

Hi, my name is Manoj Naik and I am from IBM Research. We are here at SuperComputing 2007 demonstrating a technology preview of using GPFS for WAN caching. The basic idea is to provide a mirror of the remote file system using PNFS as the protocol to exchange data from the remote server to the local cache cluster. PNFS is an emerging standard for valid data transfer. It is based on open standards using NFS technology. The basic idea here is to take data that is remotely available and bring it in whenever a user tries to access data from the cache, if it is not locally available. Once the data is available in the cache, then the user has complete control. He can access the data from the local cache and get the maximum disk bandwidth from the local cache cluster. Once the user tries to access some data that is not available in the cache, the data is fetched over the network using PNFS and writes that the user performs are first written to the cache locally and then streamed back to the remote server. Here we are demonstrating that we can get the maximum bandwidth out of the network which in this case is 3 – 6 gigabits per second. And once the data is cached locally, we can get the maximum local bandwidth which is greater than a gigabyte per second.

We are also showing the GPFS is a heterogeneous cluster file system. We have a BladeCenter here that is running different operating systems, Linux, AIX and Windows. We are showing that GPFS can have a single file system that can be accessed from different operating systems.