Making Everything Easier!™

Avere Systems Special Edition

NAS Optimization FOR DUMMIES®

A Wiley Brand

Learn to:

- Scale NAS performance separate from storage capacity
- Eliminate the three major sources of storage latency
- Gain the advantages of storage in the cloud

Brought to you by

AVERE®

Allen G. Taylor

About Avere Systems

Avere Systems brings to the market NAS Optimization solutions designed specifically to scale performance and capacity separately and take advantage of Flash storage media using real-time tiering. Avere's FXT Series Edge filers allow organizations to achieve unlimited application performance scaling, free applications from the confines of the data center by eliminating latency, and cut storage costs by more than half.

In a broad range of industries from media to financial services, from software development to oil and gas, and from life sciences to the Web, Avere customers are using its products to optimize their NAS infrastructure and increase the throughput and scalability of key applications and workflows, resulting in productivity gains that translate to higher revenues and happier customers.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

NAS Optimization FOR DUMMIES®

AVERE SYSTEMS SPECIAL EDITION

by Allen G. Taylor

WILEY

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

NAS Optimization For Dummies®, Avere Systems Special Edition

Published by
John Wiley & Sons, Inc.
111 River St.
Hoboken, NJ 07030-5774
www.wiley.com

Copyright © 2013 by John Wiley & Sons, Inc., Hoboken, New Jersey

Published by John Wiley & Sons, Inc., Hoboken, New Jersey

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without the prior written permission of the Publisher. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at http://www.wiley.com/go/permissions.

Trademarks: Wiley, the Wiley logo, For Dummies, the Dummies Man logo, A Reference for the Rest of Us!, The Dummies Way, Dummies.com, Making Everything Easier, and related trade dress are trademarks or registered trademarks of John Wiley & Sons, Inc. and/or its affiliates in the United States and other countries, and may not be used without written permission. All other trademarks are the property of their respective owners. John Wiley & Sons, Inc., is not associated with any product or vendor mentioned in this book.

LIMIT OF LIABILITY/DISCLAIMER OF WARRANTY: THE PUBLISHER AND THE AUTHOR MAKE NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE ACCURACY OR COMPLETENESS OF THE CONTENTS OF THIS WORK AND SPECIFICALLY DISCLAIM ALL WARRANTIES, INCLUDING WITHOUT LIMITATION WARRANTIES OF FITNESS FOR A PARTICULAR PURPOSE. NO WARRANTY MAY BE CREATED OR EXTENDED BY SALES OR PROMOTIONAL MATERIALS. THE ADVICE AND STRATEGIES CONTAINED HEREIN MAY NOT BE SUITABLE FOR EVERY SITUATION. THIS WORK IS SOLD WITH THE UNDERSTANDING THAT THE PUBLISHER AND AUTHOR ARE NOT ENGAGED IN RENDERING LEGAL, ACCOUNTING, OR OTHER PROFESSIONAL SERVICES. IF PROFESSIONAL ASSISTANCE IS REQUIRED, THE SERVICES OF A COMPETENT PROFESSIONAL PERSON SHOULD BE SOUGHT. NEITHER THE PUBLISHER NOR THE AUTHOR SHALL BE LIABLE FOR DAMAGES ARISING HEREFROM. THE FACT THAT AN ORGANIZATION OR WEBSITE IS REFERRED TO IN THIS WORK AS A CITATION AND/OR A POTENTIAL SOURCE OF FURTHER INFORMATION DOES NOT MEAN THAT THE AUTHOR OR THE PUBLISHER ENDORSES THE INFORMATION THE ORGANIZATION OR WEBSITE MAY PROVIDE OR RECOMMENDATIONS IT MAY MAKE. FURTHER, READERS SHOULD BE AWARE THAT INTERNET WEBSITES LISTED IN THIS WORK MAY HAVE CHANGED OR DISAPPEARED BETWEEN WHEN THIS WORK WAS WRITTEN AND WHEN IT IS READ.

For general information on our other products and services, please contact our Business Development Department in the U.S. at 317-572-3205. For details on how to create a custom *For Dummies* book for your business or organization, contact info@dummies.biz. For information about licensing the *For Dummies* brand for products or services, contact BrandedRights&Licenses@Wiley.com.

ISBN 978-1-118-48765-5 (pbk); ISBN 978-1-118-48829-4 (ebk)

Manufactured in the United States of America

10 9 8 7 6 5 4 3 2 1

Publisher's Acknowledgments

Some of the people who helped bring this book to market include the following:

Acquisitions, Editorial, and Vertical Websites
Senior Project Editor: Zoë Wykes
Editorial Manager: Rev Mengle
Acquisitions Editor: Kyle Looper
Business Development Representative: Sue Blessing
Custom Publishing Project Specialist: Michael Sullivan

Composition Services
Senior Project Coordinator: Kristie Rees
Layout and Graphics: Christin Swinford
Proofreader: Dwight Ramsey

Avere Systems contributors:
Bernhard Behn, Randy Quinn, Jeff Tabor, Rebecca Thompson

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

Table of Contents

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

Introduction

As the demand for faster access to ever-increasing amounts of data continues unabated, time — which translates directly into money — is wasted with every fraction of a second spent waiting for data. And the globalization of many enterprises, along with centralization of data centers, is placing more users farther from their organizations' NAS servers. In many cases, having all your data in one place is no longer feasible, because those who need access to it may be thousands of miles away. This book describes effective yet economical actions you can take to dramatically improve the performance of existing NAS systems in your local data center or in sites located across the world. It also introduces a new storage architecture designed to meet the challenges of the modern business. If you're wondering how to optimize storage and leverage new technology, this book is for you.

About This Book

This book consists of seven short chapters, individually wrapped and written to stand on their own.

Chapter 1: Storage Performance Is Critical. Today's organizations need to store and operate on more data than ever before. Happily, increasing storage capacity at a constant cost has enabled them to keep up with capacity. This chapter talks about the challenges created by that increased capacity.

Chapter 2: Optimizing Performance, Storage Capacity, and Cost in a Dynamic Environment. You want storage that has high capacity, high performance, and low cost. Unfortunately, these admirable goals conflict, so where's the spot that has enough of all three requirements? There must be a pony in there somewhere.

Chapter 3: A New Paradigm: Core Filers and Edge Filers. Since costs increase dramatically when you try to increase both capacity and performance of a storage system, it makes sense to separate the high-capacity portion of your storage

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

infrastructure from the high-performance portion. Breaking NAS into Core and Edge components optimizes capacity and performance at a reasonable price and allows storage resources to move to where it costs the least.

Chapter 4: NAS Optimization for Data Center Applications. Data center administrators face a lot of challenges, particularly if they are trying to deliver high performance and high capacity at low cost. Adding an Edge filer to a data center's Core filer can deliver all three.

Chapter 5: NAS Optimization for Virtualized Environments. Virtualization is the latest hot topic in the storage arena. However, it can seriously degrade performance due to the insidious I/O blender effect. With proper design though, it really *can* ease the task of maintaining a large heterogeneous storage resource and *boost* performance at the same time.

Chapter 6: NAS Optimization Streamlines the Cloud. Storing your organization's data in the cloud rather than locally has a lot of potential advantages. However, if you don't do it right, poor performance can relegate the cloud to backup and archive uses only. This chapter shows you how to do it right and enjoy sparkling performance when you do.

Chapter 7: Ten Helpful Storage Tips. Finally, in that familiar *For Dummies* style, I include a chapter with tips to help you put in place a NAS system that meets all your needs.

Icons Used in This Book

Throughout this book, special icons call attention to important information. You'll definitely want to take note!

This icon points out information that may well be worth committing to memory, at least that's my intention!

This icon points out helpful suggestions and useful nuggets of information.

Danger! These helpful alerts offer practical advice to help you avoid making potentially costly mistakes.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

Chapter 1

Storage Performance Is Critical

In This Chapter

- Looking at the increasing need for storage
- Putting fast storage where it is needed most
- Scoping out different storage technologies for different needs

If there's one thing you can count on today, it's change. While that's always been true, the changes are now coming faster than ever before. In order to survive, let alone grow, organizations must be sensitive to changes in the business environment and be able to react to those changes in a timely fashion. Some of the most dramatic changes in recent years have been in computational resources and the scale of the jobs we can tackle with those resources.

This chapter speaks to the changes in the computing environment and the impact those changes have had on the way data is stored and processed on computer networks.

The Ever increasing Demand for Storage

Perhaps the most amazing aspect of change in the computer business has been the continued applicability of Moore's Law, which states that the number of transistors that can be placed on a silicon chip doubles every two years.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

Their exponentially increasing computational power, fueled by advances in technology, has enabled computers to be applied to problems previously considered to be intractable — so difficult as to not even be worth attempting.

Another factor in the inexorable increase of processing power is the fact that hard disk drives have been increasing in capacity, also at an exponential rate. This is fortunate because having a faster processor won't do you much good if you don't have the capacity to store the results of the computations the processor makes.

Tackling bigger problems

Applications previously undreamed of are now routinely tackled by modern computers — which means that they are generating undreamed of amounts of data. The rapid growth in the amount of data that must be stored is putting unprecedented demands on storage systems.

REMEMBER

Business applications today require a high-powered processor with high-speed access to repositories of data, which reside on storage systems that are physically separated from the processor that is doing the computations. These storage systems are connected to that processor either through a local area network (LAN) or a more geographically dispersed wide area network (WAN). Widely separated computational nodes require shared access to centralized data stores, causing contention and delays.

One example of a business application that requires a lot of computing and a lot of network bandwidth is animation in the motion picture industry. Another is adding computer generated (CG) special effects to a live-action movie. A short while ago, these applications didn't even exist. They *couldn't* exist with the limited compute power available at the time. Now, block-busters such as *Avatar* or *Ironman* require huge render farms to generate the movies frame by frame.

Solving bigger problems

The demand for ever higher performance and ever more storage space has kept pace with the exponential improvements in

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

processors and storage systems. Users have come to expect exponential increases in capability. One way to attack part of the problem is to use *network attached storage,* commonly known as NAS.

Network Attached Storage

A *NAS system* is storage attached to a network. Because the NAS system is on a network, it can act as a repository for files that are used by multiple computers rather than just one. Furthermore, the computers that are accessing its files don't all have to be from the same vendor. Any computer that uses the same file-based protocol that the NAS unit uses will work with it. This is a big advantage at sites that have grown over time, where legacy systems are still in use, alongside more up-to-date equipment.

A NAS system has a processor and an operating system that have been designed to perform the task of serving files to computers on the network without taking on any extra tasks that might interfere with that primary function.

Classic NAS architecture has been around for more than 20 years. During that span, networks have become larger and more complex, processors have become faster, and applications have become more resource hungry, but NAS architecture hasn't changed with the times. Organizations using classic NAS are trying to address twenty-first century problems with a twentieth century architecture. As a result, new performance demands are not being satisfied by systems that can no longer be scaled up economically.

Climbing the Tiers of the Storage Pyramid

The function of a computer memory is to store values so they can be retrieved later and used in a computation of some sort. Different storage technologies are available in large numbers, but they all have one thing in common: they can distinguish between two states. A current is either flowing or it's not. An electric charge is either present or it's absent, and a magnetic

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

domain is either polarized north to south or south to north. The fundamental unit of a memory device holds a single digital bit, and one of the bit's two states represents a "logic one" and the other represents a "logic zero." In this way, regardless of the underlying technology, the contents of a memory device can always be interpreted in the same way.

Some of the memory devices that designers have in their arsenal are faster than others, some are cheaper per bit, and some are inherently more reliable. As you might expect, the fastest devices are the most expensive on a per bit basis. Devices that are a little slower are a little cheaper, and devices that are a lot slower are a lot cheaper per bit.

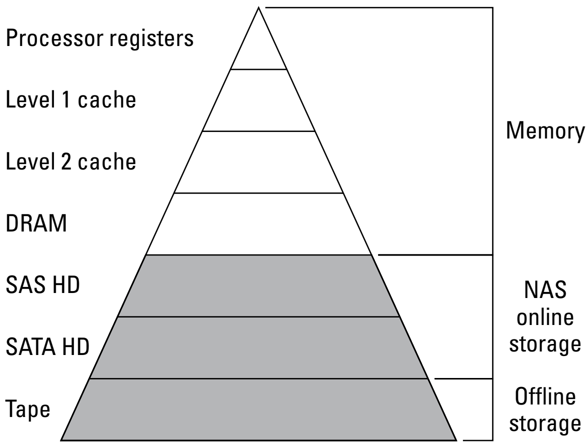
If you need a lot of storage capacity, you want as much of it as possible to be cheap. However, what if you need high performance in addition to large storage capacity? This situation is the most challenging and is confronted by more and more organizations as their requirements inexorably grow. The first solution that was implemented for this problem was the classic tiered storage model.

The classic tiered storage model

The classic tiered storage model is based on the observation that some of the data stored in a memory system is accessed more often than other data. If you know how much data you'll need to access often, theoretically you can buy just enough expensive, fast memory to handle that and put the rest of your data in much cheaper (but also much slower) memory. Such a system will perform almost as well as one made up entirely of expensive, fast memory. This ideal is a little too simple though, because the data in demand — and therefore the need to access it — can change at warp speed. This means that an efficient mechanism must exist for moving data back and forth between your fast tier and your slow tier.

To have high capacity *and* high performance at low cost, you want to have just enough of the fastest, most expensive memory to hold the data that is being used right now. This results in what is called a *memory pyramid,* with the peak being composed of the fastest, most expensive memory. A larger, slower, less expensive tier lies below the peak, and progressively larger, slower, and less expensive tiers lie below that (see Figure 1-1).

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.



Shaded portion is nonvolatile.

Figure 1-1: The classic tiered storage model.

Memory

As shown in Figure 1-1, the tiers at the top of the pyramid are either an integral part of the computer's processor or are close to it. These tiers use the fastest technology, and because of their close proximity to the site of computation, they suffer minimal delays in the time it takes for a signal to travel to the processor. The registers are literally a part of the processor, so you can't get any closer to the action than that. A little farther away is the processor cache, which may exist in several levels, each one a little bigger, a little slower, and a little farther away from the processor than the preceding level.

Even farther away and quite a bit slower is the computer's dynamic RAM (DRAM). Beyond the DRAM is the online storage.

REMEMBER

An important characteristic of memory, whether it be processor registers, cache, or DRAM, is that such memory is *volatile*. This means that if power is ever removed from the device, the data in the memory evaporates. This can be a serious, even catastrophic problem, so a lot of effort has been placed into backing up power sources. Such backup plans work great, unless something really extreme happens, such as a Richter 9 earthquake followed by a tsunami, or a really big hurricane. Since extreme events do sometimes happen, it's

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

wise to shuttle data from volatile to nonvolatile memory as soon as it is feasible to do so.

Online storage

Online storage is typically *nonvolatile.* That means that if power to the system is lost, data is retained and is ready to be used when power is restored.

Offline storage

As time passes, some data in a storage system isn't accessed any more. For example, after a company has produced its annual report, there isn't a need to retain all the detailed data that went into it in online storage. To do so would waste expensive high-speed storage with essentially "dead" data and slow down system operations.

But, you can't simply delete such data to free up resources. A question could come up at a later time that requires that information. The solution is to transfer old data to an archival system, which has a much lower cost per bit and of course slower performance. However, since access to archived information is rare, slow performance of retrieval isn't a major concern. In the past, the primary archival medium was magnetic tape. Now, as disk capacities have continued to increase, low cost SATA drives are being used for archival purposes also.

Today's expanded tiered storage model

The classic tiered storage model shown in Figure 1-1 performed well for a number of years. However, as workloads have increased and the need for speed has become ever more acute, in many applications NAS performance hasn't kept up. Wholesale replacement of older disk drives with newer, faster drives can be prohibitively expensive and may deliver only a marginal improvement anyway.

A cost-effective solution to this dilemma is to add some intermediate tiers to the storage pyramid, as shown in Figure 1-2.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

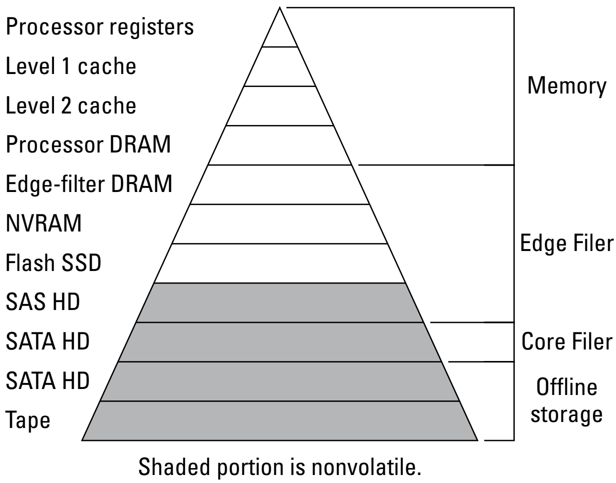


Figure 1-2: The expanded tiered storage model.

The principle of the expanded tiered storage model is to put faster storage closer to the processor, minimizing the number of times it must access the slower online storage. It essentially consists of adding another level of cache to the system — a level of cache that is much larger than the DRAM in the processor but potentially just as fast. The larger the amount of working set that can be kept in the intermediate level of storage, the better the performance.

Edge filers versus Core filers

The core of a network attached storage system is a file server that provides access to data stored on hard disks. This is the online storage tier in the classic tiered storage model. In the expanded tiered storage model, it is called the *Core filer*. What the expanded tiered storage model does is put another, quicker responding file server between the compute processor and the Core filer. This *Edge filer* is located at the edge of the NAS environment, adjacent to the processor requesting the data. The storage media in the Edge filer are faster than the storage in the Core filer. Since the cost per bit of storage in the Edge filer is greater than the cost per bit in the slower Core filer, economics dictate that it have a lower capacity. However, as long as its capacity is sufficient to hold the *working set* (the part of the data being operated on that is

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

being accessed right now), processing proceeds essentially at Edge filer speed. Only rarely does an operation occur that requires access to data stored on the Core filer.

Different Edge filer configurations for different needs

Some applications require the fastest possible response but have a relatively small working set. Others may have the opposite situation, which is a huge working set but a tolerance for slower response. You can expect applications that require both top performance and the space to hold a huge working set to require resources that cost significantly more.

The optimum mix of relatively expensive high-speed storage and relatively cheap lower-speed storage should be available in a good Edge filer system.

- **DRAM:** Highest speed, highest cost, volatile
- **NVRAM:** High speed, high cost, nonvolatile
- **Flash:** High speed, high cost, nonvolatile
- **SAS hard disk:** Lower speed, lower cost, nonvolatile
- **Tape:** Sequential access, lowest speed, lowest cost, nonvolatile

The largest capacity, slowest, and cheapest random access online storage option is SATA hard disk, which has no place in the Edge filer. SATA is relegated to the Core filer. The fact that the Edge filer handles the overwhelming majority of the accesses to storage means that the Core filer can be built up with massive capacity entirely from SATA drives. There is no need to include more expensive storage media within the Core filer, which replaces the classic NAS system. Classic NAS systems, lacking an Edge filer, must include more expensive storage media types, raising their prices substantially.

With flexibility in how an Edge filer is populated with different types of storage media, it can be optimized for the applications that run on it, boosting performance and minimizing overall cost.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

Chapter 2

Optimizing Performance, Storage Capacity, and Cost in a Dynamic Environment

In This Chapter

- Overcoming the three latencies
- Adjusting to changing workloads
- Scaling to match workload growth
- Controlling costs
- Taking a look at global namespace
- Addressing latencies

This chapter discusses the different types of latencies that cause performance bottlenecks in a NAS environment, including the challenges that distributing data across a wide area network (WAN) bring, and how an Edge/Core architecture addresses each of these challenges.

Taking In the Three Latencies

The biggest drag on performance of NAS systems today is *latency,* the delay between when an action is initiated and when it's completed. Three main kinds of latency exist in a NAS system: hard disk drive (HDD) latency, storage filer CPU latency, and network latency. These latency classes are inherent in classic NAS architecture. Here's a look at each of these types of latency and what can be done to eliminate them.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

HDD latency

As an organization's data grows and the capacities of disk drives become bigger and bigger, the amount of time it takes to access a given file increases. Existing solutions to this problem have been less than ideal — NAS systems have moved to higher-speed (10K and 15K RPM) disk drives and sometimes, for an extra increment of speed, *short stroke* them (that is, use only a fraction of their capacity) in order to read and write data off them faster. This type of disk overprovisioning takes up a lot of physical space in the data center, plus all that rotating mass generates heat and consumes a lot of electricity.

The latency associated with hard disk drives comes from the fact that HDDs are mechanical devices with moving parts that can move only so fast. The obvious solution is to use nonrotating memory consisting of semiconductor chips. Such memory, whether DRAM, NVRAM, or Flash, is more than a thousand times faster than hard disk storage. The only problem is that it is also a lot more expensive per bit. To replace all the hard disk storage in a NAS system with semiconductor memory would be entirely too expensive.

This is where Avere's Edge filers really shine. A relatively small amount of semiconductor memory in an Edge filer can shoulder the bulk of the processing load at high speed, while the high-capacity rotating memory back at the Core filer holds the bulk of the data that isn't currently being accessed.

Storage filer CPU latency

In classic NAS systems, all file-access requests pass through the storage filer, and the CPUs resident in such filers often bottleneck causing another form of latency. If more I/O requests hit the filer than its CPU(s) can handle, it slows application performance to a crawl. The only reliable way to ensure that you get good performance with the classic NAS architecture is to have very high-performance CPUs (also known as *storage heads*) within these systems. In a sense, you are overprovisioning the CPUs in a system. This is just like putting massive processing power in an older nondistributed supercomputer. The supercomputer gives you a performance boost today, but it doesn't scale.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

Storage systems incorporating Avere Edge filers avoid storage filer CPU latency by offloading the bulk of the computational burden from the storage filer. The Edge filers at each remote site handle the traffic from their site, leaving relatively little for the storage filer CPU to do.

Network latency

In today's fast-paced business environment, access to the latest information is critical. With remote offices that may be hundreds or even thousands of miles away from the organization's data centers, the transmission delay (latency) due to the distance can put you at a competitive disadvantage.

The problem is caused by the fact that the needed information is often at the data center, while the people who need it are at another physical location. If this is such a big problem, why not locate the users close to the data center, solving the problem? That might not be feasible for several good reasons.

Operation from multiple locations

For any number of good reasons, you may want to locate your offices far away from your data center, such as:

- The most economical location for a data center is likely to be far from talent pools.
- Mergers or acquisitions may require integrating dissimilar resources.
- Pools of talent (for example, software developers) tend to be scattered across the globe.
- You're running out of real estate at your data center location

Such decentralization makes the functioning of the organization's WAN a critical part of its infrastructure and brings to the forefront the problem of WAN latency.

The WAN latency problem

A communication passing between a computer and a NAS system must travel along the path between the two. If that path is hundreds or thousands of miles long, even though photons can travel through a fiber and electrons can travel

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

through a wire at nearly the speed of light, significant delays can accumulate. This problem is made even worse because a certain number of messages must pass back and forth before the transmission is considered successfully completed.

The delays occur because the data is stored far from the location where it's used. One way to solve this problem is to place the data that's needed most often close to where it's needed, leaving the data that's not often needed far from the point of use. Since such data is rarely needed, on those few occasions when access is required, performance isn't degraded much at all.

Edge filers reduce WAN latency

Edge filers sit in the communication path between the user and the NAS at a location that is nearest to the user. If the user is located at a remote site and therefore at a distance from the data center that holds the NAS, you can largely eliminate WAN latency by placing the Edge filer physically close to the user. Since the Edge filer contains the *hottest* data, that which is being accessed most frequently, the round trip transit time from the user to the data is minimized. The long trip back to the NAS for the *cold* data needs to be taken only a small fraction of the time, and overall performance is almost as good as it would be if the data center were right next door to the remote facility.

Dynamic Tiering Adjusts to Changing Workloads

Optimizing the tiering of your Edge filer (refer to Chapter 1) to the application you're running is all well and good, but what if you want to run a different application with different resource needs? Do you have to settle for seriously degraded performance? For simple architectures you do, but with systems such as the Avere Systems FXT Series Edge filers, it's possible to adjust how storage is allocated on the fly as the system workload shifts from one application to another, or from one section of a running application to another.

A system that does dynamic storage tiering automatically moves data from one storage tier to another, based on

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

changes in the level of demand for data in a given block within a file. Data that is being accessed more frequently will be moved to a faster tier, and data that is being accessed less frequently will be moved to a slower tier. All of this activity happens without any manual intervention and has no effect on the running application other than the fact that performance doesn't degrade when the application moves to operating on different data.

Balancing the Tradeoffs

In order to meet the needs of an organization, a storage system must strike a delicate balance. The things that you care most about (performance, capacity, and cost) affect each other. Improving one generally has a negative effect on one or more of the others. With automatic storage tiering within your Edge filer, you can achieve the performance you need at an affordable cost.

Higher overall capacity can now be achieved with high-density, low-cost SATA disks in your Core filers. Additional cost savings include reduced floor space and less electricity for powering and cooling storage.

REMEMBER

Delivering higher capacity and higher performance without an unacceptable jump in cost calls for a new approach to the problem. Avere's FXT Series Edge filers add a new element to classic NAS systems that provide the optimal combination of capacity, performance, and cost.

Building Storage Systems That Scale

One of the biggest challenges faced by information technology departments is maintaining an acceptable level of performance as the workload of a growing organization increases. One of IT management's biggest fears is that the equipment they have will top out with no path forward other than scrapping everything and buying all new equipment, at tremendous cost — not to mention disruption of ongoing operations.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

Using a Core/Edge architecture rather than a traditional NAS system where everything is in the Core makes scalability a lot more feasible. Multiple Edge filers can be added to an existing NAS system without requiring any changes to the Core. Within an Edge filer cluster, additional nodes can be added seamlessly, without disrupting the flow of running applications.

Anyone charged with installing a new NAS system needs to carefully consider future as well as current needs and select an architecture that will handle anticipated, and even a certain amount of unanticipated, growth. A configuration that includes Edge filers, particularly in a distributed, WAN-based environment, is likely to prove more effective than one whose assets are all concentrated in the data center.

Slashing Storage Costs

Cost is a primary concern of any manager responsible for the storage infrastructure of an organization. There are multiple ways of providing the level of service that the organization requires, but they don't all cost the same. The manager who can provide the needed service, but at a lower than expected cost, is a hero who can look forward to a bright future. One way to work toward that hero status is to be aware of what specific aspects of a system are raising costs the most and finding ways to replace them with more cost-effective solutions.

One way to address the basic conflict between the desire for high performance and the simultaneous desire for low cost is to make sure that the hardware you have is being used in the most effective way. To maximize that effectiveness, you must be able to monitor the demands that the running application is making on the system and dynamically reallocate the location of specific sets of data from one tier to another, based on how actively it is being accessed. Avere Systems has followed this philosophy with its implementation of dynamic storage tiering.

Avere's dynamic storage tiering matches storage resources with the demands of the application that is running on those resources. Data is automatically moved across classes of storage, depending on how actively it is being accessed by the applications that are running at the current time and what type of access if requested.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

The primary benefit of dynamic storage tiering is economic. If you can always find your data in the storage resource whose performance is most appropriate for the level of that data's activity, you will save money. Very active data will always be located in the fastest tier, which, being the most expensive, is also typically the smallest. At any given time, most of your data probably is *not* being actively accessed. Dynamic storage tiering will assure that this data is relegated to the slowest and thus cheapest storage available.

For a system with Avere Edge filers, the faster tiers would likely be implemented with DRAM and solid state disk (SSD). The slowest tier, implemented with SATA hard disk drives, corresponds to the traditional NAS system. Avere refers to this as the Core filer (refer to Chapter 1). Dynamic storage tiering assures that you are always getting the highest performance at the lowest cost that your hardware can provide.

Simplifying System Operation with Global Namespace

One of the headaches suffered by IT managers, particularly in large organizations that have gone through a merger or an acquisition, is having a multiplicity of servers from different vendors composed of differing hardware and software. When some of the files an application needs are on server1 with others on server2 and still others on server3, administration can get pretty complicated.

Implementing a *global namespace (GNS)* is an effective solution to the problems caused by a system comprised of heterogeneous hardware and software. In a GNS configuration, the servers can be different models from different vendors running under different operating systems and serving data from different file systems. Those physical differences are all hidden, and the only thing visible to the application is a single logical structure that appears as if it consists of files on a single server.

The essence of a global namespace is that administration is greatly simplified. Users' access is consolidated through a single mount point. Admin is freed from mundane tasks.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

WARNING! Global namespace has become a popular buzzword, and storage vendors like to claim they offer it. When evaluating competing proposals for NAS, make sure that the global namespace being claimed actually works across vendor boundaries rather than being restricted to the offerings of a single vendor. Even if you are currently a single vendor shop, you never know when your management might acquire a company with different hardware from what you already have, which will need to be integrated into your existing structure.

Dealing with the Three Latencies

Traditional storage vendors have tried to address the latency problems that have gotten progressively worse as the performance demanded by customers with widely spread out installations has increased. The nondistributed, single-system design characterizing classic NAS architecture is intrinsically unable to scale to the levels required today. The solutions offered by those vendors, although expensive, are not effective in dealing simultaneously with all three major classes of latency. A fundamental architectural shift is needed, that shift being the new paradigm exemplified by the Avere Edge/Core storage architecture. Latency is eliminated by putting the right amount of the right kind of storage where it is needed most.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

Chapter 3

A New Paradigm: Core Filers and Edge Filers

In This Chapter

- Getting an overview of NAS optimization
- Understanding the path to NAS optimization
- Appreciating Edge filer benefits

IT managers responsible for geographically dispersed networks are primarily looking for three things from their systems: performance, flexibility, and efficiency. Traditional NAS architecture has had a hard time delivering all three.

Avere's Core/Edge architecture, combined with state-of-the-art clustering technology, addresses performance by automatically maintaining the most active data on the fastest media possible. It streamlines operations with proprietary data migration and mirroring technologies that simplify storage system management through a single global namespace. Efficiency is delivered by optimizing Core storage for data protection, high capacity, and low cost, while building an Edge appliance that delivers fast response in information requests while minimizing network traffic.

This chapter goes into detail about Edge filers, which are new elements that when added to a conventional NAS (network attached storage), significantly boost performance while being completely transparent to the running applications.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

Why NAS Needs to Be Optimized

As the demand for ever faster access to ever larger storage continues to increase, the three latencies (HDD latency, storage filer CPU latency, and network latency) become ever larger problems. (See Chapter 2 for more on the latencies.) Time and money are wasted while waiting for data. The globalization of many enterprises often places users far away from their organizations' NAS servers, so having all your data in one place may no longer be a good thing since those who need it may be miles and miles away.

How to Optimize NAS

You can optimize your NAS environment by building it into Core filer and Edge filer components. Locate the Core storage where hosting costs are low and high-performance network bandwidth is available. Provision it with high-capacity, low-cost hard disk drives. Locate Edge filers close to the users, populated with high-performance storage that will always hold the data that is being most actively accessed. Place all the NAS resources under a single global namespace to hide any complications that a heterogeneous hardware or multi-filer environment might create.

The Edge Focuses on Performance

The Edge filer is a NAS file server architected to efficiently harness limited quantities of solid state memory, flash, and SAS disk to handle read and write requests from clients at the network edge. It processes reads and writes without incurring high latencies. As time permits or demand requires, it updates the Core filer so that it and all the other Edge filers remain in sync. See Figure 3-1.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

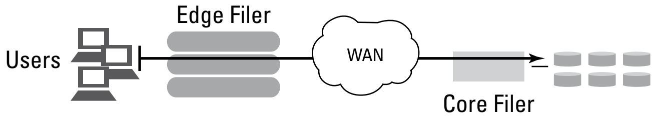


Figure 3-1: The Edge filer can be close to the users, so it doesn't matter how far away the Core filer is.

Performance

Storage engineers have tried a variety of methods to improve performance. Solid state computer memory and flash are the fastest forms of storage currently available, but they are much too expensive per bit to hold all of an enterprise's data. Luckily, it isn't necessary that all your data reside in the fastest possible storage. The only data that needs to be quickly accessible is the data that you currently need or may need soon. The rest can reside on much slower media back at the Core filer.

Moving file data blocks between tiers is the job of Avere's dynamic storage tiering technology. Typically, in a traditional system, data is never moved after it's been saved — even though it may only be heavily used for a few weeks before the usage tails off and it becomes relatively inactive. If the data is still taking up space in high-performance storage, then it serves to slow down everything else. Dynamic tiering assures that every block of file data is located in storage where its performance matches its level of activity.

REMEMBER

Dynamic tiering assures that just as soon as a block of data's activity level changes, its location in the tier hierarchy changes too.

You can do several things to boost performance by reducing network latency. By offering remote sites access to enterprise files, you eliminate the need to copy entire sets of data between sites. Handling backup centrally at the Core removes a major headache from those responsible for remote sites, allowing them to concentrate on business rather than administration tasks. Reducing the need to send whole data sets back and forth also saves money by lowering the amount of network bandwidth needed.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

Flexibility

In today's rapidly changing business environment, flexibility can be just as important as performance. A system that meets your needs right now may not serve you well next year. Since Avere's Edge filers are designed for scalability and availability, you can prevent network bottlenecks by attaching them directly to networks of server farms. Local NAS traffic stays on the network edge rather than suffering the delays inherent in transmission over the network.

One problem that troubles many organizations — often as a result of mergers and acquisitions — is trying to operate smoothly with a heterogeneous collection of filers that operate differently. Implementing the Avere Edge/Core architecture with a single global namespace, offering access to various file system exports from an assortment of multi-vendor filers, administrators no longer need to manually copy or cache local replicas of important data sets at each remote site.

With Edge filers, you can migrate and mirror data between Core filers without having to shut down or interrupt your clients during the process. This makes normally painful hardware upgrades transparent. It also enables you to actively manage data life cycles, using background migration that is invisible to clients. Data that has cooled off can be seamlessly moved to archival Core filer storage, while older archived data that may be needed for a new project can be dusted off and moved to a production Core filer.

Efficiency

Edge filers add to the efficiency of a NAS system in several important ways. One big money saver is that many scattered data centers can now be collapsed into a few strategic locations. On the other end of the pipe, remote offices can enjoy direct access to centralized storage without deploying remote Core filers or high-speed WAN links.

Edge filer architecture minimizes network traffic, while at the same time giving clients the fastest access possible for their shared data. Core filers back at the data center can use low-cost, high-capacity SATA drives, and maintenance operations can largely be confined to the data center as well, saving a lot on staffing and travel.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

Chapter 4

NAS Optimization for Data Center Applications

In This Chapter

- ▶ Understanding NAS challenges in the data center
- ▶ Riding the performance treadmill
- ▶ Looking at the footprint
- ▶ Appreciating minimized hardware
- ▶ Taking a look at mirroring
- ▶ Turning an eye toward best practices

To be successful in their business, modern enterprises depend on their Core applications being online and running with good performance. It doesn't matter whether the enterprise is designing and manufacturing semiconductor devices or creating the next blockbuster movie.

Enterprises run many different types of applications that are critical to the operations of the company. Some are horizontal applications that apply to many types of companies in many different industries. Others, for example media rendering, software build, semiconductor design, and seismic processing, are vertical applications that apply to a specific industry. While these specific applications are very different, they have two things in common. All these applications are used to access and process an enterprise's most important asset — its data. All face similar IT challenges as the amount and richness of the data grow over time. This chapter examines these challenges.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

NAS Challenges in the Enterprise Data Center

Traditional NAS solutions create four major pain points that impact performance, cost, and management. And things are only getting worse as the amount of data increases:

- Overprovisioning of disk drives in order to improve performance is expensive and wastes lots of power and space — and data centers typically don't have a lot of space.
- Scaling performance is difficult and typically comes with expensive Core filer upgrades, more disks, and yet more application downtime.
- Providing global access to the data is not simple.
- NAS sprawl, which results from the scaling limitations of monolithic, legacy NAS, greatly complicates management.

The Performance Treadmill

The demand for performance in the enterprise data center is always increasing, driven by more users, new applications, and richer data. For the administrator, this is like being on a never-ending treadmill that is constantly speeding up. To be successful, you cannot just try to run faster; you have to run smarter as well.

In today's data center, this starts with recognizing that hard disk drives (HDDs) *are getting bigger but not faster*. Gone are the days when it made sense to throw a large number of HDDs at a performance problem. The smarter approach is to optimize your NAS by using Flash storage.

Don t Break the Bank

So, how do you stay within your budget? This is an important question when working to increase the performance of your data center applications. Performance demands can increase

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

30 percent, 50 percent, or even 100 percent year over year and typical IT budgets are increasing much less, if at all.

The answer is to adopt the Edge filer/Core filer architecture. Because the Edge filer is filled with the fastest storage — RAM, DRAM, and Flash, it provides the performance you need in a footprint that minimizes the space, power, and cooling required to maintain it.

1.6 Million IOPS in the Smallest Equipment Footprint

At the Supercomputing 2011 conference, the Avere Systems FXT Edge filer took the top performance spot as measured by the SPECsfs2008 benchmark. Avere posted throughput of 1,564,404 ops/sec, which is the highest ever posted for a single file system namespace in the long history of the NFS benchmark. In addition, this throughput was achieved with an ORT (overall response time or latency) of just 0.99 msec, which is 35 percent better than other competing Core filer solutions.

In addition, the Avere FXT Edge Filer set that record with a physical package that is 79 percent smaller than a major competitor's best NAS solution and 65 percent smaller than another competitor's best.

An Avere Edge filer in the data center uses far less physical space and less electrical power for operation and cooling than these competitors as well.

Minimize Hardware to Minimize Headaches

An additional cost that must be considered when designing the infrastructure for your data center apps is the administrative cost of managing all the equipment in the data center. As performance and capacity grow in traditional NAS environments, one of the most significant tasks is that of managing and maintaining all the equipment so that it operates properly.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

Minimizing the total amount of hardware operating as part of your storage system — using an Edge filer, for example — can alleviate management headaches down the road.

Transparent Data Migration

One of the challenges that IT managers are continually faced with is change. As business needs grow, new servers and disk drives must be added to the existing installation. As older drives reach the end of their useful lives, they must be traded for new ones. As the usage of a chunk of data changes, it may need to be transferred to a different tier of storage in order for the system to continue to provide the performance level needed. All these changes can be terribly disruptive to day-to-day operations, which must be halted while such major surgery to the system is performed.

Using the FXT edge filer's FlashMove capability, data can be migrated easily from one system to another. Whether it's the decommissioning of an older system, the changing from one vendor to another, or the migration of data from data center hardware into the cloud, FlashMove makes the migration task much easier. See Figure 4-1.

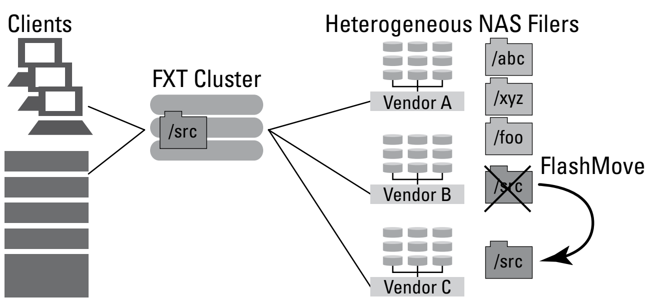


Figure 4-1: The FXT cluster coordinates the movement of the /src folder from Vendor B's Core filer to Vendor C's Core filer without affecting client operations.

Semi synchronous Mirroring

Mission-critical data deserves to be protected from all kinds of threats, even those that seem to be highly unlikely. One

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

good practice for protecting data from the inevitable disk failure is to configure drives in a RAID array. This can protect you from any single drive failure or, in the case of RAID 6, from two drives that fail at the same time.

However, all the drives in a RAID array are going to be located physically near each other. That means that a disaster that takes out an entire building, which occurred in many places when Hurricane Sandy hit the mid-Atlantic and Northeastern United States, will show that there are limits to the protection RAID provides.

Avere addresses this problem with FlashMirror software, running on an FXT cluster. Figure 4-2 shows how it works.

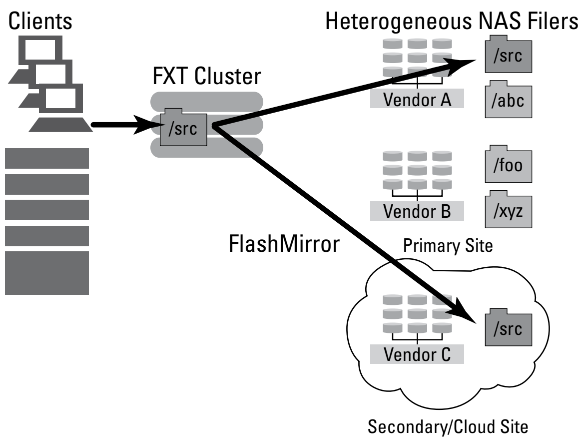


Figure 4-2: FlashMirror, running on an Avere FXT cluster, takes write operations from client applications and directs them to two destinations.

TIP

FlashMirror mirrors the write to the local storage with an identical write to secondary storage, which could be in the cloud or at some remote site. Since updates are made to both the primary and secondary storage at essentially the same time, the process is called *semi-synchronous mirroring.* If the secondary copy should fail, FlashMirror could rebuild it from the primary. If the primary copy should fail, a single mouse click will redesignate the secondary as the new primary, until the primary copy can be restored. Normal operation continues unabated while recovery efforts take place in the background.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

Instituting Best Practices

When your network grows to the point where you are confronted with the need to virtualize your storage, here are a few points to keep well in mind.

- Keep your infrastructure simple. The more complicated a system is, the harder it is to scale up as needs grow.
- The Edge filer/Core filer architecture helps you to scale with that growth by letting the Edge filer handle performance while the Core filer handles capacity.
- Latency can be defeated — in all three of its forms — through the Edge filer/Core filer architecture.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

Chapter 5

NAS Optimization for Virtualized Environments

In This Chapter

- Looking at how virtual servers and desktops pose challenges for storage
- Making the most of storage performance and cost in virtualized environments
- Tuning in to the "blender effect"
- Protecting your storage server from virtualization's "boot storms"

Virtualization is revolutionizing the IT world by bringing great efficiency to the data center. This chapter discusses two broad categories of virtualization: server virtualization and virtual desktop infrastructure (or VDI).

Server virtualization enables the delivery of high-performance servers with less hardware, power, and space than ever before. This is accomplished since many *virtual* servers are placed on a single *physical* server platform. Virtualization software from companies like VMware and Citrix ingeniously slices up the server resources (for example, CPU, memory) across all the virtual servers running on the platform.

VDI provides an efficient and easy way to manage and deploy desktop computers. Virtualization software enables placing many, even hundreds, of desktops on the same physical hardware, delivering the same efficiencies as previously discussed for virtual servers. In addition, VDI simplifies management since all the desktops can share a single *golden* version of the desktop software. This makes training, upgrades, and distributing new software a snap.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

Discussing Storage Challenges

Virtualization provides nothing but good news, right? Not so fast! Virtualization places great strains on the storage. Here's how.

Prior to the advent of virtualization, servers typically placed a predictable load on the storage. Each server ran one application, and the resulting load placed on the storage consisted of many large sequential data accesses to closely related data stored on the system's disks. Rotating hard disks handled sequential data accesses well, and everyone was happy.

Enter server virtualization. With many virtual servers running on a single physical server, the large sequential data accesses of each virtual server get broken up and mixed with those of all the other virtual servers. The end result of this *blender effect,* which I discuss in more depth later in this chapter, is that the storage sees a workload that is very random, and random access patterns put great strains on rotating hard disks.

VDI creates a separate, unique challenge for storage called *boot storms,* a particularly frustrating performance robber that I discuss later in this chapter as well.

Optimizing Performance and Cost

FXT Series Edge filers from Avere use intelligent tiering to automatically place *active* virtual server and desktop data on high-performance RAM, SSD, and SAS storage media and ensure the fastest response times for guest applications. *Inactive* data is managed on high-density SATA storage, delivering the maximum cost, power, and space efficiency. Clustering ensures that your data is always available and provides the flexibility to simply scale the performance of the solution as demand grows.

The Avere FXT Series provides performance acceleration that is purpose-built for virtualized environments. FXT Edge filers automatically tier the active portions of the virtual disk files at the block level to provide the most efficient use of the internal high-performance storage media. During times of high read

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

access (for example, boot storms), Avere FlashRead automatically replicates shared blocks across all FXT nodes in a cluster to provide parallel access and maximum throughput to the hot blocks. Virtualized workloads are often write-heavy and Avere FlashWrite provides unique write acceleration that offloads HDD-based storage and provides a cost-effective alternative to deploying many HDDs to handle the write workload.

Avere solutions are highly efficient and typically require 80 percent less total equipment than traditional NAS deployments, resulting in dramatic savings in storage capital expenses and ongoing operational expenses for power, cooling, and rack space. For existing NAS deployments, the FXT Series enables administrators to dramatically increase the performance of their systems without expensive controller upgrades or adding storage. For new deployments, the FXT Series enables administrators to meet their performance requirements with the highest possible storage density on the Core filers, using low-cost, high-capacity SATA drives.

Exploring the Blender Effect

Virtualization's blender effect turns large sequential disk accesses into many small, random disk accesses that thrash rotating hard disks. With the entire IT world moving to virtualization, what's a storage administrator to do to avoid taking a thrashing?

REMEMBER

The answer is — solid-state storage media in the form of DRAM, NVRAM, and SSD.

Hard disks use rotating *platters* for storing data and mechanical *arms* to read/write data from/to the platter. Random access patterns require the arm to physically move to many different locations on the platter and are limited by the mechanics of the motion.

On the other hand, because solid-state storage has no moving parts, it excels at random accesses patterns and therefore as storage for virtualization environments.

The Avere FXT Series Edge filer provides a solid-state tier in front of third-party disk storage. In this way, Avere provides a unique and compelling way to modernize existing storage for virtualized servers and desktops.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

Booting Up on a Network Isn't Like Booting Up on a Home Computer

A network with hundreds or thousands of client desktop machines can save a lot of money if those machines are thin clients. *Thin clients* provide a keyboard and screen, RAM, and a network connection. Rather than each machine having its own hard disk, VDI enables these machines to operate entirely with files accessed on the network's virtual server.

In normal usage, this arrangement works great. Even though hundreds of different desktops all access files on the same physical storage in the course of a workday, it's rare that all the desktops would need service at exactly the same time.

There are times, however, when contention for service is not so rare.

Rush hour traffic

Have you ever tried to take the main highway into a major city at eight o'clock on a weekday morning? What about the main highway out of that major city at five in the afternoon? If so, you are familiar with that bane of motorists everywhere, *rush hour traffic.* Rush hour traffic can reduce progress to a crawl, and small children riding tricycles on the neighboring sidewalk will leave you in their dust.

Why does this happen? It happens because the road you're on was designed to handle the amount of traffic it was likely to see throughout the day. Most of the time that design is adequate, and the traffic flows smoothly. However, during rush hour, normal conditions don't apply. Everybody is trying to go to the same place at the same time. The transportation infrastructure is overwhelmed by the load.

The same thing can happen to a computer network whose client computers are configured for VDI.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

All together now!

If everyone arrives at work at 8 a.m. and the first thing they do is boot their desktop computers, there's going to be a traffic jam. Most people need their computers to be functioning before they can start to work, so booting up right away is natural. The result is a *boot storm.* The storage controller is simultaneously hit with many more service requests than it can handle. Consequently, workers could find themselves waiting 15 to 20 minutes for their machines to boot up. This wastes a lot of employees' time, but it also impacts their attitudes because the delay is frustrating.

Boot storms aren't just a problem at the start of the day. They can also happen after lunch, during a virus scan, or when a system update forces a system-wide reboot. A raft of simultaneous logouts could cause a similar performance drop. If an organization is suffering a significant productivity loss due to boot storms, something must be done.

One idea is to replace the existing storage system with a faster, but more expensive, one. All too often this results in only a marginal reduction of boot time during a storm, but at a substantial cost.

Distributing the load

Rather than spending a lot of money on a whole new storage system, it may be possible to solve the problem much less expensively — while retaining your existing storage system. The solution is to take the load of responding to the boot storm off your storage and put it instead on an Avere Edge filer.

The Avere Edge filer protects the storage server from the storm coming in from the network, responding to all those boot requests in an interval of time that isn't much longer than it would take to respond to a boot request from a single desktop. Not only does the addition of an Avere Edge filer cost a whole lot less than a major upgrade of your main storage system, but it also provides better performance and data management functionality.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

Chapter 6

NAS Optimization Streamlines the Cloud

In This Chapter

- Understanding public clouds versus private clouds and compute clouds versus storage clouds
- Sidestepping WAN latency with Edge filers
- Looking at efficient scaling of application performance with cloud computing
- Putting primary data into the cloud

Clouds in the sky are fuzzy, fleecy things with ill-defined boundaries that move as they are blown about by winds. In the computing context, clouds also have fuzzy, fleecy boundaries, and you can't be quite sure where a particular piece of cloud is physically located. There are public clouds and private clouds, as well as computing clouds and storage clouds. It's important to know what differentiates one type of cloud from another, which is what this chapter is about.

Public Clouds

Essentially, public clouds provide IT resources, both in the form of compute processing for running your applications and storage capacity for storing your data, as a service. The enterprise no longer needs to invest in hardware beyond desktop, laptop, or even mobile devices, thus substantially reducing capital expenditures (CAPEX) and replacing them with operating expenditures (OPEX). This saves the user organization from having to make a large upfront outlay (a

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

CAPEX), replacing it with a modest monthly fee (an OPEX). By renting services from a cloud services provider, users can buy the exact capability they need for only as long as they need it.

A number of benefits accrue to enterprises that operate in the public cloud.

- **Cost savings:** IT resources are moved to the cloud, freeing up valuable real estate and lowering power and cooling expenses at the enterprise.
- **On-demand scaling:** More IT resources are rented during peak times and less during slow times without the concern for being stuck as either under- or overprovisioned.
- **Simplified management:** IT resources are consolidated and centrally located in the cloud with much of the system administration outsourced to the cloud provider.

Aside from the aforementioned advantages, you also have potential disadvantages to entrusting your organization's valuable data to a public cloud. Such organizations can accrue the many benefits of a public cloud with a *private cloud.*

Private Clouds

Functionally, a private cloud is just like a public cloud. The difference is that the organization whose applications are running on a private cloud owns and controls all the equipment in the cloud. Many of the benefits that a public cloud offers (see the preceding section) are available on a private cloud as well. In addition, you have the security of knowing that control of your data is entirely kept within your organization. Running a private cloud is a natural fit for enterprises that are spread out geographically. In fact, while the term *private cloud* is somewhat new, many global enterprises have been operating this way and accruing the benefits for years.

Private clouds can be operated at *collocation facilities*, or *colos*, which are data centers where equipment, space, power, cooling, and physical security are rented from a service provider. The equipment at a colo is not owned by the service provider, as would be the case with a public cloud, but rather it is owned by the enterprise renting the space. Colos can be

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

especially useful for enterprises located in major cities since colos are often located in remote areas where the cost of the space, power, and cooling is substantially less.

Compute Clouds

Whether you're dealing with a public cloud or a private cloud, there is another way of looking at the cloud. That way is to consider what the cloud actually does for you. For example, one thing it might do is computation. In this case, enterprises run their applications (for example, databases, e-mail, product development, scientific research) on computers that are located in the cloud. You can access a compute cloud from a thin client consisting of little more than a keyboard, mouse, screen, and a network connection — and then have all computation done in the cloud. This reduces the cost of user workstations, servers, and compute farms at the enterprise and makes sure that all applications are always at the latest revision level and synchronized across the organization.

REMEMBER

Most public compute clouds charge an hourly rate for the use of their compute platforms. Typically, these platforms are available in a variety of CPU and memory tiers with a higher price charged for tiers with more CPU cores and more memory. An additional data transfer charge is assessed for all data transferred *out* of the cloud. There is no charge for data transferred *into* the cloud.

Storage Clouds

Storage clouds leave the computation to the users' workstations, servers, and compute farms where the applications run but store the data in the cloud. Enterprises create private storage clouds because consolidating data in a centrally located data center is more efficient and safe than spreading it across the enterprise. The economy of scale benefits comes from having all the storage equipment and system administration staff in a single location. Data is safer since a well-polished data protection procedure can be run in the centrally located data center as opposed to running ad-hoc procedures in many separate locations that have minimal staff.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

Most public storage clouds charge a monthly rate for each TB of storage used. Pricing is typically capacity based, so the more data you store, the lower the cost per TB. Additional *per TB* charges are made for transferring data out of the cloud, as well as *per request* charges for running operations (for example, PUT, GET, COPY) on the data. Today, public storage clouds are used mostly for *backup* and *archival data*. The adoption of public storage clouds for *primary data* has been slow. Storage clouds are not well suited for primary data because the latency of the WAN between the enterprise and the cloud has resulted in slow response times and poorly performing applications.

WAN Latency, the Achilles Heel of Clouds

The cloud, in all the usage models (public, private, compute, and storage) involves separating the computation platforms that run the applications from the storage systems that store the data. For today's enterprises, this separation is usually over a WAN and, at the low end, spans a couple hundred miles and, at the high end, spans many thousands of miles. It's a fact of nature that geographic distance adds latency due to the propagation delay of signals travelling hundreds or thousands of miles. Latency, in turn, kills performance. So how can you use clouds for performance-sensitive primary data and applications? Avere tiering hides the latency of the WAN and makes the cloud usable for primary data and applications. Avere automatically holds active data on FXT Series Edge filers nearest to the users of the data. The following sections explore this further.

Efficient Scaling of Application Performance with Compute Clouds

Figure 6-1 shows a typical configuration of an application using a compute cloud. Here, the application runs on the compute cloud on the right and the application data is stored

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

in the enterprise on the left. As the applications runs, data at the enterprise is constantly being accessed from the compute cloud. Every data access operation (for example, read, write, or metadata) must traverse the WAN and suffer WAN latency.

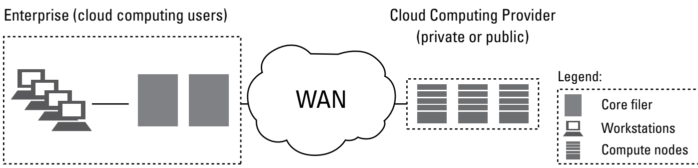


Figure 6-1: Traditional cloud computing is a prescription for high WAN latency.

Now, consider a situation in which the compute cloud is in the eastern USA. For enterprises also in the eastern USA, or within several hundred miles of the compute cloud, response times are roughly 25 msec. For enterprises in the western USA, or a couple thousand miles away, response times are roughly 100 msec. Both result in very slow application performance. So slow that the compute cloud cannot be used.

In the Avere architecture (Figure 6-2), FXT Edge filers are collocated with the compute cloud and hold the application's active data as close as possible to the compute cloud. The FXT cluster locally handles all read, write, and metadata operations and reduces average latency to 0.1 msec. This reduces the latency by a factor of 250–1,000 times what it would be without the Avere solution, making the compute cloud a cost-effective and high-performance solution for demanding applications.

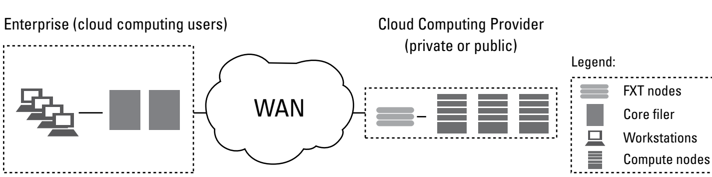


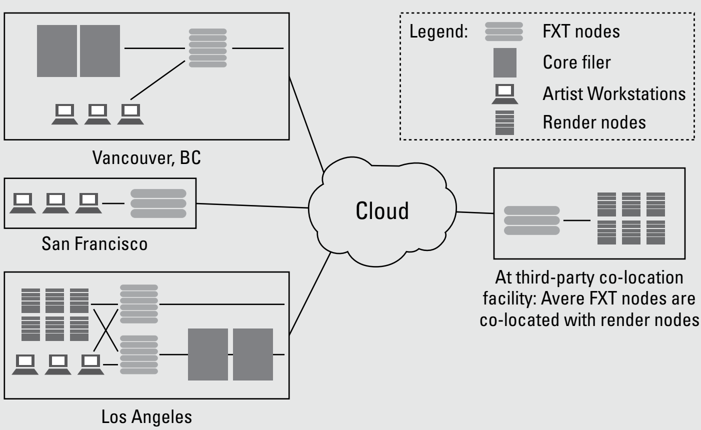
Figure 6-2: An Avere FXT Edge filer at the compute cloud localizes computation to the cloud, avoiding WAN latency.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

Case study: Digital Domain Productions

WAN latency can be an expensive problem for any business that has facilities distributed across distances measured in miles. One industry that is severely affected by this problem is the digital production segment of the motion picture industry. Digital Domain Productions, co-founded by celebrated director James Cameron, is a leader in the industry. A big part of Digital Domain's job is to seamlessly combine live action with virtual characters and scenery that's added digitally. Each frame of a motion picture requires massive computing at some centrally located render farm, in constant communication with digital artists located elsewhere. For cost reasons, Digital Domain's render farm is located in Las Vegas, but their artists are in Los Angeles, San Francisco, and Vancouver, BC. Even with powerful computers, it takes quite a while to render a single movie frame. Motion pictures are typically shot at 24 frames per second. Multiply that by 60 seconds per minute, 60 minutes per hour, and 2 hours per feature film, and you get a sense of the magnitude of the workload.

Digital Domain couldn't afford to locate their render farm near the high-rent districts where their artistic talent lives, which is why they liked the idea of locating it in Las Vegas. However, they couldn't afford to locate it in Las Vegas either because of WAN latency. The travel time from Las Vegas to San Francisco or Vancouver was just too long, until they discovered the Avere Edge filer solution. Take a look at the figure to see how the Digital Domain system is connected.



These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

Due to the nature of Digital Domain's work, the Avere Edge filer solution sped up operations by 250 times, turning a totally infeasible situation into something that met the company's needs. Using Edge filers probably won't speed up every business's operations by a factor of 250, but even a fraction of that time would make a big difference for just about anybody.

Enabling Storage Clouds for Primary Data

Figure 6-3 shows a typical configuration of an application using a storage cloud, which is the mirror image of the situation previously discussed for compute clouds. Here, the application runs in the enterprise shown on the left and the application data is stored in the storage cloud shown on the right. As the application runs, data in the storage cloud is constantly being accessed from the enterprise. Every data access operation (for example, read, write, and metadata) must traverse the WAN, suffer the latency of the WAN, and yield a response time, as is the previous case, in the 25–100 ms range. This makes the storage cloud unusable for this application. That is, until Avere enters the picture.

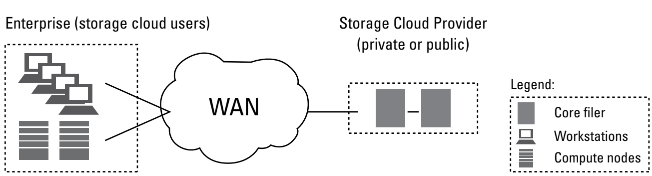


Figure 6-3: Latency is high because users must access storage across the WAN.

In the Avere architecture (Figure 6-4), FXT Edge filers are placed in the enterprise and hold the application's active data as close as possible to workstations, servers, and compute farms accessing the data. The FXT cluster locally handles all

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

read, write, and metadata operations and reduces the latency by a factor of 250–1,000 times, making the storage cloud a cost-effective and high-performance solution for the application.

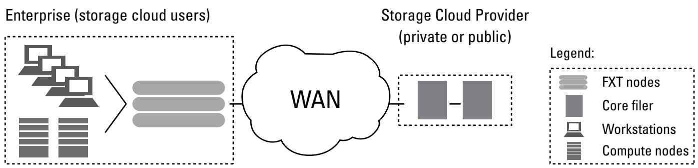


Figure 6-4: Both computation and the most frequently accessed storage are at the enterprise end, minimizing WAN traffic.

Case study: Sony Pictures Imageworks

Sony Pictures Imageworks is a visual effects and animation house that creates photo-realistic CG characters in blockbuster films such as the *Spiderman* movies, *Superman Returns,* and the surrealistic 3D *Alice in Wonderland* that starred Johnny Depp and Helena Bonham Carter. Remote artists and animators are scattered across many locations, all funneling their content to company headquarters in Los Angeles.

The company spent a lot of money, trying to slay the network latency monster. They tried WAN accelerators, Wide Area File Services (WAFS), and caching products. After all that expense, performance still wasn't satisfactory. Then they found out about Avere's FXT Edge filers. Avere's demand-driven operating system intelligently and automatically moves data to the optimal storage tier based on demand. Active data is immediately available on the Avere cluster's high performance FXT Edge filers, while inactive data is quickly shuffled to the traditional NAS Core filer. Administration at the remote sites is a breeze, being done remotely from headquarters. The company's heterogeneous environment is not a problem, since the Avere units don't care what product is used as the back-end Core filer, just as long as it talks industry-standard NFS.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

Chapter 7

Ten Helpful Storage Tips

In This Chapter

- Waxing wise about storage

This chapter offers you ten tips that will help you optimize the performance and scalability of your storage infrastructure while driving down costs, even as your storage challenges grow more complex with the introduction of new storage media and cloud services.

- **Don't let latency bog you down:** Organizations that need rapid response from remote sites need to make sure that distance doesn't get in the way of productivity. Strategically placed Edge filers can eliminate hard-drive, Core-filer-CPU, and network latency.
- **Think about scaling:** You know what storage you need now, but how well do you know what your needs might be three or five years from now? Is there a growth path from where you are now to where you might be then? Think about it.
- **Tier your storage:** Some of the items kept in your online storage are accessed frequently while others are accessed rarely if at all. Put the frequently accessed items in faster storage on Edge filers, but save money by putting the rarely accessed items behind Core filers on cheaper, slower storage.
- **Simplify with a single global namespace:** If parts of your storage system come from different vendors, or if this might become true in the future, getting everything to work together can be challenging. Integrating multiple Core filers behind a single Edge filer and global namespace could mask the difference, easing operations.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

- **Save money by optimizing storage:** Automatically place NAS data in the most appropriate storage tier for its activity level.
- **Separate capacity from performance:** Use Edge filers for only the amount of expensive, fast storage that you need. Put your rarely accessed data in cheap, slower storage devices.
- **Get the most for your dollar with edge filer storage:** Maximize performance delivered to remote sites by giving them an edge filer containing the data they are most likely to need next.
- **Alleviate problems caused by a lot of users doing the same thing at the same time:** Boot storms result when too many users try to access the same file resource at the same time. Spread out demand if you can. Failing that, install an edge filer to solve your problem.
- **Reduce capital expenditures with cloud computing:** If your data is in the cloud, you don't have to buy and maintain a lot of expensive storage. A small, relatively inexpensive edge filer to eliminate latency may be all that you need.
- **Provide robust data protection while retaining high productivity:** Avere's FlashMove and FlashMirror enable you to distribute copies of your critical data across multiple machines in multiple places. No single disaster will take you out. Productivity remains high because of the easy accessibility of the data no matter where you are on the network.

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

Give your NAS storage a big performance boost!

Today, more and more organizations need to deal with massive quantities of data. Accessing that data without frustrating and costly delays — latency — becomes harder as the amount of data stored grows larger. This book explores what you can do to eliminate latency and maintain high-speed operation, even as the amount of data you store multiplies.

- ***Go beyond the limits of the conventional Core filer* — understand how Edge filers use Flash memory to handle performance scaling**
- ***Implement dynamic tiering* — install storage that dynamically reconfigures itself to match the processing load**
- ***Take advantage of the cloud* — reduce costs and enhance the accessibility of your data for remote users**
- ***Locate storage near computation* — Edge filers place high-performance storage where it will do the most good, close to the processors doing the computation**

Open the book and find:

- **How to configure NAS storage for optimal performance**
- **Ways to reduce the cost of upgrades as your needs for storage grow**
- **How to minimize the impact of multiple users hitting the storage system at the same time**

Go to Dummies.com®
for videos, step-by-step examples, how-to articles, or to shop!

Allen G. Taylor is the author of more than 30 books, including *SQL For Dummies, SQL All-in-One For Dummies, Database Development For Dummies,* and *Crystal Reports 2008 For Dummies*. He resides in Oregon City, Oregon.

WILEY

978-1-118-48765-5
Not for resale