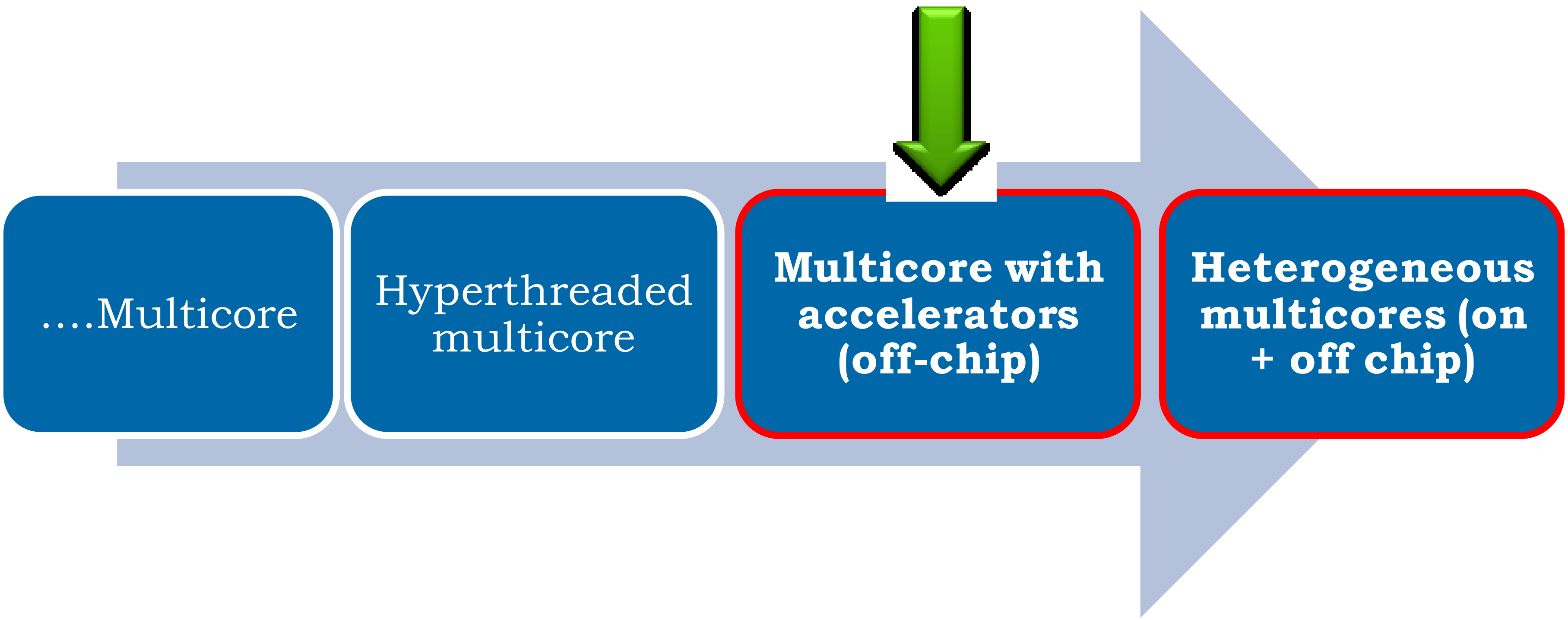
GViM: GPU-accelerated Virtual Machines

Vishakha Gupta, Ada Gavrilovska, Karsten Schwan, Harshvardhan Kharche @ Georgia Tech

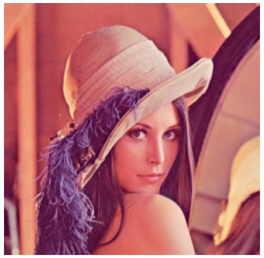
Niraj Tolia, Vanish Talwar, Partha Ranganathan @ HP Labs

Trends in Processor Hardware



Possible Application Domains

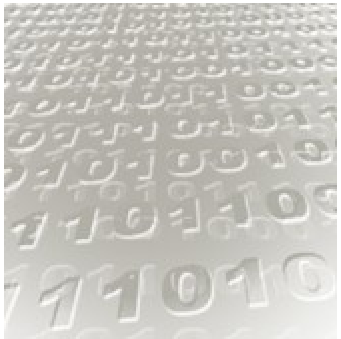
Media and image processing:
- **Snapfish-like image processing suite**
- Image compression/decompression
- Denoising
- Filtering
- Transforms



Scientific computing, Visualization:
- Electromagnetic simulation
- Ray tracing

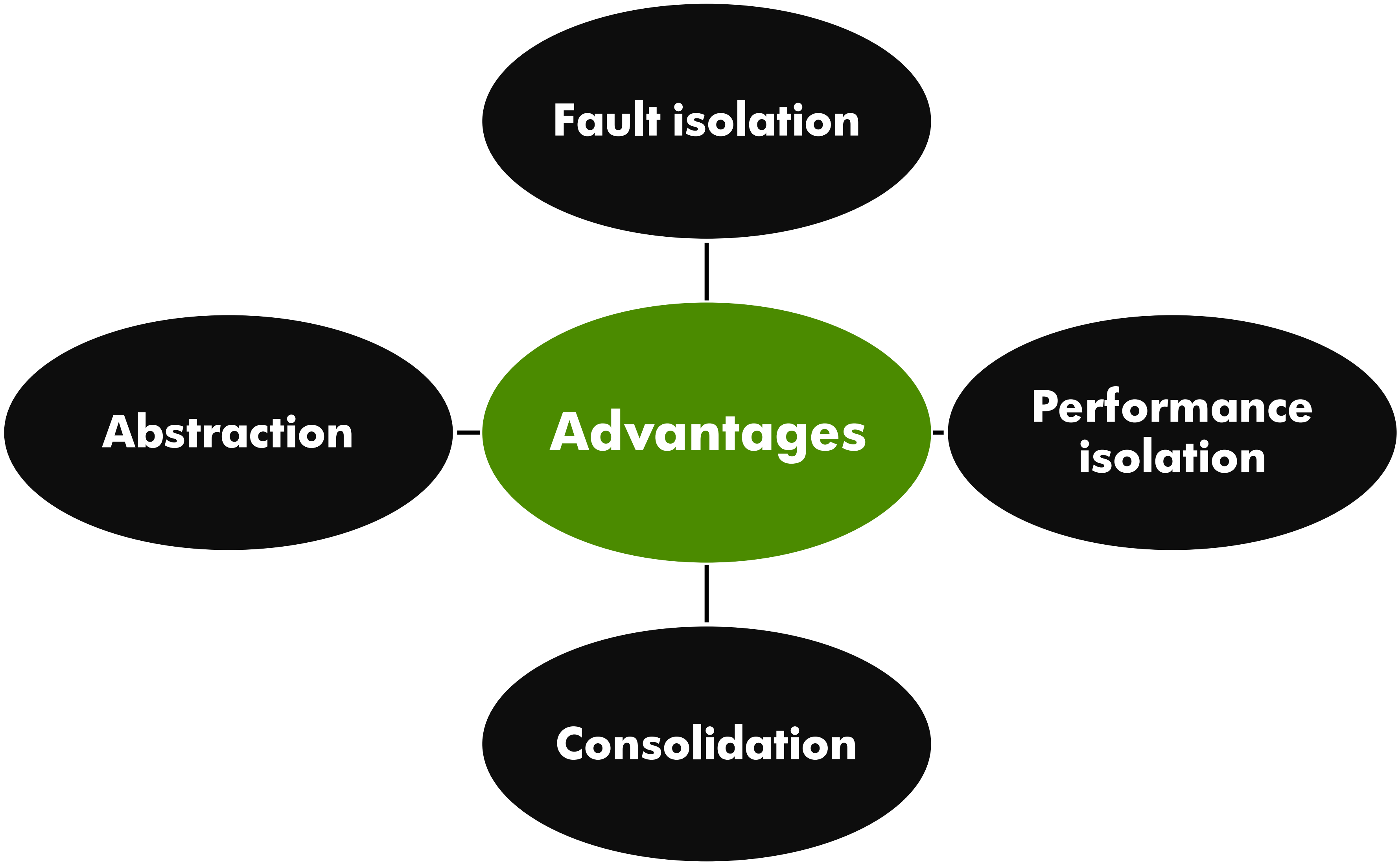
Financial analysis:
- Black Scholes
- **Stock Portfolio Risk Analysis**
- **Credit derivatives processing**
- Random number generators

Linear algebra and others:
- **Dense Matrix Multiplication**
- **AES256 encryption/decryption**
- **Large data sorting**
- **Database operations**

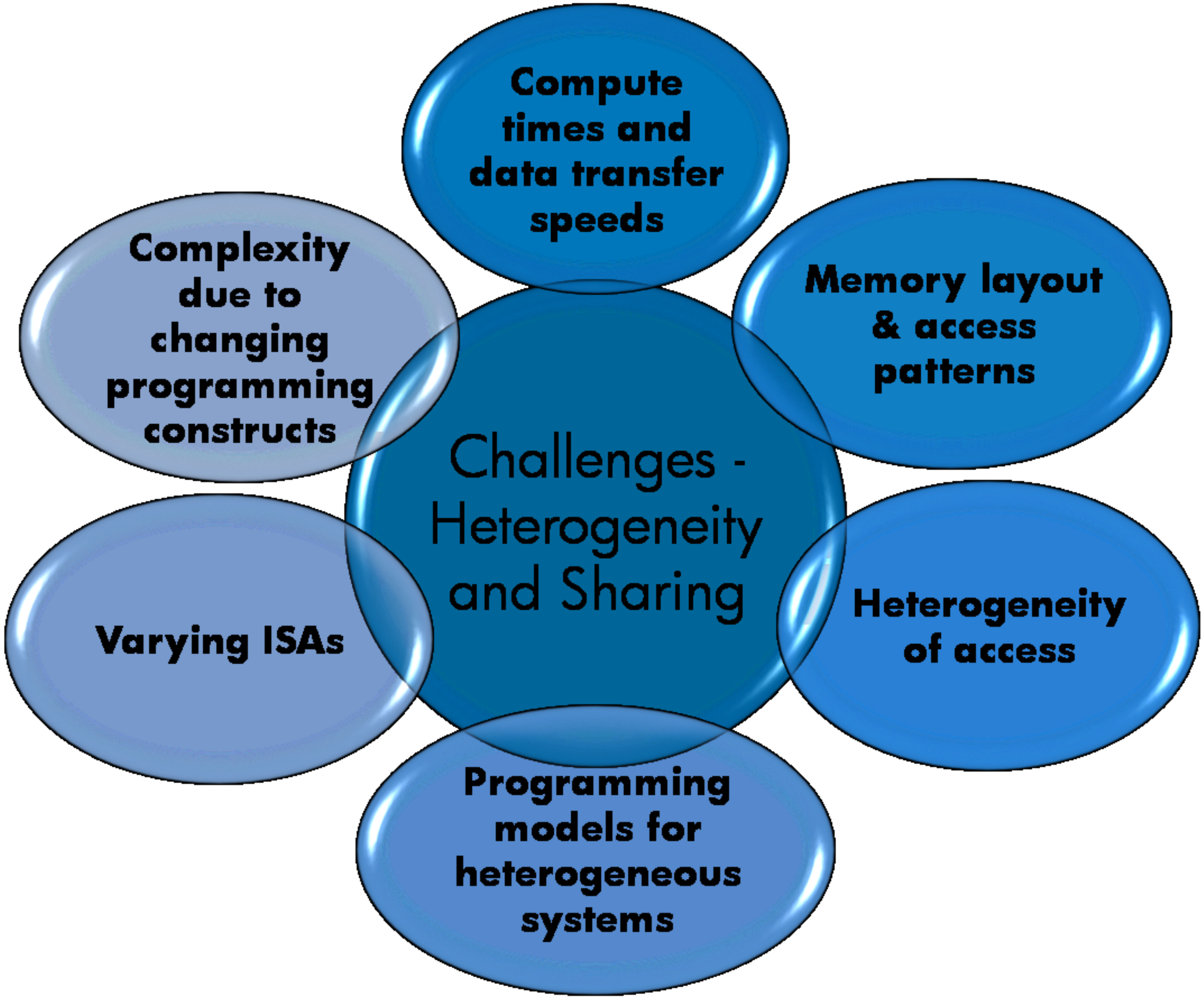


Georgia Tech College of Computing LABS hp

Virtualized Heterogeneous Multicores



Virtualizing Heterogeneous Multicores



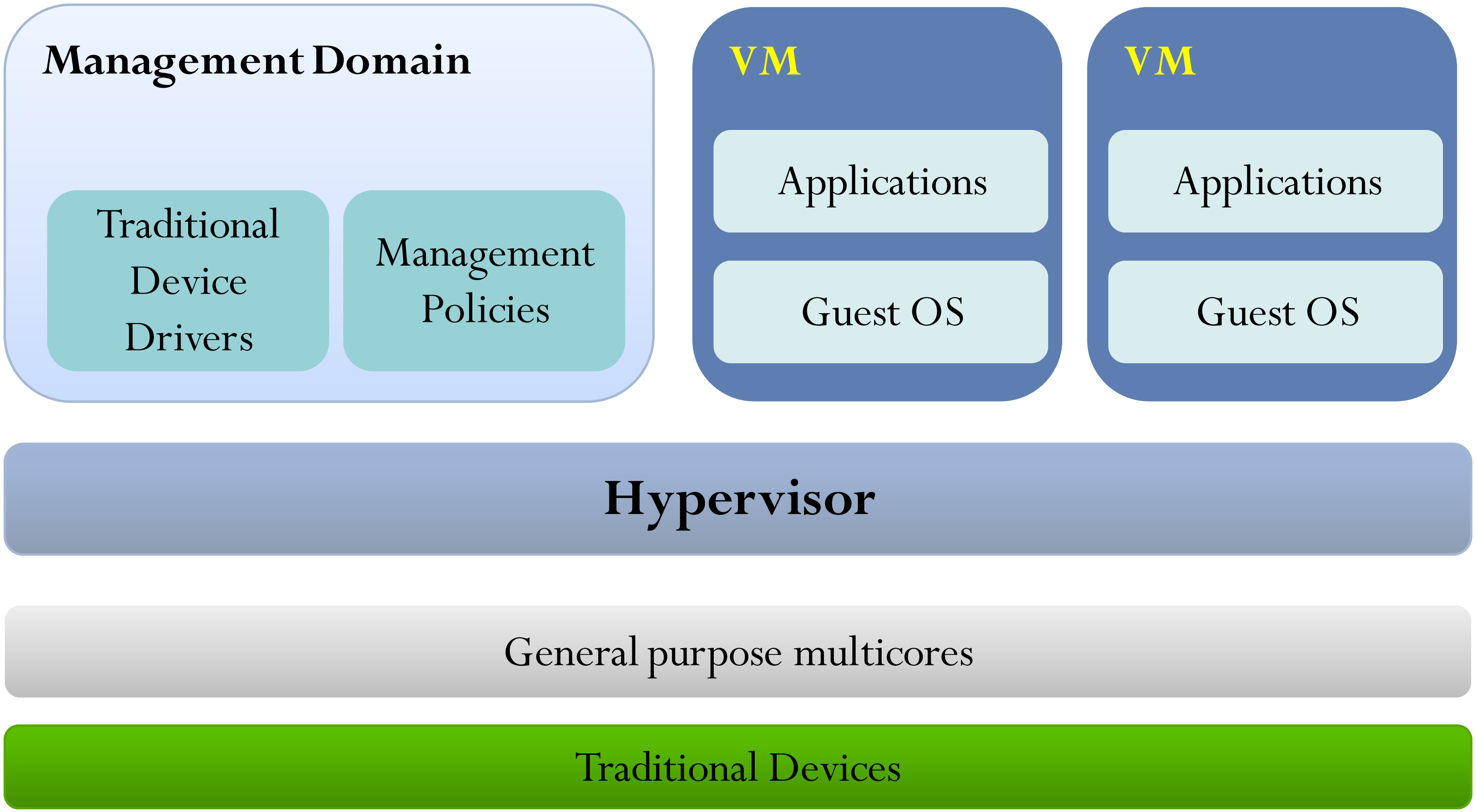
GViM: GPU-accelerated Virtual Machines

Heterogeneity of access →	'Amorphous' images of machine resources to applications
Complexity due to changing programming constructs →	Improved programmability and portability
Memory layout & access patterns →	Efficient accelerator virtualization
Compute times and data transfer speeds →	Coordinated resource management (Future work)
Programming models for heterogeneous systems →	Beyond scope
Varying ISAs →	Beyond scope

Outline

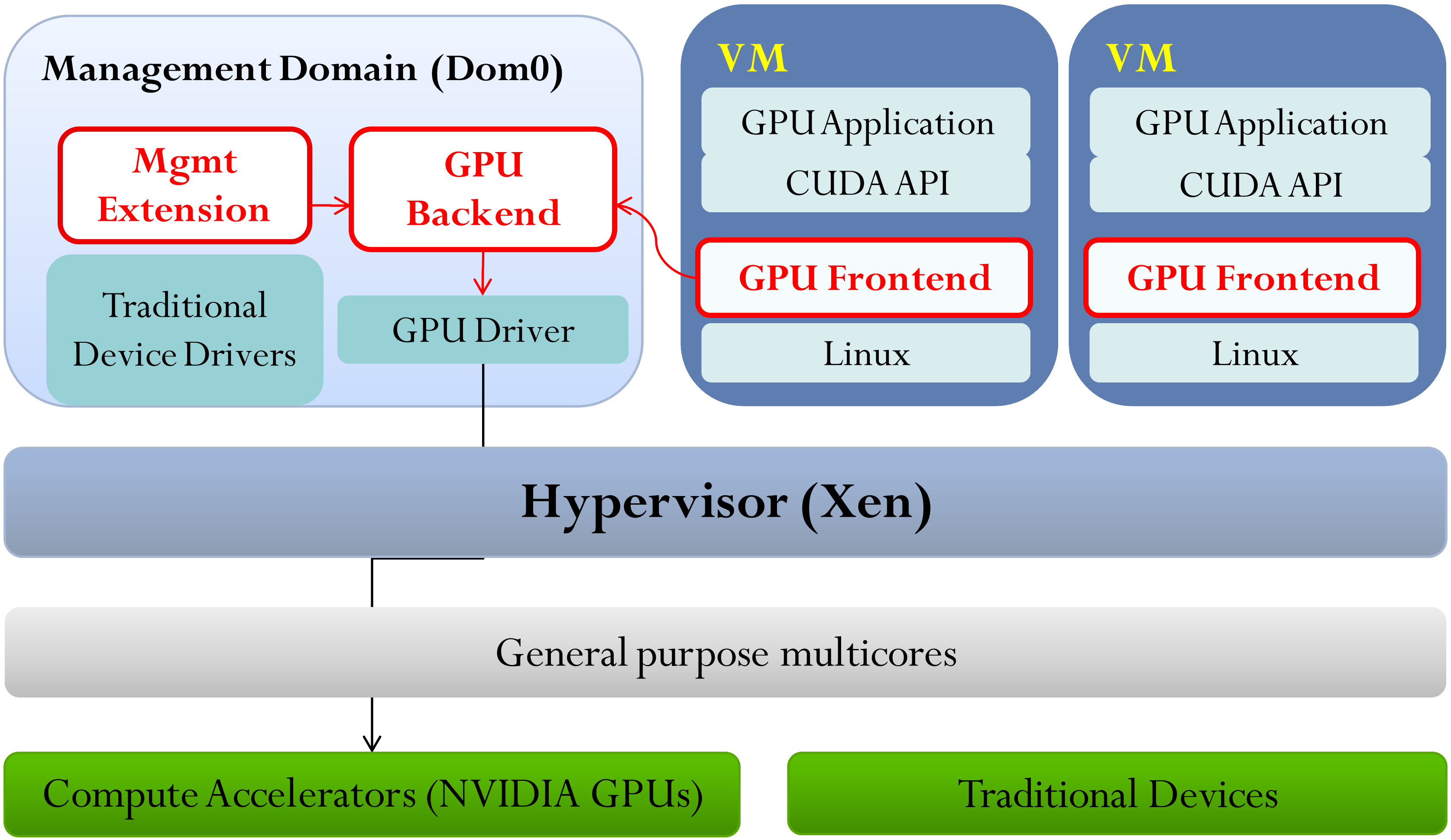
- Motivation
- System architecture
  - Virtualized Homogeneous Multicore Systems
  - Virtualization of Accelerator based Systems
  - GPU Virtualization
- Management extension
- Evaluation
- Related work
- Future work and conclusion

Virtualized Homogeneous Multicore Systems



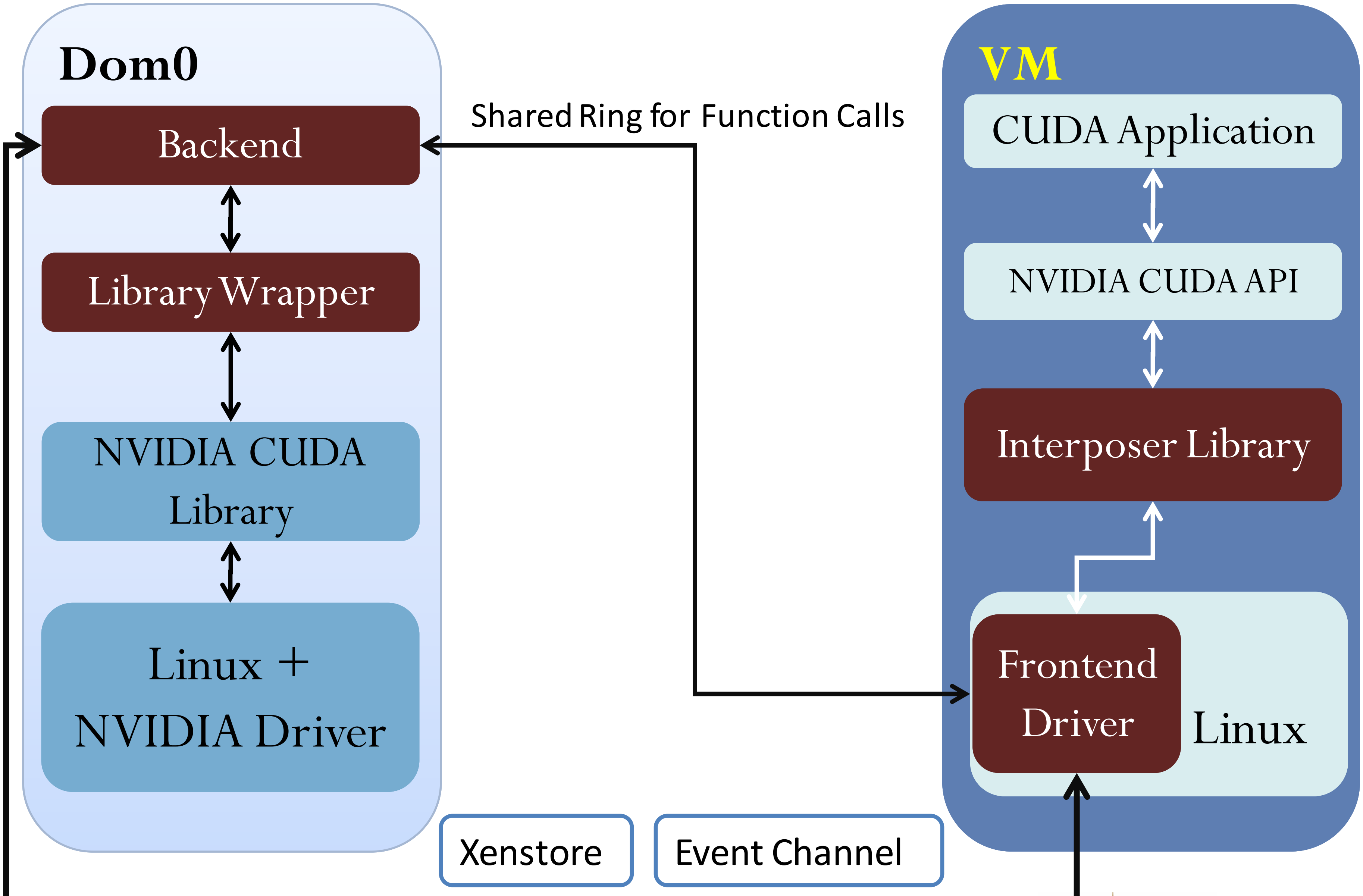
Georgia Tech College of Computing
LABS hp

Virtualization of Accelerator based Systems
Extending Xen for GPU

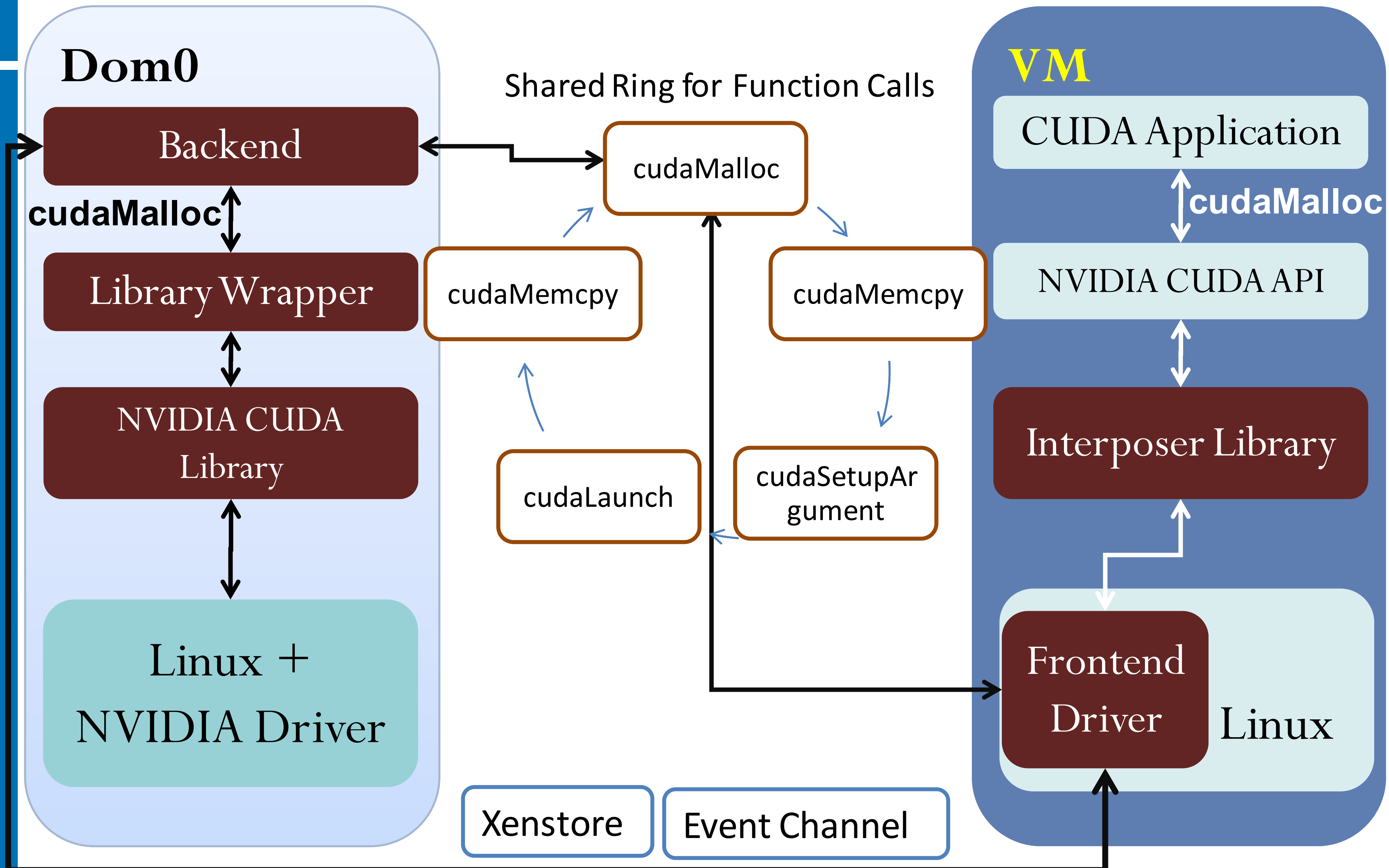


NVIDIA's CUDA – Compute Unified Device Architecture for managing GPUs

GPU Virtualization – Components



GPU Virtualization – Execution Path



Georgia Tech College of Computing | hp LABS

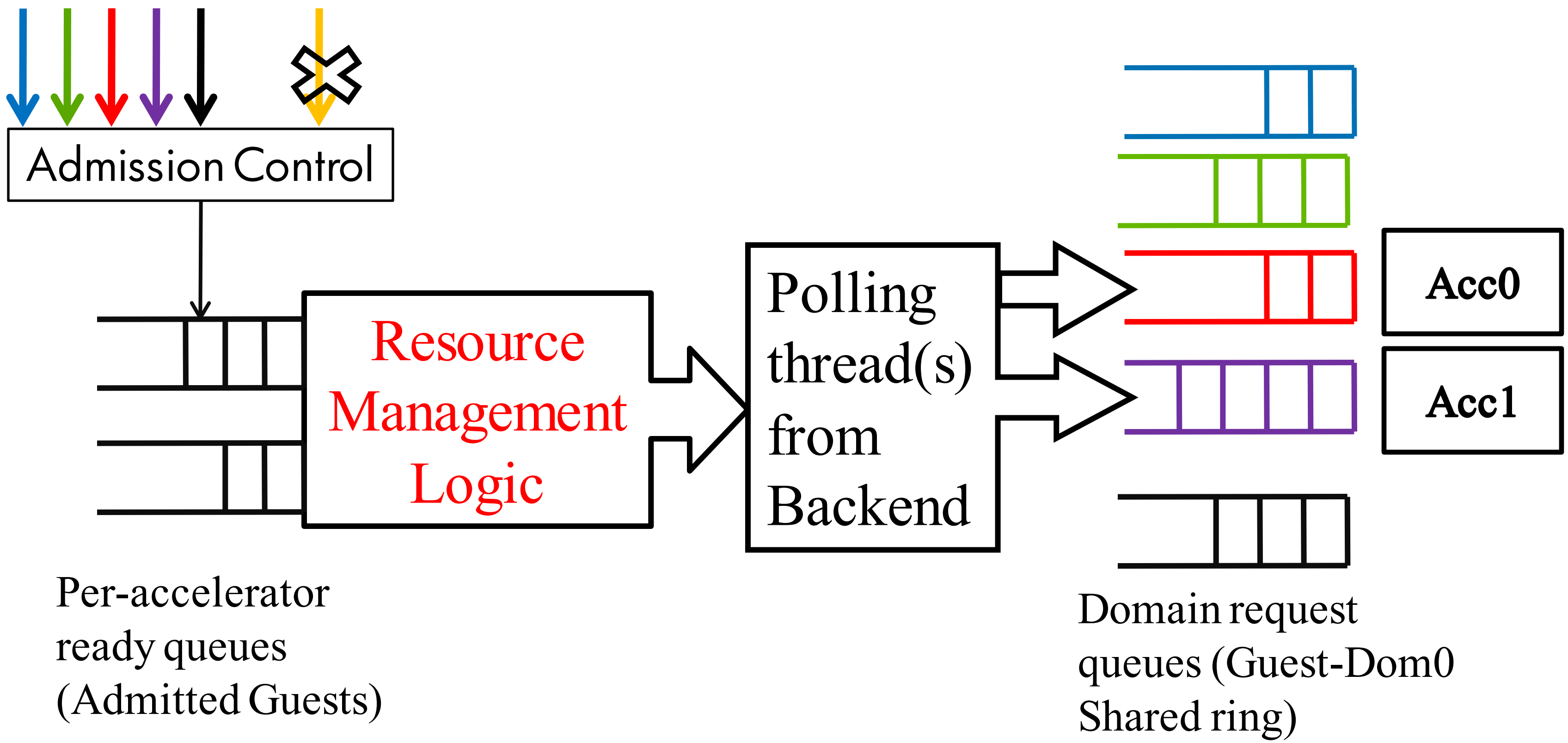
GViM Memory Management

- ‘Closed’ accelerator – GViM’s smart data movement
- In non-virtualized case (application in Dom0)
  - Pageable memory - allocate in Dom0 using malloc
  - Pinned memory - use mapped memory from GPU
- From guest domain to backend
  - 2-copy - allocate in guest using malloc
  - 1-copy - mmap pre-shared memory from Frontend
  - Bypass - mmap the Dom0 pinned memory from Frontend
    - Desirable (future work)
- More details in the paper

Outline

- Motivation
- Tasks for summer
- System architecture
- Management extension
  - Scheduling in Dom0
  - Algorithms
- Evaluation
- Related work
- Future work and conclusion

Scheduling in Dom0 – Management extension



Scheduling Algorithms

- Round robin (RR)
  - Some guest domain chosen every period
  - Poller monitors its queues for the period
  - Moves to the next domain

- XenoCredit
  - Guests get credits assigned to them at boot time
  - Use these credits to calculate proportions
  - Poll domains for time proportional to their credits

Outline

- Motivation
- Tasks for summer
- System architecture
- Management extension
- Evaluation
  - Testbed details
  - Benchmarks
  - Results
  - Discussion
- Related work
- Future work and conclusion

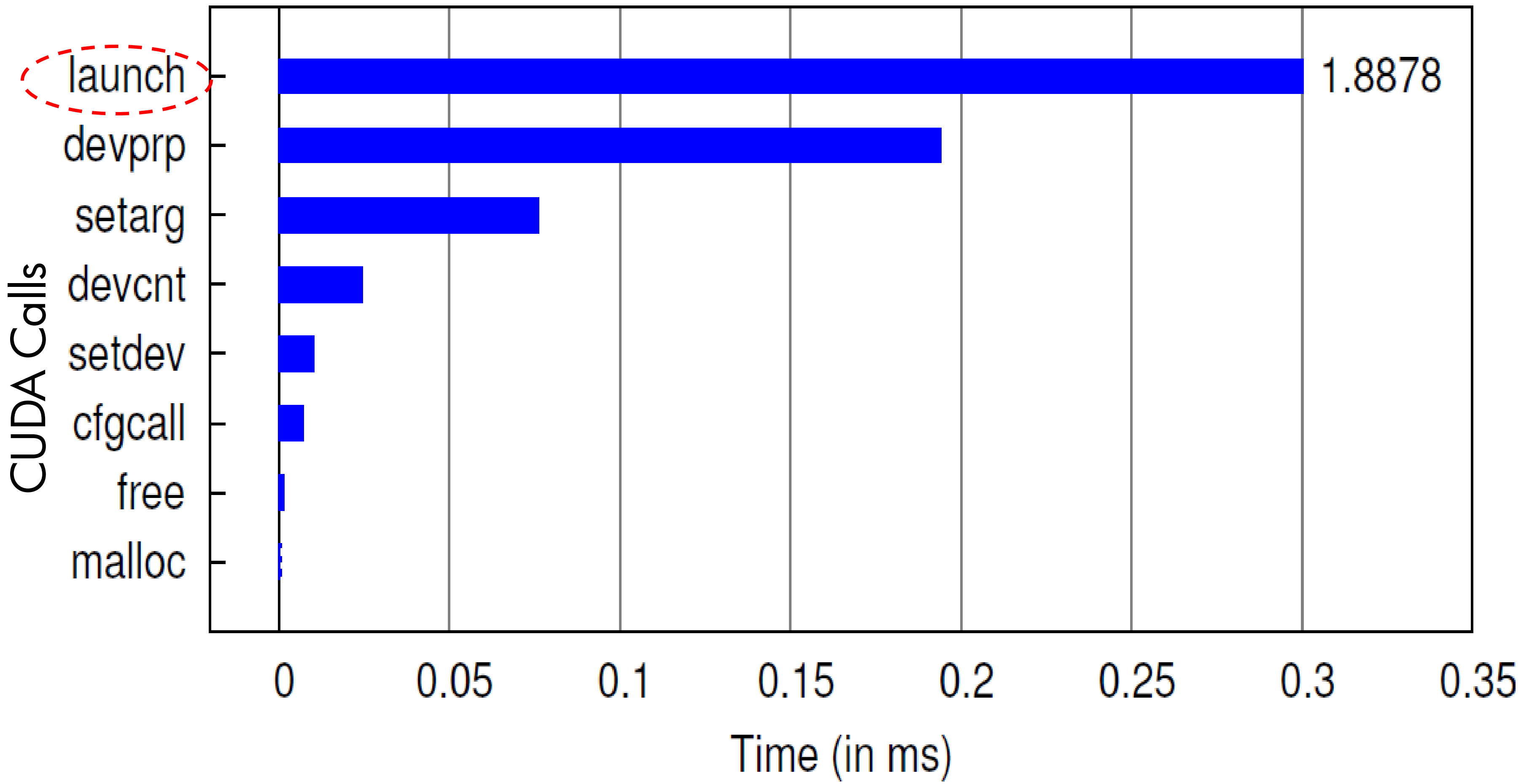
Testbed Details

- Hardware configuration:
  - Xeon quad-core @ 2.5GHz and 2GB memory
  - NVIDIA 8800 GTX PCIe card
  - NVIDIA 9800 GTX PCIe card
- Software configuration
  - Xen 3.2.1 running 2.6.18 Linux kernel
  - CUDA SDK 1.1 with gpu driver 169.09

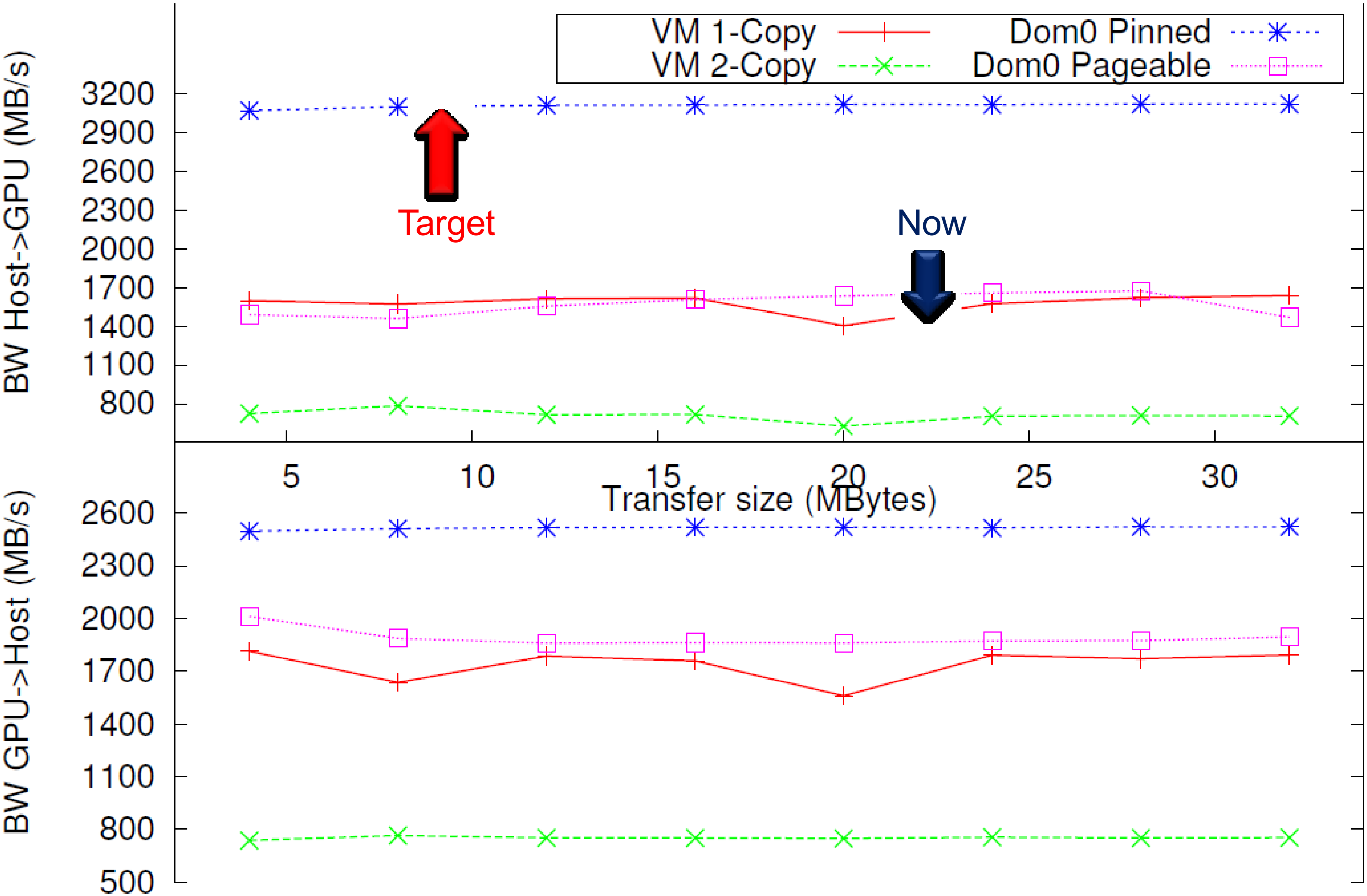
Benchmarks

- Matrix multiplication (MM[2K])
  - Used 2048x2048 floating pt. matrices for the numbers
- BlackScholes (BS[1m,512])
  - Financial algorithm for calculating call and put option prices (1million for testing)
  - Configurable number of iterations
- FastWalshTransform (FWT[64])
  - Class of generalized Fourier transformations
  - 64MB in input and output buffers
- Microbenchmark and bandwidth test from CUDA SDK

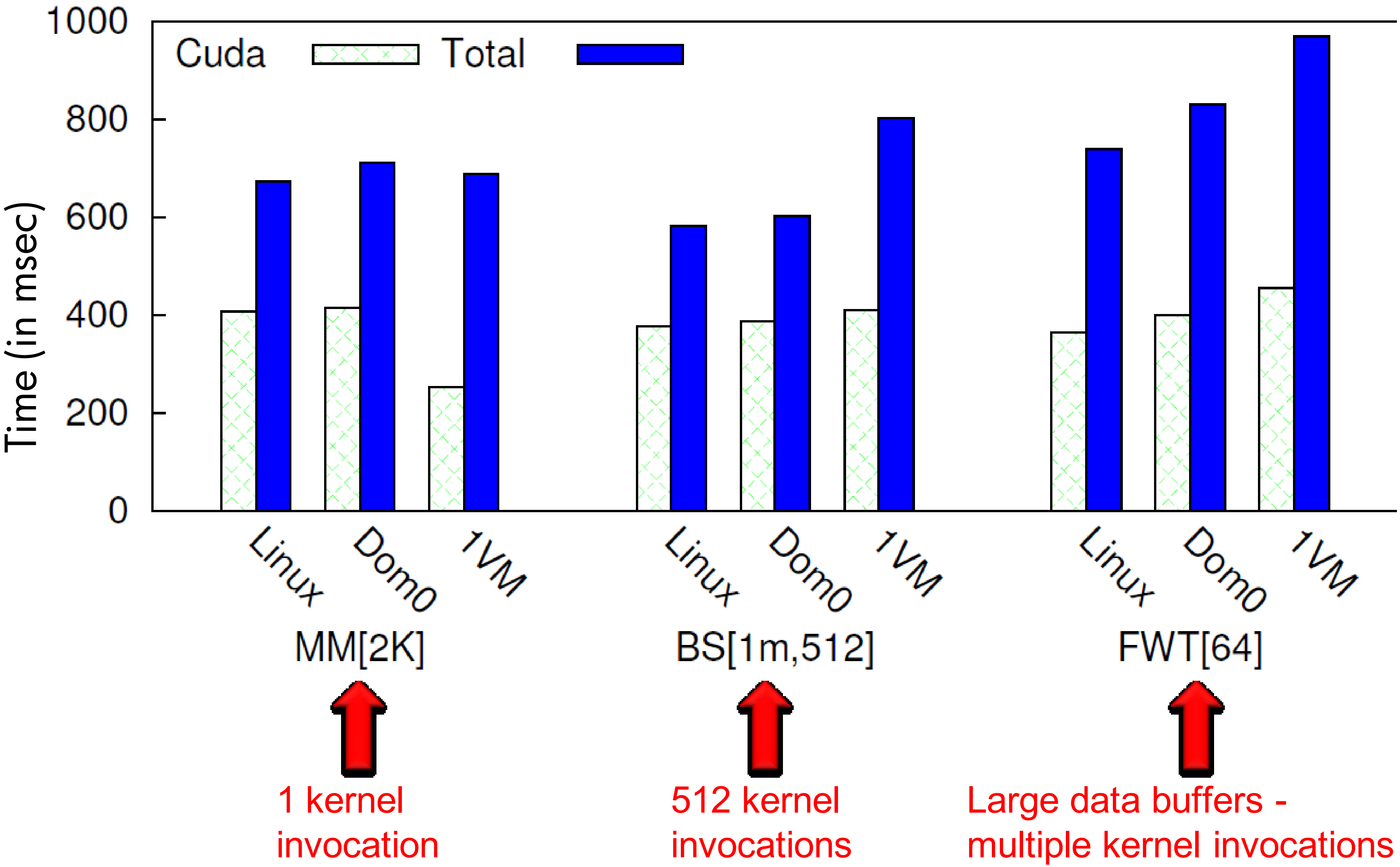
CUDA Calls Virtualization Overhead



Optimized GViM – Increased Bandwidth (1-copy)



Benchmark Evaluation



Georgia Tech College of Computing LABS hp

BlackScholes: XenoCredit vs. Round Robin

- BlackScholes with 60000 options and 4096 iterations
- Host with 2 GPUs and 4 guests

VMs	Credits	Expected iter/ms	RR	XC
VM1	512	1.8	2.632	1.867
VM2	256	2.4	2.567	3.12
VM3	256	2.2	2.804	3.213
VM4	256	2.54	2.901	3.45

Motivates new methods for resource management in accelerator-based systems

Summary

- Trend towards heterogeneous multicores
- Virtualization has advantages
- Efficient GPU virtualization solution
- Proposed scheduling extension
- Evaluation

Ongoing and Future Work

- Ongoing
  - Memory bypass solution
  - Scheduling in Dom0

- Future
  - Heterogeneous multicore scheduling
  - SLA management policies
  - Scalability and stability models/analyses
  - Power-awareness in the scheduler