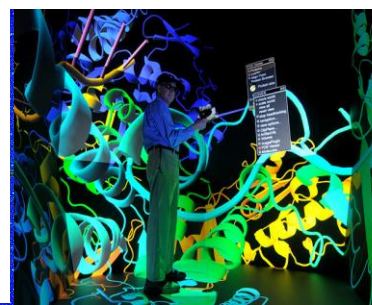
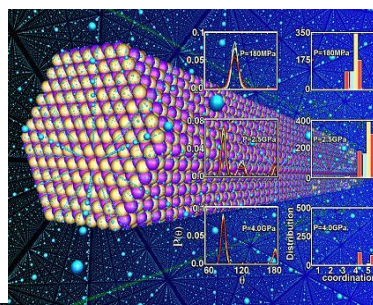
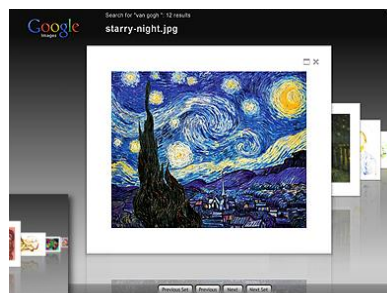


# Gdev: First-Class GPU Resource Management in the Operating System

**Shinpei Kato**  
**Nagoya University**

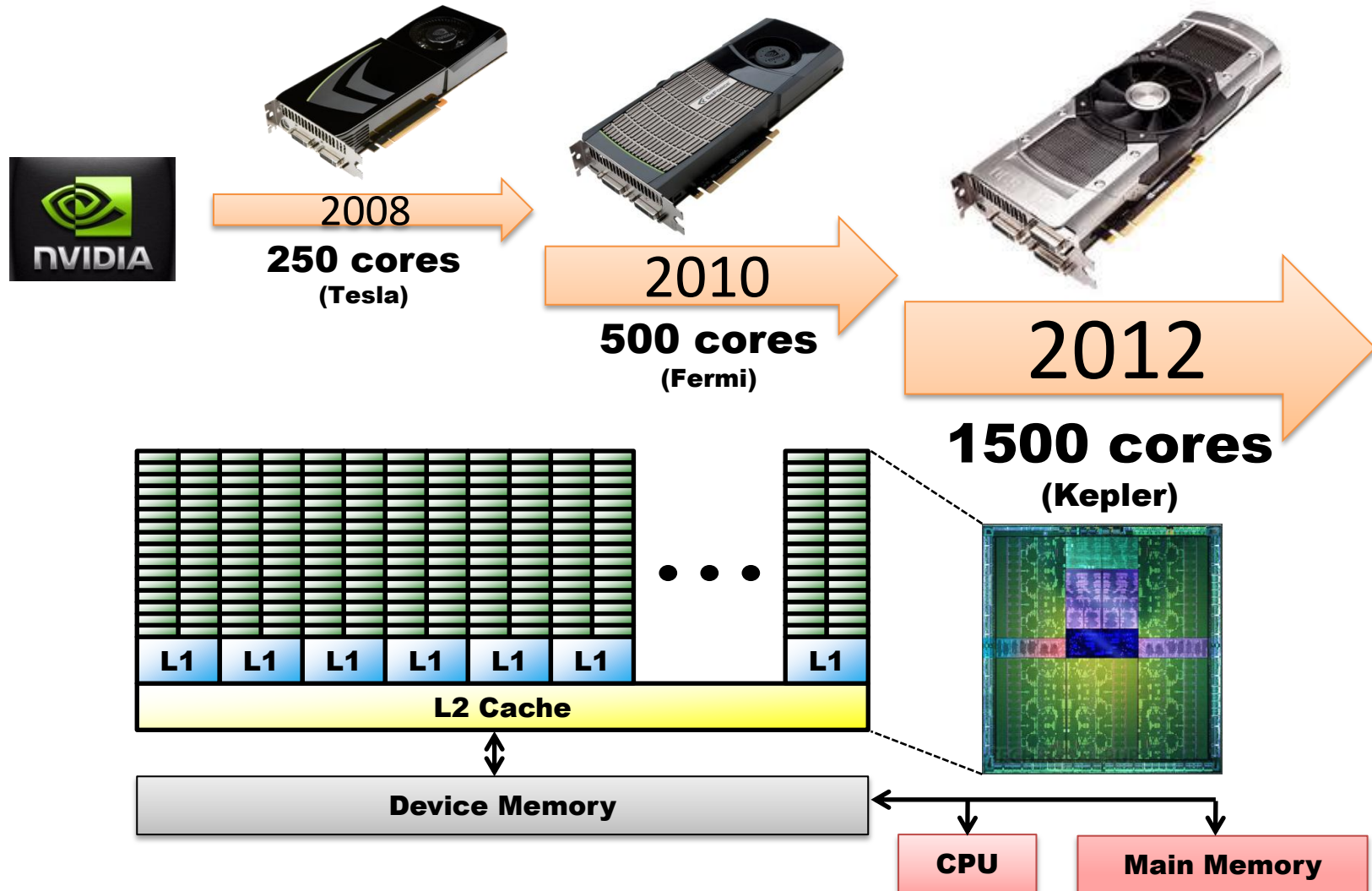
**Michael McThrow Carlos Maltzahn Scott Brandt**  
**UC Santa Cruz**



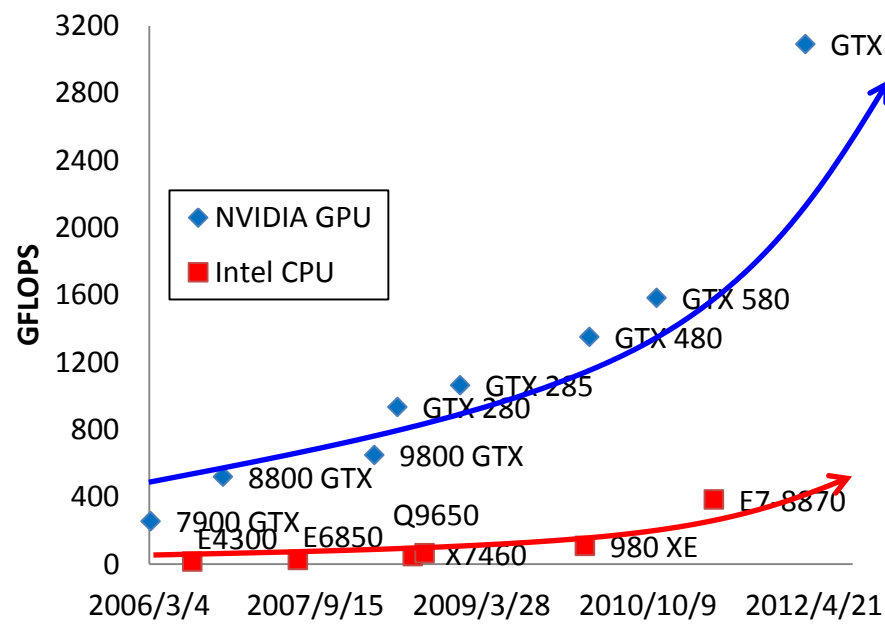
# GPUs embrace “many cores”.



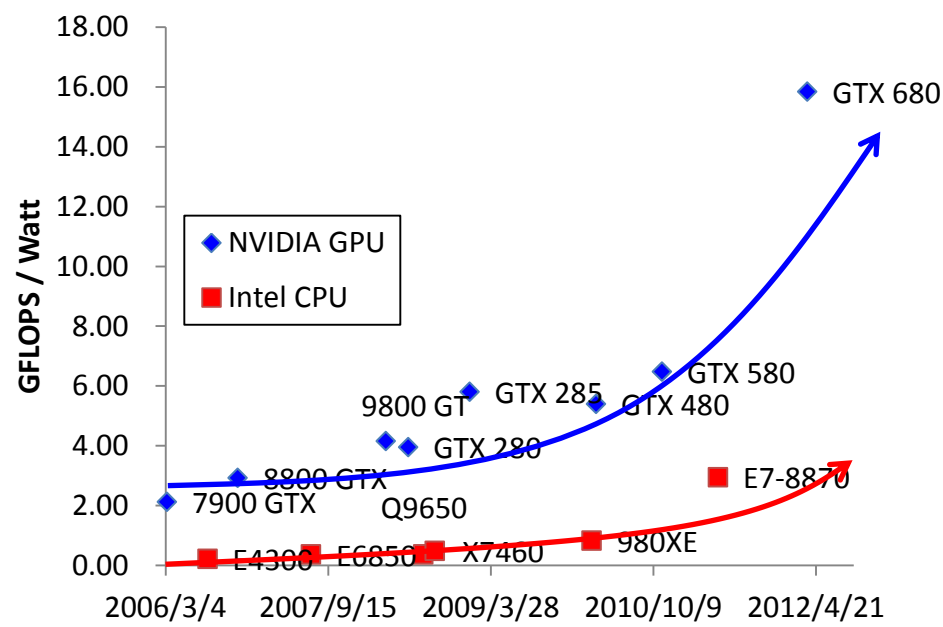
# Graphics Processing Unit (GPU)



# Performance Trend

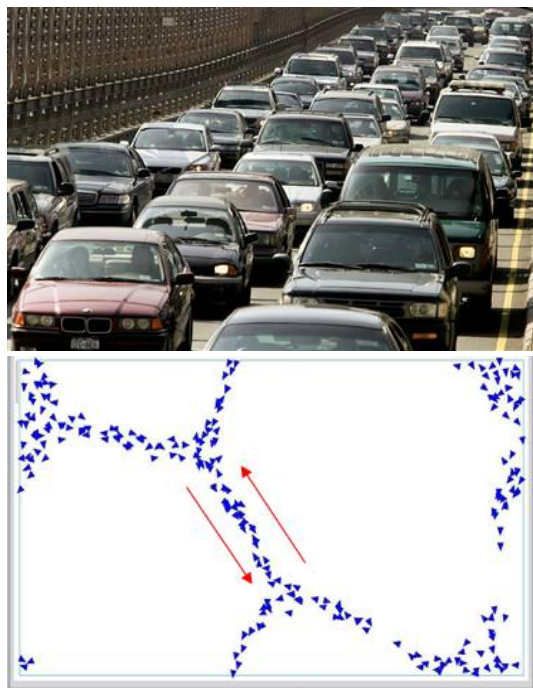


Single Precision Performance

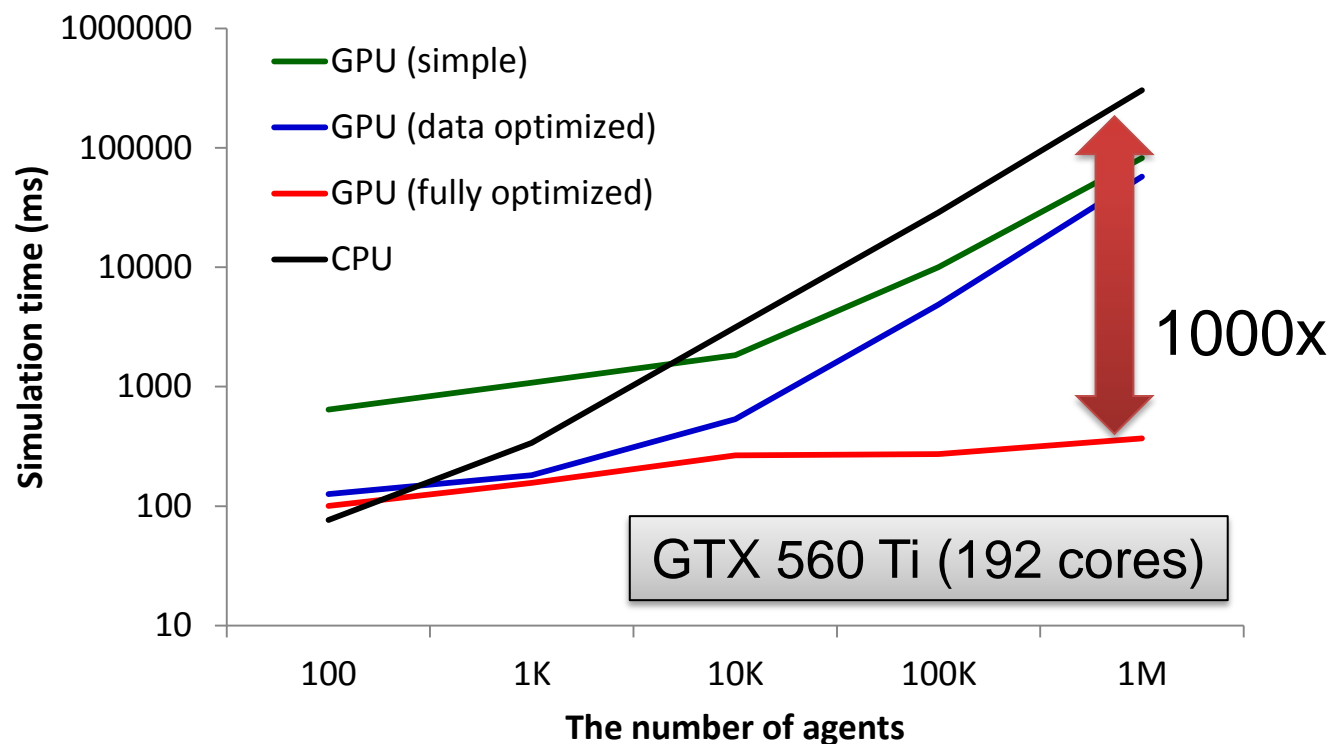


Performance per Watt

# GPUs Suit Science

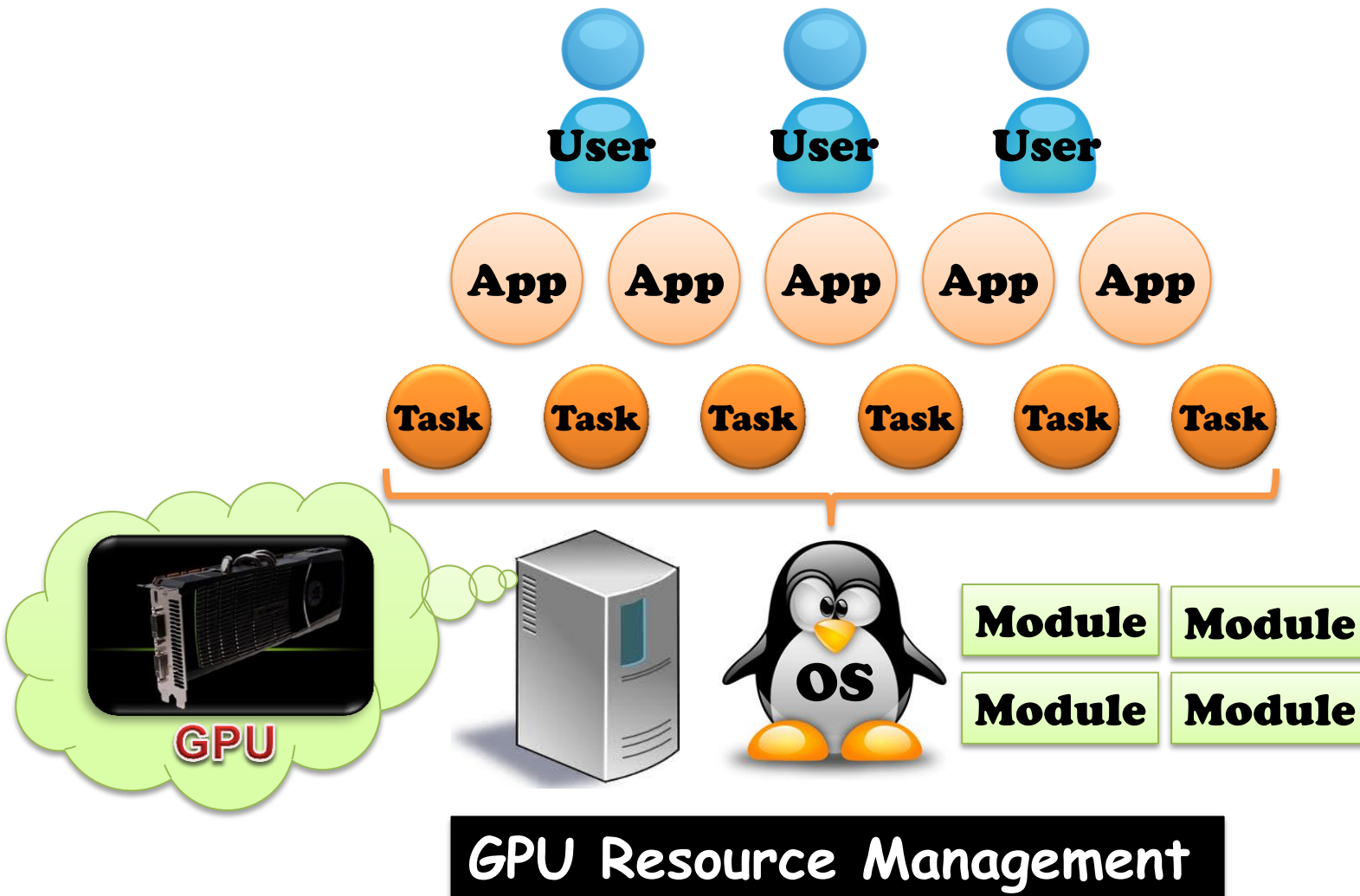


**Traffic Simulation**





# Not Yet “General-Purpose”



# Gdev

- **New approach** to GPU resource management
  - Allows the OS as well as user-space applications to use GPUs.
- **New functions** of GPU resource management
  - Shared device memory (IPC)
  - Data swapping
  - System-level virtualization
- **Open-source implementation**

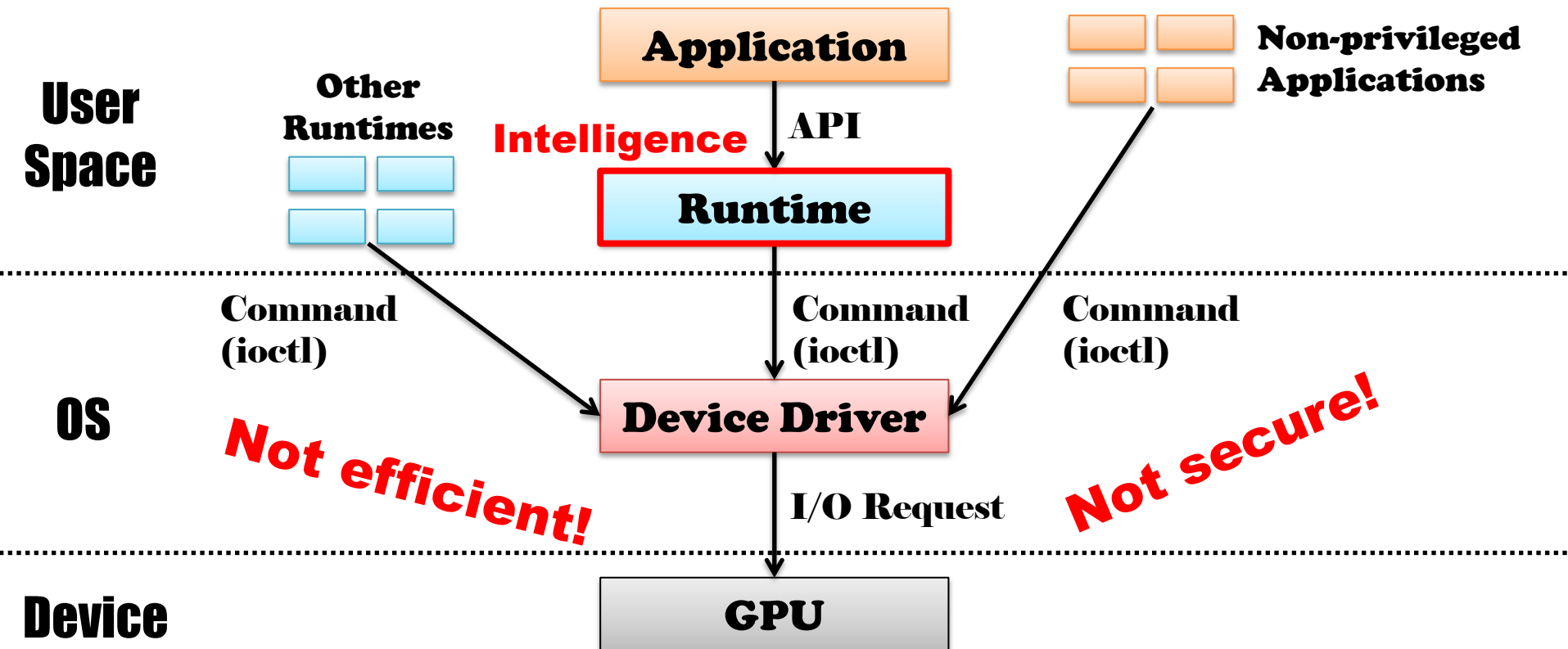
# Outline

---

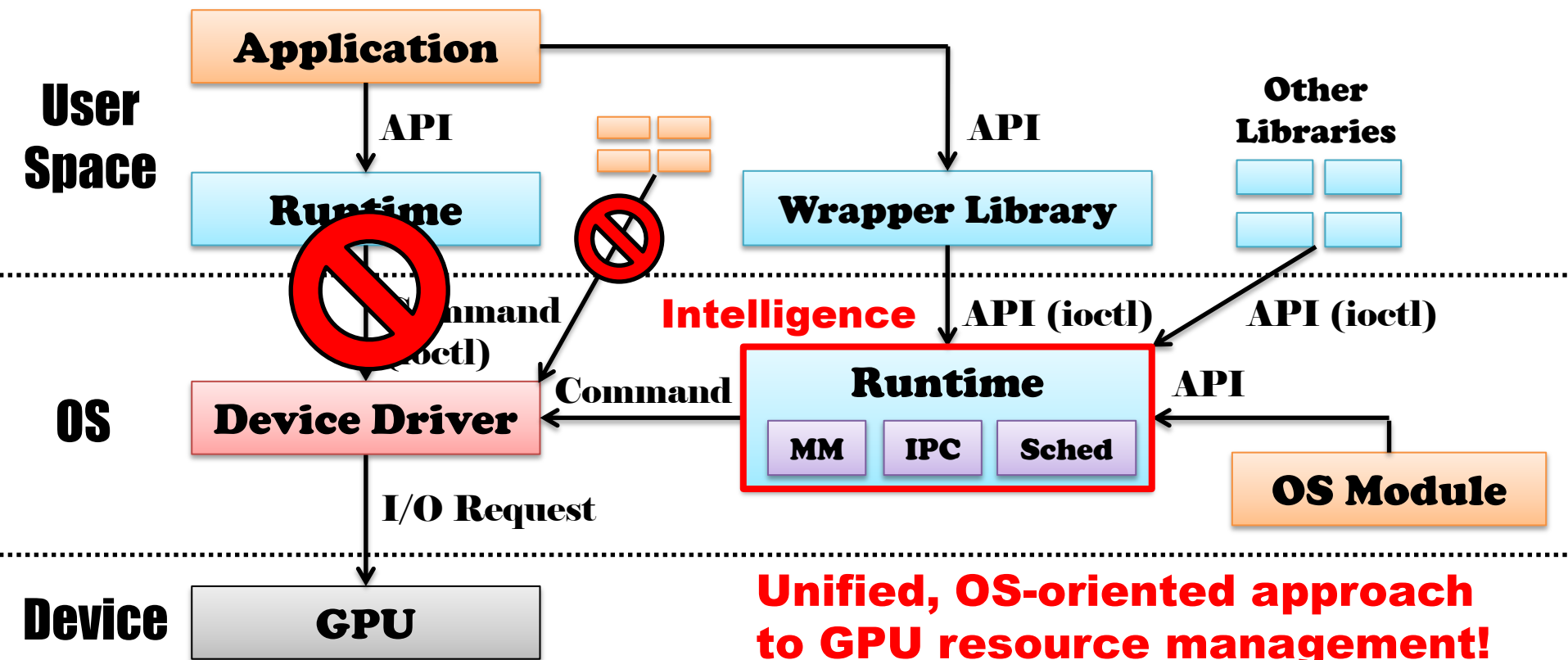
- **Motivation**
- **Approach**
- **GPU Resource Management**
- **Evaluation**
- **Conclusion**



# Traditional Naïve Approach



# Gdev Approach

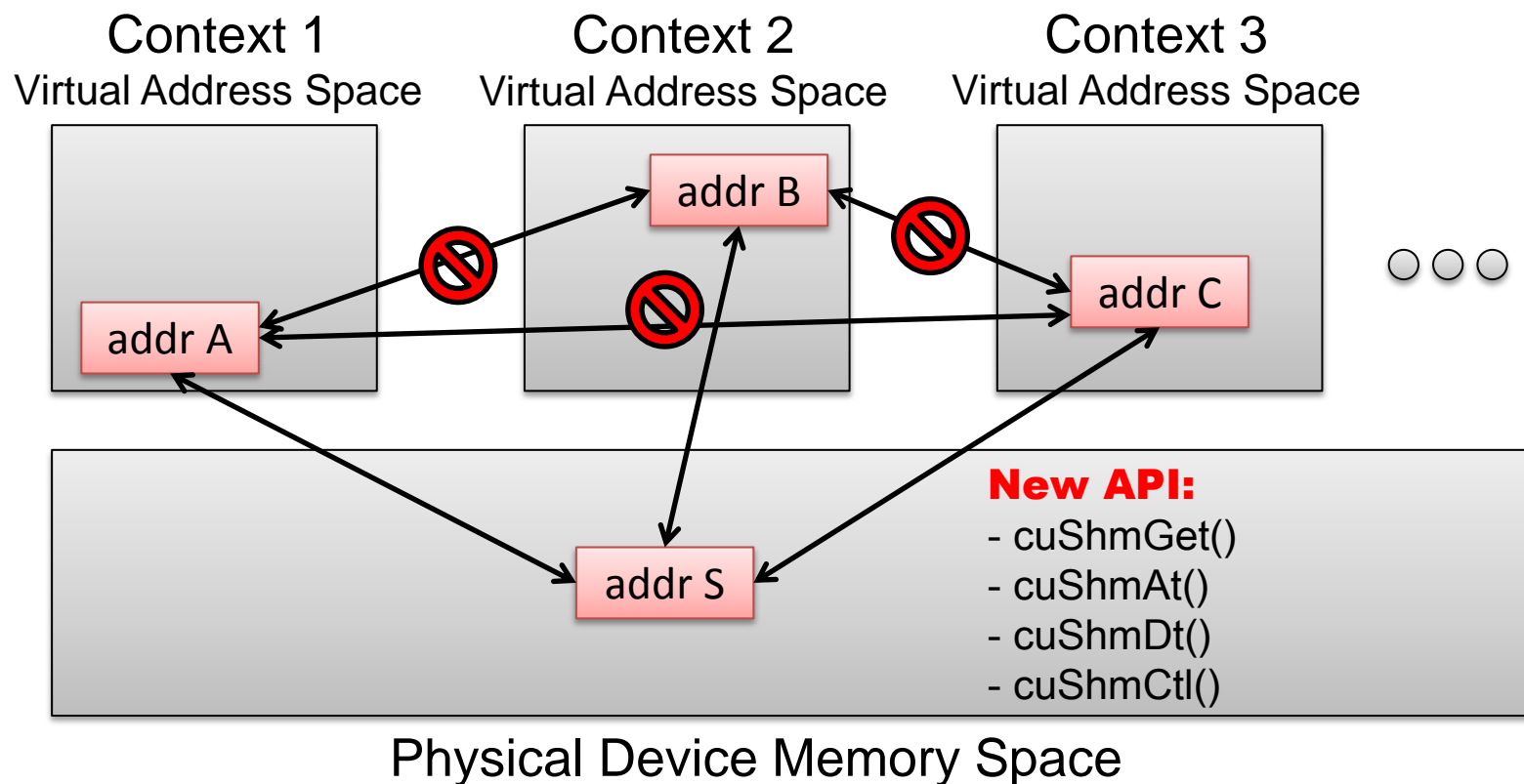


# Outline

---

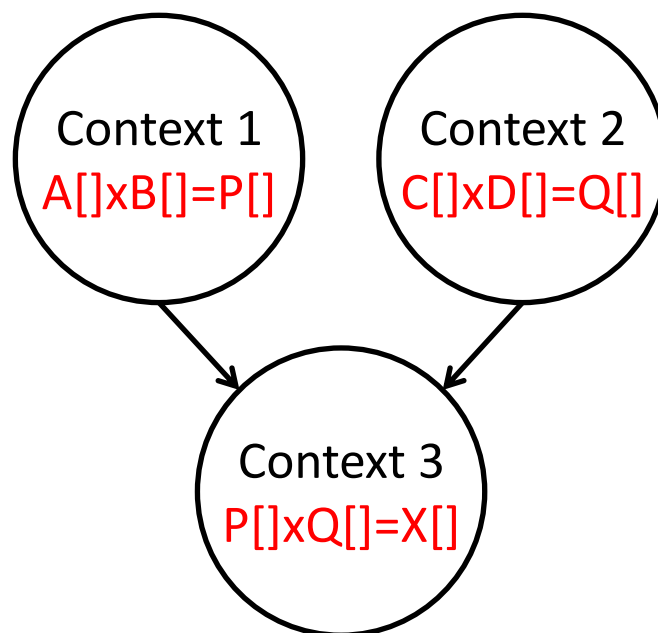
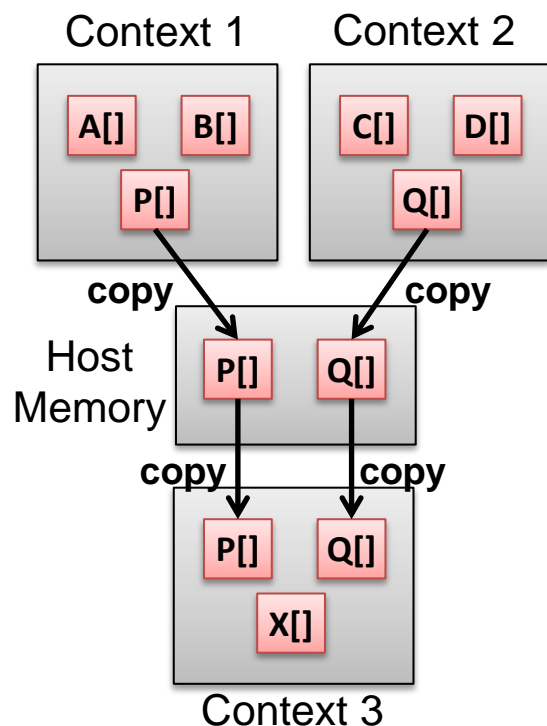
- Motivation
- Approach
- **GPU Resource Management**
- **Evaluation**
- **Conclusion**

# Shared Device Memory

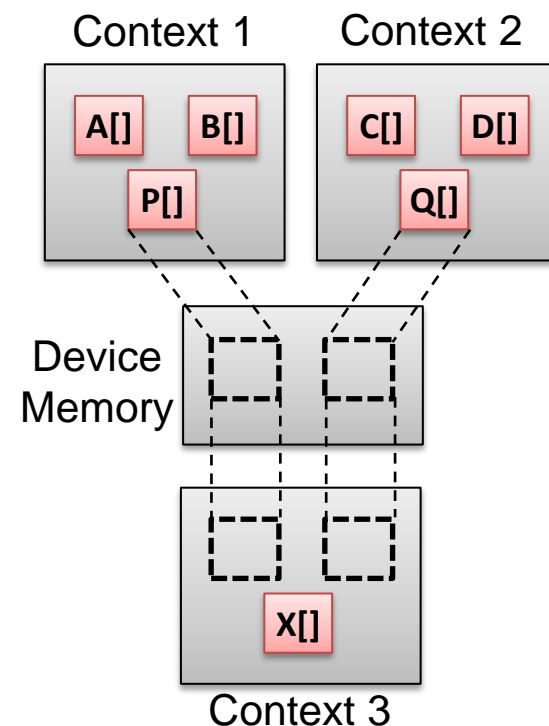


# E.g., Dataflow (2x2 Tree)

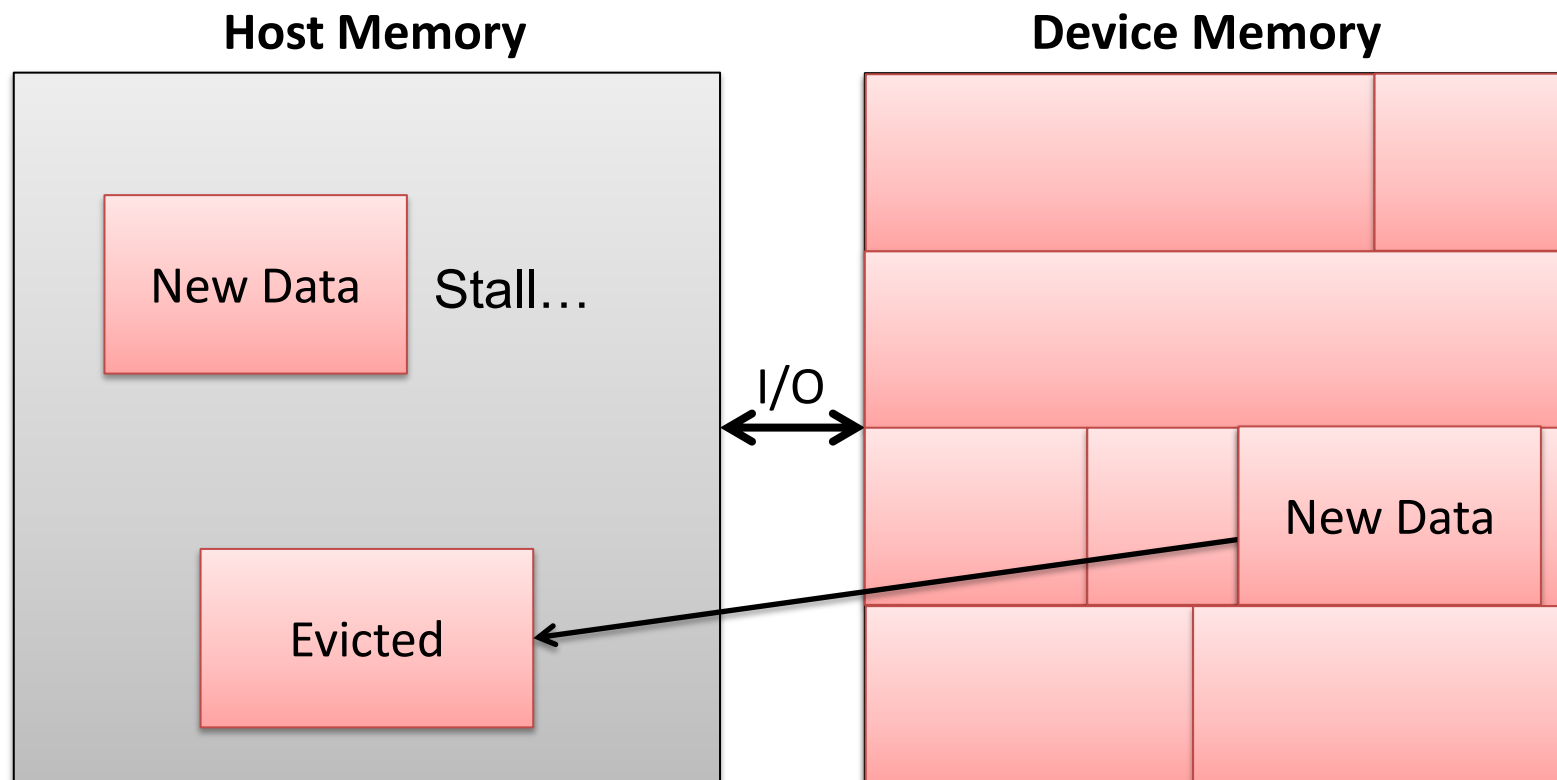
## No Shared Memory



## Shared Memory IPC

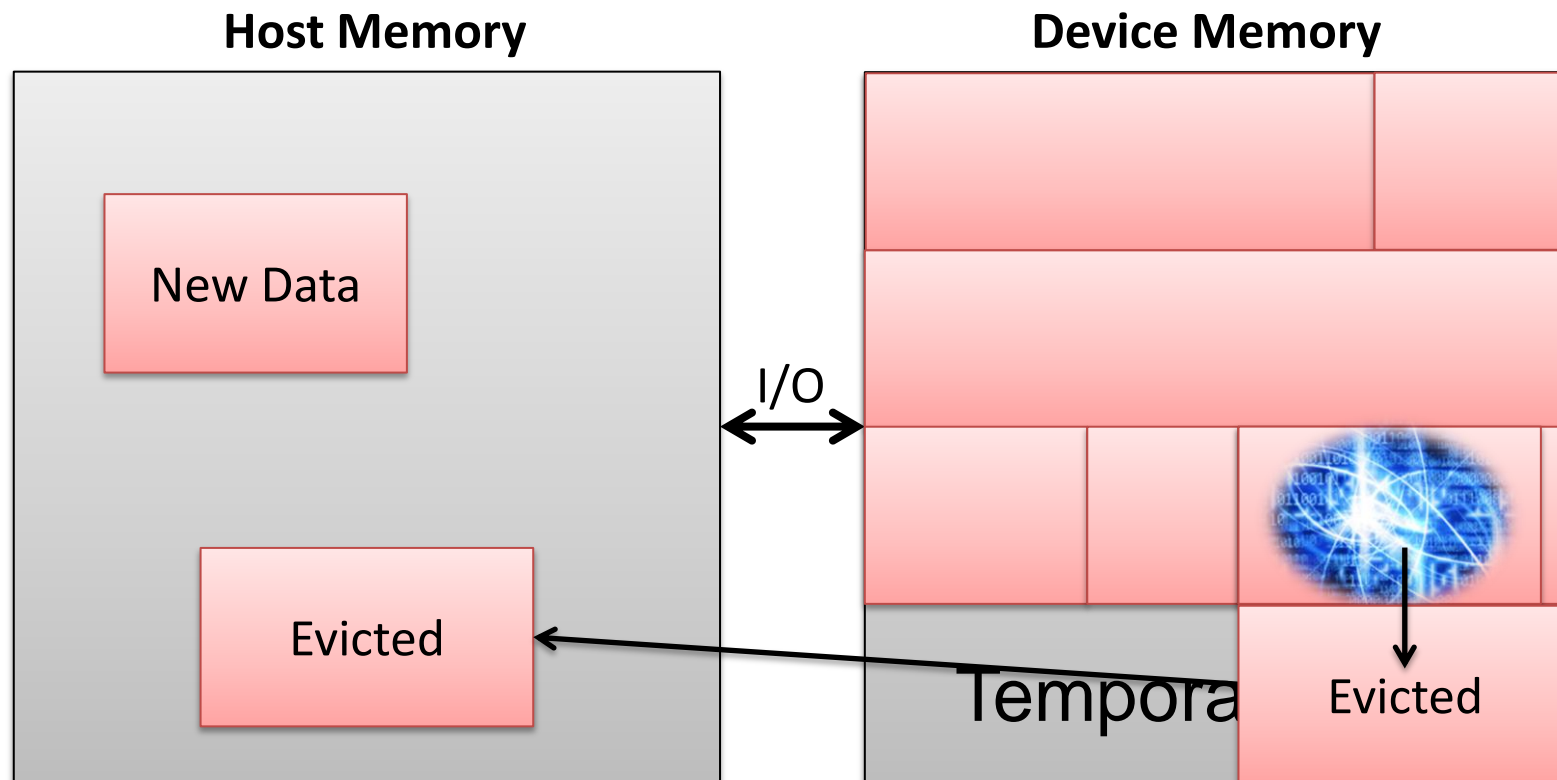


# Data Swapping

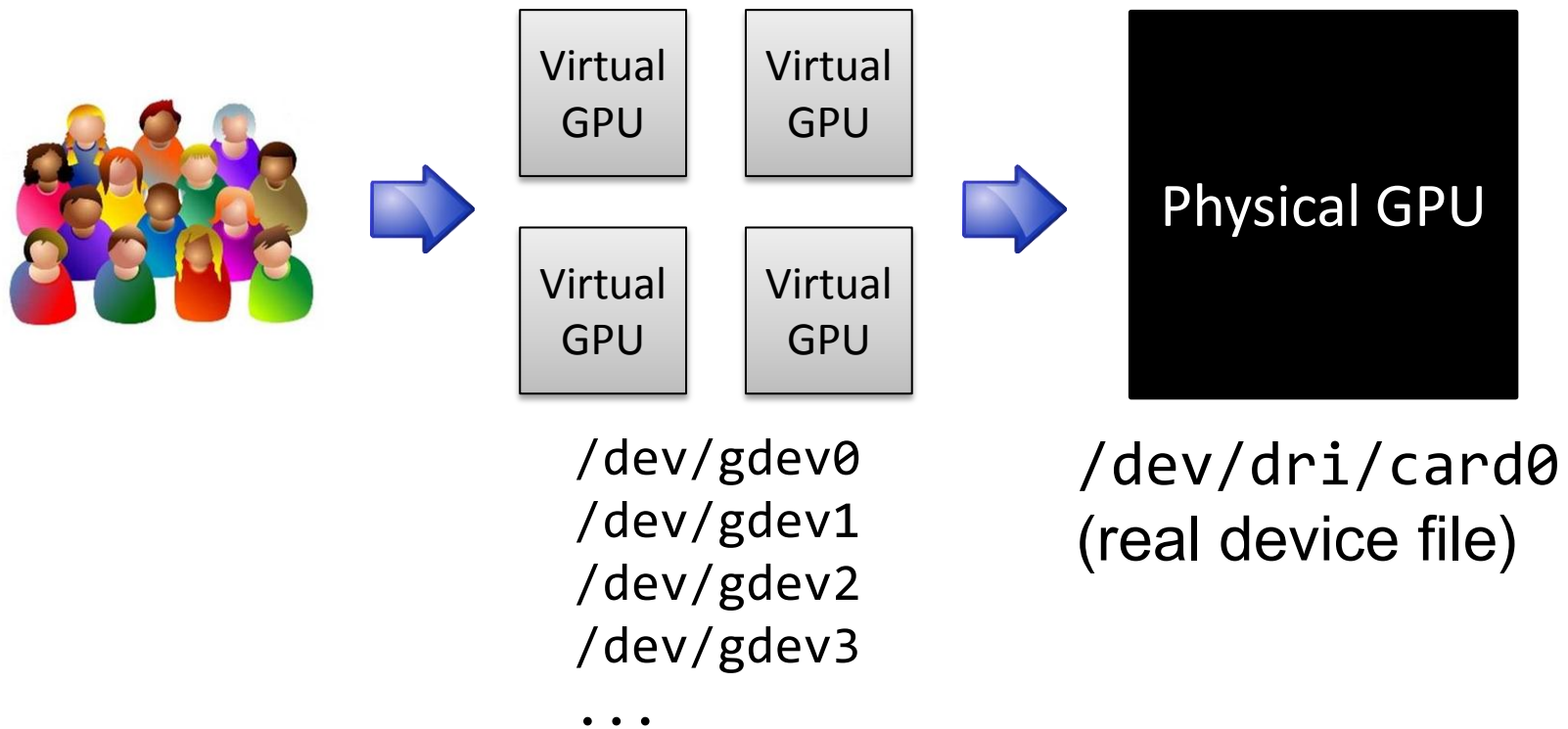




# Data Swapping (Enhanced)

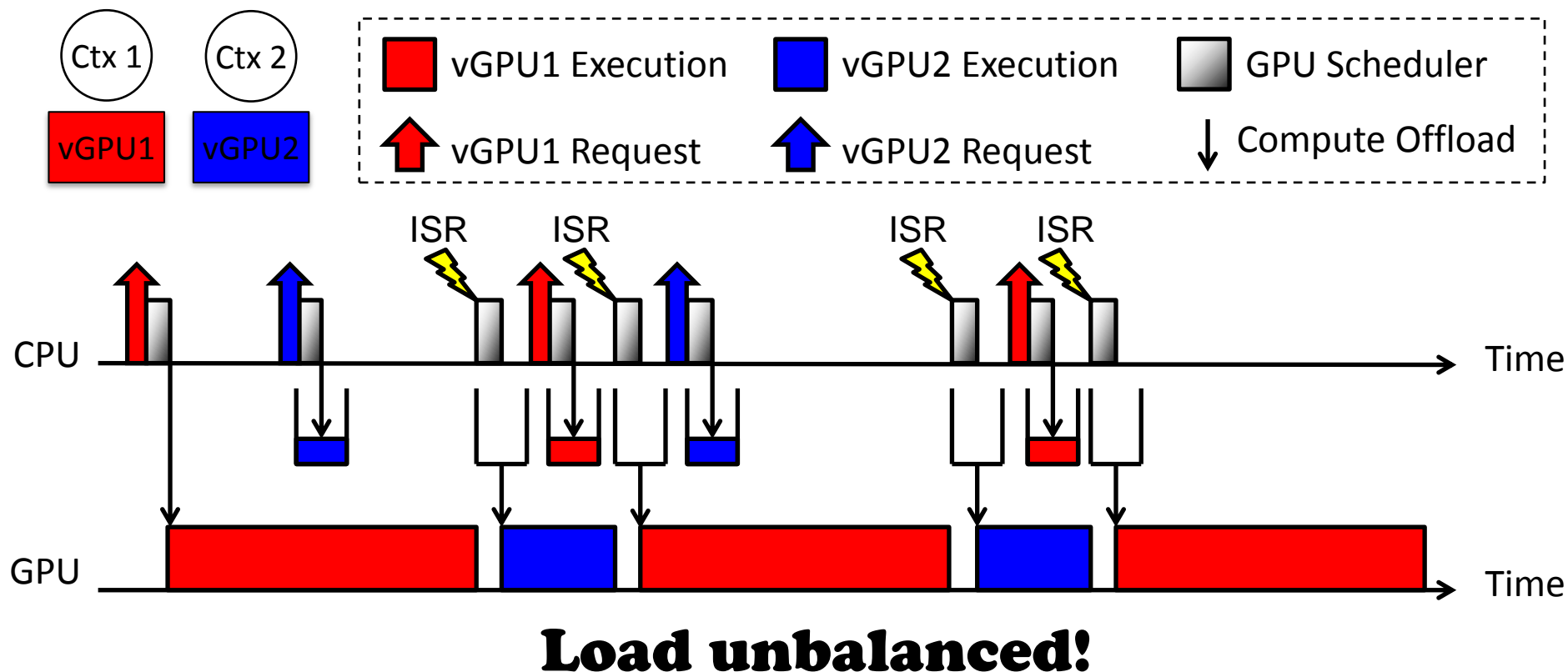


# GPU Virtualization

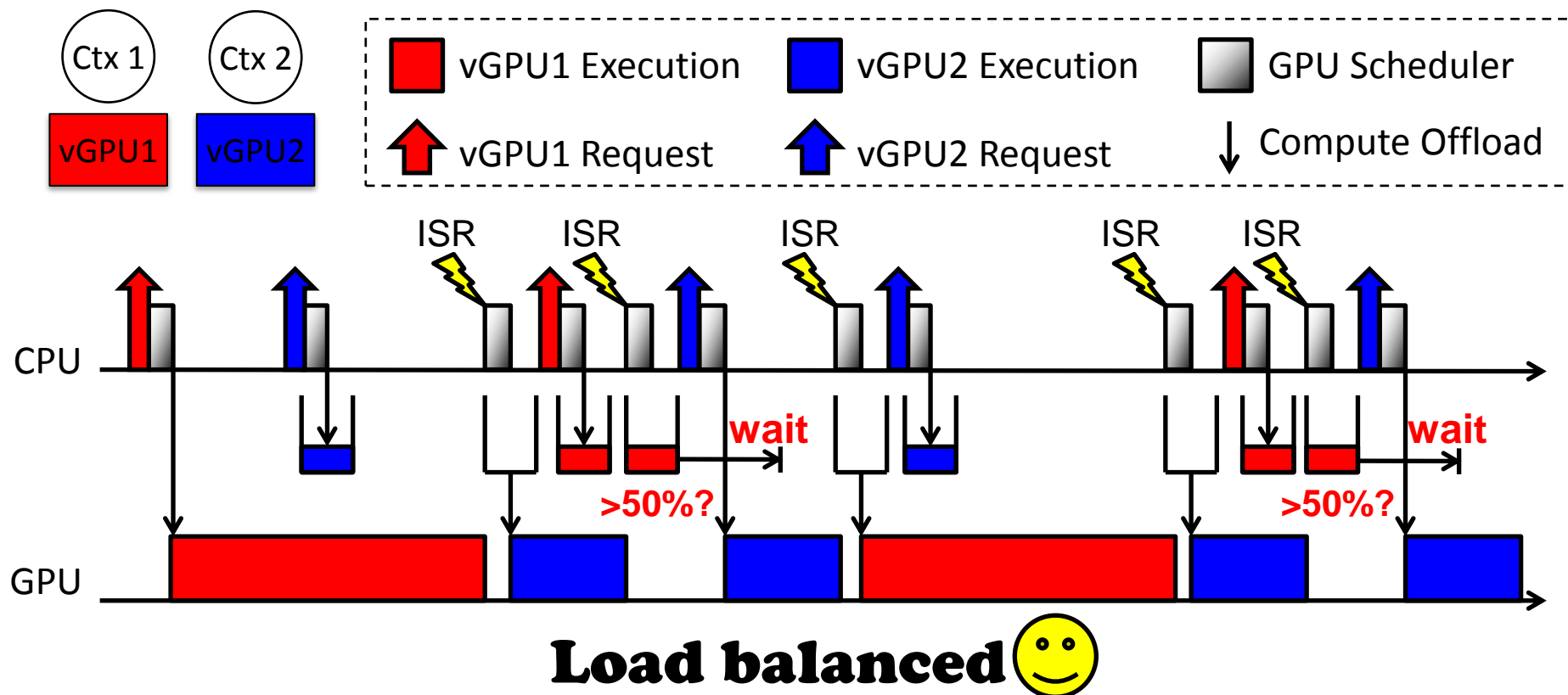


# Existing GPU Schedulers

Queue and dispatch [Kato ATC11] [Kato RTSS11]



# Bandwidth-aware non-preemptive device (BAND) Scheduler



# Outline

---

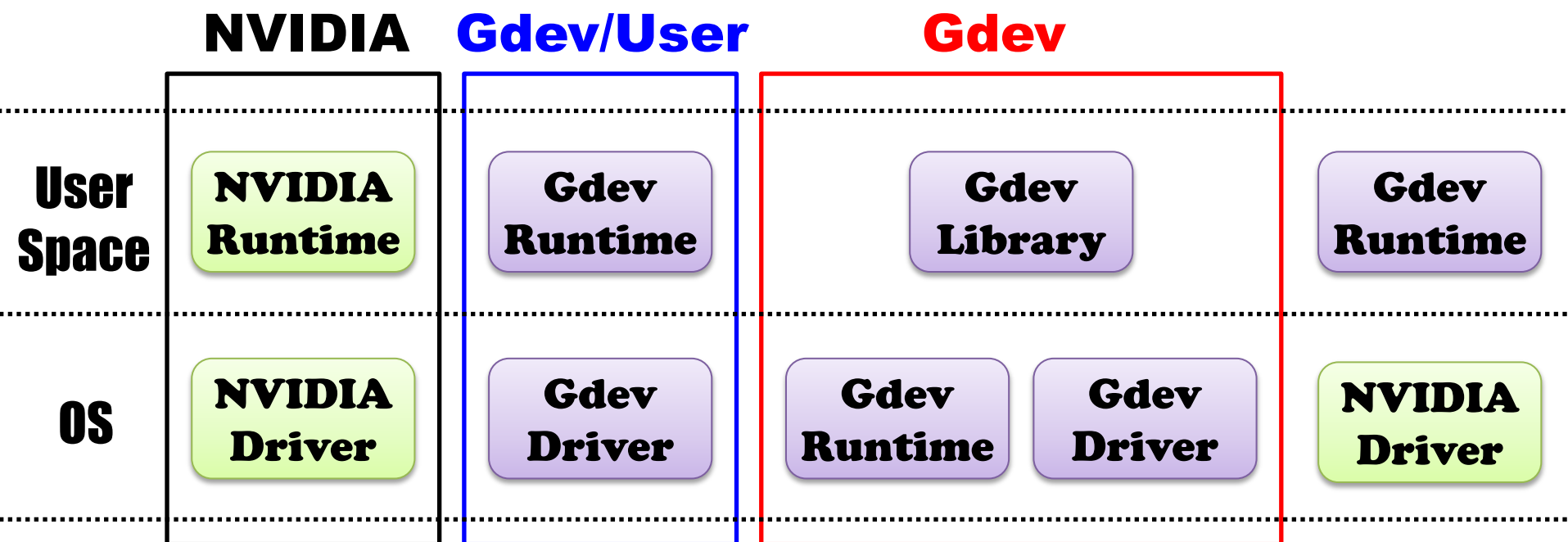
- Motivation
- Approach
- GPU Resource Management
- **Evaluation**
- **Conclusion**

# Experimental Setup

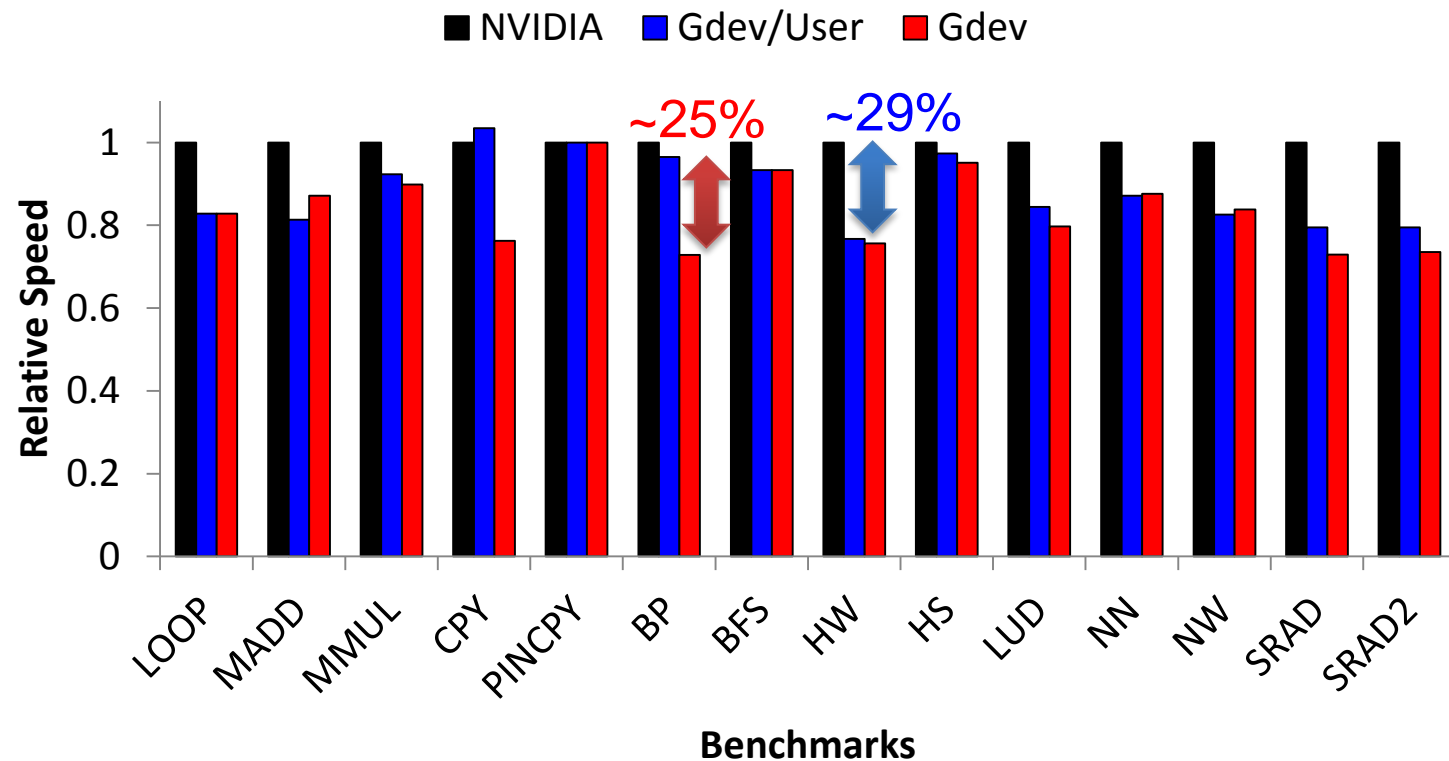
- Linux kernel 2.6.39
- NVIDIA GeForce GTX 480
- Intel Core 2 Extreme QX9650
- NVIDIA CUDA Compiler 4.0 and GCC 4.4.6
- Benchmarks & Applications:
  - Rodinia benchmark [Che *et al*, IISWC'09]
  - eCryptfs encrypted filesystem
  - FAST database search [Kim *et al*, SIGMOD'10]
  - PTask dataflow benchmarks [Rossbach *et al*, SOSPP'11]



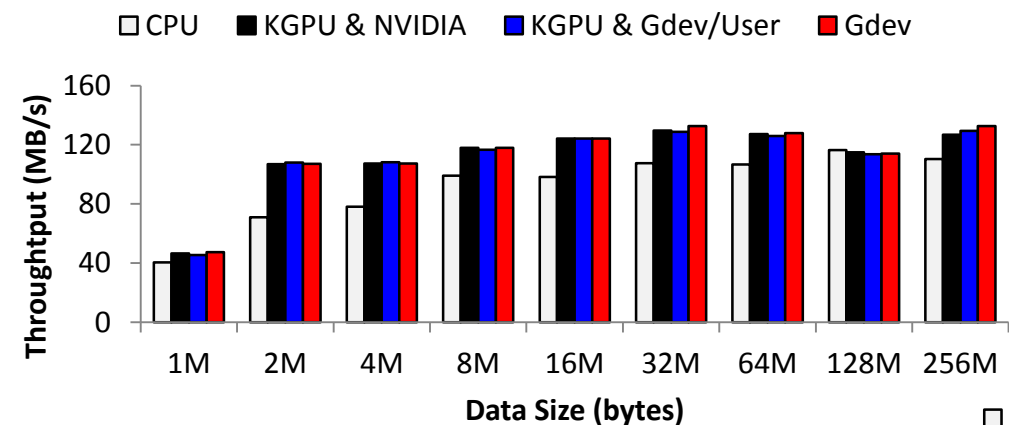
# Runtime and Driver Choice



# Basic Performance

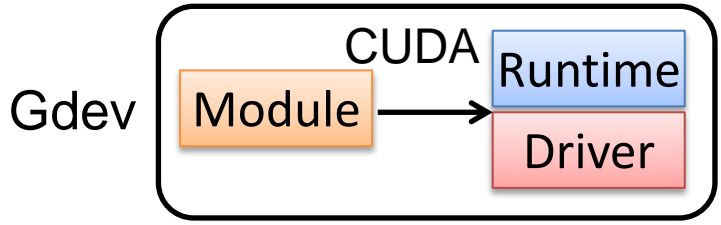
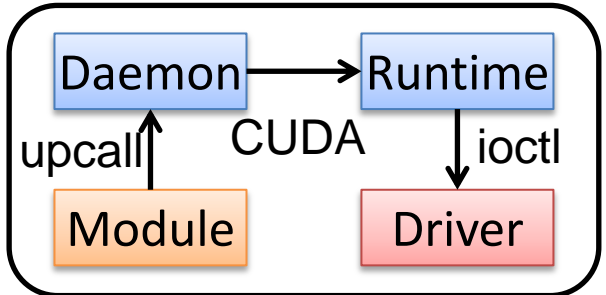


# eCryptfs Read&Write Throughput

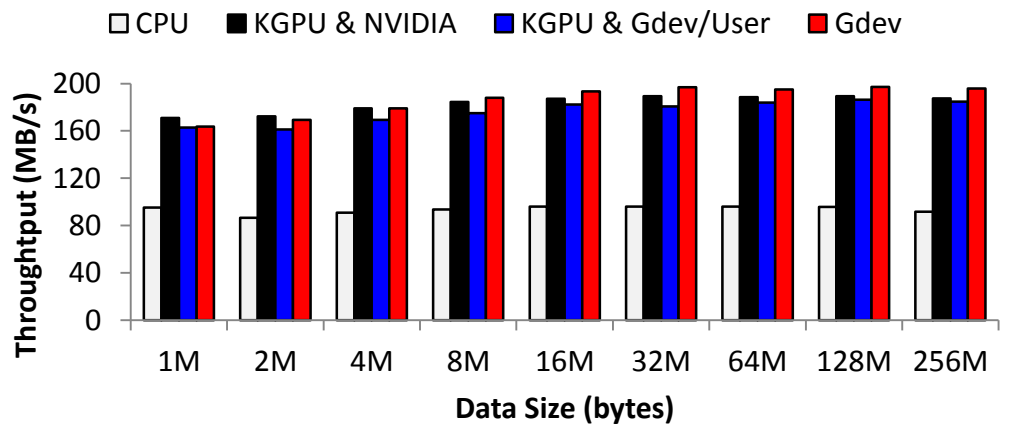


Read throughput

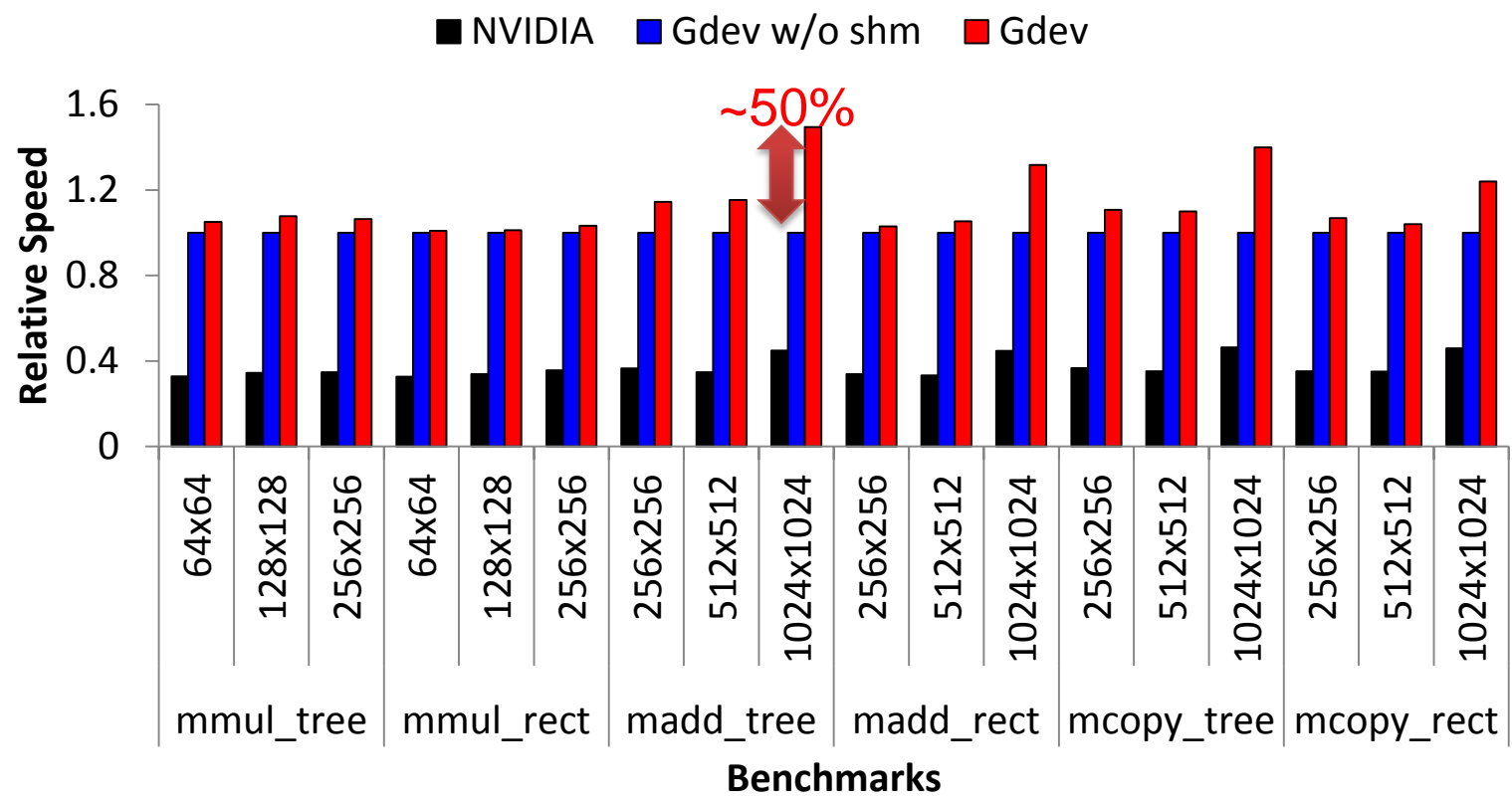
KGPU  
Sun et. al.  
2012



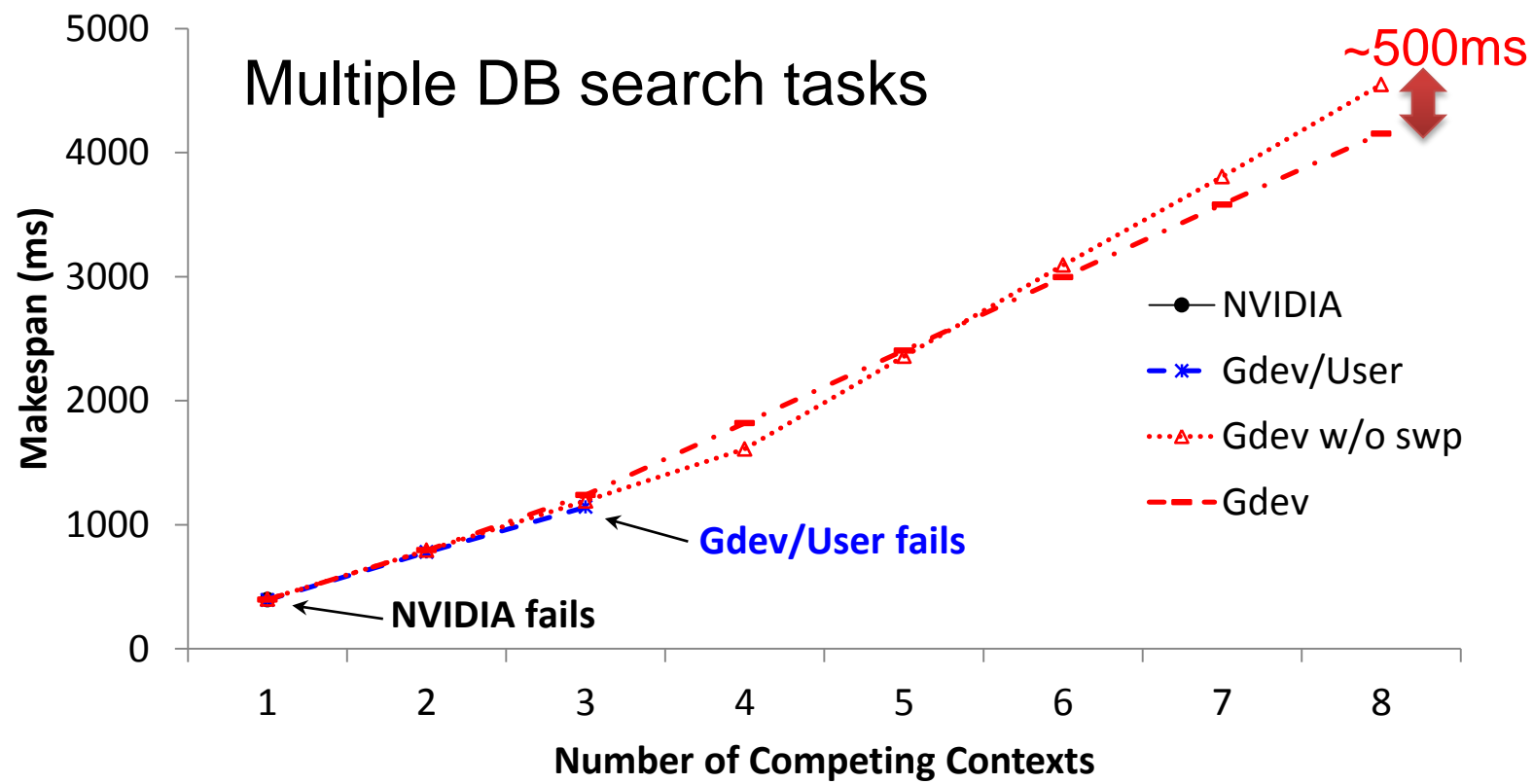
Write throughput



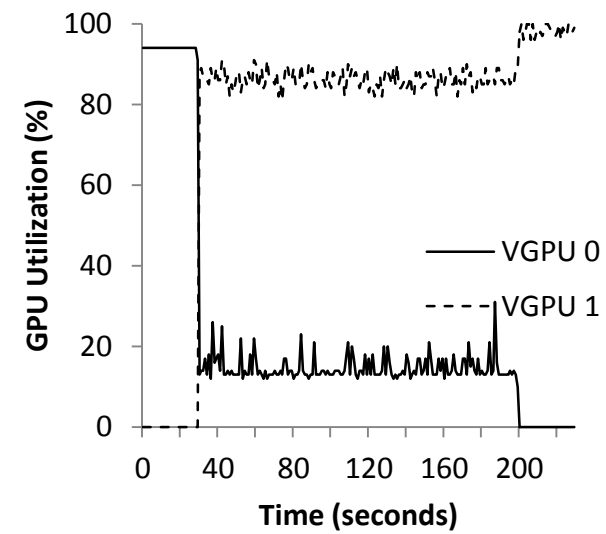
# Impact of Shared Device Memory



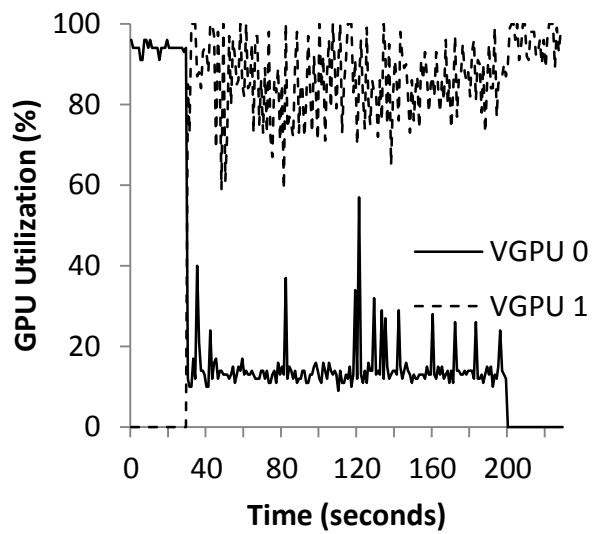
# Impact of Data Swapping



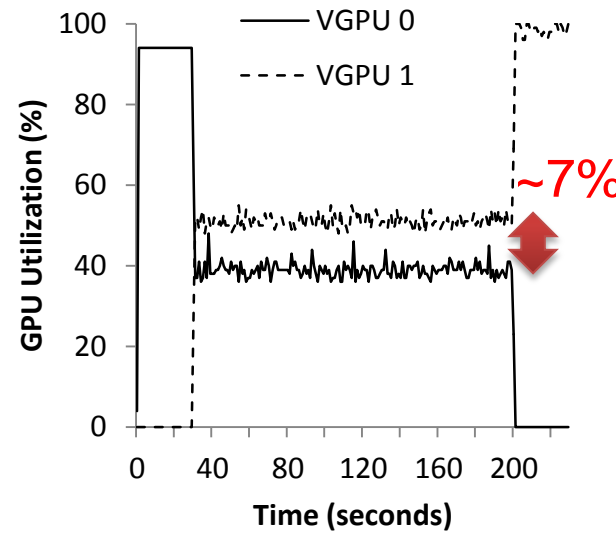
# Virtual GPU Isolation



No scheduling  
(FIFO)



Xen VM Policy  
(Credit)



Gdev Policy  
(BAND)



# Outline

---

- Motivation
- Approach
- GPU Resource Management
- Evaluation
- **Conclusion**

# Concluding Remarks

Gdev is an OS approach to **first-class** GPU resource management.

GPUs can be **used by the OS**.

GPUs can be **protected by the OS**.

GPUs can be **multi-tasked by the OS**.

Compromising basic performance to some extent.

# Concluding Remarks

Gdev is **open-source**.

Facilitate systems research.

Visit <http://sys.ertl.jp/gdev/>.

What's up-to-date:

- RAID6 erasure coding acceleration.
- Dynamic power management.
- Zero-copy between I/O devices and GPUs.



# Thank You!

## Questions?

