

The Entropia Virtual Machine for Desktop Grids

Brad Calder, Andrew Chien, Ju Wang and Don Yang

UCSD

Presenter:
Michael R. Hines

Entropy Setup (1)

- Entropia: specialized distributed computing solution
 - Instead of large mainframes or supercomputers
 - Examples: Condor, SETI @ Home
- Runs on networks of end-user PCs
- Supports Windows-based, unmodified W32 distributed applications
- Meant for Enterprise-level deployment

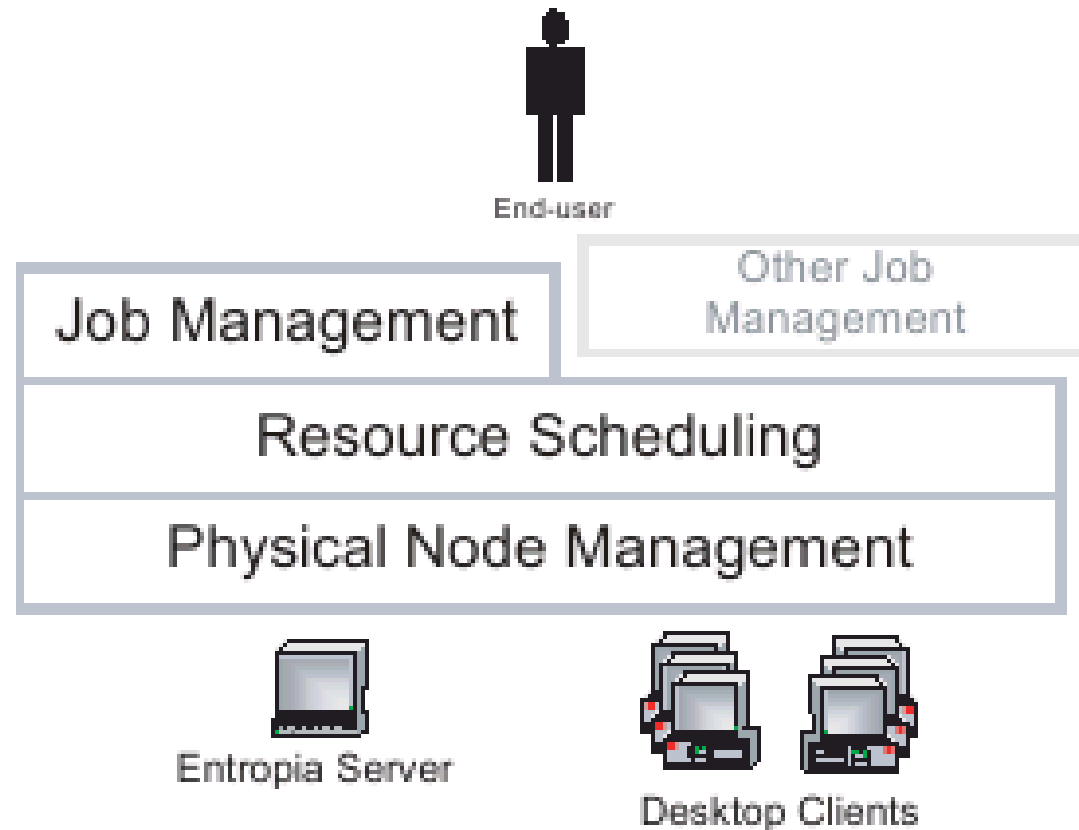
Quick VM Review

- Process VMs:
 - Normal multiprogrammed processes themselves
 - Binary optimization
 - Translators: Java, .NET
- System VMs:
 - Hosted: VMM and VM have same ISA
 - Codesigned: Different ISAs all virtualized

Entropia Setup (2)

- An Entropia cluster has:
 - 1) Jobs to be run on the distributed system
 - 2) A central management server to start them
 - 3) Nodes containing VMs to run them in conjunction with normal user-activity
- The VM is off a different kind:
 - It runs modified, sand-boxed sub-jobs
 - It monitors sub-job activity
 - It manages sub-job interaction with the host

Entropy Setup (3)



Submitting a Job

1. A job is submitted to the central server
2. The job is broken up into sub-jobs
3. The binaries are modified to allow for system calls (only) to be intercepted
4. Sub-jobs are sent to VMMs on various PCs
5. VMM runs, monitors, and virtualized the sub-jobs, intercepting resource-intensive system calls
6. VMM reports back to central server upon completion with results

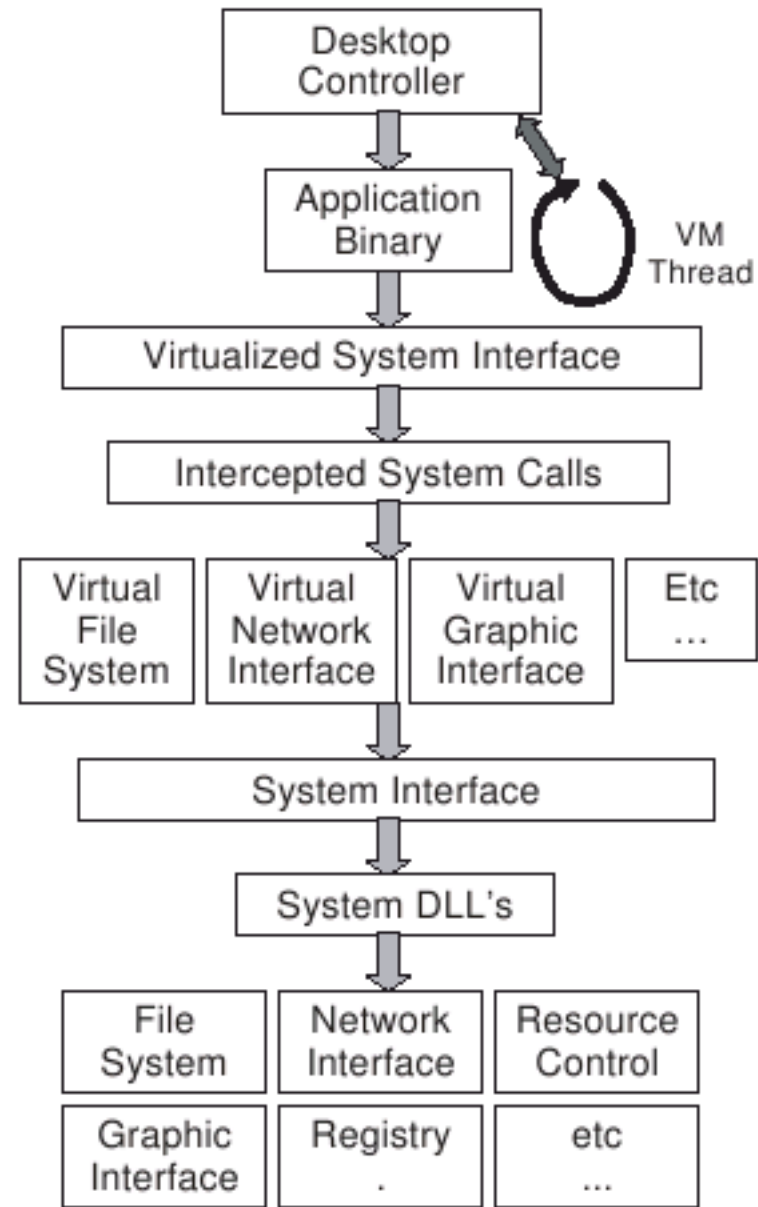
Virtualization Capabilities (1)

- Does **not** do:
 - Checkpointing or job migration
 - Full-OS virtualization or translation
 - Parallelize source jobs for distribution
 - inter-subjob networking is the job's responsibility
- Does do:
 - System call interception and mem/disk/network/file monitoring by modifying the kernel symbol table
 - Security: binary checksums and sub-job termination when out of control
 - Distribution of already-parallelized binaries

Virtualization Capabilities (2)

- A device driver modifies the symbol table to intercept relevant system calls
- Non-virtualized processes that invoke calls are allowed to go through
- Adds small overhead to normal system operation
- Advantages of modified virtualizing:
 - Eliminates memory management virtualization
 - Eliminates ISA simulation
 - Secures normal user activity from subjobs

Virtualization Capabilities (3)



Job Scheduling (1)

- Central server:
 - Gets reports of resource availability on participating PCs
 - Sends sub-jobs to PCs only if enough resources are available
 - Can instruct PCs to pause or terminate sub-jobs at whim
- VMM: monitors job activity. Pauses or stops during:
 - keyboard/mouse/resource usage
 - Exceeded performance constraints or resource/execution violations

Job Scheduling (2)

- Users submit jobs with detailed parameterized specifications of memory, cpu, and execution time limits.
- Entropia can schedule subjobs on thousands of clients by checking each one for availability
- Entropia subjobs are broken up for specialized applications (mentioned later)

Distributed & Parallel Computing

- Entropia Jobs must be designed for distribution. To do so requires:
 - Parallel Programming Strategies
 - Decompose the problem's computation and data into pieces that run simultaneously
 - Add concurrency control:
 - Secure critical sections
 - Enforce dependencies between threads
 - Communication Primitives:
 - Gather, scatter, all gather, broadcast
 - Be able to transmit computation results fast

Resource Monitoring

- Disk I/O, paging, network, memory, CPU usage, # processes/threads all monitored to stay within a configurable limit
- Violations result in throttling (pausing) of the sub-job
- Only one sub-job per machine (but with potentially many processes or threads)
- Sub-jobs may contain intensive inter-job communication protocols, for example

Entropia Usefulness

- Deployed at 12 industry sites
- Used for over 50 applications
- Applications include:
 - large 3D-renders
 - java-based applications (w/ sand-boxed JVM)
 - Prime numbers and cryptographic operations
 - Monte-carlo and database search jobs
 - Pharmaceutical Virtual Screening/Sequence Analysis - the simulation of drugs on target proteins, their potency and interactions)

Performance of the VM

- A test of pharmaceutical applications:
 - A single 1.6 Ghz w/ 1 GB memory, XP Pro
 - compute-intensive simulation of the interactions of drug molecules on a protein that can be broken up into sub-jobs
- Results: differences w/ and w/o a VM
 - Execution time speedups differ by 1% at most
 - I/O activity increases by 0 – 50 KB
 - Context of results unclear, probably for business
 - Virtualization methods are far less than System VMs
- They hold records for prime number discovery

Questions?

Thanks.

Bye.