Motivation
Applications in Grid environment are required to share distributed large blocks of data among distributed multiple sites. Network throughput for data transfer affects total processing time as well as the task processing.

The predicted network throughput would be a useful parameter on scheduling tasks to improve the total processing performance.

Purpose of our research
Improving precision of the throughput prediction method called “Network Weather Service” [1]. Adapting the prediction method to virtualized hosting environment, which shows anomalous behavior more frequently than physical nodes.

Connection pair
A network throughput prediction has been a challenging issue due to the dynamics of network traffic and no guarantee for bandwidth reservation. Connection pair uses a small size of probe transfer to predict the throughput of a large size of data transfer. In previous work [1], the restricted sets of pairs on probe and data size were examined on limited network environment.

Traffic anomalies
The evaluation results were affected by oversize packet spacings, which are caused by CPU scheduling latency. The packet spacing which is larger than the TCP transmission period involves packet transmissions, which results in severe throughput[2].

Experimental settings
We used PlanetLab nodes, equipped with a virtualization mechanism called V-Server (http://www.linux-vserver.org). Various sizes for both probe and data transfer are used. Correlation between both probes is evaluated by Spearman’s rank correlation coefficient, one of non-parametric metrics.

Original result
Smaller-size probes had better conditions than larger-size probes.

Results without anomalous cases
We re-evaluate the results without the anomalies, and a larger-size probe is required for improving predictability. If throughput is decreased by the anomalies, we should carefully review measurement results.

Conclusion and future work
Anomalies from virtualized hosting environment have great impacts on the prediction results. We re-evaluate the results without the anomalies, and found that larger-size probe is required for improving predictability.
Our future works are to measure throughput with various probe size and to devise an anomaly estimation method.

References