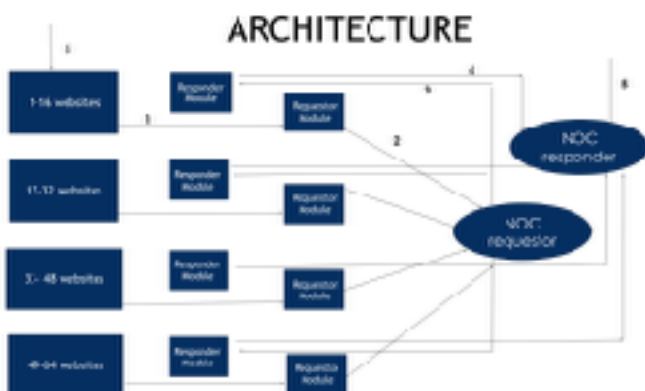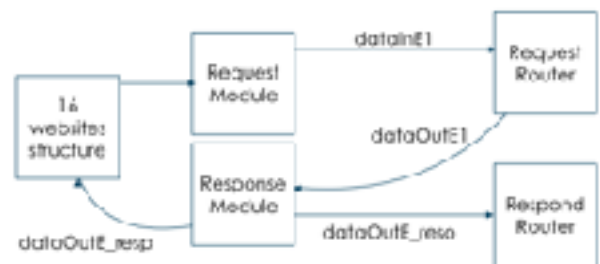
the equation to calculate the scores for all websites. The equation that we have used is :

Packet Formats :

The figures shown below depicts the packet format for our code. as shown in figure 3, the packet format for requests , the 0th bit is a valid bit. bits 2 and 1 depicts destination

structure of the packet. Bits 4 and 3 show the source structure of request and bits 6 and 5 are register Ids. The bits from 20 to 7 are empty bits.
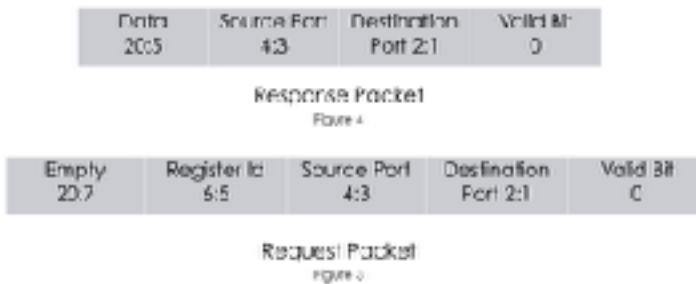
Data 20:5	Source Port 4:3	Destination Port 2:1	Valid Bit 0

Response Packet

Figure 4

Empty 20:7	Register Id 6:5	Source Port 4:3	Destination Port 2:1	Valid Bit 0

Request Packet

Figure 3

Figure 4 shows the packet format for Respond packet. It has data in the bits from 20 to 5. Bits 4 and 3 show the source port for responder packet and bits 2 and 1 show the destination port for the responder packet. The 0th bit is a valid bit.

III. DESIGN CHOICES

While designing the page rank we had to make some design choices. First we decided on number of structures we used. We decided to use four structures of 16 websites each. Then we decided to use one router for the communication in between different structures. This choice of structure was not good as it has a major drawback of deadlock situation to occur in the routers. To overcome this failure, we opted for a design that has two routers, wherein each router would serve the purpose for individual process. One of the router is to request the

data from request module and the other to accept the data that has to be sent to respective destination.

Another design change we adopted is for responder packet where the baseline is of 16 bit format. In order to match the design we have used in implementation, we enhanced the data packet length to 21 bits. The reason for this change is that the data bit is 16 bit format and when concatenated with source and destination bits along with a valid bit constitutes for this length.

Considering the output format we have adopted, the
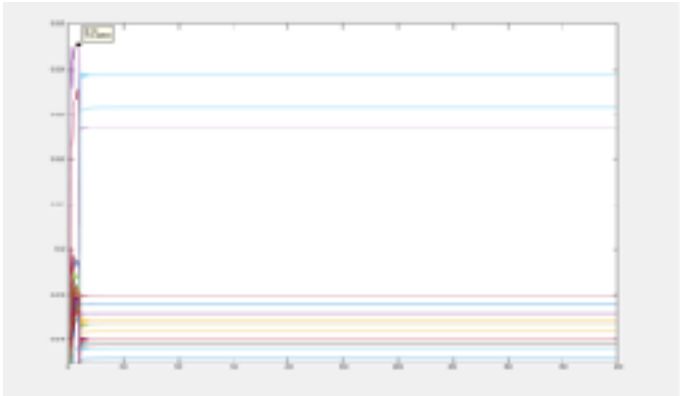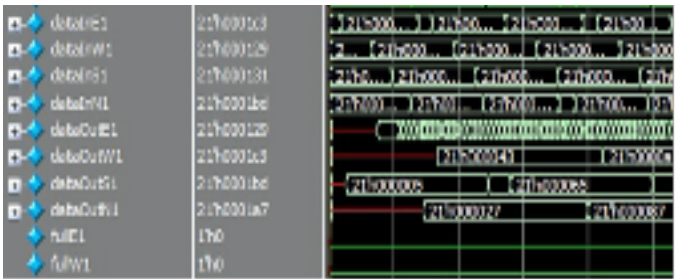
Figure 5

Figure 5

convergence rule we opted, converged the output after 10 iterations. For deciding on the convergence rule we first run the file on MATLAB which gave us the results shown in figure 5. the MATLAB results showed us that the scores converges after 10 iterations.

Another change that we have made was first we had fifo width as 4 when we had data of 16 bits, but then we changed the data width to 21 bits for the reasons stated above. Hence we had to change the fifo depth to 5 bits.

Results

After simulating the code written in VSIM with the test

bench provided, we got following simulation result. The signals used for the requester router input are

dataInE1 and for output are dataOutE1. Figure 6 shows that the Request router is working.

The convergence mechanism which we are using should converge the code after 10 iterations. Figure 7 shows that the code is converging after 10 iterations.

The main results we expect are the top 10 values and top 10 Ids of the websites. The top 10 values gives us the
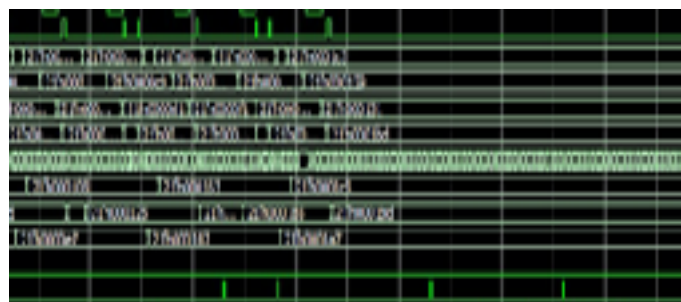

Figure 7

scores of websites having highest importance. Top 10 Ids gives us the Ids of those top 10 web sites. Figure 8 shows us that the module which we have implemented is giving us the top 10 values as well as the top 10 Ids.
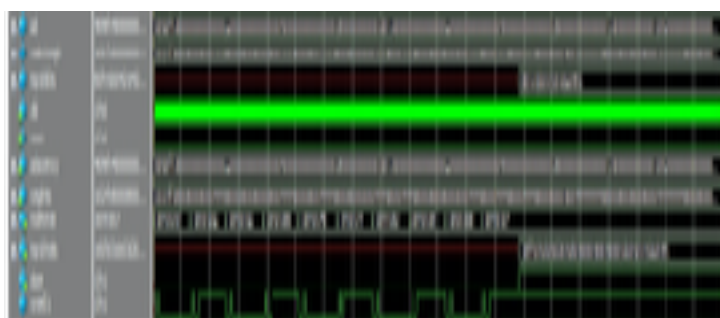

Figure 8

FOR BETTER DESIGN

The design we have implemented has drawbacks that we may overcome through implementations to mitigate those flaws. The option for better design would be to implement it through a better NOC having a better router design that can avoid any kind of dead locks and can handle the complete design in less number of clock cycles.

CONCLUSION

Thus we have implemented a page rank algorithm which gives us the top ten scores and IDs of the most important websites.

ACKNOWLEDGMENT