

SDH Processing of Large Scale Data  
University of South Florida  
CIS 4930 - Summer 2016  
Programming on Massively Parallel Systems

Kevin Wagner (wagnerk1@mail.usf.edu)  
Elijah Malaby (emalaby@mail.usf.edu)  
John Casey (jcasey2@mail.usf.edu)

July 8, 2016

## Contents

## **1 Abstract**

This can be seen as an extension of your Project 2. In particular, you will design memory management mechanisms for GPU global memory, implement CUDA kernels and evaluate the performance of such kernels. Your main task is to write the kernel to compute SDH with the assumption that input data are too large to place in global memory. The straightforward method to handle this is dividing data into small block. Then, load some of data block into GPUs memory and start the computation. However, the efficient way to load block of data in to global memory is required for this project. Because loading data to global memory is a costly operation, you need to reduce the total amount of data transmitted as much as possible. To evaluate your kernel, you may use an analytical model that quantifies the number of access to global memory or empirically run the CUDA program and plot the running time of your kernel under different data sizes.[1]

## **2 Introduction**

## **3 Strategies**

## **4 Code Examples**

## **5 Conclusion**

## **6 References**