

Wyreless.NET Solutions Case Study

The Wrong File System

The Original Assignment:

A leading computer manufacturer was expanding their product line from workstation products to server products. The higher margins and burgeoning server business was attractive since they were approaching market saturation in the workstation business. They were concerned about the long reboot times associated with “global consistency checking” the file system and how it would adversely affect the necessary high availability. With our track record in high performance file systems, we were recruited to rectify the problem.

What We Discovered:

After reviewing the file system architecture of their workstation products, we discovered that the problems were much deeper. The products’ file system technology did not scale from single user workstations (for which it was designed) to multi-access servers. The core on-disk structure was not well suited to heavy duty “industrial-strength” data processing applications. Global consistency checking was time consuming and required exclusive access, which prevented any application access and reduced availability. Space allocation was highly fragmented increasing seek times and reducing performance. Algorithms were not tuned for large data applications primarily due to the sequential nature of storage structures. Although marginal improvements were possible, it would continue to be more and more problematic with each succeeding generation servers, larger disk subsystems and more demanding applications.

What We Did:

We proposed and prototyped a next generation file system that relied on database transaction technology. This new file system provided the features required for the new target markets: high availability and performance. High availability was enhanced with very fast crash recovery, on-line maintenance, and zero data corruption. System crashes due to file system errors were eliminated. Performance was optimized with contiguous allocation support, more aggressive I/O clustering, improved random I/O (critical in database applications), disk striping, and a multiprocessor-ready design.

The Results:

The prototype was well received by the company’s core software team. Many of the notions contained in it were adopted and designed into their next system version. The company, in conjunction with the leading database company, continues to set performance and high availability benchmark records.