Service Guaranteed Scheduling Model of the MPMS Packet Switching Fabric

Xiangjie Ma
Air Force Engineering University, Beijing, P.R. China
Email: maxiangjie100@163.com

Xiaozhong Li and Xinglong Fan
Air Force Engineering University, Beijing, P.R. China
Email: {lxz, fxl}@gmail.com

***Abstract*—In this paper, the service guaranteed scheduling model was proposed for the MPMS switching fabric. Firstly, the service distribution characteristic of each cell flow was analyzed in the MPMS switching fabric. Then, the service guaranteed distributed architecture was designed, and the scheduling model is proposed for demultiplexors, multiplexors, input switching units, middle switching units and output switching units in the MPMS fabric, respectively. Finally, we evaluated the proposed service guaranteed scheduling model (*SGS-MPMS*) under *ON-OFF* burst traffic. Simulation results show that *SGS-MPMS* can overcome the problem of delay divergence for traditional *iSLIP*, and *PPS* method, and provide delay guaranteed performance under *ON-OFF* burst traffic, and converge to the optimal delay curve with only four switching planes and four iterations.**

***Index Terms*—MPMS, switching fabric, service guaranteed, scheduling model**

I. INTRODUCTION

Currently emerging QoS sensitive services in Internet, such as IPTV[1], video on demand[2], e-science[3] and so on, are proposing new demands for service guarantee in network equipments. However, traditional switching and scheduling techniques[4,5] tend to maximize average switching performance, such as switcing throughput and resource utilization, which lead to extensive management and average distribution of network resources. These switching and scheduling service modes are sufficient to guarantee most service needs, but when network congestion occurs, QoS sensitive guarantee services will be affected or even become unavailable[6,7]. In order to achieve service guarantee in network, service resources must be effectively managed and efficiently allocated, and fair allocation of reserved service among packet flows must be realized, so as to achieve the purpose of guaranteed services in network [8].

In this paper, based on our previous work on the multiple-plane and multiple-stage (MPMS) packet switching fabric[9], towards supporting requirement of universal service in network, regarding both effectiveness and efficiency, we studied the cell flows and the service distribution in the MPMS packet switching fabric, and designed the MPMS distributed scheduling structure with service guarantee, and proposed the service guaranteed scheduling model for the MPMS demultiplexor, multiplexor, input switching unit, middle switching unit and output switching unit. Finally, based on the scheduling model of the MPMS fabric, we evaluated the proposed *SGS-MPMS* model under *ON-OFF* burst traffic, which proved its validity to realize service guaranteed scheduling in the MPMS fabric .

II. ABBREVIATIONS AND SYMBOLS

$f_{v_i^0 v_j^6}^{mpms}(i,j=1,2,\dots,nk)$: packet flows from the input vertex v_i^0 to the output cell vertex v_j^6 in MPMS fabric.

$R_{v_i^0 v_j^6}^{mpms}(i,j=1,2,\dots,nk)$: the reserved service of cell packet flows $f_{v_i^0 v_j^6}^{mpms}$ in MPMS fabric, in unit of cells/S.

$R(\vec{e})$: the cell service of directed edge $\vec{e}$ in MPMS fabric, in unit of cells/S.

$R(\overline{v_x v_y v_z})$: path cell service switching through v_x, v_y, v_z in MPMS fabric.

$FPL^*(\vec{e})$: the priority list of directed edge $\vec{e}=\overline{v_x v_y}$ FPL, with "*" equal to 1 or 2.

$NID^*(\vec{e})$: the number of vertices *NID* in the priority list $FPL^*(\vec{e})$ of the directed edge $\vec{e}=\overline{v_x v_y}$.

$NID^*(v_z,\vec{e})$: the number of the priority list $FPL^*(\vec{e})$ identity v_z of the directed edge $\vec{e}=\overline{v_x v_y}$ in MPMS fabric.

$P^*(\vec{e})$: the priority pointer of the service priority list $FPL^*(\vec{e})$ of the directed edge in MPMS fabric.

$VOQ(q,r,t,u)$: virtual output queue of vertices v_q^1 in IMU, and satisfy: $v_r^2 \in N^+(v_q^1)$, $v_t^4 \in N^-(v_{(p,u)}^5)$.

$VIQ(q,r,t,u)$: the virtual input queue of vertices v_u^5 in OMU, and satisfies: $v_r^2 \in N^+(v_q^1)$, $v_t^4 \in N^-(v_u^5)$.

Manuscript received October 17, 2012; revised November 18, 2012; accepted November 25, 2012.

This work was supported in part by the grants from Natural Science Foundation of China with No.61003252

doi:10.4304/jcp.8.7.1826-1830

III. SERVICE GUARANTEED AND DISTRIBUTED SCHEDULING ARCHITECTURE

In order to facilitate distributed implementation, the service guaranteed scheduling architecture in MPMS switching fabric is designed as shown in Figure 1. Schedulers are set to be responsible for the link service allocation and traffic scheduling in *DEM*'s input links, *ISU*/*MSU*/*OSU*'s input directed edges and output directed edges, and MUX's output links; at the same time, all the schedulers adopt the round-robin scheduling mode in its service priority list (*FPL*). In practice, the number of vertex identifiers in *FPL*(e_y) is calculated by the corresponding service proportion of *R*(v_x) in the flow service, and then the packet traffic information is transformed into the number of identifiers in *FPL*.

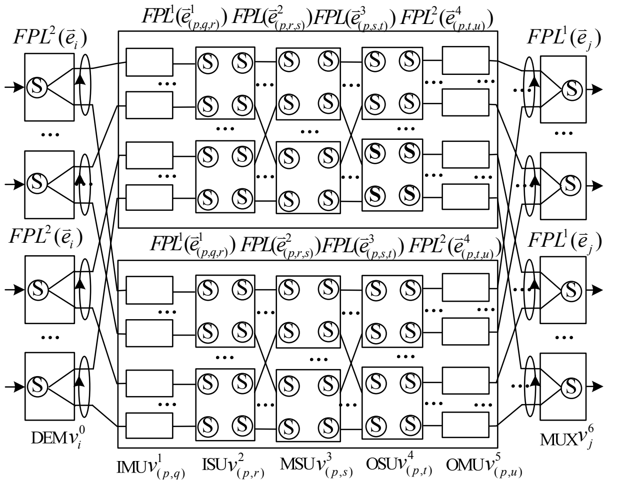

Figure 1. General topological graph of MPMS packet switching fabric.

IV. SERVICE GUARANTEED SCHEDULING MODEL

To take full advantage of the switching resources in MPMS fabric, cell switching and scheduling load is evenly shared among all vertexes, vertex buffers and internal links. The packet flow service and in the path and the directed edge depends on the following three factors: (1) the packet flow service of the starting vertex of the path and directed edge; (2) the packet flow service and of the terminating vertex in the path and the terminating vertex in the directed edge; (3) the number *P* of switching planes and the number *m* of the middle switching unit MSU.

Therefore, path $\overrightarrow{v_xv_yv_z}$ and directed edge $\vec{e}=\overrightarrow{v_xv_y}$ should have flow service $R(\overrightarrow{v_xv_yv_z})$ and $R(\vec{e}=\overrightarrow{v_xv_y})$ as follows:

$$R(\vec{e}=\overrightarrow{v_xv_y})=\begin{cases}\dfrac{1}{P}\sum\limits_{v_j^6\in M^+(v_y)}\sum\limits_{v_i^0\in M^-(v_x)}R^{mpms}_{v_i^0v_j^6}, v_x\notin V^3 \text{ and } v_y\notin V^3\\ \dfrac{1}{Pm}\sum\limits_{v_j^6\in M^+(v_y)}\sum\limits_{v_i^0\in M^-(v_x)}R^{mpms}_{v_i^0v_j^6}, v_x\in V^3 \text{ or} v_y\in V^3\end{cases}\tag{1}$$

$$R(\vec{e}=\overrightarrow{v_xv_yv_z})=\begin{cases}\dfrac{1}{P}\sum\limits_{v_j^6\in M^+(v_z)}\sum\limits_{v_i^0\in M^-(v_x)}R^{mpms}_{v_i^0v_j^6}, v_x \text{ and } v_y \text{ and } v_z\notin V^3\\ \dfrac{1}{Pm}\sum\limits_{v_j^6\in M^+(v_z)}\sum\limits_{v_i^0\in M^-(v_x)}R^{mpms}_{v_i^0v_j^6}, v_x \text{ or} v_y \text{ or} v_z\in V^3\end{cases}\tag{2}$$

By equation (1) and equation (2), cell flow service of all paths and directed edges in MPMS switching fabric can be obtained as shown in Table 1, which is used in the service guaranteed scheduling (*SGS-MPMS*) model.

TABLE I.
CELL SERVICE OF PATHS AND DIRECTED EDGES IN MPMS

Service of Edges	Value	Service of Paths	Value
$R(\vec{e}^0_{(p,i,q)})$	$\frac{1}{P}\sum_{j=1}^{kn}R^{mpms}_{v_i^0v_j^6}$	$R(\overrightarrow{v_i^0v^1_{(p,q)}v^2_{(p,r)}})$	$\frac{1}{P}\sum_{j=1}^{kn}R^{mpms}_{v_i^0v_j^6}$
$R(\vec{e}^1_{(p,q,r)})$	$\frac{1}{P}\sum_{j=1}^{kn}R^{mpms}_{v_i^0v_j^6}$	$R(\overrightarrow{v^1_{(p,q)}v^2_{(p,r)}v^3_{(p,s)}})$	$\frac{1}{Pm}\sum_{j=1}^{kn}R^{mpms}_{v_q^0v_j^6}$
$R(\vec{e}^2_{(p,r,s)})$	$\frac{1}{Pm}\sum_{i=(r-1)n+1}^{rn}\sum_{j=1}^{kn}R^{mpms}_{v_i^0v_j^6}$	$R(\overrightarrow{v^2_{(p,r)}v^3_{(p,s)}v^4_{(p,t)}})$	$\frac{1}{Pm}\sum_{i=(r-1)n+1}^{rn}\sum_{j=(t-1)n+1}^{tn}R^{mpms}_{v_i^0v_j^6}$
$R(\vec{e}^3_{(p,s,t)})$	$\frac{1}{Pm}\sum_{i=1}^{kn}\sum_{j=(t-1)n+1}^{tn}R^{mpms}_{v_i^0v_j^6}$	$R(\overrightarrow{v^3_{(p,s)}v^4_{(p,t)}v^5_{(p,u)}})$	$\frac{1}{Pm}\sum_{i=1}^{kn}R^{mpms}_{v_i^0v_u^6}$
$R(\vec{e}^4_{(p,t,u)})$	$\frac{1}{P}\sum_{i=1}^{kn}R^{mpms}_{v_i^0v_u^6}$	$R(\overrightarrow{v^4_{(p,t)}v^5_{(p,u)}v_j^6})$	$\frac{1}{P}\sum_{i=1}^{kn}R^{mpms}_{v_i^0v_j^6}$
$R(\vec{e}^5_{(p,u,j)})$	$\frac{1}{P}\sum_{i=1}^{kn}R^{mpms}_{v_i^0v_j^6}$	——	——

In the distributed structure of the MPMS fabric, if the total number of vertex in the priority list $FPL^*(\vec{e})$ is denoted by $NID^*(\vec{e})$, and the service of directed edges can be denoted by $R(\vec{e})$, then the number of identifiers $NID^*(v_z,\vec{e})$ of vertex v_z satisfies equation (3):

$$NID^*(v_z,\vec{e})=\frac{R(\overrightarrow{v_xv_yv_z})}{R(\vec{e})}NID^*(\vec{e})\tag{3}$$

A. Demultiplexor and Multiplexor SGS Scheduling Model

The service guaranteed scheduling model of demultiplexor and multiplexor is shown in Figure 2. If the service priority list of input link $\vec{e}_i$ of demultiplexor vertex v_i^0 is denoted by $FPL^2(\vec{e}_i)$ in MPMS switching fabric, the distribution method of identifier $NID^2(v^1_{(p,i)},\vec{e}_i)$ of vertex $v^1_{(p,i)}$ in the priority list in MPMS switching fabric is calculated as equation (4):

$$NID^2(v^1_{(p,q=i)},\vec{e}_i)=\frac{R(\vec{e}^0_{(p,i,q=i)})}{R(\vec{e}_i)}NID^2(\vec{e}_i)=\frac{NID^2(\vec{e}_i)}{P},\forall p\in[1,P]\tag{4}$$

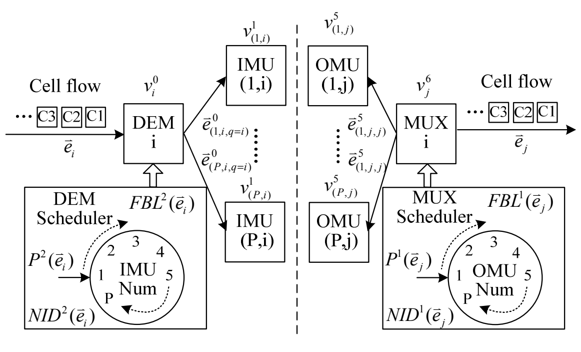

Figure 2. The service guaranteed and scheduling model for demultiplexor and multiplexor.

If the service priority list of output link $\vec{e}_j$ of multiplexor vertex v_j^6 is denoted by $FPL^1(\vec{e}_j)$ in MPMS switching fabric, The distribution method of identifier $NID^1(v^5_{(p,j)},\vec{e}_j)$ of vertex $v^5_{(p,j)}$ in the priority list is calculated as equation (5):

$$NID^1(v^5_{(p,u=j)},\vec{e}_j)=\frac{R(\vec{e}^{\,5}_{(p,u=j,j)})}{R(\vec{e}_j)}NID^1(\vec{e}_j)=\frac{NID^1(\vec{e}_j)}{P},\forall p\in[1,P] \tag{5}$$

B. SGS Scheduling Model of Input Switching Unit

There are *n* input directed edges $\vec{e}^{\,1}_{(p,q,r)}((r-1)n+1\le q\le rn)$ and *m* output directed edges $\vec{e}^{\,2}_{(p,r,s)}(1\le s\le m)$ in the input switching unit $ISU\,v^2_{(p,r)}$ of MPMS switching fabric. The *ISU* vertex *SGS* scheduling includes *SGS* input scheduling and *SGS* output scheduling, and scheduling method is shown in Figure 3.

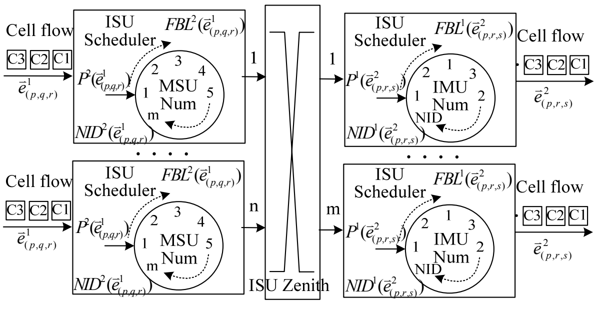

Figure 3. *SGS* Scheduling model with guaranteed service in the input switching unit of the MPMS switching fabric.

1) SGS Input scheduling of the input switching unit

The service priority list of the input directed edge $\vec{e}^{\,1}_{(p,q,r)}$ in ISU vertex is denoted by $FPL^2(\vec{e}^{\,1}_{(p,q,r)})$. If the total identification number is expressed by $NID^2(\vec{e}^{\,1}_{(p,q,r)})$, the identification number $NID^2(v^3_{(p,s)},\vec{e}^{\,1}_{(p,q,r)})$ in $FPL^2(\vec{e}^{\,1}_{(p,q,r)})$ of *MSU* vertex $v^3_{(p,s)}$ is calculated as follows:

$$NID^2(v^3_{(p,s)},\vec{e}^{\,1}_{(p,q,r)})=\frac{NID^2(\vec{e}^{\,1}_{(p,q,r)})}{m},\forall s\in[1,m] \tag{6}$$

2) SGS Output scheduling of the input switching unit

The service priority list of the output directed edge $\vec{e}^{\,2}_{(p,r,s)}$ in *ISU* vertex is denoted by $FPL^1(\vec{e}^{\,2}_{(p,r,s)})$. If the total identification number is expressed by $NID^1(\vec{e}^{\,2}_{(p,r,s)})$, the identification number $NID^1(v^1_{(p,q)},\vec{e}^{\,2}_{(p,r,s)})$ in $FPL^1(\vec{e}^{\,2}_{(p,r,s)})$ of *MSU* vertex $v^1_{(p,q)}$ is calculated as follows:

$$NID^1(v^1_{(p,q)},\vec{e}^{\,2}_{(p,r,s)})=\frac{\sum_{j=1}^{kn}R^{mpms}_{v^0_q v^6_j}NID^1(\vec{e}^{\,2}_{(p,r,s)})}{\sum_{i=(r-1)n+1}^{rn}\sum_{j=1}^{kn}R^{mpms}_{v^0_i v^6_j}},\forall q\in[(r-1)n+1,rn] \tag{7}$$

As $R(\overrightarrow{v^1_{(p,q)}v^2_{(p,r)}v^3_{(p,s)}})$ is not always an integer, the identification number $NID^1(\vec{e}^{\,2}_{(p,r,s)})$ is calculated as follows:

$$NID^1(\vec{e}^{\,2}_{(p,r,s)})=\sum_{q=(r-1)n+1}^{rn}\left[\!\left[\frac{1}{Pm}\sum_{j=1}^{kn}R^{mpms}_{v^0_q v^6_j}\right]\!\right]\ or\ \sum_{q=(r-1)n+1}^{rn}LCM_n\bullet\frac{1}{Pm}\sum_{j=1}^{kn}R^{mpms}_{v^0_q v^6_j} \tag{8}$$

Where $[\![*]\!]$ is the the approximate integer of parameter $*$, LCM_n is the least common multiple of all n identification count of $\frac{1}{Pm}\sum_{j=1}^{kn}R^{mpms}_{v^0_q v^6_j}(q=(r-1)n+1,...,rn)$.

C. SGS Scheduling model of middle switching unit

There are k input directed edges $\vec{e}^{\,2}_{(p,r,s)}(1\le r\le k)$ and *k* output directed edges $\vec{e}^{\,3}_{(p,s,t)}(1\le t\le k)$ in the middle switching unit MSU $v^3_{(p,s)}$ of MPMS switching fabric. The *MSU SGS* scheduling includes input scheduling and output scheduling, and the *SGS* scheduling method of middle switching unit is shown in Figure 4.

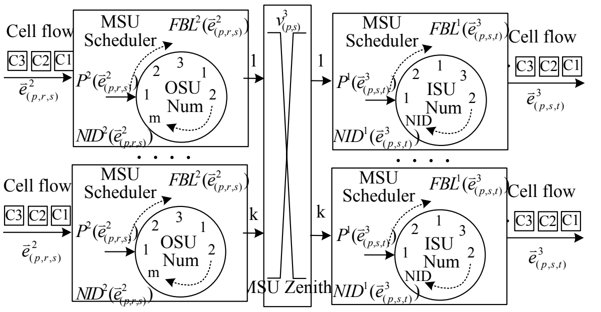

Figure 4. *SGS* Scheduling model with guaranteed service in the middle switching unit of the MPMS switching fabric.

1) SGS Input Scheduling of the middle switching unit

The service priority list of the input directed edge $\forall\vec{e}^{\,2}_{(p,r,s)}$ in *MSU* vertex is denoted by $FPL^2(\vec{e}^{\,2}_{(p,r,s)})$. If the total identification number is expressed by $NID^2(\vec{e}^{\,2}_{(p,r,s)})$, the identification number $NID^2(v^4_{(p,t)},\vec{e}^{\,2}_{(p,r,s)})$ in $FPL^2(\vec{e}^{\,2}_{(p,r,s)})$ of OSU vertex $v^4_{(p,t)}$ is calculated as follows:

$$NID^2(v^4_{(p,t)},\vec{e}^{\,2}_{(p,r,s)})=\frac{\sum_{i=(r-1)n+1}^{rn}\sum_{j=(t-1)n+1}^{tn}R^{mpms}_{v^0_i v^6_j}NID^2(\vec{e}^{\,2}_{(p,r,s)})}{\sum_{i=(r-1)n+1}^{rn}\sum_{j=1}^{kn}R^{mpms}_{v^0_i v^6_j}},\forall t\in[1,k] \tag{9}$$

2) SGS Output scheduling of middle switching unit

The service priority list of the output directed edge $\vec{e}^{\,3}_{(p,s,t)}$ in *MSU* vertex is denoted by $FPL^1(\vec{e}^{\,3}_{(p,s,t)})$. If the total identification number is expressed by $NID^1(\vec{e}^{\,3}_{(p,s,t)})$, the identification number $NID^1(v^2_{(p,r)},\vec{e}^{\,3}_{(p,s,t)})$ in $FPL^1(\vec{e}^{\,3}_{(p,s,t)})$ of *ISU* vertex $v^2_{(p,r)}$ is calculated as follows:

$$NID^1(v^2_{(p,r)},\vec{e}^{\,3}_{(p,s,t)})=\frac{\sum_{i=(r-1)n+1}^{rn}\sum_{j=(t-1)n+1}^{tn}R^{mpms}_{v^0_i v^6_j}\bullet NID^1(\vec{e}^{\,3}_{(p,s,t)})}{\sum_{i=1}^{kn}\sum_{j=(t-1)n+1}^{tn}R^{mpms}_{v^0_i v^6_j}},\forall r\in[1,k] \tag{10}$$

D. SGS Scheduling Model of the output switching unit

There are *m* input directed edges $\vec{e}^{\,3}_{(p,s,t)}(1\le s\le m)$ and *n* output directed edges $\vec{e}^{\,4}_{(p,t,u)}(1\le u\le k)$ in the output switching unit $OSU\,v^2_{(p,r)}$ of MPMS switching fabric. *OSU SGS* scheduling includes input scheduling and output scheduling, and the *SGS* scheduling method of the output switching unit is shown in Figure 5.

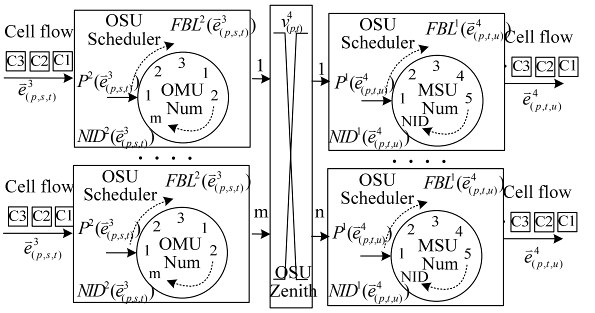

Figure 5. *SGS* Scheduling model with guaranteed service in the output switching unit of the MPMS switching fabric.

1) *SGS Input schedulingof the output switching unit*

The service priority list of the input directed edge $\vec{e}^{3}_{(p,s,t)}$ in *OSU* vertex is denoted by $FPL^2(\vec{e}^{3}_{(p,s,t)})$. If the total identification number is expressed by $NID^2(\vec{e}^{3}_{(p,s,t)})$, the identification number $NID^2(v^5_{(p,u)},\vec{e}^{3}_{(p,s,t)})$ in $FPL^2(\vec{e}^{3}_{(p,s,t)})$ of *MSU* vertex $v^5_{(p,u)}$ is calculated as follows:

$$NID^2(v^5_{(p,u)},\vec{e}^{3}_{(p,s,t)})=\frac{\sum_{i=1}^{kn} R^{mpms}_{v_i^0 v_u^6} NID^2(\vec{e}^{3}_{(p,s,t)})}{\sum_{i=1}^{kn}\sum_{j=(t-1)n+1}^{tn} R^{mpms}_{v_i^0 v_j^6}},\forall u\in[(t-1)n+1,tn] \qquad (11)$$

2) *SGS Output scheduling of the output switching unit*

The service priority list of the output directed edge $\vec{e}^{4}_{(p,t,u)}$ in *OSU* vertex is denoted by $FPL^1(\vec{e}^{4}_{(p,t,u)})$. If the total identification number is expressed by $NID^1(\vec{e}^{4}_{(p,t,u)})$, the identification number $NID^1(v^3_{(p,s)},\vec{e}^{4}_{(p,t,u)})$ in $FPL^1(\vec{e}^{4}_{(p,t,u)})$ of *MSU* vertex $v^3_{(p,s)}$ is calculated as follows:

$$NID^1(v^3_{(p,s)},\vec{e}^{4}_{(p,t,u)})=\frac{NID^1(\vec{e}^{4}_{(p,t,u)})}{m},\forall s\in[1,m] \qquad (12)$$

Therefore, *MSU* vertex $v^3_{(p,s)}$ has identical identification number in $FPL^1(\vec{e}^{4}_{(p,t,u)})$, and all *MSU* vertex $v^3_{(p,s)}$ have the same priority. If the total identification number is m in $FPL^1(\vec{e}^{4}_{(p,t,u)})$, that is $NID^1(v^3_{(p,s)},\vec{e}^{4}_{(p,t,u)})=\frac{m}{m}=1$, and therefore each *MSU* vertex $v^3_{(p,s)}$ has a unique identification.

V. SIMULATION ANALYSIS

We studied the performance of the proposed SGS algorithm based on under *ON-OFF* burst traffic. The status length of *ON-OFF* burst traffic is subject to the geometric distribution with parameter of (p,q), and the burst length and traffic load are denoted by L_{ON} and λ, respectively, which satisfy equation (13):

$$\lambda=\left(L_{ON}\right)^{-1},\lambda=\frac{\lambda\times p}{1-\lambda+\lambda\times p} \qquad (13)$$

We suppose the fixed length of all traffic cells to be 512 bytes, and the parameter of burst length L_{ON} to be 24, and the traffic intensity to satisfy $5\%\le\lambda\le100\%$. The segmentation and re-assembly delays were not considered in our simulation expriments. In our simulations, the *iSLIP* algorithm adopts 16×16 switching unit, and the number of iterations is equal to three. The *PPS* switching fabric adopts three planes with 16×16 switching units and parameters n=8, k=8, m=8.

The *iSLIP* and *PPS* algorithms are typical single and parallel switching fabric, separately. Delay performance comparison of *iSLIP*, *PPS* and *SGS-MPMS* under *ON-OFF* burst traffic is shown in Figure 6. All three methods can guarantee cell delay performance with low offered load under the burst traffic, but *iSLIP* with $\lambda\ge0.8$ and *PPS* with $\lambda\ge0.8$ fall into the delay divergence area, and no longer guarantee any delay performance. However, SGS-MPMS delay is always locked in convergence area even under the maximum traffic intensity (λ=1.0).

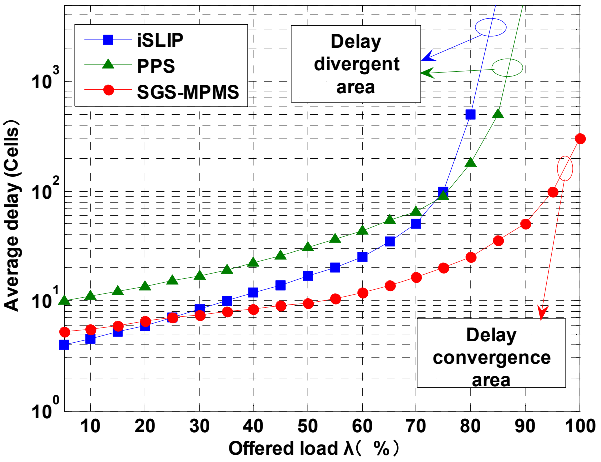

Figure 6. Delay performance comparison of *iSLIP*, *PPS* and *SGS-MPMS* under *ON-OFF* burst traffic.

The delay convergence property of *SGS-MPMS* is shown in Figure 7 under different number of switching planes. For *ON-OFF* burst traffic model, the *SGS-MPMS* method is always kept in delay convergence area with any switching plane count, and the delay performance of *SGS-MPMS* obtains the optimal delay convergence curve when the switching plane count is equal to four.

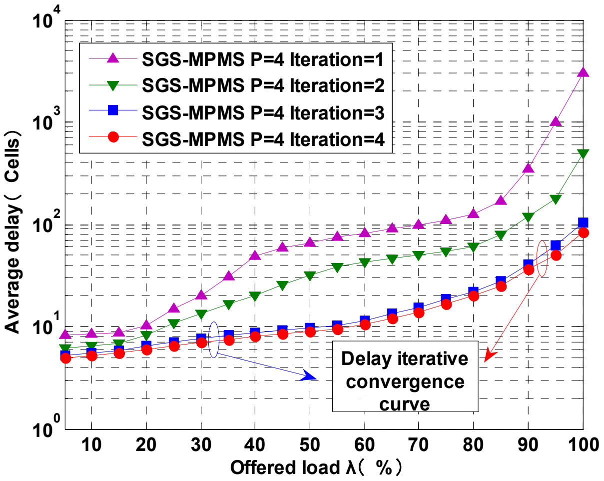

Figure 7. The switching plane convergence property of *SGS-MPMS* under *ON-OFF* burst traffic.

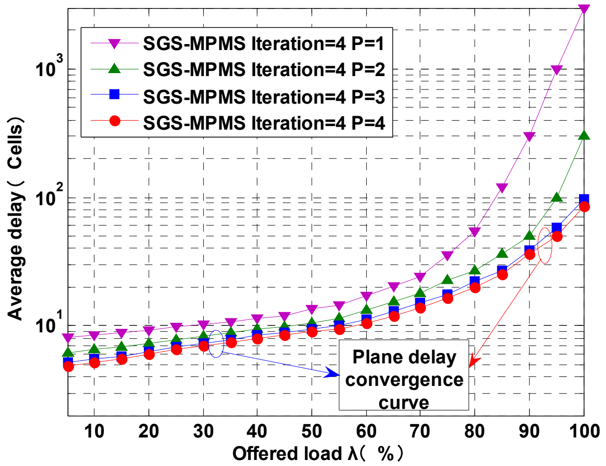

Figure 8. The switching delay convergence property of *SGS-MPMS* under *ON-OFF* burst traffic.

In order to testify the iterative delay property of the proprosed service guaranteed scheduling algorithm, we conducted the simulation experiments shown in Figure 8. For any value of iterations of the *SGS-MPMS* method, it is always within the delay convergence area under *ON-OFF* burst traffic, and converge to the optimal delay curve delay property when *P* is eual to four.

VI. CONCLUSIONS

The multiple-plane and multiple-stage switching fabric (MPMS) is a novel packet switching technology of Internet and next generation information network. In this paper, to support more and more universal services of the information network, regarding both effectiveness and efficiency, we studied service distribution of packet flows in the MPMS switching fabric, and designed distributed service guarantee scheduling structure of the MPMS switching fabric, The scheduling model of MPMS demultiplexor, multiplexor, input switching unit, middle switching unit and output switching unit is proposed. Simulation results show that the proposed *SGS-MPMS* model can overcome the delay divergence problem traditional *iSLIP*, and *PPS* method, and provide delay guarantee performance under *ON-OFF* bursty traffic, and converge to the optimal delay curve with only four switching planes and four iterations.

ACKNOWLEDGMENT

This work was supported in part by the grants from Natural Science Foundation of China with No.61003252. We thank several members of National Digital Switching Fabric Engineering & Technological R&D Cinput (NDSC) for their technical suggestions, including Julong Lan, Yuxiang Hu and Junpeng Mao, and the anonymous reviewers for their constructive comments and suggestions.

REFERENCES

[1] Cherry S, "The Battle for Broadband (Internet protocol television)," *IEEE Spectrum*, vol.42(1), pp.24-29, 2006.

[2] Liu Hengzhu, Chen Xucan and Chen Fuie, "Study of algorithm about storage allocation of video stream in video server," *Chinese Journal of Computers*, vol.21(4), pp.289-295, 1998.

[3] H. B. Newman, M. H. Ellisman and J. A. Orcutt, "Data-intensive Escience Frontier Research," *Commun. of the ACM*, vol.46(11), pp.68-77, 2005.

[4] N. McKeown, "iSLIP: a scheduling algorithm for input-queued switches," *IEEE/ACM Transactions on Networking*, vol. 7(2), pp.188-201, Apr. 1999.

[5] I. Radusinovic and M. Pejanovic, "Impact of Scheduling Algorithms on Performances of Buffered Crossbar Switch Fabrics," *IEEE International Conference on Communications*, pp.2416-2420, April 2002.

[6] S. Jiang, G. Hu, S.Y. Liew, and H. J. Chao, "Scheduling algorithms for shared fiber-delay-line optical packet switches, Part II: The 3-stage Clos-Network case," *IEEE Journal of Lightwave Technology*, vol.23(4), pp.1601-1609, Apr. 2005.

[7] Nikos Chrysos, Manolis Katevenisz, "Scheduling in Non-Blocking Buffered Three-Stage Switching Fabrics," *IEEE Globecom*, pp.6-13, Francisco, 2006.

[8] F. Wang and M. Hamdi, "Scalable Central-stage Buffered Clos Network Packet Switches with QoS," *IEEE HPSR Workshop*, pp.455-468, Poland, 2006.

[9] Ma Xiangjie Li Xiuqin Lan Julong Zhang Baisheng, "Graphic Model and Performance Analysis of a Novel Scalable Multiple-Plane and Multiple-Stage Packet Switching Fabric," *Journal of Electronics & Information Technology (China)*, 31(5), pp.1026~1030, May, 2009.

Xiangjie Ma was born in Hebei province of China on Oct. 20, 1977. He gained his bachelor degree in computer science from Xi'an Telecommunication Engineering College, Shanxi province, China, in 2001. He gained his master degree in computer science from Xi'an Telecommunication Engineering College, Shanxi province, China, in 2004. He gained his doctor degree in computer science from Zhengzhou Information Engineering University, Henan province, China, in 2008. His main research area includes packet switching fabric, packet buffering mode and distributed scheduling mechanism.

Xiaozhong Li was born in Shanxi province of China . 1962. He gained his bachelor degree in computer science from Zhengzhou Information Engineering University, Henan province, China, in 1986. His major field of study includes computer network and. Computer Security.

Xinglong Fan was born in Hebei province of China in 1964. He gained his bachelor degree in computer science from Zhengzhou Information Engineering University, Henan province, China, in 1987. His major field of study includes Computer Networks and Data Communication.