

High-speed File Transfer Tool with the Gfarm File System

WATANABE, Hidenobu^{1*} ; KUROSAWA, Takashi² ; MURATA, Ken T.¹

¹National Institute of Information and Communications Technology, ²Hitachi Solutions East Japan, Ltd.

A distributed storage system of scale-out type is gradually being used in the High Performance Computing (HPC) to store large scale data. NICT is also running an about 3 petabyte-scale (PB) distributed storage system with the Gfarm file system and a 10Gbps Layer 2 network (JGN-X) in Japan. Gfarm is open source software of a distributed file system for a petabyte-scale grid computing, and has been adopted as a shared storage of the High Performance Computing Infrastructure (HPCI).

When Gfarm copies data between storage servers in long-distance network, it uses a multiple TCP streaming technique to transfer data faster because TCP single streaming is known to produce a low network throughput in a long distance network. However, efficiency of high-speed by the technique becomes low as more distant.

We developed a high-speed file transfer tool worked with Gfarm. The tool adopts the UDT protocol as a data transfer protocol and has a control function for a parallel data transfer. UDT is a reliable UDP based application level data transport protocol over wide area high-speed networks, and uses UDP protocol to transfer bulk data with its own reliability control and congestion control mechanisms. In fact, UDT can provide a high network throughput than TCP in a long distance network.

We explain our tool and report the performance results of the tool in basic evaluation.