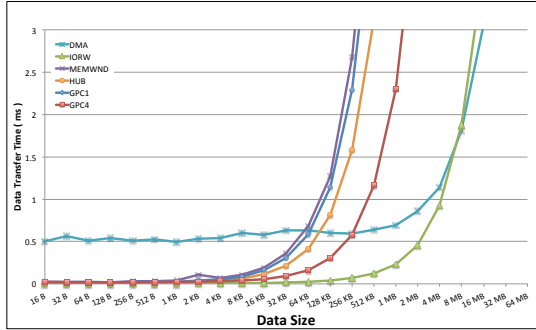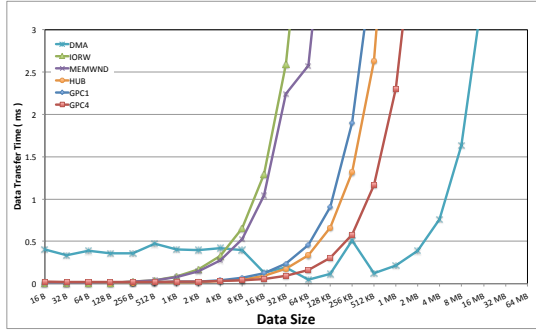# Data Transfer Matters in Low-Latency GPU Computing

Yusuke Fujii, Takuya Azumi, and Nobuhiko Nishio
College of Information Science and Engineering
Ritsumeikan University

Shinpei Kato and Masato Edahiro
Department of Information Science
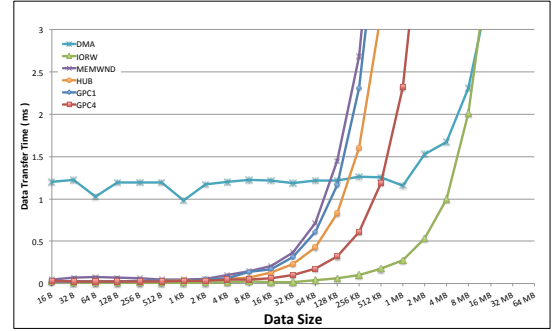*Nagoya University*

(a) Host to Device

(b) Device to Host

Figure 1.   Average data transfer times of single streams.
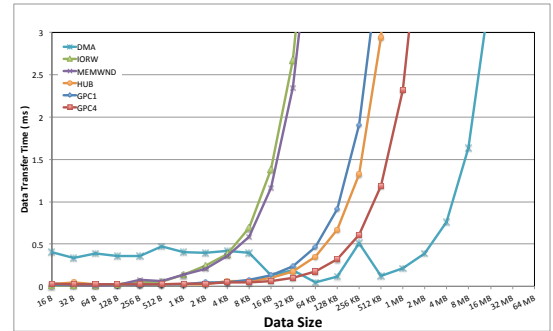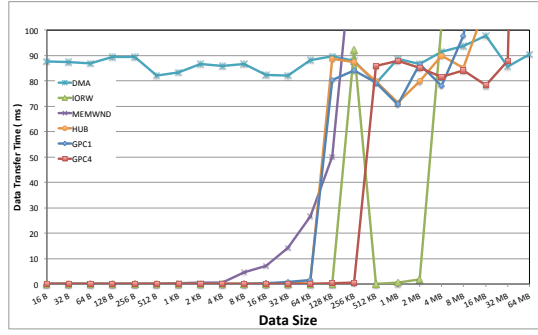

(a) Host to Device

(b) Device to Host

Figure 2.   Worst-case data transfer times of single streams.

Abstract—

Keywords-**Low Latency; Data Transfer; GPGPU**
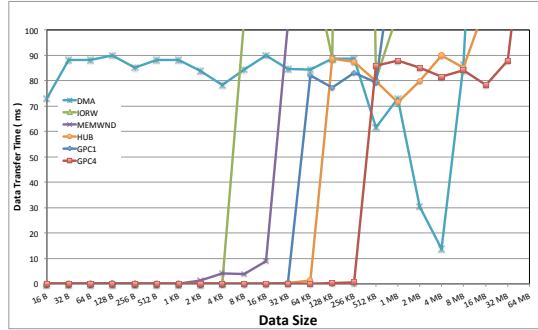
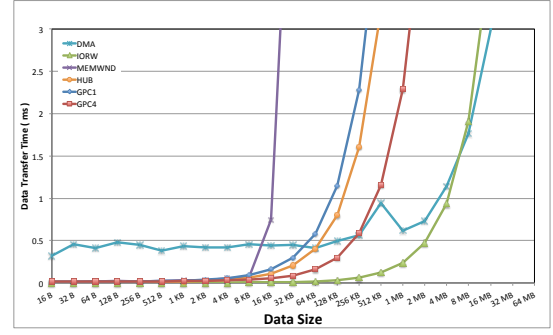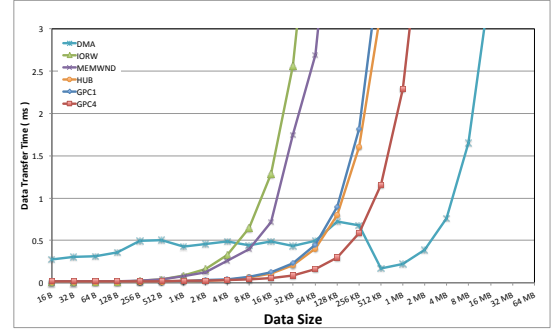## I. EMPIRICAL COMPARISON

(a) Host to Device

(b) Device to Host

Figure 3.   Average data transfer times of single streams under high CPU load.

(a) Host to Device

(b) Device to Host

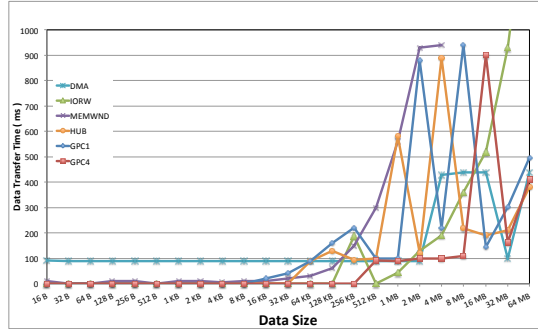Figure 4.   Worst-case data transfer times of single streams under high CPU load.
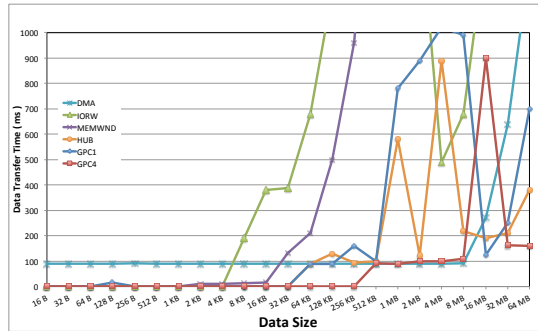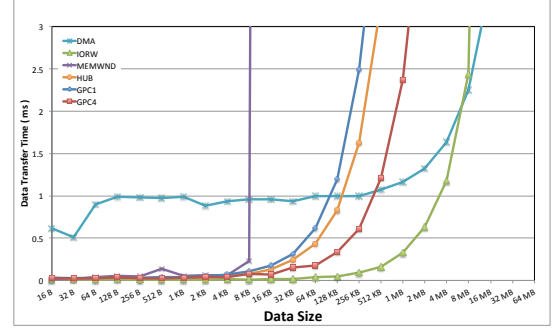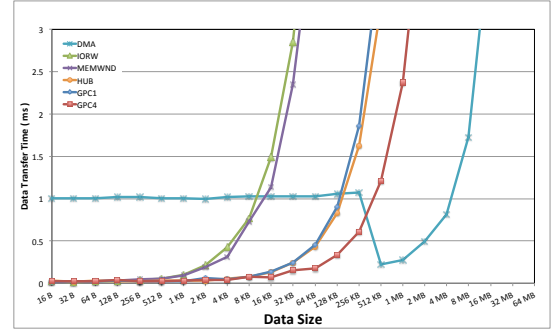

(a) Host to Device

(b) Device to Host

Figure 5.   Average data transfer times of single streams under high memory access load.
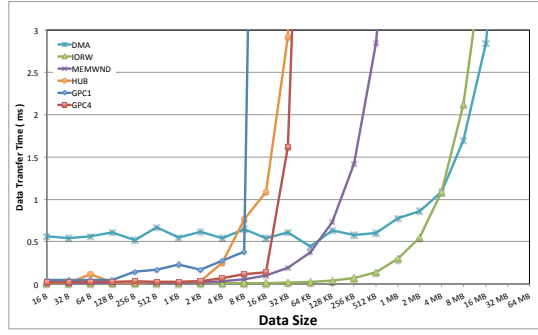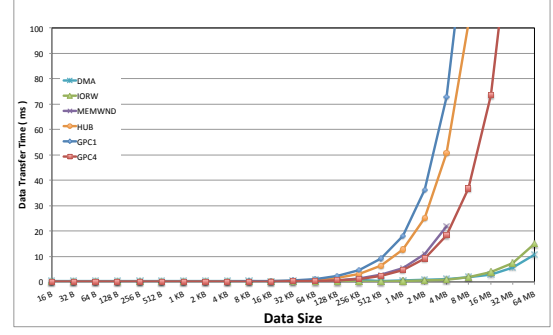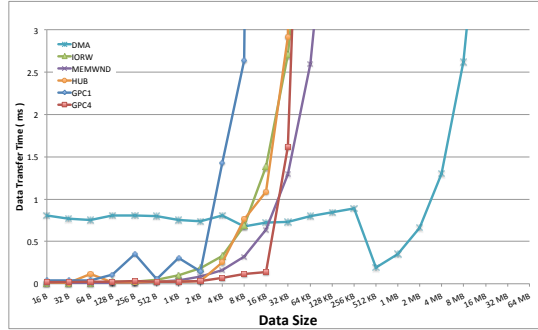

(a) Host to Device

(b) Device to Host

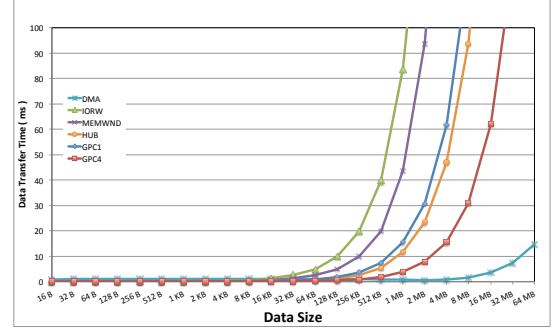Figure 6.   Worst-case data transfer times of single streams under high memory access load.
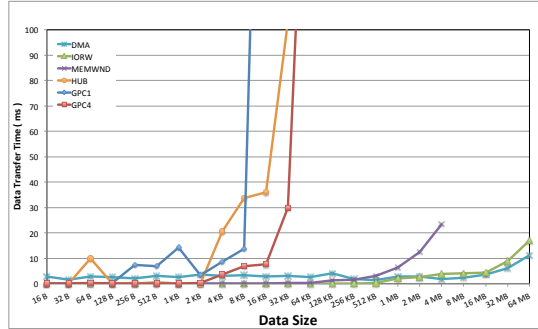
(a) Host to Device

(b) Device to Host

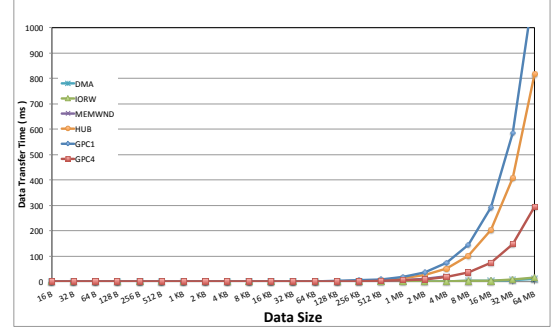Figure 7. Average data transfer times of single streams in the presence of hackbench.

(a) Host to Device

(b) Device to Host

Figure 8. Worst-case data transfer times of single streams in the presence of hackbench.
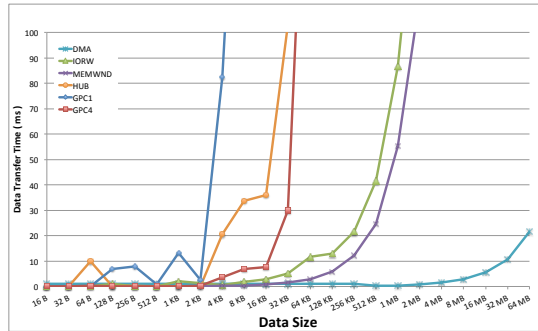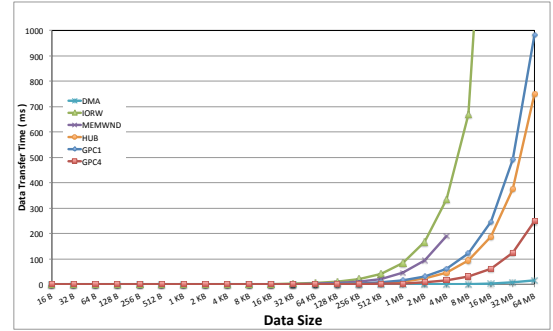

(a) Host to Device

(b) Device to Host

Figure 9. Average data transfer times of real-time single streams under high CPU load.

(a) Host to Device

(b) Device to Host

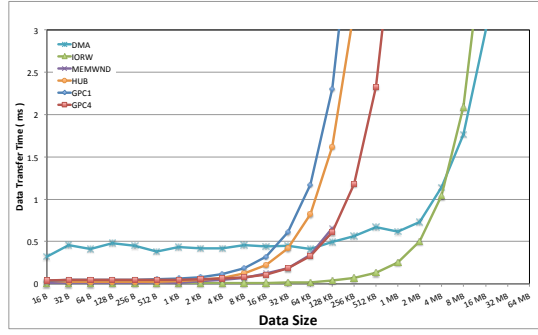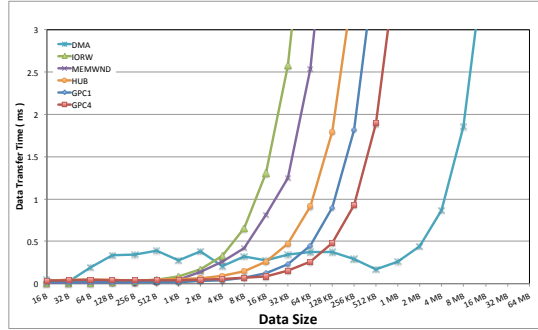Figure 10. Worst-case data transfer times of real-time single streams under high CPU load.
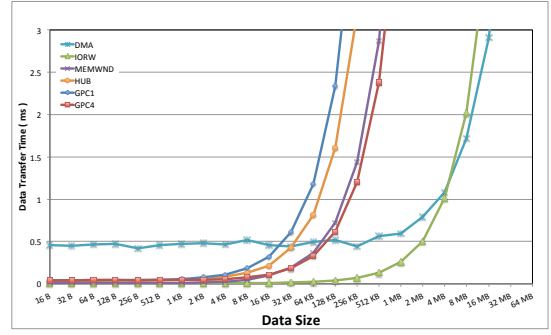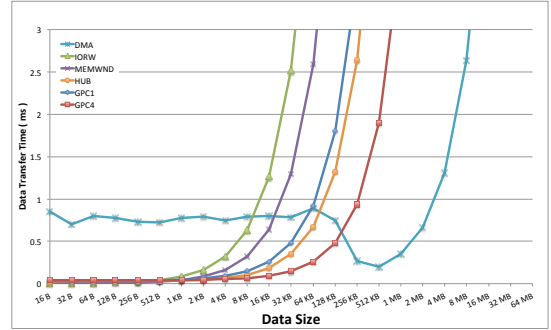
(a) Host to Device

(b) Device to Host

Figure 11. Average data transfer times of real-time single streams under high memory access load.

(a) Host to Device

(b) Device to Host

Figure 12. Worst-case data transfer times of real-time single streams under high memory access load.
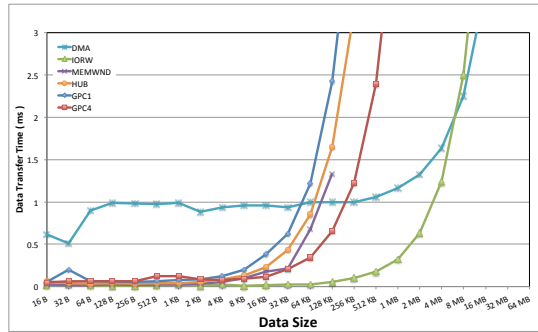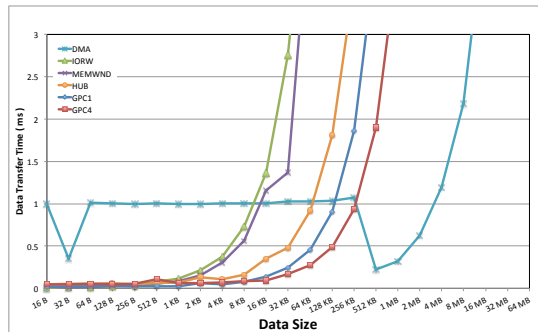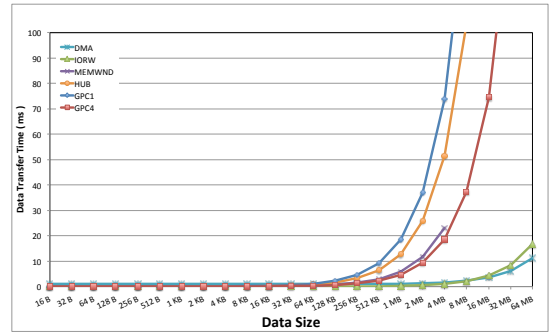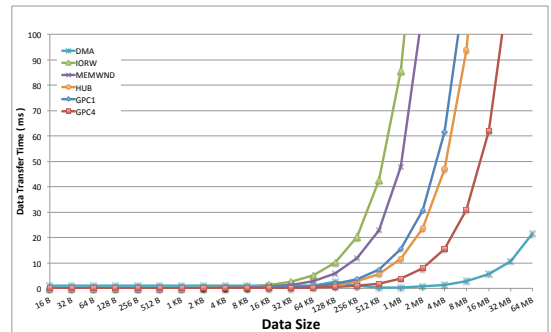

(a) Host to Device

(b) Device to Host

Figure 13. Average data transfer times of real-time single streams in the presence of hackbench.
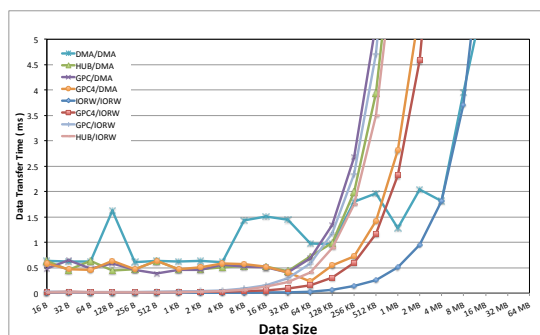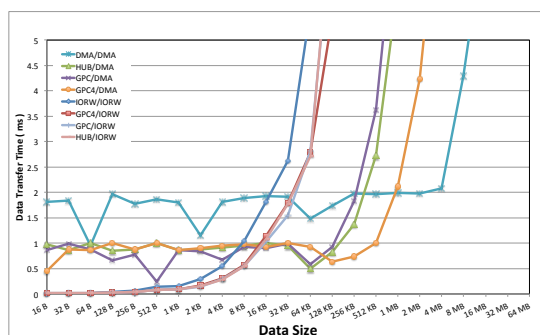

(a) Host to Device

(b) Device to Host

Figure 14. Worst-case data transfer times of real-time single streams in the presence of hackbench.

(a) Host to Device

(b) Device to Host

Figure 15.   Average data transfer times of concurrent two streams.