Version 3.1

Technical White Paper

Table of Contents

Introduction

Virtual Iron provides enterprise-class software solutions for creating and managing virtual infrastructure in the data center. The software provides advanced server virtualization and management capabilities that take advantage of industry standards and open source economics and enable companies to dramatically reduce the cost and complexity of enterprise service delivery.

Companies use Virtual Iron's software to:

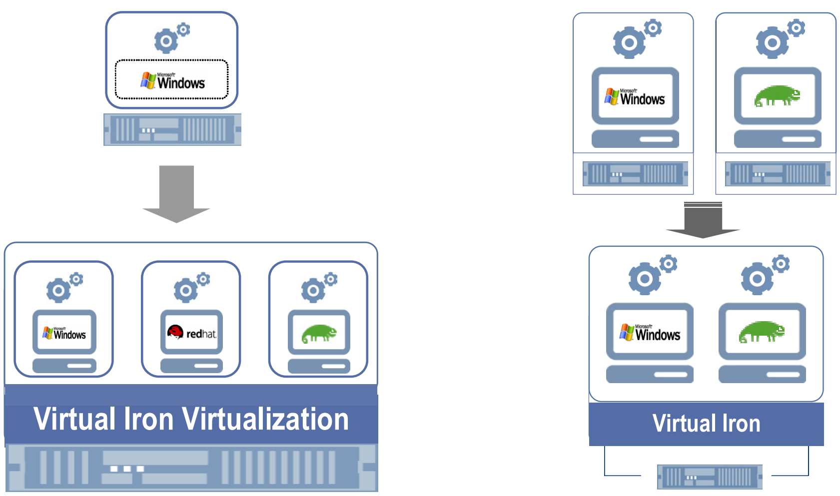
- Virtualize enterprise-class workloads.
- Improve the utilization of current systems and get more out of today's industry-standard hardware systems through server consolidation.
- Quickly set up development, test, staging and production environments.
- Recover from hardware failures quickly, reliably and cost-efficiently.
- Match resource capacity to workload demands via capacity management capabilities.
- Reduce human labor and errors via policy-based automation.

This document provides a detailed discussion of the Virtual Iron product and technology and is intended to complement and provide further technical detail to the Virtual Iron *Product Overview*.

What is Virtualization?

Virtualization is used to describe many different technologies and approaches to abstract operating systems from hardware. Server virtualization presents a virtual view of hardware to an operating system to allow multiple operating systems to share the same physical resource in complete isolation from each other - see Figure 1.

Figure 1



The key benefits of virtualization are:

- Isolation: A virtual server's state is unaffected by the state of other virtual servers on the same physical hardware
- Encapsulation: The state of a virtual server can be captured and files representing a virtual server are portable
- Hardware-independence: Virtual hardware doesn't have to be identical to the underlying physical hardware

The x86 architecture was not originally designed for virtualization. This created tradeoffs in early implementations in terms of both performance and complexity. Historically there have been two approaches to virtualize x86 architecture – Full Virtualization with binary patching and Paravirtualization. Although both approaches create the illusion of physical hardware to achieve the goal of operating system independence from the hardware, there are significant differences between the approaches:

- Full Virtualization with binary patching, rewrites x86 instructions at run-time that cannot be trapped and converts them into a series of instructions that can be trapped and virtualized. Full Virtualization is capable of running existing, legacy operating systems without modification, however it has significant costs in complexity and run-time performance.

- Paravirtualization modifies an operating system to replace non-trappable x86 instructions with a series of calls directly into a hypervisor (a virtual machine monitor). It achieves high performance with less complexity in the virtualization layer but requires the guest operating system to be substantially modified and tied to a particular version of the hypervisor.

Native Virtualization® – A New Approach

Native virtualization is a new approach that takes full advantage of new hardware-assist capabilities from Intel and AMD processors to eliminate the need for operating system changes while simultaneously providing the highest levels of performance. This allows customers to run any operating system version and the broadest variety of workloads.

Virtual Iron leverages this approach to deliver advanced virtualization and management software capabilities that are designed for production-class performance and scalability. The platform can support hundreds of industry-standard (x64) physical servers and thousands of virtual servers.

The Virtual Iron solution includes comprehensive capabilities to manage virtual infrastructure and addresses a number of data center and virtualization initiatives:

1) Server Consolidation - Run a large number of operating systems and applications on a single physical server.

2) Rapid Provisioning – Leverage software reference stacks (golden images) to quickly and easily install and deploy software environments. This might include managing and orchestrating a large number of operating systems, software updates, configuration files, and patches associated with each

3) Business Continuity – Deliver high availability and disaster recovery more efficiently and cost effectively by enabling multiple virtual servers to fail over to a single physical server that does not have to be physically identical.

4) Capacity Management - Optimally match data center capacity to workload demands. Align computing resources with business initiatives and priorities by configuring, reconfiguring, moving and allocating capacity on-the-fly, as needed, to any/all applications in the system. Automate common administrative operations related to managing virtual infrastructure. This reduces human labor costs and errors.

Virtual Iron Virtual Infrastructure Architecture

Virtual Iron virtualizes all data center resources (CPU, I/O and storage) to create virtual infrastructure. The components described in the table below provide the foundation to create virtual servers. A virtual server consists of the same components found in a physical server: 32 or 64-bit CPUs, memory, disks, network adapters, fibre channel adapters, keyboard, video, and mouse. A virtual server can run standard Linux and Windows operating systems and applications.

Physical Resource	Virtual Resource
Server: Industry-standard Intel and AMD servers upon which the virtualization layer is automatically deployed.	Virtualized Node: Collection of CPUs and RAM that can be allocated to a virtual server.
Ethernet network: Each server can have multiple gigabit Ethernet cards (NICs) to provide required throughput and availability.	Ethernet network: Virtual servers connect through virtual NICs to physical or virtual networks.
Storage: SAN and NAS storage technologies are used for reliable persistent storage.	Data Store: A collection of storage resources that can be partitioned and allocated to virtual servers using raw mappings or virtual hard disks.

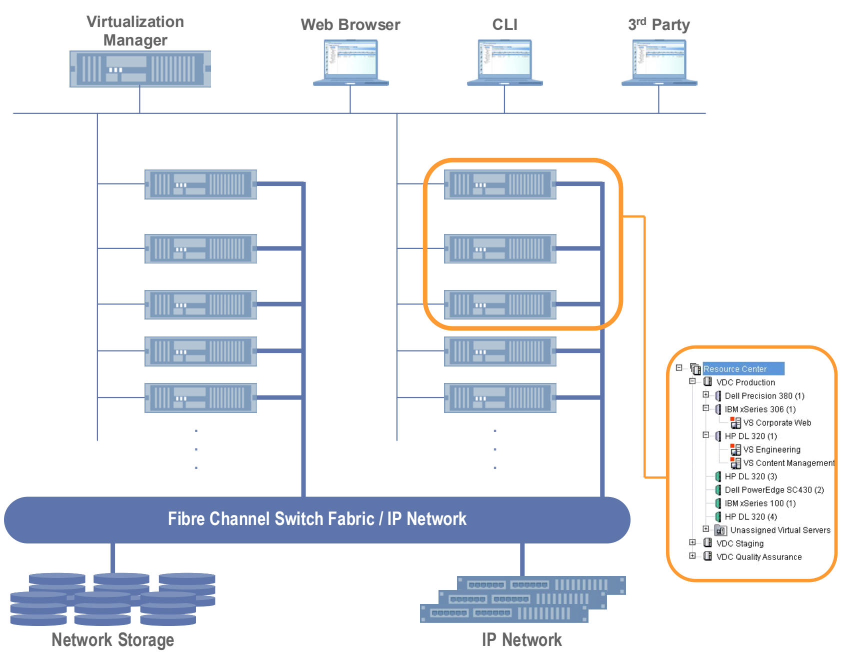
Components of a Virtualization Solution:

Virtualization solutions typically consist of the components identified in the table below. Virtual Iron's license terms for each component are also shown.

Component	License	Function
Hypervisor	GPL	First software loaded when physical server boots. Manages all hardware resources.
Service Partition	GPL	Second software loaded when physical server boots. Manages virtual server creation and configuration and all I/O.
Virtualization Manager	Commercial	Controls virtual servers through an agent in the service partition.
Guest operating systems	Varies	Operating systems that are fully virtualized on a physical server.

These components are shown in **Figure 2**. The Virtualization Manager communicates over a secure network to discover physical hardware resources and allow users to create, configure and manage virtual infrastructure (e.g. create virtual servers that consist of resources such as CPUs, memory, network connectivity, and disks).

Figure 2



Virtualization Services

Virtualization Services manage virtual infrastructure running on a physical server. They are designed to be:

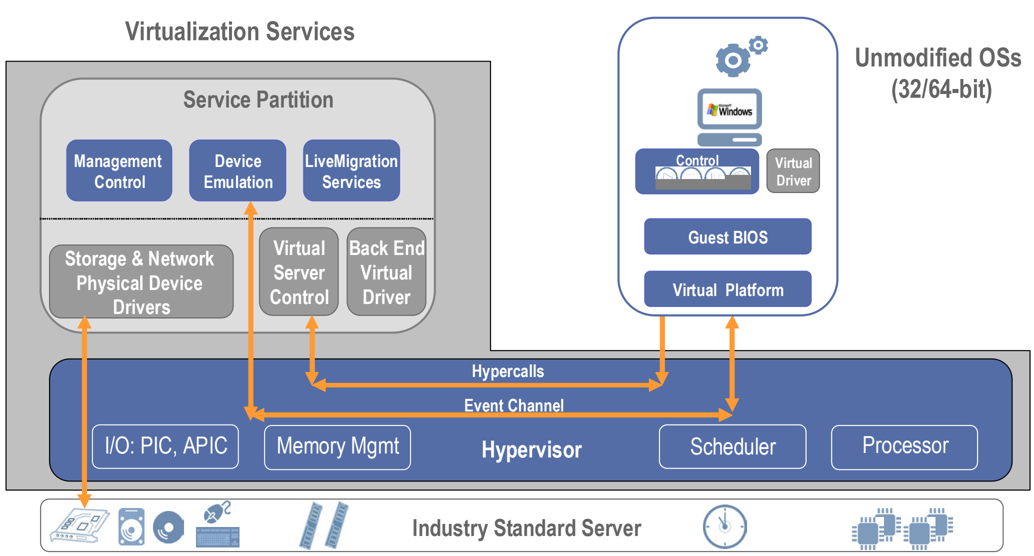
- Secure: Extremely compact with no external access except through the Virtualization Manager.
- Reliable: Based on an industry-standard Linux kernel and device drivers.
- Scalable: Automatic deployment option; no software to install or manage on each server.

This approach is different than other virtualization solutions that rely on a full host operating system or console operating system.

Figure 3 provides an architecture overview of the components in Virtual Iron's Virtualization Services. An open source hypervisor derived from the open source project is the first software loaded when the physical server boots. It manages all hardware resources - such as registers, memory and I/O devices. The hypervisor shares resources between virtual servers to ensure that they each receive a designated allotment of processing time and isolates resources to provide each guest operating system with the same degree of protection as if it was deployed on a separate physical host.

When the hypervisor starts, it launches the service partition as a privileged domain (also known as Domain-0) that can touch all hardware and control the hypervisor. The service partition consists of device drivers and modules that provide functionality such as remote management and accelerated I/O. It is completely hidden from user access and has no console login or persistent state. This means there are no additional components to manage or patch, resulting in greater uptime and less administrative costs. The service partition is controlled by the Virtualization Manager through an embedded management agent.

Figure 3



Network and storage I/O are virtualized to allow virtual servers to share physical connections to external networks while preserving security. The service partition exports a subset of the devices to the unmodified guest operating systems based on the virtual server's configuration. Guest operating systems can use either emulated devices or accelerated drivers for I/O.

Accelerated drivers have a back-end device in the service partition that is connected to the physical device (e.g. NIC) using a standard device driver and through memory to each guest operating system's front-end for that device. All messages between the front- and back-end drivers are queued to the designated guest OS. Device drivers come from the hardware manufacturer along with certification to the hardware. This ensures that operating systems and applications in the virtual environment maintain certifications and operate as intended.

An event channel through the component labeled Virtual Server Control in Figure 3 relays messages from the Virtualization Manager to the guest operating systems. Example communications include start, stop and pause requests, and the collection of performance statistics for CPU, memory, and I/O.

Virtualization Services Capabilities

Features	Benefits
Open source virtualization foundation	Industry standards provide rapid innovation and prevent vendor lock-in.
Bare-metal virtualization services	Virtualization interfaces directly with the hardware to provide the highest performance.
Automatic virtualization services deployment	No software is installed on physical hardware, simplifying deployments and upgrades.
Efficient server partitioning (multiple multi-processor guest operating systems can run concurrently on the same physical hardware).	Consolidation improves physical server utilization.
Support for up to 8 CPUs, 32 and 64-bit operating systems, 96 GB RAM and 4 vNICs per virtual server	Supports everything from small to large multi-processor virtual servers – key for handling enterprise-class workloads.
LiveMigrate™: The ability to move a virtual server between physical servers without any application downtime.	Gain additional processing capacity to seamlessly handle peaky workloads, perform maintenance on physical hardware without impacting maintenance windows.
Fault isolation	Prevents a crash in one virtual server from impacting other virtual servers.
Security isolation	A virtual server can never access the memory or I/O operations of another virtual server.
Resource isolation	Runaway applications in one virtual server do not cause other virtual servers to starve.
Hardware-assisted native virtualization	Low virtualization overhead delivers fast performance on resource-intensive workloads.
Resource monitoring	Heartbeat and virtual server CPU, memory, network and storage monitoring to support auto recovery of virtual servers in event of failure and make workload balancing decisions.

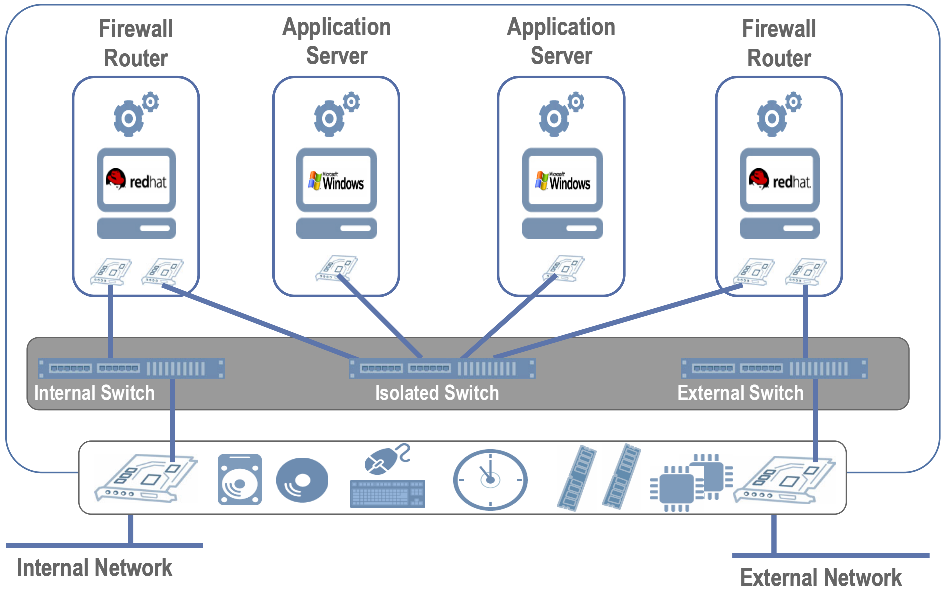
Virtual Networks

The Virtual Iron Virtualization Manager supports the creation of multiple virtual network switches that allow virtual servers to connect to internal (virtual) and external (physical) networks. Virtual switches are set up once and can be used on any physical server that is connected to the correct physical network. Each virtual server can support up to 4 virtual network cards, with unique MAC addresses, that connect to a designated network. Virtual NICs migrate with the virtual server, allowing network security settings to remain constant even when running on different physical servers.

Virtual Iron supports two virtual network devices. The default network device emulates the AMD PCnet32 PCI network adapter. VSTools provides a high-performance driver that connects to a back-end driver in the service partition. This approach provides lower latency and uses less CPU overhead than the default network device.

Internal and external switches allow network configurations such as in Figure 4, where the application server network does not exist in the physical world, guaranteeing no external network connectivity.

Figure 4



Virtual Storage

Virtual Iron supports SAN, NAS, and local storage through two methods, shown in Figure 5:

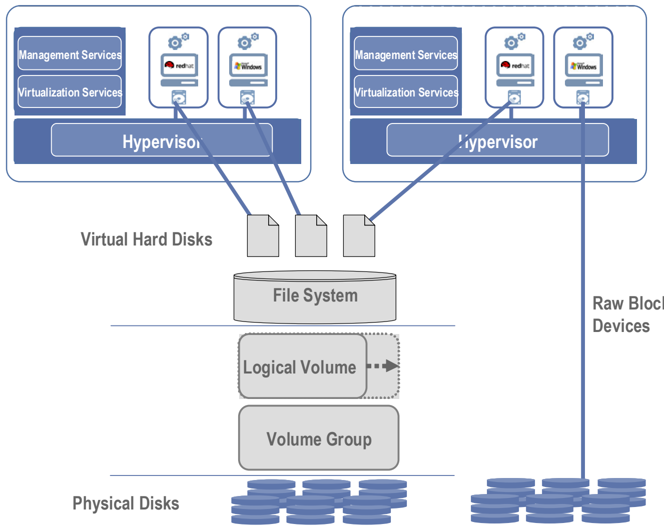
- Direct access to SAN LUNs, shown on the right side, provides very low virtualization overhead, supporting applications with high I/O requirements, such as databases.
- Virtual hard disks, shown on the left side, look like real disks to the unmodified guest operating system but are files that can be copied, moved, archived and backed up as easily as any other file.

Virtual Iron allows administrators to create dynamic storage as shares of physical volumes. Virtual hard disks, representing actual hard disks to a virtual server, are created in the logical volume. Virtual disks can be imported and exported to a standard file system. A virtual server, with all its virtual disks, can be archived, restored, or cloned on the Virtualization Manager. Virtual hard disks have the following benefits:

- New virtual disks can be created without requiring a physical volume to be partitioned. This allows disks to grow and shrink without impacting other virtual servers.
- Virtual disks are sparse files that take up only the space being used at the current time. This supports a pay-as-you-go model by allowing virtual disks to grow without requiring all the physical space to be allocated ahead of time.

Virtual servers access storage as a block device, supporting boot and data access. A virtual server can use either a default storage device or VSTools. The default storage device emulates an IDE device and requires no additions to the guest operating system, but provides lower performance and higher CPU overhead than the accelerated driver provided in the VSTools. VSTools provides a high-performance virtual block device that should be used for high performance applications.

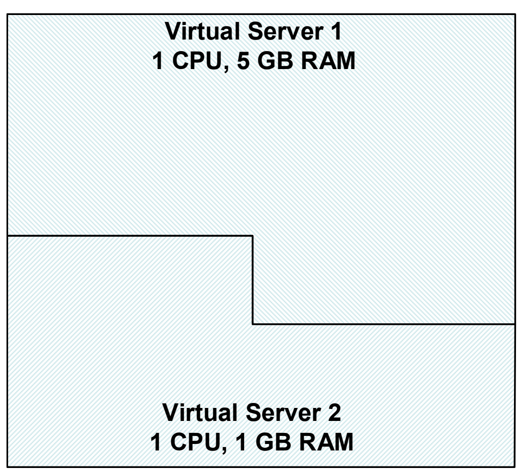
Figure 5



Virtual Processors and Memory

Users can specify virtual server resource requirements, such as number of CPUs, memory size, and priority. Virtual servers are scheduled by the hypervisor to time-share physical processors based on the priority of each virtual server on the virtualized node. A virtual server's priority determines the percentage of CPU time it receives. The hypervisor credit scheduler provides time slice guarantees as well as weighted CPU sharing to each virtual server. CPU affinity is supported in the virtualization services but is not currently exposed in the Virtualization Manager. The scheduler optimizes virtual server performance by not scheduling virtual servers that are idle, allowing their time slots to be "stolen" by other virtual servers.

Figure 6



While virtual servers can share CPUs and network resources, they must have dedicated memory and storage. Virtual Iron partitions memory as requested by the virtual server's configuration. The memory can be from 250MB to the total memory available in the physical server, minus overhead for the Virtualization Services. Multiple virtual servers can run on one physical server in complete isolation. This means that if a physical server has 1 CPU and 6 GB RAM, two virtual servers can share the resources as shown in **Figure 6** above.

Virtualization Manager

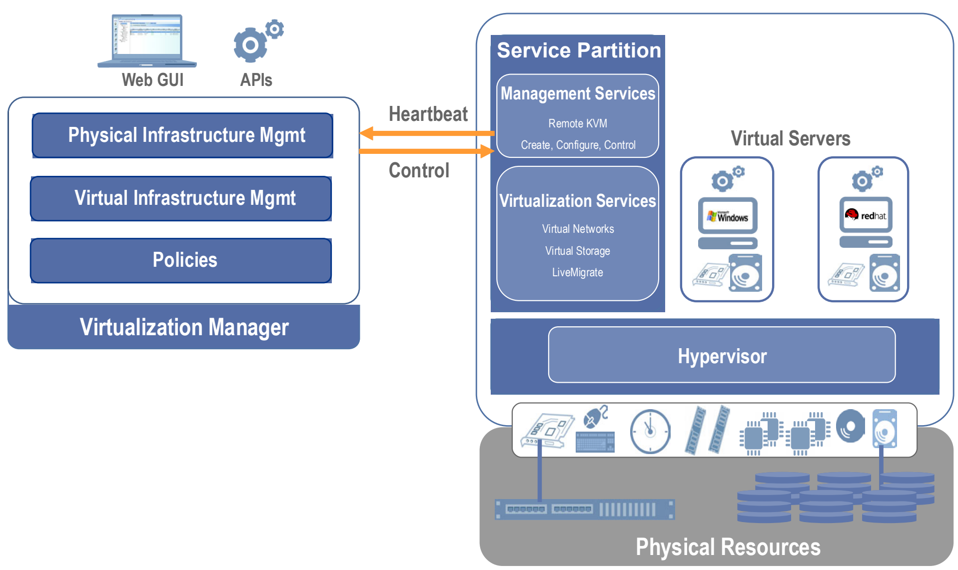
Virtual Iron's Virtualization Manager provides a central place to control and automate virtual resources. It is a Java application with a client-server architecture and a high performance distributed object oriented database. The user interface uses a transactional job-based model to provide fault tolerant workflows with rollback. The Virtualization Manager's built-in policy engine and event monitor allow users to customize the environment to optimize application performance, ensure availability, and simplify resource management. A remote virtual desktop provides graphical console, keyboard and mouse without client or server-side additions.

Virtualization Manager provides the following capabilities:

- Physical infrastructure: Physical hardware discovery, bare metal provisioning, configuration, control, and monitoring
- Virtual Infrastructure: Virtual environment creation and hierarchy, visual status dashboards, access controls
- Virtual Servers: Create, Manage, Stop, Start, Migrate, LiveMigrate™
- Policy-based Automation: LiveCapacity™, LiveRecovery™, LiveMaintenance™, Rules Engine, Statistics, Event Monitor, Custom policies
- Reports: Resource utilization, System events

These capabilities streamline tasks that are normally highly manual and time-intensive and significantly reduce data center complexity. For example, users can consolidate many physical servers using virtual servers, quickly create new virtual servers to deploy operating systems and applications using templates and cloning, reduce the cost and complexity of business continuity, and automate the data center using policies.

Figure 7



Virtualization Manager consists of the components shown in Figure 7 above. It is connected to virtualized nodes, shown on the right, to control virtual servers and receive status and statistics. The Virtualization Manager can be controlled by a web-based client, command line, or third party utilities using the API.

The Virtualization Manager user interface, shown in Figure 8 below, is organized into three panes. On the far left is the control pane that allows the user to choose between a view of virtual resources, historical view of jobs, policies and reports, hardware, and user accounts. The middle pane provides a tree view of physical and virtual servers, organized by virtual data centers. The right pane shows details of what has been selected in the tree.

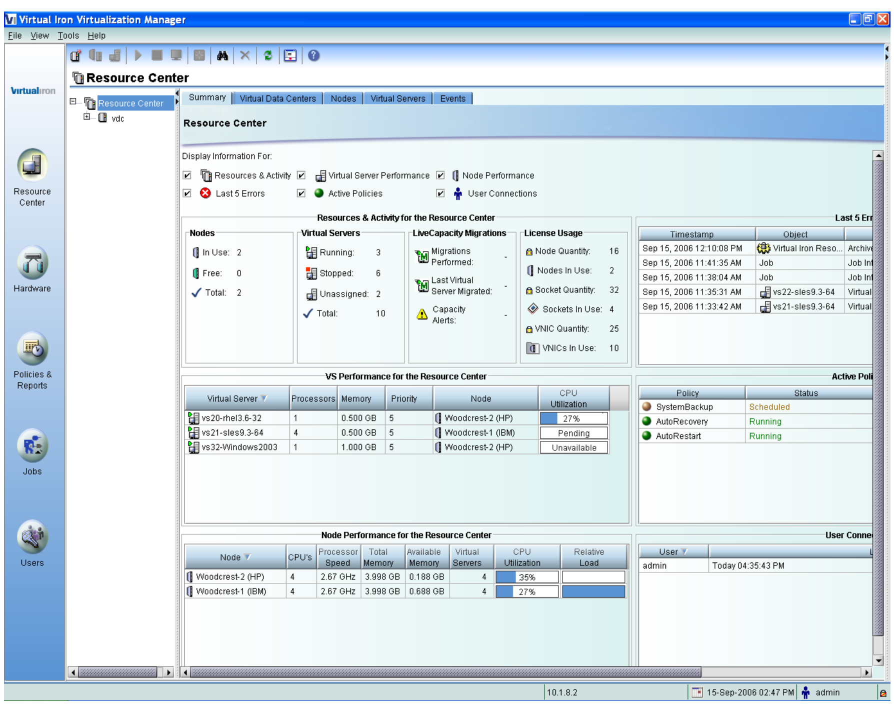
Users discover the hardware resources that will be used in their virtual environment. The hardware view displays a list of servers, NICs and their associated networks, and HBAs and their associated disks.

The Resource Center provides the interface for creating, configuring, and managing virtual resources. This view also includes dashboards that allow users to monitor physical and virtual resource activity. It is organized into Virtual Data Centers that consist of physical servers and virtual servers. It is in this view that a virtual server can be created and deployed. Resources can be added or removed from virtual servers depending on the needs of the workload. Resource changes can be automated using policies to reduce system administration.

Virtualization Manager includes policy-based automation to reduce administration. Many policies are based on LiveMigrate, which enables workloads to move from one server to another without stopping or impacting the running application.

- LiveCapacity monitors virtual server CPU utilization or other application needs to determine when a workload needs additional capacity. When a user-defined threshold is met, the virtual server is LiveMigrated to a physical server that has the necessary resources.
- LiveRecovery monitors the status of physical resources and moves virtual servers to maintain uptime in the event of a hardware failure.
- LiveMaintenance moves virtual servers to alternative locations without downtime when a physical server is taken offline for maintenance. This allows physical server maintenance to be performed outside of scheduled maintenance windows without incurring any downtime.

Figure 8

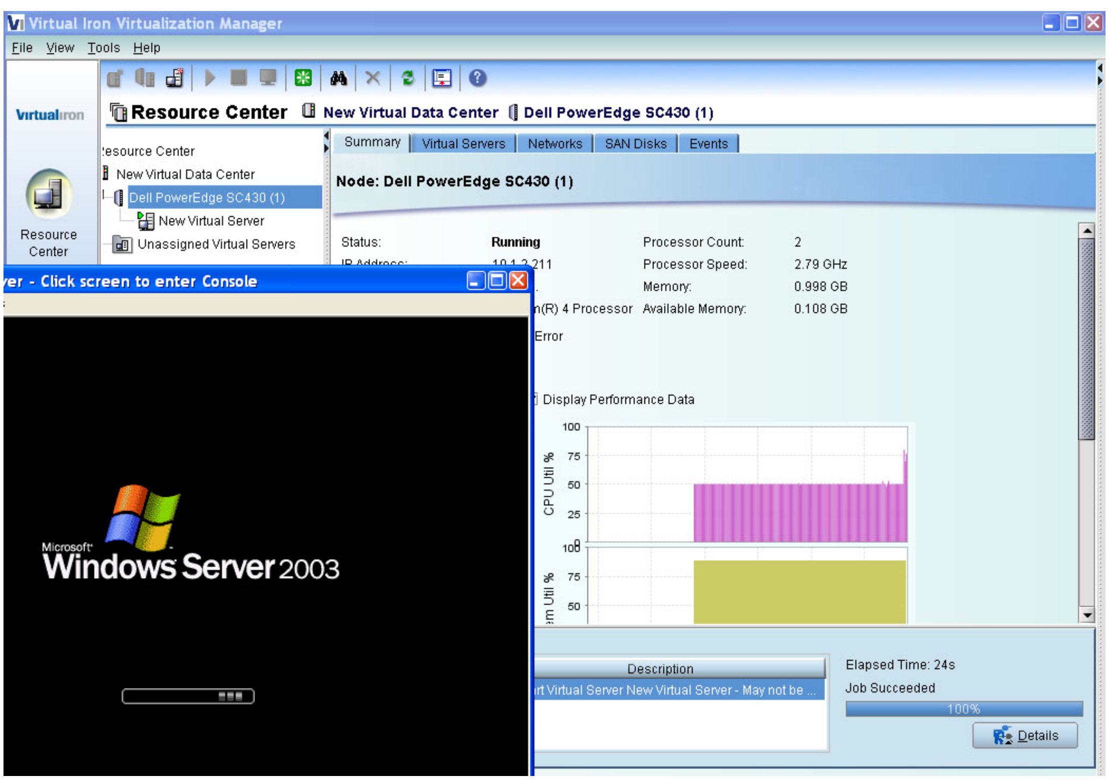


Virtualization Manager Capabilities

Features	Benefits
Centralized resource management	Unified management of both virtual and physical resources.
Hardware Discovery	Automatic discovery of physical resources such as servers, network and storage.
Policy-driven resource and workload management	Automates application management to SLAs and reduces administrative overhead.
Performance and availability reports	Provides capacity planning, what-if scenarios, and post-incident analysis.
Virtual server cloning	Use operating system templates to simplify server deployments.
LiveRecovery	Move virtual servers to alternate physical servers in the event of a hardware failure. Reduces the hardware required for business continuity.
LiveCapacity	Provide additional capacity to operating systems and applications by migrating to servers that have additional processing capacity.
Jobs and Alerts	Uses transactional management to provide audit trail and notifications of data center reconfigurations.
User management	Access control for delegated administration.
Virtual switches	Manage network and fibre-channel settings centrally.

The screenshot in Figure 9 shows a virtual server dashboard with a current and historical view of resource utilization. The console operates in VGA mode with remote mouse and keyboard.

Figure 9



Virtual Server Configuration

Virtual servers run unmodified guest operating systems. VSTools can be installed in the operating system to provide accelerated drivers for network and storage as well as operating system control from the Virtualization Manager.

Virtual servers have a number of configuration options:

- Number of processors
- Amount of memory
- Priority for scheduling
- Virtual NICs for Ethernet network connections
- Virtual disks
- Boot source: CDROM, PXE, RAW SAN disk, virtual disk (from SAN or local drive), network block device

Virtual Iron's partners provide utilities that convert physical servers to virtual servers.

LiveMigration and Dynamic Scaling

LiveMigration moves a running virtual server from one physical server without pausing or impacting running applications. LiveMigration, combined with dynamic scaling, allows virtual servers to take advantage of different amounts of resources on destination physical servers to enable capabilities such as load balancing and high availability.

LiveMigration consists of several steps:

Step 1: Virtual server resources reserved on destination server. The CPU resources do not need to be identical, but there must be sufficient memory and access to the same Ethernet subnets and storage on the destination server.

Step 2: Copy virtual server memory over Management Network. The disk remains on network storage. The length of time it takes to copy the virtual server's memory depends on a number of factors such as how much memory is in use, the speed and latency of the network connection, and how much the memory is changing. In general, this process takes minutes, however the virtual server continues to run the entire time on the original server, meaning application availability is unaffected. This operation does not impact the network traffic on the public network.

Step 3: Complete migration. This step stops the virtual server on the origin, copies any changed memory, and starts the virtual server on the destination. This process takes milliseconds.

LiveMigration has the following requirements:

- Both physical servers must be able to access the same networks and storage as the virtual server is using
- Both physical servers must be in the same virtual data center
- Both physical servers must have processors in the same family (i.e. Intel or AMD)

Policies & API

Virtual Iron provides an extensive API that can control any aspect of the virtual data center. The API is accessible within the Virtualization Manager and through an external interface. The API may be used for writing policies or integrating with third party applications.

Policies provide the ability to automate any Virtual Iron capability and can be scheduled to run once or on a recurring basis. Example policies include historical performance reports, or virtual server LiveMigration on a nightly basis to provide additional overnight processing capacity.

Virtual Iron also provides a Common Information Model (CIM) interface for integrating with third party management applications. Developers can use the CIM standard to explore and manage the virtual environment.

Supported Configurations

Feature	Support
Operating systems	• 32 and 64-bit Red Hat Enterprise Linux 4 • 32 and 64-bit SUSE Linux Enterprise Server 9 • 32-bit Windows XP • 32-bit Windows 2003
Processors	• Intel Xeon with Intel VT • AMD Opteron with AMD-V
Virtualized nodes	Up to 1000 per virtual data center
Processors per Virtual Server	Up to 8
RAM per Virtual Server	Up to 96GB
Virtual servers per physical server	Up to 5 per CPU
Virtual NIC adapters per Virtual Server	Up to 4

About Virtual Iron

Virtual Iron provides enterprise-class software for creating and managing virtual infrastructure. Its software enables companies to dramatically reduce the complexity and cost of managing and operating their data center. Virtual Iron delivers advanced virtualization capabilities that leverage industry standards, open source economics and processors with virtualization technology built-in. Organizations use Virtual Iron's software for server consolidation, rapid provisioning, business continuity, capacity management and policy-based automation to deliver significant improvements in utilization, manageability and agility. For more information, visit www.virtualiron.com or email info@virtualiron.com.

Virtual Iron Software
900 Chelmsford Street
Tower I, Floor 2
Lowell, MA 01851
978.849.1200
www.virtualiron.com