— Title & problem statement: (a) Implement parallel reduction using min, max, sum & average operators. (b) write a CUDA program that given an N element vector find the maximum, minimum elements as well as arithmatic mean & standard deviation.

— Objectives: - To understand parallel reduction operators - to understand vector operations.

— Outcomes: Understood the parallel reduction operations as well as vector operations.

— S/W H/w requirements: OPEN MP (C++ library), g++, Google colab, CUDA. 8 GB RAM.

— Theory:
(a) CUDA (computer Unified Device Architecture) is a parallel computing platform and application programming interface model created by NVIDIA. It allows software developers & engineers to use CUDA enabled graphics processing unit for general purpose processing an approach termed GPU. Languages c, c++, fortan can be used with CUDA. This accessibility makes it easier for the specialist in parallel programming to use GPU resources in contast

to prior APIs like direct 3D & openGL which required advanced skills in graphic programming. CUDA also supports programming frameworks

(b) Min-Max operations:

(i) Max method returns the er largest element out of a, b compares function can be omitted By default the compare functions is used to determine which object is larger in case they are non-numeric otherwise the operator is used

Syntax : max(object-type a, object-type b, compare())

(ii) Min method: Returns smaller element of out a & b Same rule applies for comparison as max fun.

(iii) Arithmatic mean : This value is found by taking the sum of all individual data elements and then dividing this sum by the total no. of individual data elements

(iv) Standard Deviation: represented by the greek letter σ (sigma), standard deviation is a measure that is used to quantify the amount of variation or dispertion of a set of data values. This number is used to determine how measurements for a group of data are spread out from mean value on either side.

- **Parallel Reduction:** Reduction operations are those which reduce a single collection of values to a single value. Operations which are associative & commontetive can be reduction operations.

- Some of them are : addition, manipulation multiplication, bitwise AND, OR. XOR, logical AND, OR, XOR finding minimum/maximum amongst a given set of numbers.

- computation complexity likely to be (ogcn)

- below is an example of sum of an array using parallel reduction.

13	27		15	14		33	2		24	6

⓪   ①   ②   ③

40		29			35		30	

⓪   ①

69

65

⓪

134

↑
sum
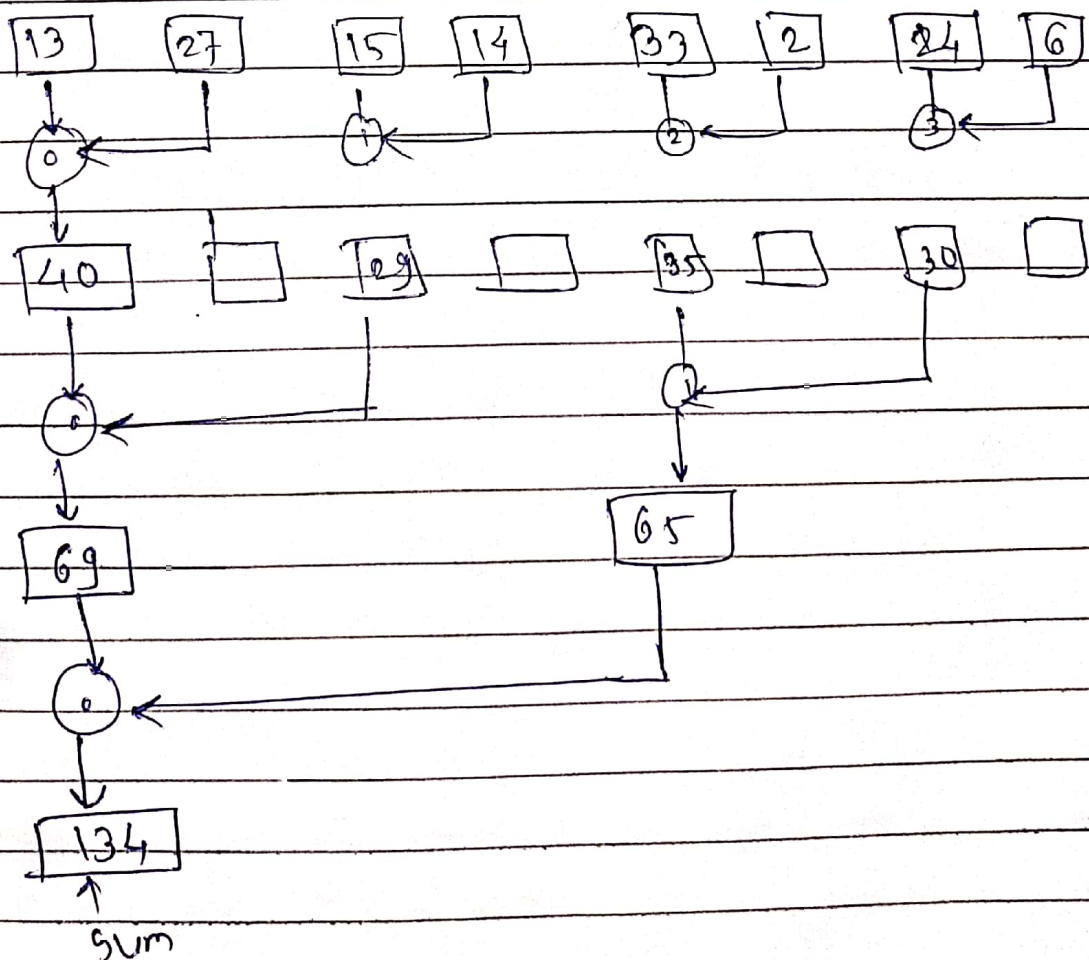
- conclusion: I have studied parallel reduction using min, max, avg, sum & CUDA program that given an N element array finds max, min, mean, SD parallelly & serially.

Both programs executed successfully & gave expected results.