Highwinds Software

HARDWARE

Sizing Guide

Document Version 1.1
April 2003

Highwinds Software LLC

Highwinds Software

Introduction

Managing Usenet is brutal on hardware. As of this writing, Usenet ranges from 350 to 450 Gigabytes of new articles every day meaning Usenet feeds deliver between 20 and 40 new articles every second. Ensuring proper hardware specifications will assure you have a Usenet system that can handle the current load of Usenet as well as be able to grow with Usenet.

Overview

Tornado is comprised of two separate Usenet server products. The first, Tornado Back End, manages spool objects and forwards overview and active information to Tornado Front Ends. The second, Tornado Front End, manages client connections and delivers overview and active information to them. Tornado effectively off-loads the bulk of random I/O operations from the Back End to the Front Ends and allows significant hardware leverage by allowing spool sharing .

Tornado was developed to facilitate architectures of many inexpensive machines (Tornado Front End) communicating with only a few costly machines (Tornado Back End). While Tornado Front End will work phenomenally on high-end servers, it's designed to save the Usenet provider from having to invest as heavily in storage and servers.

The Tornado product is flexible and lends itself to many different network architectures. Below are some general guidelines to use in sizing Tornado systems.

Storage Options

A current full feed is approximately 350-400 GB a day. The majority of a Usenet feed consists of binary articles pertaining to the Music, Movies and Software groups. In a close second, are the binary articles that compose the hundred or so erotic groups. Out of the total Usenet volume, text groups consist of 25 – 50 GB worth of volume per day.

To determine how much storage should be allotted for a Tornado Back End, refer to Table 1.0.

Rough % of Usenet Traffic	GB/day	Groups
93%	372	binary groups
7%	28	text groups

Table 1.0

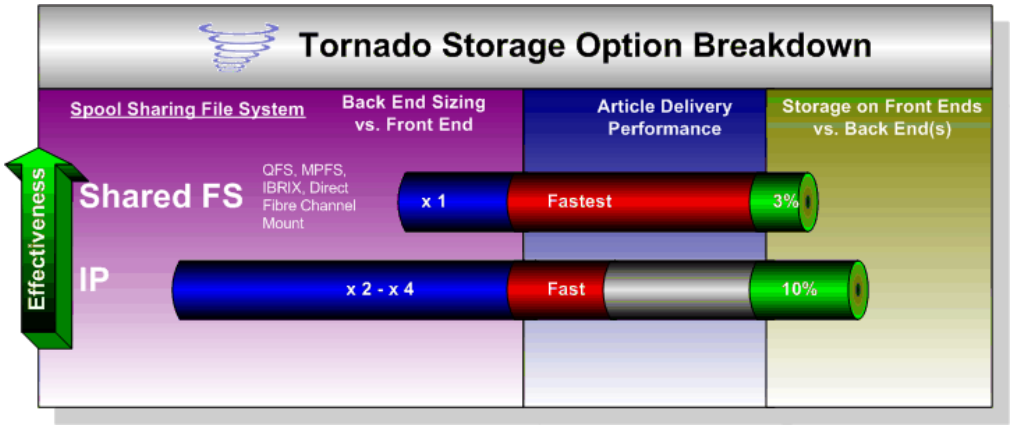

Figure 1.0

Once a Tornado Back End's storage requirements are decided, Tornado Front ends can be sized according to the comparison given in Figure 1.0. Critical for Tornado Front End performance is spreading data over as many disks as possible to alleviate I/O demands. A typical Tornado Front End should have 10-12 disk drives to spread overview data over. The drives need only be 9 GB, so opt for many smaller drives when building machines.

Usenet is potentially the most I/O intensive service per user that exists on the planet. Disk vendors cringe at the thought of what Usenet does to their hardware... So be sure to keep this in mind when designing your storage architecture. What matters **is not** total I/O throughput, but latency, seek time, and true random read performance. Seek times of disk drives have not made the same technology leaps as other tech hardware, their improvement has been mostly linear since the early 1990's, so while big bad SAN's have massive sequential I/O capability, the underlying disks still have a relatively mediocre seek time performance.. Whether using RAID (software RAID is not recommended) or JBOD (Just a Bunch Of Disks), maximize for random read performance.

When building an I/O system, you have many options: SAN, NAS, Direct Attached using SCSI, Ultra Wide SCSI, Fibre Channel, etc. Fancy and costly storage architectures are sometimes abysmal choices for Usenet, so you don't necessarily get what you pay for. Direct attached Fibre Channel is Highwinds' typical recommendation, but consider the following when choosing:

- What is the latency of an I/O request?
- What is the path of the data delivery and does it contend with other resources in short supply (e.g. Network in the case of NAS)?
- When the system gets bombarded by random read requests and sequential write requests (incoming feed) simultaneously, how will it respond?

File System Choice

Great storage with a poor file system choice will hamper the overall quality of the Usenet architecture. No matter what file system you choose, **always** 'newfs' the storage with a brand new file system so fragments and thrashing are kept to a minimum.

Tornado introduces the idea of *directly shared spools* which allows Tornado Front End to have a local file system that is really a remote mount from the Tornado Back End. Many Tornado Front Ends can share the spool of a Back End using this method. File system choice (coupled with OS choice) when using directly shared spools can make an order of magnitude swing in the performance of the I/O system as a whole.

The most performant option (at the time of this writing) is to use a file system that can be shared using direct

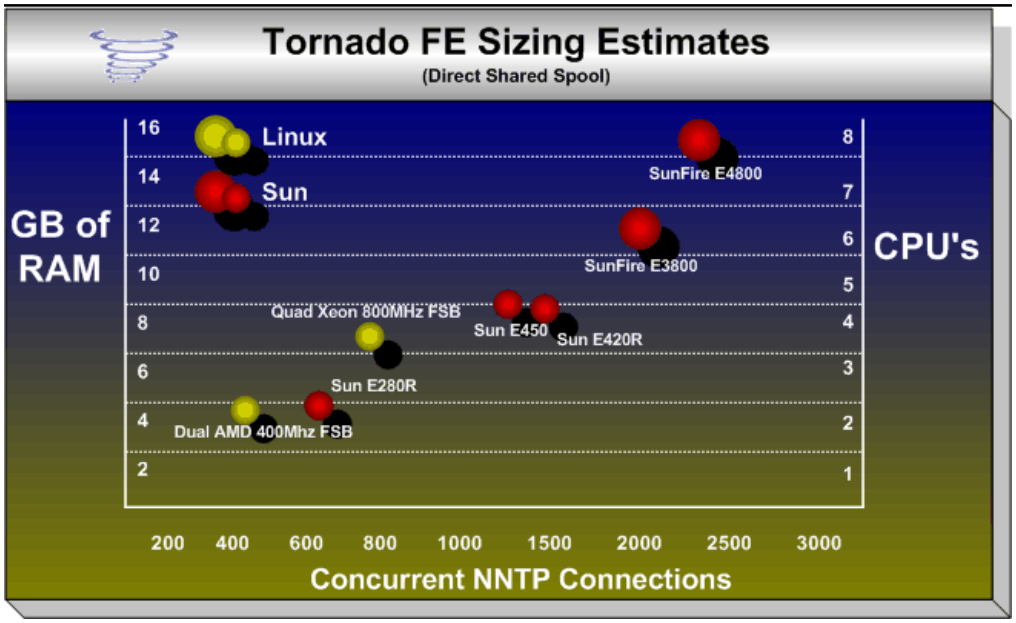

Figure 1.1

Fibre Channel mounts. Sun QFS has been tested and performed very well in this environment. Other options include EMC Highroad, Veritas VXFS, IBRIX, SANergy, etc. Using a shared file system with dedicated I/O channels to the storage will yield the best possible performance (up to a 30-50% increase over spools shared with Tornado's IP mechanism) when deploying Tornado.

Another option is NFS using a NAS. NFS is not flawed in it's design, but has proved to be in several implementations and should be used with caution and testing. Solaris 9 introduced a new NFS client and server that has increased performance significantly. NFS version 2 is also preferred over NFS version 3. Deployment of NFS should stray from other implementations where clients could "share" a connection to the server for data. Usenet usage is too demanding for such a configuration and the clients and servers should be in great enough quantity to dedicate at least one per concurrent Usenet connection.

Operating System Requirements

As of this writing, Highwinds' products run on the following operating systems:

- Solaris 9, 8, 7, 2.6 (Sparc)
- Solaris x86
- Linux
- Free BSD
- Tru64 Unix

Unfortunately (from the administrator perspective) the above operating systems are not created equally. Highwinds' products are massively multi-threaded and require high performance threading libraries to operate well and use the hardware to it's fullest potential. Second to storage architecture, the OS is the next major component in determining an architecture's scalability.

Experience has shown that Solaris on Sparc (later versions are more perfomant) and Tru64 Unix provide the best threading support. The other operating systems don't have the maturity in threading and will cost about **2x to 3x times the resources** (i.e. for CPU, RAM and Concurrent Connections figures below, estimate

accordingly) of the higher performance operating systems. The Linux 2.6 kernel is reported to have a totally redesigned thread scheduler that should help the performance of Linux installations.

Network Requirements

Supporting large amounts of NNTP traffic requires a significant amount of network resources. Inbound & outbound Usenet feeds consume approximately 33Mbps of bandwidth each. The Usenet feed volume is doubling in size every 12-16 months so such growth should be considered. Gigabit Ethernet is recommended for reader servers. At least 100Mbps for low volume Cyclones and 400Mbps or 1Gbps for higher volume Cyclones.

CPU and Memory Requirements

Highwinds products efficiently use CPU resources given the task of managing Usenet. The addition of more disk and memory to a system can greatly increase the machines CPU requirements. Other factors, such as customized authentication programs will also influence your CPU usage.

Memory on a Usenet server is often in short supply and thus you can never have enough RAM in a machine. On a Reader server, broadband readers consume about 256kb each and dial-up users about 512kb each (more i/o buffering is required with dial-up). Under Solaris, all the system RAM is always being used. If applications aren't using the memory, Solaris will use the memory to cache file system data. The Solaris caching behavior can greatly increase news server efficiency as well as decrease the I/O load on the disk subsystems. Figure 1.1 shows some sample hardware choices for a required number of concurrent NNTP connections.

Additional Questions

Please direct any sizing questions to support@ticket-highwinds-software.com.
Given the above metrics, Highwinds generally recommends about 1 CPU and 1 GB of RAM per 250-300 broadband users and about 1CPU and 1 GB of RAM per 350 dial-up users.