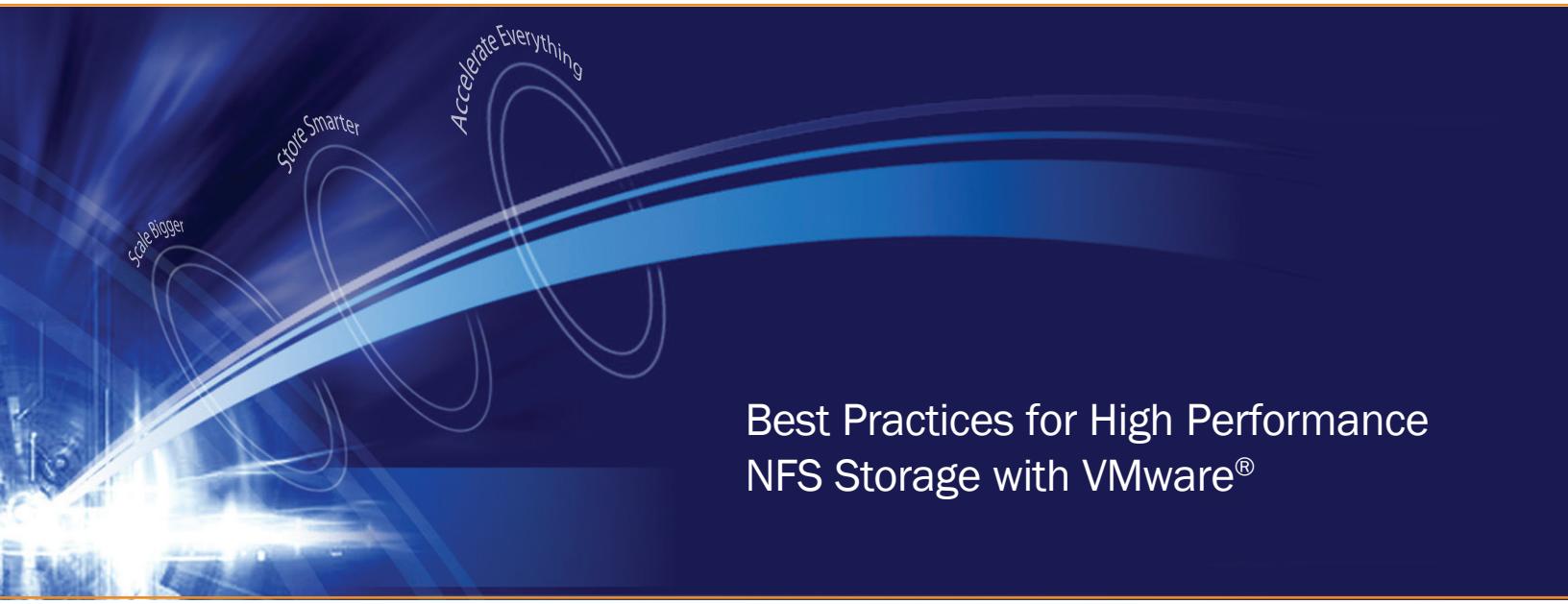
WHITE PAPER



Best Practices for High Performance NFS Storage with VMware®

BLUE ARC®

Executive Summary

The proliferation of large-scale, centralized pools of processing, networking, and storage resources is driving a virtualization revolution in the data center. These changes demand a reassessment of the assumptions underlying traditional datacenter design. One of the most glaring examples of this phenomenon is the erroneous belief that high-performance, flexible, scalable, centralized data storage can only be achieved through the deployment of a Fibre Channel Storage Area Network (SAN). BlueArc® Network Attached Storage using SiliconFS® renders this notion obsolete.

BlueArc NFS technology offers a storage solution for the virtualized data center that meets or exceeds the level of performance, manageability, and scalability of a SAN, without requiring the purchase and deployment of expensive and complex Fibre Channel fabric. BlueArc storage, built on proven IP and NFS technology, is an excellent storage platform for VMware® vSphere™. Data center managers deploying vSphere should consider BlueArc NFS as a cost-effective solution that meets the performance requirements of even the most demanding VMware environments.

The Data Center Is Evolving

Few events have been as truly disruptive to the practice of Information Technology (IT) as the introduction of virtualization to the data center. Virtualization is the inevitable result of the proliferation of extremely powerful central processors, massive computer memory, faster networks, and vast amounts of centralized data storage. This technology is driving a fundamental shift in data center design and operations.

Of the virtualization platforms, VMware's vSphere is the most widely adopted. With ~89% of the virtualization market, VMware is also the most mature virtualization technology. VMware vSphere achieves unparalleled scalability, supporting up to 200 virtualization hosts and 4500 virtual machines (VM) per vCenter™ management server instance.

VMware virtualization improves the operational agility and efficiency of data centers through:

- workload consolidation
- automatic resource optimization/efficiency (DRS)
- VM mobility (vMotion™)
- distributed power management (DPM)
- improved availability and quality of service (HA)
- achievable Business Continuity and Disaster Recovery (BC/DR)
- dynamic provisioning and de-provisioning
- increased scalability
- improved automation

Given these benefits, it is clear why the spread of virtualization is accelerating.

The adoption of virtualization has brought with it a fundamental change in the way information services are managed and provisioned. Virtualization has shortened the provisioning cycle greatly. Where it once took weeks or months to provision additional servers, equivalent VMs can be instantiated or cloned in minutes.

The files that constitute a VM image (VMDK) can reach up to two terabytes in size. As they are simply files, VMDKs can be moved, copied or deleted. A VM can be cloned by copying

its files, and each clone becomes an operating VM of its own. VM clones are a powerful means to add service capacity quickly. However, each VMDK copy will consume the same amount of disk space as the original VM. The ease of provisioning has led to a substantial increase in the amount of data that must be stored and managed.

As one can imagine, the amount of disk space required to house these images can quickly grow to astronomical proportions. The rise of large, complex storage has collided with a fundamental change in the behavior of storage traffic. The pooling of data center resources, the ability to instantiate additional VMs dynamically, and the freedom to migrate a VM from host to host are all contributing to a fundamental change in storage scale, at the same time that the nature of storage traffic is morphing due to virtualization. The proliferation of virtualization and its complexity is driving the need for powerful, intelligent storage.

Centralized Storage Is Key

The operational benefits of virtualization are widely understood. However, it should be noted that the most powerful features of virtualization technology are enabled only by the use of centralized, shared storage. Features such as VM Migration (vMotion) Distributed Resource Scheduler (DRS), and High Availability (HA), rely on a centralized storage infrastructure that can be shared and managed by all hosts in the resource pool or HA cluster. Business Continuity and Disaster Recovery (BC/DR) solutions are built upon the replication of the VM images from one site to another. This technology also relies on centralized storage. If an organization wishes to…

- achieve operational efficiency and agility
- manage the complexity and diversity of virtualized workloads
- support the massive scalability of VMware resource pools
- make use of power of vMotion, DRS, HA, and BC/DR
- leverage the real power of virtualization

… centralized storage infrastructure is essential.

Storage Performance is Critical

While scalable, centralized storage is necessary to enable the real power of virtualization, the performance of that storage is equally critical. Depending on the type of storage, a single centralized storage volume in a VMware environment can house hundreds of VM images. Each active VM will generate disk reads and writes (I/O) nominally equivalent to those of an active physical host. A storage volume housing a number of VMDK images must be able to support the aggregated disk activity of all of the VMs running simultaneously. The higher the number of VMs stored on a storage volume, the more disk activity that volume will experience. A centralized storage solution must not only offer scalability, it must also support an extremely high level of performance VMs require.

Re-Evaluating NFS Storage with VMware

Historically, for reasons primarily tied to VMware's development and release processes, the vast majority of VMware deployments have been implemented using a SAN. However, while VMware initially deemed development of support for SAN storage a higher priority, subsequent VMware releases have brought support for NFS to a level of parity with SAN. Not only are SAN and NAS equally supported by VMware, using NFS with VMware actually offers benefits unavailable from SAN technology.

An obvious advantage of NFS is the level of familiarity and expertise many data center administrators have with NFS technology. Based on time-tested Ethernet/IP networking and NFS protocols, NFS technology offers administrators a familiar management paradigm, greater visibility, and easier troubleshooting than VMFS/SAN storage. Administrators understand NFS and know how to manage it.

NFS storage I/O performance and availability can be readily optimized. Storage traffic is more likely to be impacted by latency and congestion than by bandwidth limitations. NFS storage reliability and performance can be enhanced using familiar network optimization techniques, such as redundant network connections, 802.3ad link aggregation, and network Quality of Service (QOS). Network engineers are well versed in designing and optimizing Ethernet/IP networks to prevent latency and congestion, and have well established techniques to manage bandwidth. This ability serves NFS storage very well.

Using NFS storage with VMware allows the ESX host to leverage the flexible I/O afforded by Ethernet networking. NFS traffic is not subject to the I/O constraints of VMware's SAN I/O queuing. Unlike a SAN volume, NFS I/O uses any of the numerous network I/O queues, and not a SAN volume's specific I/O queue. A single NFS export can also have multiple network addresses to support clustering. This flexibility allows the host and the network to work cooperatively to dynamically optimize the flow of storage traffic, allocating bandwidth as needed, and assuring that the storage traffic will reach its destination reliably and efficiently.

NFS storage provisioning with VMware is exceedingly simple. NFS exports can be created and presented to VMware virtualization hosts both quickly and non-disruptively. Unlike SAN storage, NFS volumes can be grown or shrunken without requiring a reconfiguration of a VMware host's storage profile.

VMware's default "thin provisioning" on NFS recognizes the inherent efficiency and flexibility of NFS storage. By default, VMware does not provision the total disk space allocated to accommodate a VMware VMDK on an NFS data store. Instead, a portion of the space is made available immediately, with the remainder being allocated as needed. As a result, unused space in an NFS volume can be dynamically assigned to the VMs that are in need, assuring that storage resources flow where they will be used most efficiently. Thin provisioning greatly improves the efficiency of disk use and maximizes the return on investment in NFS storage.

NFS storage delivers a 10-fold increase in efficiency per volume. VMware recommends that no more than 10-40 VMDKs be hosted on a given SAN volume using VMware's VMFS file system. In the case of an NFS volume, the SAN/VMFS limitations are not encountered. As a result, the recommended number of VMDKs per store increases to 100-400; a number that provides room for growth. Clearly, NFS storage offers much greater scalability than a SAN.

NFS storage offers very compelling benefits for a VMware environment. When optimally configured, NFS will deliver performance and scalability that is far superior to that of a SAN. NFS benefits include:

- Simple, rapid storage provisioning
- Efficient disk utilization through VMware's default VMDK "Thin Provisioning" on NFS
- Transparent, non-disruptive volume management
- Powerful performance optimization and redundancy leveraging Ethernet\IP networking
- Better storage scalability and support for more VMs per volume

BlueArc Advantages

BlueArc's network storage system technology offers administrators a set of features that makes NFS an even better storage choice for virtualized environments. Leveraging the power of SiliconFS, BlueArc storage systems offer manageability, scalability and performance that exceed the capabilities of NFS products from other manufacturers. BlueArc's pedigree as a provider of storage solutions for High Performance Computing (HPC) environments assures administrators that BlueArc can meet the storage requirements of even the most demanding virtualized data centers.

BlueArc's specialized storage processors and unique SiliconFS file system are optimized to work together to service storage requests with unparalleled efficiency. The I/O performance of NFS products using traditional file system technologies and multi-function processors will be limited by the ability of those processors to service storage requests. As the volume of storage traffic increases, traditional processors will bog down, degrading storage performance rapidly. Increasing storage I/O capacity will require the purchase of additional hardware. BlueArc's storage processors are specially designed to meet the unique demands of storage I/O, even at scale. Not only will BlueArc storage provide greater throughput per storage server, the performance of BlueArc storage will be sustained as the demands on each server intensifies. Virtualization administrators will realize greater efficiency and more predictable performance from each BlueArc storage server, improving ROI, even as the number of VMs on a volume increases.

BlueArc's "Cluster Namespace" (CNS) technology offers storage administrators namespace scalability that is unique in the industry. CNS can place up to sixteen petabytes of storage within a single namespace. Using CNS, storage administrators can present all the hosts in an environment with the same view of even large amounts of complex storage. Powerful VMware features such as HA and DRS require that storage be presented consistently to all the hosts in a pool or cluster. The ability to manage and present complex storage as a single, logical structure greatly simplifies storage provisioning, allowing all hosts to access storage in a consistent manner, and providing manageable, scalable storage that meets the needs of large virtualization resource pools.

The unique scalability of BlueArc storage prevents the creation of isolated "data silos". As storage deployments increase in scale, many customers encounter the scalability limits of NFS storage products from other vendors. The only way to overcome the limitations of these products is to purchase additional storage hardware. Adding hardware raises the cost of that storage solution, raises the size of its data center footprint, and greatly increases managerial complexity. For example, when the namespace limit of a competing NFS product is reached, added storage must be placed behind a new storage server, creating an isolated island of storage that cannot be integrated into the existing storage fabric. BlueArc's 16 petabyte namespace assures that administrators will not encounter namespace size limitations. Storage added to BlueArc managed infrastructure can be integrated into the existing fabric seamlessly, without requiring the purchase of additional storage hardware and the creation of segregated storage silos.

BlueArc storage offers a uniquely expandable file system that can be grown, nondestructively, to provide space for expanding VMs. The ability of SiliconFS to expand the file system on-the-fly is especially compelling when combined with VMware's default "thin provisioning" on NFS stores. In the event that a set of thinly provisioned VMs starts to approach the allocated space in a data store, the BlueArc volume can be easily expanded to accommodate the growth of those VMDKs. BlueArc's expandable file system provides assurance that a VMware data store will always be able to grow as the number and size of VMs in the environment grows.

BlueArc's SiliconFS offers a 16X increase in file system size above that which is offered by other NFS vendors. Products from most NFS vendors support a file system of only 16 terabytes. In contrast, BlueArc NFS supports the creation of file systems of up to 256 terabytes. The ability to create file systems of this capacity assures virtualization administrators that they will not run into the size limits of any BlueArc storage volume, no matter how many VMDKs that volume may need to host, or how large a virtualized environment may grow.

BlueArc's "Enterprise Virtual Server" (EVS) multi-node NFS clustering capability provides optimized, flexible, highly reliable access to VMware data stores. Using BlueArc's EVS technology, storage being managed by BlueArc Mercury or Titan storage management can be "virtualized" to provide:

- flexible, multi-tiered disk provisioning across disparate levels of storage
- cooperative load balancing and redundancy between BlueArc storage nodes
- powerful, automated optimization assuring that highly-active VMs can meet the demands for critical business services

BlueArc nodes can share IP addresses and storage labels, presenting a single, easy-to-manage view of the storage layer. Storage access can be shared by multiple nodes to optimize throughput and availability. BlueArc EVSs provide VMware hosts with a set of broad, reliable, high-performance paths to critical VMDK data stores.

Given these capabilities

- Unique namespace scalability up to 16 petabytes
- Unique volume scalability up to 256 terabytes
- Unique nondestructive, expandable file system
- Flexible virtualization and management of the storage layer
- Sustained high performance, even at scale

BlueArc NFS is clearly a powerful storage solution that is very suitable for use in a VMware virtualized data center. BlueArc NFS offers flexibility, scalability, and performance advantages over a SAN, and SiliconFS improves the NFS value proposition even further. BlueArc storage maintains scalability and performance, at scale, that is an order of magnitude greater than that available with any other NFS product.

Optimizing NFS with VMware

While using NFS storage with VMware offers many advantages, it is important to understand how to configure the two technologies to work together to deliver optimal performance. Below are some architectural recommendations and configuration parameters that should be considered to successfully implement BlueArc NFS storage with VMware.

Network Configuration

- **Segment NFS Storage Traffic -** Isolate standard network traffic from NFS storage traffic. This isolation can be accomplished using dedicated network interfaces, virtual LAN segments (VLANs), and/or dedicated network switches. Dedicated network interfaces are preferred, where possible.
- **Use Gigabit Ethernet Networking -** Use network interface cards (NICs) with a minimum speed of one gigabit. Ten gigabit Ethernet is supported by ESX and BlueArc's NFS hardware and may be appropriate for especially large or busy environments.
- **Enable Inter-Switch Links -** If possible, consider using network switches that support inter-switch linking (ISL, Etherchannel or Virtual Port Channeling, for example). This

technology can be used to interconnect the switches through which storage network traffic passes. ISL offer the possibility of additional throughput, congestion avoidance, and increased reliability in the event of a network outage

- **Avoid Network Over-Subscription in Storage Networking –** Although network oversubscription is a standard and very necessary practice in network management, network connections between the host and the storage infrastructure should not be oversubscribed. Storage traffic is particularly sensitive to latency, and packet loss. Network retransmission cause by congestion should be avoided, as packet loss can have a detrimental impact on storage reliability and performance.
- **Use Network Link Aggregation –** Link aggregation (802.3ad) allows multiple network interfaces to act as a single logical connection, sharing the traffic load. NIC teaming aggregates network bandwidth, increasing the total throughput available for storage traffic. It should be noted that NIC teaming does not provide network redundancy, as network sessions will follow an established path while that path remains available.
- **Select a NIC Teaming Session Identifier–** NIC Teaming behavior is affected by a IP Session Identifier. Consider using "IP Hashing" as the algorithm to identify each network sessions. IP Hashing is preferred as it minimizes the computation necessary to calculate session identifiers, decreasing the networking overhead experienced by the host. There are other opinions as to which algorithm setting is best, as other session algorithms may be more appropriate for your environment. For additional information, please see VMware KB 1004048.
- **Use Jumbo Frames from End to End -** Configure all network components between the VM and the storage layer to use Jumbo Frames. This technology improves network efficiency by increasing the amount of data that will be transferred by each network event. (For the specific ESX configuration parameter, see below.)
- **Enable Network Flow Control -** Flow Control allows network components to notify their neighbors of congestion events, allowing them to adjust the flow of traffic accordingly. This feature decreases the likelihood that the network will experience congestion, and lowers the probability that a packet will require retransmission.
- **Disable Spanning Tree on Links to VMware Host Interfaces –** The Spanning Tree protocol is intended to prevent the introduction of bridging loops in a local area network. Bridging loops cannot be created in a VMware environment unintentionally. This feature is not necessary on links to VMware hosts as hosts do not forward packets to VMs unless those VMs are present and running on that host. Spanning Tree can be disabled on network connections to virtualization hosts with no risk or impact.

VMware Configuration

- **Increase the Number of NFS Data Stores -** It is recommended that the number of NFS stores a VMware host can mount simultaneously be increased from the default value of "8" stores. While a VMware host can support up to "64" stores, the value selected should reflect the needs of each environment. The configuration change can be made by raising the "NFS_MaxVolumes" parameter found in the vCenter® advanced networking settings dialogue.
- **Increase Network Buffers -** Administrators should adjust the size of the buffers allocated for network I/O in the VMware vmkernel to support the demands of the higher volumes of network traffic generated by the use of NFS storage. These changes are particularly important when increasing the maximum number of NFS data stores a VMware host can mount. Two configuration variables must be increased to affect this change.
    - » Increase the Net.TcpIpHeapSize value to 30

  - » Increase the Net.TcpIpHeapMax value to 120

- **Open NFS ports in the VMware Firewall –** By default, the service ports for the NFS client in VMware are not open, blocking the passage of NFS storage traffic. The service ports (TCP and UDP ports 111 and 2049) must be opened to allow passage of this type of traffic. These ports can be enabled by activating the "NFS Client" settings in the Security Profile Properties dialogue in vCenter

- **Enable Jumbo Frames -** Traditional Ethernet frames have been sized at a Maximum Transfer Unit (MTU) of 1500 bytes. VMware ESX 4.0 added support for jumbo frames, with an MTU of 9000 bytes. As was mentioned above, enabling jumbo frames increases the throughput of each network event, improving network efficiency. To configure Jumbo Frames with VMware

  - » Set the "dvSwitch MTU port value" to 9000 in the advanced settings dialogue of the vCenter Networking Configuration dialogue.

- **Enable NIC Teaming -** As was mentioned above, it is recommended to enable NIC teaming between the VMware host and the adjacent network switch. This option provides load balancing, aggregated bandwidth, and increased network resiliency to support NFS storage network requirements. NIC Teaming can be enabled in the network configuration section of each virtualization host within the vCenter interface. For more details, please see VMware KB1004088.

- **Disable NIC Teaming Failback -** NIC teaming also offers a "failback" feature that, in the event of a failure and subsequent repair of a teamed NIC, causes network traffic to return to the interface from which it was redirected when the failure occurred. It is recommended that the "failback" option be set to "No" for all hosts using NFS storage. Disabling failback will prevent sessions carrying NFS traffic from being disrupted by intermittent or unstable network interfaces. This setting can be found in the vCenter NIC teaming network configuration dialogue.

- **Consider Adjusting the NFS Heartbeat Values –** A VMware host using NFS storage will send a periodic NFS Heartbeat to verify that an NFS volume it is using remains available. Adjusting the VMware NFS heartbeat values can assure that the VMware NFS client allows adequate time for an NFS cluster failover to complete before prematurely marking NFS storage as unavailable. The three variables affecting NFS heartbeat and their default values are:

  - » NFS.HeartbeatFrequency (Default value of 9 seconds)
  - » NFS.HeartbeatTimeout (Default value of 5 seconds)
  - » NFS.HeartbeatMaxFailures (Default value of 3)

  Using these values allows only 32 seconds before an NFS volume is marked unavailable. Cluster failover events of NFS stores can require additional time to complete. In order to allow adequate time for the failover, the recommended values for these variables are:

  - » NFS.HeartbeatFrequency value of 12
  - » NFS.HeartbeatTimeout of 5
  - » NFS.HeartbeatMaxFailures to 10

  Using the new values, an NFS failure will only occur after 125 seconds. A NFS cluster failover will generally complete well within this period, allow the client VMs residing on the NFS data store to continue operating. For additional information, please see the VMware KB1007909.

While the configuration parameters described will improve the performance, reliability and scal-

ability of NFS storage when used with VMware, there are additional settings and options that are worth understanding. For a more complete list, please see VMware KB 1007909.

The Simplicity of Provisioning BlueArc NFS with VMware ESX

Provisioning a BlueArc NFS export to serve as a data store for use with VMware is incredibly simple. High performance BlueArc NFS exports can be provisioned and in use by VMware ESX in a matter of minutes following these simple steps.

Provision a BlueArc NFS Export

a. Log in to BlueArc's Mercury or Titan Management Interface

   i. Create an EVS - Assign a label and an IP address to identify the resource.

   ii. Create a file system – Define the size of the volume, its "Read-Write" policy, and its block size – Assign the appropriate block size for your environment.

   iii. Assign the new volume to the EVS defined above

   iv. Create an NFS export to allow the ESX server to mount the share

      1. Name the export appropriately

      2. Select the appropriate caching value to reflect the type of storage

b. Log in to VMware's vCenter Management Server

   i. Select the ESX Host -> Configuration -> Storage -> Add Storage

   ii.Select the Network File System option

   iii.Specify the IP address of the BlueArc EVS and the name of the export defined above

   iv. Assign the export a name within vCenter

The BlueArc export is now accessible and ready to host VMware VMDK files.

Conclusion: The Power of BlueArc NFS with VMware

BlueArc systems, leveraging the power of SiliconFS, deliver simplicity, performance, and scalability to meet the needs of even the largest VMware environments. The familiarity of NFS and the power of BlueArc provide administrators and managers with a cost-effective storage solution that can support thousands of VMs. Maintaining performance even at scale, BlueArc's flexible storage virtualization can manage all types of back-end storage transparently to ensure that highly active VMDKs reside on the storage they need to perform. BlueArc NFS is an excellent storage platform for VMware vSphere. Data center managers deploying vSphere should consider BlueArc NFS as a cost-effective solution that meets the performance requirements of even the most demanding VMware environments.

About BlueArc

BlueArc is a leading provider of high performance unified network storage systems to enterprise markets, as well as data intensive markets, such as electronic discovery, entertainment, federal government, higher education, Internet services, oil and gas and life sciences. Our products support both network attached storage, or NAS, and storage area network, or SAN, services on a converged network storage platform.

We enable companies to expand the ways they explore, discover, research, create, process and innovate in data-intensive environments. Our products replace complex and performance-limited products with high performance, scalable and easy to use systems capable of handling the most data intensive applications and environments. Further, we believe that our energy efficient design and our products' ability to consolidate legacy storage infrastructures, dramatically increases storage utilization rates and reduces our customers' total cost of ownership.

BLUEARC®

BlueArc Corporation
Corporate Headquarters
50 Rio Robles Drive
San Jose, CA 95134
t 408 576 6600
f 408 576 6601
www.bluearc.com

BlueArc UK Ltd.
European Headquarters
Queensgate House
Cookham Road
Bracknell RG12 1RB, United Kingdom
t +44 (0) 1344 408 200
f +44 (0) 1344 408 202