FAQs for Developers

Listed below is information about OmniTraak in question and answer form that will provide additional information of a technical nature about the OmniTraak Software Components.

1. **What platform support is available for OmniTraak? What is involved in porting it to another platform?**

   The OmniTraak File System currently is supported and operational on NetBSD. It also runs on two proprietary operating system implementations: Conexus and FasTraak, which is available from Traakan as a reference hardware platform. In addition FreeBSD and BSD/OS will be supported. Linux, QNX, VxWorks, and Solaris are planned.

   Traakan has thus far worked with 3 very different operating systems and thus has a good understanding of the tasks involved with porting to different platforms. For those mentioned above, we estimate that approximately 8 engineer weeks are required to port the file system and NFS portions of OmniTraak.

2. **Has OmniTraak performance statistics for NAS been benchmarked against competitive offerings?**

   Over the years Traakan (and the industry) has looked at a variety of methods for testing the capabilities of file system technologies. This included tests like Netbench and IOMeter. After extensive analysis it was concluded that the SPEC SFS 3.0 was the only test suite that met all conditions of an effective NAS benchmark. In designing the test methodology, the SPEC organization gathered data on more than 1,000 servers in different application environments and found that 60 percent of these users have similar mixes of NFS operations. SPEC generated the testing model after these data patterns, which is a true representation of real world NAS workload. In order for a vendor's SFS results to get published on the SPEC website, they must go through a 2 week review process by a committee of representatives from around the storage industry to ensure adherence to SPEC guidelines. SPEC validated results allow customers to confidently evaluate two or more NAS solutions side by side. SPEC has leveled the playing field for performance evaluation.

   Traakan has posted several record setting results on the SPEC site. They can be seen at: http://www.spec.org/osg/sfs97r1/results/sfs97r1.html.

3. **How does OmniTraak dramatically increase performance on benchmarks?**

   There are literally dozens of techniques used to boost performance. Here are some major examples.

OTFS provides high levels of overlapping operations by using a lock free design. Read-ahead and sync rates are automatically tuned to maintain optimum response times as load shifts. OTFS journaling provides excellent responsiveness while meeting strict stability requirements required by the SPECsfs97r1 benchmark.

OTNFS processing is streamlined, uses no intermediate buffering, and interacts with the networking subsystem in a highly efficient manner.

OTNFS obtains file service directly from OTFS thus avoiding the costs of the VFS layer. OTNFS and OTFS work together to provide zero-copy file read, meaning file data is transmitted to NFS clients directly from the OTFS file cache.

4. **What specific things does OmniTraak have for High Availability that would be of interest to programmers?**

   Traakan has a good deal of experience with HA systems and has designed OmniTraak to exist in that arena.

   OTFS journaling assures that the on-disk image of the file volume is always in a restartable state. When a node takes over the disks of a lost node, all file operations up to the point of failure are recovered from the journal. The event much resembles a server reboot, and applications simply see a stall.

   For example, an application that is doing a sequence of 200 file operations might see a stall after doing the first 100. When service resumes it simply completes the last 100. The first 100 operations are perfectly protected by the journal.

5. **Does OmniTraak streamline development for programmers? Are capabilities provided that makes it easy to add new features in support of new markets and applications?**

   Absolutely. A key objective of OmniTraak is that it is extensible and makes adding new features straightforward without disrupting the design or an installed base.

   The OTFS inodes are organized as a list of variable sized properties. New property types are easily added while maintaining support for all existing properties. Most portions of OTFS deal with properties opaquely, so even the OTFS checker (otfs_fsck) is ready to go when property types are added. New property types may add features, or may replace prior property types as an enhancement. OTFS is adapted to add or upgrade properties on the fly, thus an installed base is not disrupted.

   The OTNFS and OTFS are built around an extensible, highly efficient, message-oriented interface called the OmniTraak File Service (OTFSVC). New file operations – message types – are easily added without disrupting existing message types.

New OTFS directory entry types may also be defined while maintaining support for existing entries. For example, Traakan has reserved directory entry types to support an NT-style entry that gives a file a true name and a DOS compatible ("mangled") name.

The OmniTraak family is completely integrated with the target operating system. All COTS products that operate under the target operating system will work with any or all components of the OmniTraak family. In addition, the filesvc API supports direct calls to OTFS for specialized applications.

6. Is OmniTraak compatible with VFS (Virtual File System) and Vnodes?

OmniTraak is compatible with VFS, but does not require it. The OmniTraak file service (OTFSVC) interface is extremely simple and lightweight. Translating between VFS and OTFSVC interfaces is straightforward and uses little processing power. The translation costs are a fraction of VFS costs. Internally, OTNFS issues OTFSVC requests directly to OTFS to avoid VFS costs and thus boost NFS performance. Further, OmniTraak may readily be used with operating systems (such as micro kernels) that do not use VFS.

7. Does OmniTraak have disk stripping and other high performance disk I/O bandwidth features?

An OTFS file volume consists of one or more segments. OTFS distributes file allocation across segments based on individual segment performance. For example, when a file volume has been expanded with a faster disk subsystem than the earlier segments, OTFS will proportionately place more allocations on the new faster disk subsystem.

Each segment may be contained by any type of disk subsystem, including simple disks, a logical volume manager such as OTVM, or by a RAID subsystem. Disk technologies may be freely mixed.

The OTFS design includes a direct I/O segment. A direct segment is used to contain a modest number (60000) of large, high-speed files. I/O to direct files is done directly between the disk subsystem and the application's I/O buffer without intermediate kernel buffering.

8. Does OmniTraak do anything special that will reduce or eliminate disk fragmentation problems compared to traditional file systems?

OmniTraak employs several effective methods to minimize fragmentation.

OTFS uses a 4096-byte disk page size. There are no fragments as with traditional file systems. A fixed-sized page provides 14% better disk utilization than a blocks/frags design. There is no need to waste a 10% reserve to assure blocks (consecutive frags) are available. And overhead for the allocation map is one-eighth the size, proportionately reducing disk space, main memory, and search costs.

Inodes are contained in dynamically allocated pages. There is no preallocated, fixed sized inode table, and so no wasted disk space for unused inodes. Over half of all files and over 90% of all directories fit within the inode itself. This reduces the amount of allocations and fragmentation, and also improves I/O efficiency since subsequent I/Os are not necessary.

OTFS uses sophisticated allocation rules that result in sequential allocation of disk space for files a high percentage of the time. This improves I/O efficiency because files are often retrieved in a single I/O.

9. **Does OmniTraak provide multiple file access methods for cross platform storage appliances?**

   OmniTraak includes state-of-the-art NFS and VFS support and is designed to fully integrate into the target operating system. OmniTraak is an accelerator for software packages that require fast and reliable file system services. OmniTraak is a good choice for a NAS server that must also provide CIFS access using SAMBA, HTTP using Apache, NDMP using Veritas , etc. In the future, OmniTraak will have specialized enhancements for these applications.

10. **What capabilities does OmniTraak provide that make it easy to administer the system post installation? Can inodes be dynamically expanded? Can the number of active files be increased? Does the volume manager allow the expansion of disk partitions?**

    OTFS inodes are dynamically allocated and dynamically shaped. There is no preallocated table of small, fixed format inodes. There are no administrative burdens or downtime required to increase the number of inodes or to apply OTFS enhancements.

    OTFS volumes can be expanded in two ways. An existing segment may be grown larger, which many RAID subsystems and logical volume managers such as OTVM support. The second way to expand an OTFS volume is to attach another segment. This flexibility works well with sophisticated RAID subsystems and also with simple one or two disk configurations. Logical volume managers are nice but are not a requirement to expand an OTFS volume.

    OTVM can create, grow, or delete a logical unit at any time. Disks may be added to an OTVM pool at any time.

11. **What sort of information is provided by OTFS to the application about errors?**

    Error and warning events are delivered by the platform specific logging facilities. The logging facilities of a platform usually have sophisticated configuration, routing, storing, aging, and notification methods such as email and beepers.

12. Does the software scale?

Yes. In addition to 64bit support and dynamic adjustment to memory availability, OTFS supports disk caching, symmetric multiprocessor configurations, and software scaling to the specifications of the I/O subsystem. In addition, file access and writing is supported in large configurations at near rated device speeds.

13. Does OmniTraak provide an embedded version of OTFS? If yes, what is different?

OmniTraak is built to be portable to a variety of operating systems and platforms. Embedded operating systems, real-time operating systems, and micro kernel operating systems are all candidates for OmniTraak. The Platform Generic Components are identical between all platforms. The only difference between platforms is the small Platform Specific Components ("glue").

14. Are the details of the OTFS snapshot data published so that application programs may make effective use of them? Or are innovative uses of snapshots only possible for Traakan and OmniTraak OEMs?

OTFS snapshots include a feature called BRAM (Backup and Replication Access Method). BRAM allows access to the snapshot database information in a straightforward way. No technical details need be published. The express purpose of BRAM is to encourage innovations in backup and replication technology by interested parties regardless of their license status.

15. Does Traakan provide after sale consulting services to HA product developers? Put stuff about SPEC testing in here.

In addition to providing the software components available in OmniTraak and on going releases and support, Traakan makes its expertise available to developers on a Non-Recurring Engineering (NRE) basis. This capability can be used to port OmniTraak to an unsupported platform or to help in the fitting of our components to the customer's application. And because of our extensive experience with performance tuning of the OmniTraak software, Traakan also will make its resources available on a contract basis to aid in configuration and testing for SPEC 97 performance runs.