

File Virtualization Performance: Understanding CIFS Create ANDX



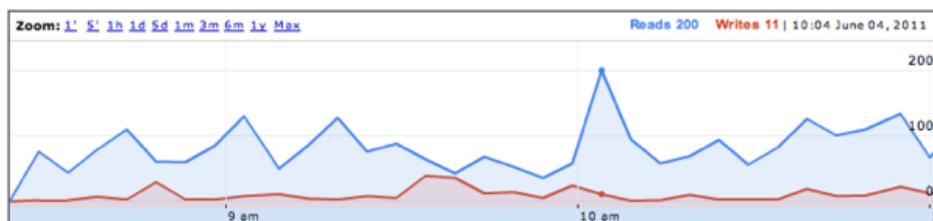
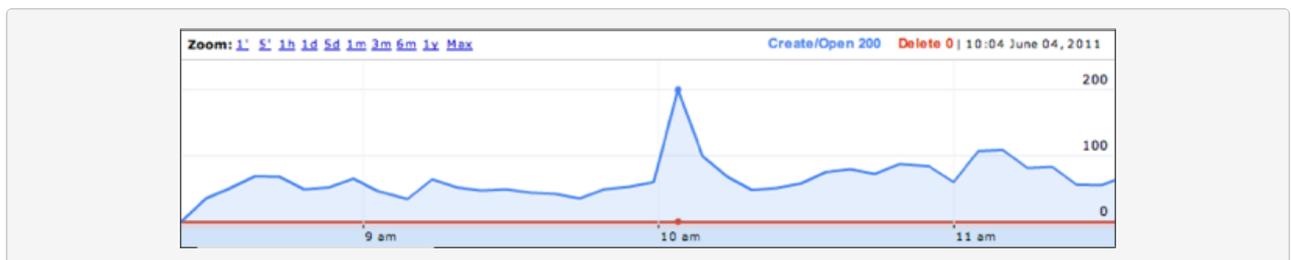
Derrick Schommer, 2011-06-09

Once upon a time, files resided on a local disk and file access performance was measured by how fast the disk head accessed the platters, but today those platters may be miles away; creating, accessing and deleting files takes on a new challenge and products like the *F5 ARX* become the Frodo Baggins of the modern age. **File Virtualization** devices are burdened with a hefty task (this is where my Lord of the Ring Analogy really beings to play out) of becoming largely responsible for how close your important files are to your finger tips. How fast do you want it to perform? Quite expectedly, “as fast as possible,” will be the typical response. File Virtualization requires you to meet the expectations of an entire office—many of which are working miles away from the data they access and every single user hates to see a waiting cursor. To judge the performance of a storage environment we often ask the question, “How many files do you create and how many files do you open?”



Fortunately, *Microsoft CIFS* allows humans to interact with their files over a network and it does so with many unique *Remote Procedure Calls (RPCs)*. One such procedure call, *Create ANDX*, was initially intended to create new files on a file system but became the de-facto standard for opening files as well. While you and I can clearly see an obvious distinction between opening a file and creating a file, CIFS liberal use of *Create ANDX*, gives us pause, as this one tiny procedure has been overloaded to perform both tasks. Why is this a problem? Creating a file and opening a file requires a completely separate amount of work with entirely different results, one of the great challenges of File Virtualization. Imagine if you were given the option between writing a book like *The Fellowship of The Ring* or simply opening the one already created. Which is easier?

Creating a file may require metadata about the file (security information, other identifiers, etc.) and allocating sufficient space on disk takes a little time. Opening a file is a much faster operation compared to “create” and, often, will be followed by one or more read operations. Many storage solutions, EMC and Network Appliance come to mind, have statistics to track just how many CIFS RPC’s have been requested by clients in the office. These statistics are highly valuable when analyzing the performance of a storage environment for File Virtualization with the *F5 ARX*. Gathering the RPC statistics over a fixed interval of time allow easier understanding of the environment but one key statistic, *Create ANDX*, leaves room for improvement... this is the “all seeing eye” of RPC’s because of its evil intentions. Are we creating 300 files per second or simply opening them? Perhaps it’s a mix of both and we’ve got to better understand what’s going on in the storage network.



When we analyze a storage environment we put additional focus on the *Create ANDX* RPC and utilize a few other RPC's to try to guess what the client's intentions so we can size the environment for the correct hardware. In a network with 300 *Create ANDX* procedures a second, we would then look into how many read RPC's we can find compared to the write RPC's and attempt to judge what the client is intending to perform as an action. For example, a storage system with 300 "creates" that then performs 1200 reads and five writes is probably spending much of its time opening files, not creating them. Logic dictates that a client would open a file to read from it and not create a 0-byte file and read emptiness, which just doesn't make much sense.

Tracking fifteen minute intervals of statistics on your storage device, over a 24-hour period, will give you a bit of understanding as to what RPC's are heavily used in the environment (a 48-hour sample will yield even more detailed results.) Taking a bit of time to read into the intentions of *Create ANDX* and try to understand how your clients are using the storage environment, are they opening files or are they creating files? Just as creating files on storage systems is a more intensive process compared to the simple open action, the same can be said for the *F5 ARX*. The *ARX* will also track metadata for newly created files for its virtualization layer and the beefier the *ARX* hardware, the more file creations can be done in a short interval of time.

Remember, while it's interesting and often times impressive to know just how many files are virtualized behind an *F5 ARX* or sitting on your storage environment, it's much more interesting when you know how many are actually actively accessed.

With a handful of applications, multiple protocols, dozens of RPC's, hundreds of clients and several petabytes of information, do you know how your files are accessed?

F5 Networks, Inc. | 401 Elliot Avenue West, Seattle, WA 98119 | 888-882-4447 | f5.com

F5 Networks, Inc.
Corporate Headquarters
info@f5.com

F5 Networks
Asia-Pacific
apacinfo@f5.com

F5 Networks Ltd.
Europe/Middle-East/Africa
emeainfo@f5.com

F5 Networks
Japan K.K.
f5j-info@f5.com