Quest-V – A Virtualized Multikernel for High-Confidence Systems

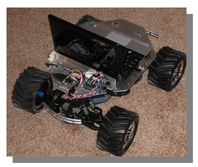
Ye Li, Eric Missimer, Richard West

BU Operating Systems and Services

Objective

- Operating system for high-confidence systems (NCO/NITRD)
  - Predictable
  - Resistant to component failures & malicious manipulation
- Self-healing system
  - Online recovery of software component failures
  - Avoid impact on other functional components

Applications

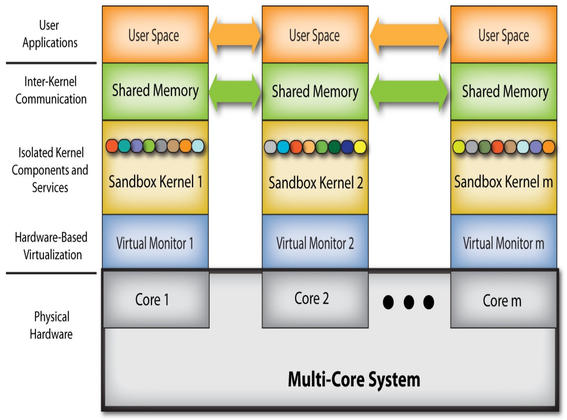


- Healthcare
- Avionics
- Automotive
- Factory Automation
- Robotics
- Space exploration
- Other safety-critical domains

Approach

- Quest-V for multicore processors
  - Distributed system on a chip
  - Time as a first-class resource
    - Cycle-accurate time accountability
  - Separate sandbox kernels for system sub-components
    - Isolation using hardware-assisted memory virtualization
      - Extended Page Tables (EPTs – Intel)
      - Nested Page Tables (NPTs – AMD)
    - Security enforcible using VT-d + Interrupt Remapping (IR)
      - Device interrupts scoped to specific sandboxes
      - DMA transfers to specific host memory

Architecture Overview



Isolation

- Memory virtualization using shadow paging isolates sandboxes and their components
- Dedicated physical cores assigned to sandboxes
- Temporal isolation using Virtual CPUs (VCPUs)

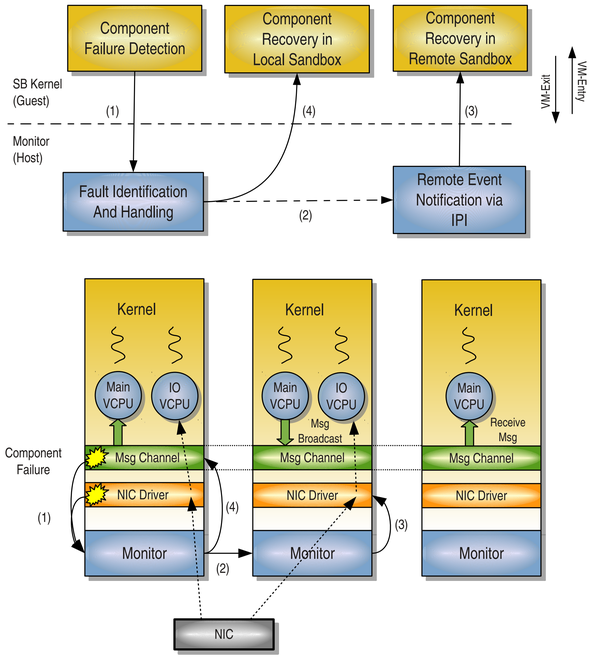
Predictability

- VCPUs for budgeted real-time execution of threads and system events (e.g., interrupts)
  - Threads mapped to VCPUs
  - VCPUs mapped to physical cores
- Sandbox kernels perform local scheduling on assigned cores
  - Avoid VM-Exits to Monitor – eliminate cache/TLB flushes

Efficiency

- Lightweight I/O virtualization & interrupt passthrough capabilities
  - e.g., VNICs provide separate interfaces to single NIC device
- Hardware performance monitoring for improved resource management
  - Cache-aware scheduling

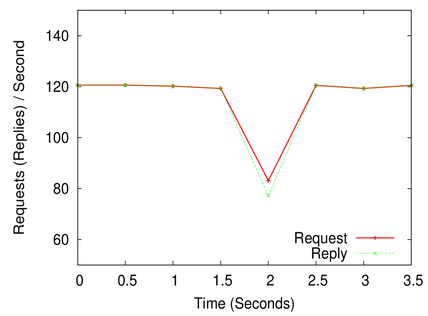
Example Fault Recovery



Fault Recovery

- Inter-Processor Interrupts (IPIs) for inter-sandbox communication and remote recovery of faulty components

Example Web Server w/ NIC Driver Fault



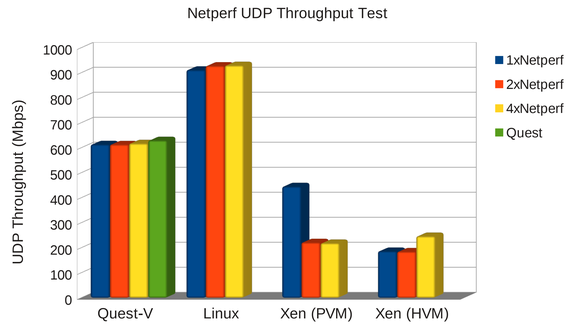
- httperf with web server in presence of faulting driver
- Requests/replies set at 120/sec under normal operation

Performance Costs

Recovery Phases	CPU Cycles	
	Local Recovery	Remote Recovery
VM-Exit	885	
Driver Switch	10503	N/A
IPI Round Trip	N/A	4542
VM-Enter	663	
Driver Re-initialization	1.45E+07	
Network Re-initialization	78351	

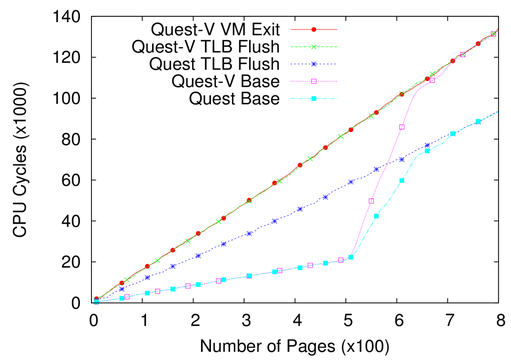
*Core i5-2500K 4-core Processor w/ 8 GB RAM

Shared Driver Costs



Virtualization Costs

- Example Data TLB overheads
- Xeon E5506 4-core Processor 2.13GHz w/ 4GB RAM



Quest Website

- http://www.cs.bu.edu/fac/richwest/quest.html

Thanks to Gary Wong and Matthew Danish for help with parts of the development of Quest.
This work is funded in part by various grants, including: NSF #0615153 and #1117025.