Carlos Sathler | cssathler@gmail.com
INFO-I 535 – MGMT ACCESS USE BIG DATA (Fall 2016)

DISTRIBUTED FILE SYSTEMS ASSIGNMENT

1. Semantics of Sharing (topic of section 3 of the reading) pertains to what? Select one answer:
   Answer 2: The expected behavior when multiple clients (user program) have the same file open for reading or writing

2. What kind of information will a distributed file system store in a cache?
   File data, including file header information and file location information, which can be used as a "hint" to locate a file in the DFS.
   What makes a cache a beneficial piece of a distributed file system to have?
   In short, it reduces network traffic. In the case of location information, it may help a client fetch file data with fewer trips over the network to properly locate the file. In the case of file data proper, a cache will normally contain more file data than initially requested by the client; it will contain the requested data plus data that is in its vicinity. Once file data is cached in the client file system (or memory), there is a good chance that many read requests for file data will be met locally. As a result, a cache will greatly reduce the need for frequent transfers of file data from servers to clients, in response to read requests. Conversely, when it comes to updating file data (write requests), many DFS implementations will delay updates to the file data in the server (e.g. with Session Semantics, until the file is closed in the client), so that transfers of modified file data from clients to servers are also less frequent. All in all, a cache will greatly benefit a DFS by significantly reducing network traffic.

3. Drawing on the framework provided by the list of characteristics of a distributed file system given in Section 1 of the reading, compare and contrast how a distributed file system and its use differs from the web and app based file storage used by a modern smart phone.
   I identified the following main characteristics of a DFS according to Section 1:
   1. Looks like a conventional, centralized file system.
   2. Offers "network transparency" – no distinction between local and remote files and file copies.
   3. Offers "user mobility" – users land on the same "home" directory and access files through the same paths regardless of how they connect to the network.
   4. Fault tolerant – A DFS is an aggregation of multiple components so that if one fails, another can act as a backup to deliver the same functionality
   5. Scalability – A DFS can increase its capacity by incremental aggregation of new system components; e.g. addition of new file servers.
   Lastly, I think this quote from the text is very important: "The fact that a DFS manages a set of dispersed storage devices is its key distinguishing feature" (p 325).

   Web and app based file storage used by modern smart phones is implemented through cloud computing [1], understood as "a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources" [1, p2], in our case, storage, "that can be rapidly provisioned and released with minimal management effort or service provider interaction" [1, p2]. Specifically, when it comes to web and app file storage used by smart phones, the cloud solution is implemented under the Infrastructure as a Service (IaaS) service model [1, p

3]. Examples include iCloud, DropBox and Google Drive. These cloud solutions exhibit the key DFS distinguishing feature of managed dispersed storage devices and offer a significant improvement over traditional DFS systems when it comes to user mobility, fault tolerance and scalability, thanks to some of the inherent characteristics of cloud computing, namely "broad network-access" (the cloud is available to a smart phone everywhere, as long as data service is available) and "rapid-elasticity" which is the ability of cloud software to allocate new resources as needed for added capacity, improved performance or to respond to failure of system components [1, p2].

However, cloud solutions for smart phones file storage do not offer "network transparency" since it is always clear to users accessing a file through their smartphone whether the file is in the cloud or in their local device, as a downloaded copy. Consequently, the web and app based file storage used by a modern smart phone does not resemble a conventional, centralized file system.

1 - Mell, Peter, and Tim Grance. "The NIST definition of cloud computing." (2011).