論 文 の 内 容 の 要 旨

論文題目　Low Latency On-Chip Networks through Compression and Multicasting
（圧縮とマルチキャストを用いた低遅延オンチップネットワーク）

氏　名　和 远

The inevitable advent of the multi-core era has driven an increasing demand for low latency on-chip interconnection networks (or NoCs). Being a critical part of the memory hierarchy for modern chip multi-processors (CMPs), these networks face stringent design constraints to provide fast communication. Modern NoC's first order concern is clearly its latency, so we present three low latency techniques in this thesis. One is through traffic compression while the other two are low latency router designs based on multicast-able crossbar switches.

Firstly, an adaptive traffic compression scheme is proposed, taking account of the vertical bandwidth limitation of 3D NoCs and traffic compressibility. It is found that the compressibility based adaptive compression is very useful against incompressible traffic while the location-based adaptive compression is more effective with more layers.

Secondly, an improvement is made on Prediction Router, which routes packet based on predicted destinations. This improvement makes use of multiple prediction algorithms (so-called the Predict-more Router) at the same time for one packet. It helps increasing the prediction accuracy by 15% and outperforms the best PR by 3.5% in speeding-up the system.

Thirdly, to further lower the latency and to get rid of predictions, a novel low latency router (McRouter) is proposed through multicasting a packet to all possible outputs. This design allows a single cycle transfer of flits when having enough bandwidth within the router. So it may behave like an always-hit prediction router. Evaluation shows that McRouter helps achieving system speed-ups of 1.28, 1.17 and 1.05 over the conventional router, the VSA router and the best prediction router, respectively.