Course Code		Course Name	GPU PROGRAMMING	Course Category	O	Open Elective	L	T	P	C
							3	0	2	4

Pre-requisite Courses	Nil	Co-requisite Courses	Nil	Progressive Courses	
Course Offering Department	Computing Technology	Data Book / Codes/Standards	Nil		

Course Learning Outcomes (CLO):		Learning			Program Learning Outcomes (PLO)														
		1	2	3	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
		Level of Thinking (Bloom)	Expected Proficiency (%)	Expected Attainment (%)	Engineering Knowledge	Problem Analysis	Design & Development	Analysis, Design, Research	Modern Tool Usage	Society & Culture	Environment & Sustainability	Ethics	Individual & Team Work	Communication	Project Mgt. & Finance	Life Long Learning	PSO - 1	PSO - 2	PSO – 3
CLR-1 :	Understand the difference between different parallel programming architectures.	3	80	70	H	H	H	H	M	L	L	L	M	M	L	H	H	H	H
CLR-2 :	Knowledge of GPU aware programming using CUDA and OpenAcc frameworks.	3	85	75	H	H	H	H	M	L	L	L	M	M	L	H	H	H	H
CLR-3 :	Design and develop GPU accelerated real-world simulations and applications	3	75	70	H	H	H	H	M	L	L	L	M	M	L	H	H	H	H
CLR-4 :	To understand basic GPU processor & memory architectures	3	85	80	H	H	H	H	M	L	L	L	M	M	L	H	H	H	H
CLR-5 :	To evaluate the performance analysis, host- and device-side memory and _optimization	3	85	75	H	H	H	H	M	L	L	L	M	M	L	H	H	H	H

Duration (hour)		15	15	15	15	15
S-1	SLO-1	*Introduction – Traditional Computing*	GPU Programming	CONSTANT MEMORY	Directive-Based Programming	OpenACC Accelerator Model
	SLO-2	*Heterogeneous Parallel Computing*	CUDA's Programming Model	Constant memory caching	OpenACC, OpenAcc Directives	Openacc Data directives
S-2	SLO-1	*GPU Architecture*	Threads	Constant memory broadcast	OpenACC Versus CUDA C	Implied data regions
	SLO-2	*Multi-node Computing*	Blocks	TEXTURE AND SURFACE MEMORY	Execution Model	unstructured data Directives
S-3	SLO-1	*GPU Computing*	Grids	GLOBAL MEMORY	OpenAcc portability	Data synchronization
	SLO-2	*GPGPU,GPU Components*	CUDA'S EXECUTION MODEL: STREAMING MULTIPROCESSORS AND WARPS	Global memory sorting	Memory Model	Parallel Jacobi
S-4	SLO-1	*CPUs vs GPUs*	CUDA COMPILATION PROCESS	OPTIMIZATION TECHNIQUES	Basic OpenACC Programs	Loop Optimization
	SLO-2	*Parallel Programming*	PUTTING TOGETHER A CUDA	BLOCK AND GRID DESIGN	Profiling OpenAcc Code	Parallelizing loops

		Languages, Models, Directives and libraries	PROJECT A PRACTICAL EXAMPLE-HISTOGRAMS			
S-5	SLO-1	*Evolution of Graphics Pipelines*	Memory Handling with CUDA	KERNEL STRUCTURE	Speedup	Loop correctness

	SLO-2	*The logical 3D graphics pipeline*	CACHES REGISTER USAGE	SHARED MEMORY ACCESS	Parallel Construct	collapse clause
S-6	SLO-1	*Speedups Serial/Parallel Code and problems*	MEMORY HIERARCHY	GLOBAL MEMORY ACCESS	Loop Construct	Tile clause
	SLO-2	*Concurrency, Flynn's Taxonomy*	LOCAL MEMORY/REGISTERS	PAGE-LOCKED AND ZERO-COPY MEMORY	Kernels Construct	Gang worker vector
S-7	SLO-1	*Types of Parallelism*	SHARED MEMORY	UNIFIED MEMORY	Data Regions	Gang worker vector
	SLO-2	*Common Parallel Pattern*	Sorting using shared memory	ASYNCHRONOUS EXECUTION AND STREAMS	Data clauses	Asynchronous Programming
S-8	SLO-1	*NVIDIA GPU architecture*	Radix sort	DYNAMIC PARALLELISM	Data Management	Asynchronous Programming
	SLO-2	*Computational Model*	Merging lists	DEBUGGING CUDA PROGRAMS	Data Dependencies	Interoperability
S-9	SLO-1	*GPU's Multi-Threaded Multi-processor*	Parallel merging	PROFILING CUDA PROGRAMS	Jacobi iteration	Sharing data with CUDA
	SLO-2	*A case study: sorting in OpenMP*	Parallel reduction	*CASE STUDY*	Asynchronous Computation and Data Transfer	Sharing CuDA Data

Learning Resources	1. Shane Cook, CUDA Programming: —A Developer's Guide to Parallel Computing with GPUs (Applications of GPU Computing), First Edition, Morgan Kaufmann, 2012 2. David B. Kirk and Wen-mei W. Hwu, " Programming Massively Parallel Processors A Hands-on Approach" Second Edition, Elsevier Inc,2016 3. Prof. Stewart Weiss ,"GPUs and GPU Programming " Lecture Notes	4. Gerassimos Barlas , "Multicore and GPU Programming An Integrated Approach ",, Elsevier Inc,2015 5. Openacc.org,"OpenACC Programming and Best Practices Guide", 2021 6. *https://developer.nvidia.com/openacc*

Learning Assessment											
	Bloom's Level of Thinking	Continuous Learning Assessment (50% weightage)								Final Examination (50% weightage)	
		CLA – 1 (10%)		CLA – 2 (15%)		CLA – 3 (15%)		CLA – 4 (10%)#			
		Theory	Practice	Theory	Practice	Theory	Practice	Theory	Practice	Theory	Practice
Level 1	Remember Understand	*20%*	*20%*	*15%*	*15%*	*15%*	*15%*	*15%*	*15%*	*15%*	*15%*
Level 2	Apply Analyze	*20%*	*20%*	*20%*	*20%*	*20%*	*20%*	*20%*	*20%*	*20%*	*20%*
Level 3	Evaluate Create	*10%*	*10%*	*15%*	*15%*	*15%*	*15%*	*15%*	*15%*	*15%*	*15%*
	Total	100 %		100 %		100 %		100 %		-	

CLA – 4 can be from any combination of these: Assignments, Seminars, Tech Talks, Mini-Projects, Case-Studies, Self-Study, MOOCs, Certifications, Conf. Paper etc.,

Course Designers		
Experts from Industry	Experts from Higher Technical Institutions	Internal Experts
1. Mr.Hemant, NVIDIA		*1. Dr.S.Nagadevi,Assistant Professor,SRMIST*
		2. Dr.R.Vidhya, Assistant Professor,SRMIST