New switch fabric architecture for bursty traffic

Qimin Yang, *Member, IEEE* and Keren Bergman, *Member, IEEE*

I. ARCHITECTURE

A new switch fabric architecture, the Data Vortex, is designed with bufferless operation and minimum routing logic aiming for large-scale, low-latency optical implementations [1]. As seen in Fig.1, a fabric of (A,H) can be arranged as a collection of routing nodes on multiple fiber cylinders. Each routing node is labeled uniquely by the coordinates (a, c, h), where $0 \le a < A$, $0 \le c < C$ and $0 \le h < H$. As each packet is self-routed in the fashion of binary-tree decoding when it propagates from the outer cylinder towards the inner cylinder, each cylinder progress fixes a specific bit within the binary header address. Within each clock cycle, every packet propagates one angle forward along either the solid or the dashed line. To avoid packet contention, the switch architecture employs a synchronous and distributed control mechanism to properly schedule the neighboring packet flow [2]. As a result, each node encounters at most one packet at a time and no optical buffering will be necessary within the Data Vortex switch fabric. This also greatly simplifies the routing procedure at each hop and facilitates the photonic implementation of the architecture [3].

II. PERFORMANCE

It is important that the switch architecture shows robust performance for bursty traffic, in which consecutive time slots tend to send packets for the same destination. These consecutive slots are considered as an active ON period which alternates with an OFF period. The length of each period has a distribution modeled by [4]:

$$T_{on} = \left\lfloor \frac{1}{U^{1/\alpha_{on}}} \right\rfloor \quad , \quad T_{off} = \left\lfloor \frac{1}{U^{1/\alpha_{off}}} \right\rfloor \quad (1)$$

where U is a random variable uniformly distributed over [0,1], and $\lfloor x \rfloor$ indicates the floor function. The parameters α_{on} and α_{off} specify the variability of the ON and OFF time periods, resulting traffic from rather smooth cases ($\alpha >> 2$) to highly bursty cases ($1 < \alpha < 2$). We set α_{off} to generate short OFF periods while vary α_{on} for different burstiness. The bursty system performance measures are studied under various operating conditions, which include different bursty parameters (α_{on}, α_{off}) and different asymmetric I/O modes (θ). System performance is monitored over time to investigate the bursty traffic effect.

Performance under a highly bursty traffic case (α_{on}=1.05, α_{off}=2.5) is studied. In Fig.2 (a) the latency distribution is shown for the I/O mode of θ=1/5 with A=5, H=128 (dashed line). It is compared to the random case (solid line) in which each time slot is statistically independent and the same traffic load of 0.85 is used. As the results show, there's almost no degradation caused by the traffic burstiness. In Fig.2 (b) and (c), the monitored successful injection probability and latency over the time slots illustrate how the switching fabric behaves under the bursty traffic. There is slight performance degradation around the 500th time slot and 4500th time slot because more ON bursts are overlapped during the period in this specific case. In general however, the performance variation is very small over the time due to the smoothing effect provided by the switching fabric topology. The asymmetric I/O mode is thus critical in providing enough smoothing effect for the bursty traffic. For comparison, an example of θ=1/3 with A=3, H=128 is studied under the same degree of traffic burstiness. In Fig.3 (a), a much wider latency distribution and longer tail results with θ=1/3 due to the traffic burstiness. Therefore, if an insufficient asymmetry is provided between the number of input and output ports, the switch performance can be degraded severely under bursty traffic. Similarly, the performance measures over the time slots illustrate the degradations as shown in Fig.3 (b) and (c). Both the successful injection probability and the latency fluctuate widely over time, especially during the period around the 500th and the 4500th slots. The degradations are directly proportional to the number of input ports, which overlap their long ON period and to the length of the overlapped period. Thus compared with the θ=1/5 case, the I/O asymmetry of θ=1/3 provides less routing redundancy and therefore less smoothing effect on the traffic congestion.

Given the same degree of asymmetric I/O mode, different bursty modes will result in different degrees of performance degradation. In addition, we find that the above performance comparisons for bursty traffic are quite independent of the switch size. This indicates that as long as a modest I/O asymmetry (θ=1/5) is provided, the switch scalability will not be affected by the bursty nature of input traffic.

0-7803-7378-2/02/$17.00©2002 IEEE

The results show that the architecture generally maintains robust throughput and latency performance under bursty traffic conditions as long as a modest asymmetric I/O mode (θ=1/5) is provided.

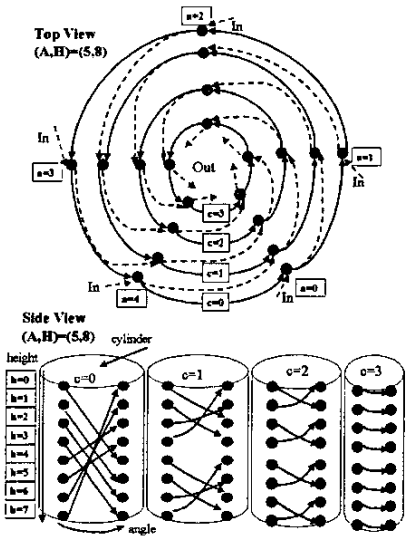

Fig.1 Data Vortex topology (A, H) = (5, 8) with routing tours seen from the top and the side.

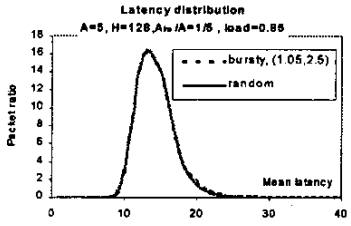

Fig.2 (a) Latency distribution for (α_{on} =1.05, α_{off}=2.5)

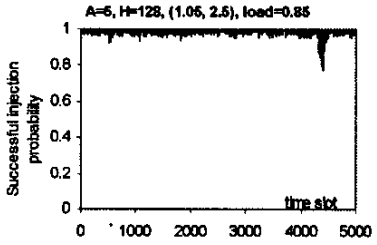

Fig.2 (b) Successful injection probability during the time slots for (α_{on}=1.05, α_{off}=2.5)

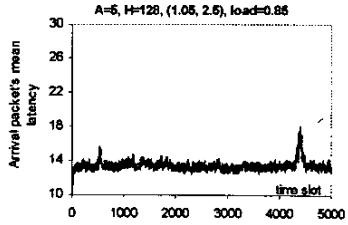

Fig.2 (c) Arrival packets' mean latency during the time slots for (α_{on}=1.05, α_{off}=2.5)

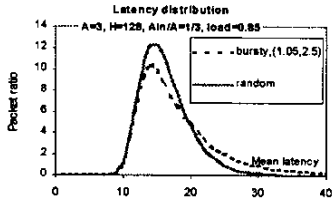

Fig.3 (a) Latency distribution for (α_{on}=1.05, α_{off}=2.5)

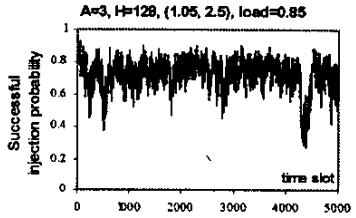

Fig.3 (b) Successful injection probability during the time slot for (α_{on}=1.05, α_{off}=2.5)

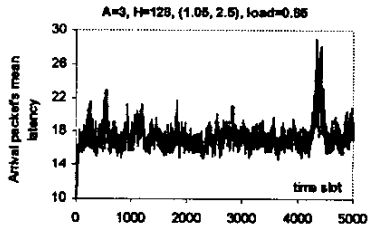

Fig.3 (c) Arrival packets' mean latency during the time slot for (α_{on}=1.05, α_{off}=2.5)

REFERENCES

[1] Coke Reed, "Multiple Level Minimum Logic Network", U.S Patent 5996020, Nov.30, 1999

[2] Q.Yang, K.Bergman, "Traffic Control and WDM Routing in the Data Vortex Packet Switch", *IEEE Photonic Technology Letters*, pp236-238, February 2002.

[3] Q.Yang, K.Bergman, G.D.Hughes and F.G.Johnson, "WDM Packet Routing for High Capacity Data Networks", *IEEE Journal of Lightwave Technology*, pp.1420-1426, October, 2001.

[4] A.Bononi, et al, "Large Power swings in Doped-Fiber Amplifiers with Highly variable Data", *IEEE Photon. Tech. Letters.*, vol. 11, pp.131, January, 1999.