

GPU Page Cache – Leveraging GPU Resources in the OS

Alexander Schmidt, Frank Feinbube
Operating Systems and Middleware Group

Herbsttreffen GI/ITG-Fachgruppe Betriebssysteme

Hybrid Computing

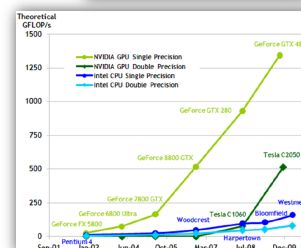
Dealing with massively multi-core:

- New architectures are evaluated (Intel SCC)
- Accelerators (APUs) that accompany common general purpose CPUs (Hybrid Systems)

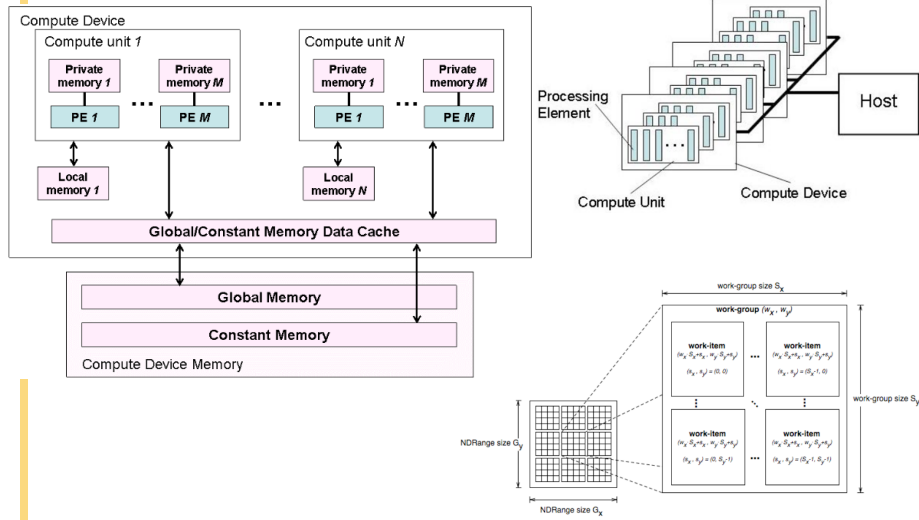


Hybrid Systems

- **GPU Compute Devices:** High Performance Computing (3 of top 5 supercomputers are GPU-based!), Business Servers, Home/Desktop Computers, Mobile and Embedded Systems
- **Special-Purpose Accelerators:** (de)compression, XML parsing, (en|de)ryption, regular expression matching



OpenCL Abstractions



Alexander Schmidt | Herbsttagung GI-Fachgruppe BS | 10. November 2011

Typical Operating System Services

- Resource abstraction
- System wide guaranties:
 - Fairness
 - Timeliness
 - Throughput
 - Isolation
- Security
- Scheduling resources

Alexander Schmidt | Herbsttagung GI-Fachgruppe BS | 10. November 2011

Operating System Perspective

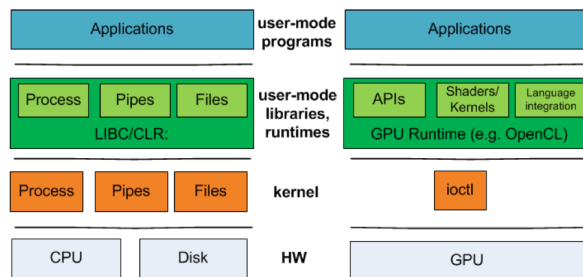
- Accelerators as black box
- No appropriate abstractions
 - ioctl based interface
 - Unknown architecture
 - Unknown capabilities
- No system-wide guarantees
 - Coarse grained scheduling
 - No isolation guarantees
 - No protection guarantees

Alexander Schmidt | Herbsttagung GI-Fachgruppe BS | 10. November 2011

Call for Action

„Operating Systems must support GPU abstractions“

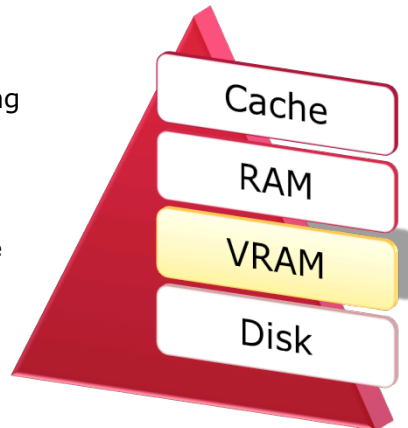
– Rossbach, Currey, Witchel [HotOS11]



Alexander Schmidt | Herbsttagung GI-Fachgruppe BS | 10. November 2011

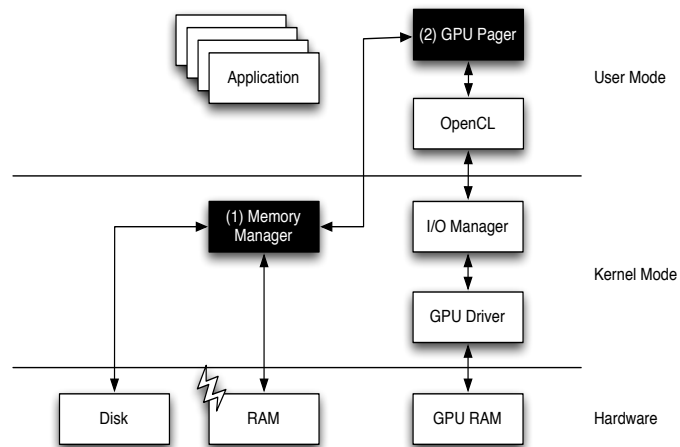
GPU Page Cache

- Use Global Memory as paging backing store
 - Typical desktop systems with VRAM of 1+ GB
 - VRAM unused most of the time
- Decrease paging operation latencies



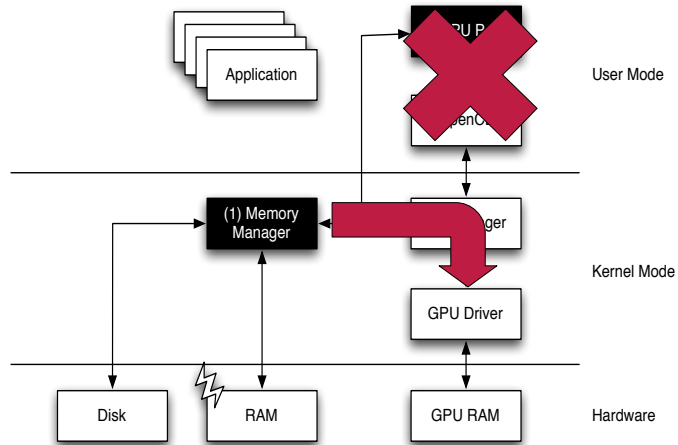
Alexander Schmidt | Herbsttagung GI-Fachgruppe BS | 10. November 2011

First Prototype Design



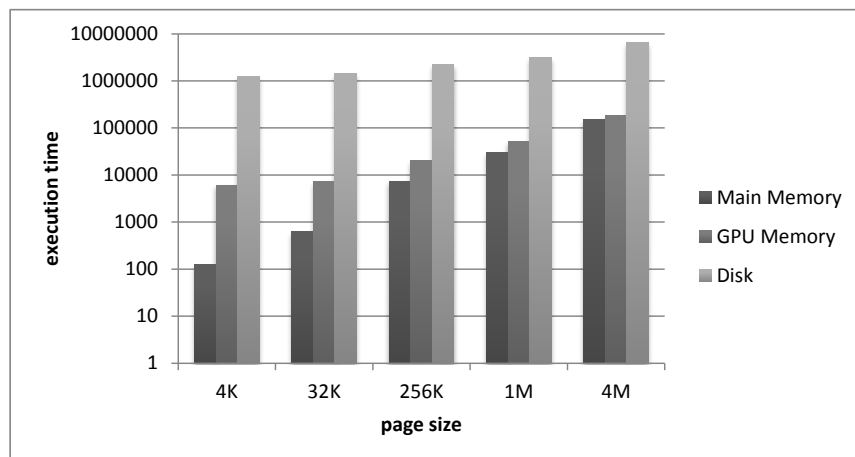
Alexander Schmidt | Herbsttagung GI-Fachgruppe BS | 10. November 2011

Design Goal



Alexander Schmidt | Herbsttagung GI-Fachgruppe BS | 10. November 2011

Access Latencies – Write Operations



Alexander Schmidt | Herbsttagung GI-Fachgruppe BS | 10. November 2011

Further Research Questions

- Optimal page transfer sizes?
 - Bandwidth limitations and communication overhead
- Optimal placement
 - When to swap pages to disk, or to GPU?
- Hardware architecture
 - Integrated GPU vs. dedicated GPU vs. Hybrid
 - Multi-GPU
 - Dynamic GPU properties

Agenda

1. GPU Pager
2. InstantLab
3. Demo
4. Conclusions

Windows Research Kernel (WRK)

- Stripped down Windows Server 2003 sources
 - Only kernel itself, no drivers, GUI, user-mode components
 - Missing components: HAL, power management, plug-and-play
- Released in 2006
- Freely available to academic institutions
- Encouraged by license:
 - Modification
 - Publication (of excerpts)

Alexander Schmidt | Herbsttagung GI-Fachgruppe BS | 10. November 2011

Structuring Experiments: The UMK Approach

- U-phase
 - Concentrate on OS concepts
 - Introduce OS interfaces
 - Systems programming
- M-phase
 - Observe concepts at run-time
 - Introduce monitoring tools
 - System measurements
- K-phase
 - Discuss kernel implementation
 - Introduce kernel source code (WRK/UNIX)
 - Kernel programming

Alexander Schmidt | Herbsttagung GI-Fachgruppe BS | 10. November 2011

Kernel Programming Experiments

- Debugging/Instrumenting the WRK
 - Boot phase
 - Process creation
 - Single-step debugging the WRK in a virtual machine
- Creating a new system call
 - Hide/Show a specified process from the system
 - Memorize hidden processes
 - Implement a system service DLL
- Memory management

Alexander Schmidt | Herbsttagung GI-Fachgruppe BS | 10. November 2011

Kernel Programming Experiments – Bottom Line

- Experiments comprise
 - Documentation
 - Source code
 - Workload generators
 - Measurement/visualization tools
- Experiment setup:
 - Install and configure test operating system
 - Build and deploy the sources
 - Configure kernel debugging infrastructure
- Virtualization helps, but
 - Variety of OS platforms, virtualization vendors among students
 - Hardware requirements

Alexander Schmidt | Herbsttagung GI-Fachgruppe BS | 10. November 2011

The InstantLab Idea

■ Provision of “canned experiments”

- Virtual machine images (VMI) as foundation
- Self-contained, pre-configured experiment in one VMI
- Instantaneous execution of a lab or experiment on Cloud resources



Alexander Schmidt | Herbsttagung GI-Fachgruppe BS | 10. November 2011

Embrace The Cloud

■ Virtualize laboratory environment

- No physical machines in university, no maintenance
- Compute resources in the Cloud

■ Migrate exercises and demos into the Cloud

- Provision of VM template(s) for each exercise
- Instantiation on demand

■ Facilitate experiments through remote display session

- Run experiments in Web browser
- Support of various platforms and compute power

Alexander Schmidt | Herbsttagung GI-Fachgruppe BS | 10. November 2011

Lab Management

- Goal

- Run experiments in the Cloud
- Regardless of specific Cloud provider

- Our approach

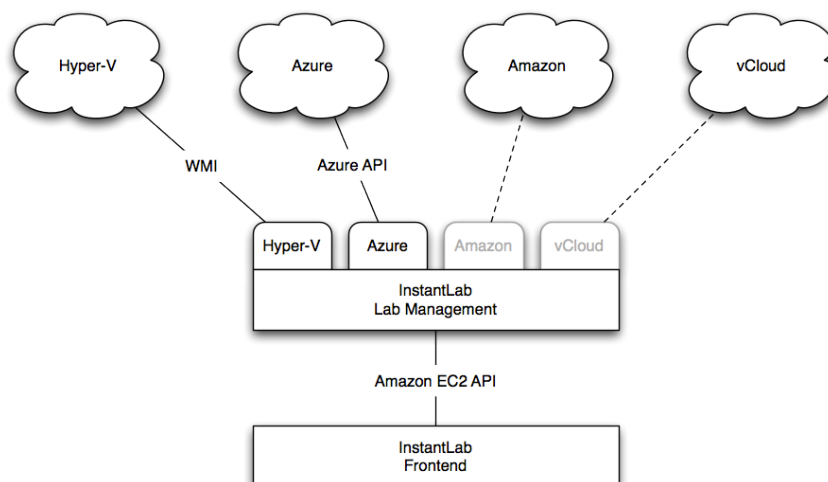
- Management services implement Amazon EC2 API (de-facto standard)
- Provide adapters for various Cloud vendors

- Adapters

- Hyper-V (private Cloud)
- Azure (public Cloud)
- Amazon (public Cloud)
- vCloud

Alexander Schmidt | Herbsttagung GI-Fachgruppe BS | 10. November 2011

Lab Management – Architecture



Alexander Schmidt | Herbsttagung GI-Fachgruppe BS | 10. November 2011

Agenda

1. GPU Pager
2. InstantLab
3. Demo
4. Conclusions

InstantLab Demo – Working Set Replacement Experiment



InstantLab Demo – Lab Management



Alexander Schmidt | Herbsttagung GI-Fachgruppe BS | 10. November 2011

Conclusions

■ OpenCL Accelerators

- New system architectures
- Operating systems lack appropriate abstractions
- New concepts to manage/handle accelerators
- Various approaches (novel OS design vs. legacy integration)

■ GPU Page Cache

- Leverage GPU Global Memory for caching non-resident pages
- Reduce paging operations latencies

■ InstantLab

- Experiments useful vehicle for research and teaching
- Leverage Cloud resources
- Reduce setup overhead

Alexander Schmidt | Herbsttagung GI-Fachgruppe BS | 10. November 2011