

# Sparse Matrix Transposition for GPUs

Massimiliano Incudini - VR433300

Michele Penzo - VR439232

**Sommario**—L'obiettivo principale di questo progetto è stato quello di implementare alcune metodologie proposte per effettuare *Sparse Matrix Transposition* su *Gpu*. Sono stati analizzati alcuni algoritmi, descritti in sezione II, partendo dall'algoritmo seriale, passando a cuSPARSE per finire con l'implementazione degli algoritmi descritti in [1]. Infine vengono esposti i risultati e le conclusioni tratte.

## I. INTRODUZIONE

### II. METODOLOGIE ANALIZZATE

In questa sezione vengono spiegate ed evidenziate le differenze tra le varie metodologie analizzate.

#### A. Trasposta seriale

#### B. Nvidia cuSPARSE

Questo toolkit è implementato all'interno nelle librerie NVIDIA CUDA runtime. Le routine delle librerie vengono utilizzate per le operazioni tra vettori e matrici che sono rappresentate tramite diversi formati. Inoltre mette a disposizione operazioni che permettono la conversione attraverso diverse rappresentazioni di matrici, ed inoltre la compressione in formato *csr* che è una delle più usate quando si vuole rappresentare matrici sparse in modo efficiente.

Il codice è stato sviluppato basandosi su due versioni di cuSPARSE a causa delle Gpu utilizzate. In fase di compilazione viene quindi controllata la versione usata: 9 o 10.

Nel caso in cui la versione usata sia la 10 vengono svolti alcuni ulteriori passi, viene effettuata l'allocazione dello spazio necessario e del buffer per il calcolo della trasposta. Per quanto riguarda la versione 9 invece questi passi non sono necessari. Infine viene chiamata la procedura **cusparsesCsr2cscEx2** che effettua il calcolo della trasposta.

Nello specifico la precedente procedura ...

#### C. Scan Trans

#### D. Merge Trans

## III. RISULTATI

## IV. CONCLUSIONI

### RIFERIMENTI BIBLIOGRAFICI

- [1] K. H. W.-C. F. Hao Wang, Weifeng Liu, "Parallel transposition of sparse data structures," *ICS '16*, 2016.