HPC User Site Census: Storage

Christopher G. Willard, Ph.D. Addison Snell Laura Segervall

August 2012

EXECUTIVE SUMMARY

This report, part of our Site Census research, provides an examination of the storage characteristics and capacities found in a sample of HPC user sites. Intersect360 Research surveyed a broad range of users about their current computer system installations, storage systems, networks, middleware, and supporting software. The resulting data is presented in a series of reports throughout the year, each describing different aspects of the HPC end user landscape. This report examines storage usage within the HPC user communities and explores how this usage varies based on categories such as storage capacity, site configuration, supplier, and network.

Key findings of the survey include the following:

- Approximately 36% of the total available storage at respondents' sites resides on compute servers. Storage available to each node (referred to as node-level storage) represents 13% of the capacity and storage available to the server (referred to as system-level storage) accounts for 23% of the total available storage. The remaining 63% of storage is found at the site level, generally on Network-Attached Storage (NAS) or Storage Area Network (SAN) systems. On average, 332TB of storage resides at the site level on a storage system.

- About 82% of the 405 sites have at least one site-level storage system installed. No vendor dominates the storage system market for HPC sites. IBM has the largest share with 14.2% but is closely followed by DataDirect Networks with 14.0%.

- Data revealed that InfiniBand is used at 27% of the sites while 10 Gigabit Ethernet is used at 23% for site-level storage system networks.

- Most storage management software (34%) in use by the survey sites was provided by the storage system vendor. Lustre and GPFS were the next most frequently mentioned storage management software with 23% and 15% of the systems, respectively.

- Clustered parallel file systems are gaining share in the HPC market and represent 46% of the storage management tools, up from 33% in the previous report. As these file systems gain greater share, the practical distinction between NAS and SAN may diminish.