Quality-Of-Service Class Specific Traffic Matrices in IP/MPLS Networks

Stefan Schnitter
Deutsche Telekom, T-Systems

D-64295 Darmstadt
+49 6151 9378521

Franz Hartleb
Deutsche Telekom, T-Systems

D-64295 Darmstadt
+49 6151 9372726

Martin Horneffer
Deutsche Telekom, T-Com
Hammer Str. 216-226
D-48165 Münster
+49 251 7985208

stefan.schnitter@t-systems.com franz.hartleb@t-systems.com

ABSTRACT
In this paper we consider the problem of determining traffic matrices for end-to-end demands in an IP/MPLS network that supports multiple quality of service (QoS) classes. More precisely, we want to determine the set of traffic matrices T_i for each QoS class i separately. T_i contains average bandwidth levels for QoS class i for every pair of routers within the network. We propose a new method for obtaining QoS class specific traffic matrices that combines estimation and measurement methods: We take advantage of the fact that the total traffic matrix can be measured precisely in MPLS networks using either the LDP or RSVP-TE protocol. These measurements can then be used in a mathematical model to improve estimation methods – known as network tomography – that estimate QoS class specific traffic matrices from QoS class specific link utilizations. In addition to the mathematical model, we present results of the proposed method from its application in Deutsche Telekom's global IP/MPLS backbone network and we show that the estimation accuracy (mean relative error) is improved by a factor of 2.5 compared to results from network tomogravity. We investigate the structure of the estimated traffic matrices for the different QoS classes and motivate in this paper why QoS class specific traffic matrices will be essential for efficient network planning and network engineering in the future.

Categories and Subject Descriptors
C.4 [**Performance of Systems**]: Measurement techniques, Modeling techniques. C.2.3 [**Computer-Communication Networks**]: Network Operations.

General Terms: Algorithms, Measurement.

Keywords: Traffic Matrices, MPLS, LDP, QoS.

1. INTRODUCTION
Traffic matrices – or origin-destination (o-d) matrices – contain end

to end traffic demands between each pair of nodes in a given IP network. In this paper we are interested in 15 minutes average traffic demands between routers. For an ISP, there are multiple reasons why the availability of good quality traffic matrices is essential: Important tasks for an ISP that require traffic matrices include network planning and traffic engineering. They are needed to perform simulations of failure scenarios and network extension scenarios as well as for (IGP) routing optimization. For some tasks – especially in network planning – traffic matrices are needed on the level of PoPs but they can generally be obtained by aggregating router level traffic matrices.

Already the task of traffic matrix generation is difficult for the total traffic within a network and may require estimation methods or complex measurement infrastructures. But as true multi-service networks become reality and traffic classes with different service requirements exist, the need for traffic matrices per class-of-service (CoS) is getting stronger: In order to plan and operate a multi-service IP/MPLS network economically, an ISP needs to incorporate QoS class specific traffic matrices into its planning and engineering tasks. Possible scenarios include:

- Different service level agreements for the QoS classes require a different bandwidth dimensioning. For each QoS class the actual traffic must never exceed the bandwidth guaranteed to that class by the scheduler. Only the best effort class can be allowed to use up to the physical link bandwidth without violating given service level agreements.
- The use of links with low delay (e.g. direct links from Europe to Asia – not via America) only for traffic from certain QoS classes.
- Failure simulations in order to validate certain QoS concepts (e.g. to make sure that non-best-effort traffic remains below a certain link utilization threshold)

While applications with high QoS requirements such as IPTV or VoIP are growing to a significant contribution to the overall traffic demand, a service provider needs to incorporate QoS class specific traffic data into its planning and engineering processes.

In this paper we consider only unicast traffic matrices. Multicast traffic will generally be divided into different QoS classes, too, and has to be measured separately. For distribution platforms with fixed traffic sources and sinks (e.g. IPTV within the backbone network) this may be possible from the knowledge of the source traffic and the multicast tree.

For (unicast) traffic matrices, a variety of methods have already been investigated, both for traffic matrix estimation and for traffic matrix measurement. One method that is widely deployed for traffic matrix measurement based on IP flow level measurements is Cisco's Netflow [1]. But the implementation is rather complex and issues that arise in practice include:

- Netflow availability and performance depend on the line card types in use – in a service provider's network there is usually a large number of different hardware types in use. Thus only a partial measurement might be possible.
- IP flow measurements generally use packet sampling, so there is a tradeoff to make between the measurement accuracy (low sampling rates) and performance (high sampling rates). Some router types are not even able to provide anything than very high sampling rates.
- Large effort for aggregation of the flows that are exported by the routers.

Well-established estimations methods for traffic matrices are *gravity estimation* [10], where a traffic demand from s to t is set proportional to the total outgoing traffic of s and the total incoming traffic of t or *network tomography*, where end-to-end demands are estimated from link utilizations – see [12]. The combination of those two estimation methods is commonly referred to as *network tomogravity* – see [13] for a survey and comparison of different estimation methods. It is obvious that those estimation methods can also be used to estimate QoS class specific traffic matrices if the necessary input data (mostly link utilizations) are available per QoS class.

In networks that use multi-protocol label switching - MPLS [2] – to forward packets there are additional methods to measure the total traffic matrix. If RSVP-TE [3] is used, a full mesh of tunnels can be deployed and counters for those tunnels exist to measure the traffic matrix. If LDP [4] is used to distribute the label information in a network, LDP statistics of the routers can be used to compute traffic matrices on a router level – see [5], [6]. The LDP method results in a very high measurement accuracy (for example, there is no sampling involved) while the measurement complexity is very low: The measurement is based on aggregated *forwarding equivalence classes* (FEC) that are introduced in MPLS/LDP and is not based on the IP flow level. However, the LDP method can only be used for the network's total traffic matrix and not per QoS class. For Deutsche Telekom's IP/MPLS backbone network the LDP method is currently used to compute total traffic matrices to support IGP metric optimization (see [9] for a theoretical survey or [8] for a discussion on the practical implementation) and network planning.

Dependent on the existing measurement infrastructure, the network topology or the deployed protocols, it may be easier to obtain traffic matrices for the total traffic within the network than QoS class specific traffic matrices. This is why we propose a model for QoS class specific traffic matrices that combines estimation methods with total traffic matrix measurements.

This paper is organized as follows: Section 2 describes the mathematical model used to estimate QoS class specific traffic matrices and section 3 gives numerical results from the application of this model to a part of Deutsche Telekom's global IP/MPLS backbone network.

2. QOS TRAFFIC MATRIX MODEL

2.1 Notations for Tomogravity Model

If our network has *n* nodes and *m* links we denote by x the vector of link utilization and t the vector representation of the traffic matrix,

i.e. $t_{i \cdot j}$ (i,j=1,…,n) contains the traffic from node *i* to node j. From the network topology and the IGP metrics in use, we can construct the network's routing matrix *A*. The entry $a_{k,i \cdot j} \in [0,1]$ (i,j=1,…,n; k=1,…,m) of *A* contains the part of the demand $t_{i \cdot j}$ on link *k*. The routing matrix *A* results from a shortest-path calculation with respect to the given IGP metrics – if the network makes no use of equal cost path splitting (ECMP) $a_{k,i \cdot j} \in \{0,1\}$ holds. The relation between traffic matrix, link utilizations and routing matrix is then described by the following equation:

(R) $$A\,t = x.$$

The tomography estimation of the traffic matrix t^* can be constructed by solving the following system of equations for a given routing matrix *A*, a given vector of measured link utilizations x and a given initial estimate t_E for the traffic matrix:

(TG)
$$\| At^* - x \| \to \min$$
$$s.t. \| t^* - t_E \| \to \min$$

The solution of (TG) is given by

$$t^* = t_E + \widetilde{A} \cdot (x - A \cdot t_E)$$

where

$$\widetilde{A} = V \cdot S^{-1} \cdot U^T$$

denotes the Moore-Penrose-Inverse (also pseudo inverse) of the routing matrix *A* and is constructed with a singular value decomposition (SVD) of *A*:

$$A = U \cdot S \cdot V^T$$

We use the SVD routines contained in LAPACK [14] for this purpose – iterative methods for large scale SVD computations are also available – see [15].

If a gravity estimation is used as an initial estimate t_E, (TG) is referred to as *tomogravity method*. The solution of (TG) can be interpreted as a (orthogonal) projection of the initial solution t_E onto the subspace of all traffic matrices that satisfy the link utilization restrictions (R), i.e. we calculate that admissible traffic matrix that is closest to t_E. For tomogravity estimation methods a gravity estimate is chosen for t_E – one possibility to improve the gravity estimate is the use of a generalized gravity model as given in [12] which takes into account the knowledge of where peering traffic enters the network. The assumption that there is no transit peering traffic in the network then restricts the number of possible traffic demand combinations and thereby improves the estimate.

One method to improve the numerical stability of the solution of the system (TG) is to remove parallel links from the network topology. Service providers often use parallel links in their networks to increase capacity before they move to a technology with higher capacity (e.g. two or three 2.5Gbit/s links before installing on 10Gbit/s link). Those parallel links have the same IGP metric so that they are utilized equally when using ECMP. In practice, the traffic is not shared entirely equal depending on which hashing algorithm the routers use for load sharing – that means that there are differences in the corresponding components of the utilization vector x. On the other hand the routing matrix *A* introduces theoretical load sharing properties and equation (R) cannot be fulfilled exactly. We therefore

combine parallel links to one link with the sum of the parallel links capacities and their average utilization.

2.2 Integrated Model with QoS class traffic

In this section we extend (TG) to a system that can be used to estimate traffic matrices per QoS class. Unlike in the previous section we now assume that the full traffic matrix t^* is given. In Deutsche Telekom's IP/MPLS backbone network we use the LDP method to obtain t^* - other methods as discussed in the introduction (RSVP-TE tunnel counters or Netflow measurements) could also be used.

If we intend to apply (TG) for different QoS classes, QoS class specific link utilizations are a prerequisite. But there may not be QoS class specific link utilizations available for all links in the network: For example in case of Cisco routers QoS-class specific link utilizations require the use of the *Modular QoS CLI* (MQC) whose availability depends on the IOS software release and the hardware (line card types) in use. For a QoS-class traffic matrix estimation method that can be implemented in practice it is reasonable to assume

- QoS-class link utilizations are available for a subset of all links and
- End-to-End loads are available for the sum of all QoS classes

If there are q QoS classes in the network and we denote by $t^{(l)}$ and $\hat{x}^{(l)}$ the traffic matrix and the link utilization vector for the QoS class l ($l =1,…,q$), and by $\hat{A}$ the routing matrix reduced to subset of links whose QoS-class specific link utilizations are available, the following set of equation holds analog to (R):

$$
\underbrace{\begin{pmatrix} \hat{A} & & 0 \\ & O & \\ 0 & & \hat{A} \\ I & A & I \end{pmatrix}}_{A_Q} \cdot \underbrace{\begin{pmatrix} t^{(1)} \\ t^{(2)} \\ M \\ t^{(q)} \end{pmatrix}}_{t_Q} = \underbrace{\begin{pmatrix} \hat{x}^{(1)} \\ M \\ \hat{x}^{(q)} \\ t^* \end{pmatrix}}_{x_Q}
\qquad \text{(RQ)}
$$

Using the model (RQ) we can now apply the same estimation method as in section II.A analog to (TG) to solve (RQ):

(TGQ)
$$
\| A_Q t_Q - x_Q \| \to \min
$$
$$
s.t. \| t_Q - t_{EQ} \| \to \min
$$

The initial estimate for the QoS class specific traffic matrix t_{EQ} is constructed from the total traffic demand between the nodes that is divided proportionally to QoS class specific link utilizations, more precisely:

$$
t_{EQ} = (t_{EQ}^{(1)}, K, t_{EQ}^{(q)})
$$

where

$$
\left(t_{EQ}^{(l)} \right)_{i \cdot j} = t_{i \cdot j}^* \cdot \frac{X_l}{\sum_{r=1}^{q} X_r} \quad, l = 1, \dots, q \text{ and } X_r = \sum_{k=1}^{m} \hat{x}_k^{(r)} .
$$

The estimation accuracy of (TGQ) will depend highly on the size of the subset of links that have QoS specific link utilizations available. Also it should be noted that (TGQ) can result in traffic matrix estimations with negative entries since (TGQ) has no condition for t_Q being nonnegative. We apply a simple iteration scheme to assure positive solutions whose convergence is discussed in the following section.

Compared to an application of the tomogravity method (TG) for each QoS class separately (TGQ) introduces further restrictions: the sum of the QoS class specific traffic demands must equal the total traffic for the respective traffic relation. These additional restrictions should improve the estimation quality of the method. Section 3 investigates the improvement of the estimation quality for a concrete example with realistic data.

Model (RQ) applies the same routing for all QoS classes but it can be easily extended to QoS specific routing schemes if we replace $\hat{A}$ by QoS specific routing matrices $\hat{A}^{(l)}$ ($l = 1,…,q$).

3. NUMERICAL RESULTS

The QoS-tomography model (TGQ) is applied to an example network with 26 nodes and 190 edges (Figure 2). We assume four QoS classes: voice, low loss, low delay, and best effort.

In the following section, we discuss two problems: estimation convergence and estimation accuracy. The third example shows estimation results based on QoS link utilization measurements from our backbone network. The numerical results focus on the differences in the demand structures of the QoS class specific traffic matrices. Therefore traffic matrices for one given (15-minute) time interval are compared.

Another difference that is not covered within this paper is the difference in the time dependent variance of the traffic volume: Different QoS classes will have different daily peak hours. As an example Figure 1 shows this difference in the daily link utilization chart for one backbone link. The time dependent behavior of traffic matrices is particularly important for the choice of the traffic matrix that is actually used for traffic engineering or network planning: For non-best-effort traffic classes the traffic profile might follow business hours and a traffic matrix from a different time interval has to be used for network planning or traffic engineering purposes.
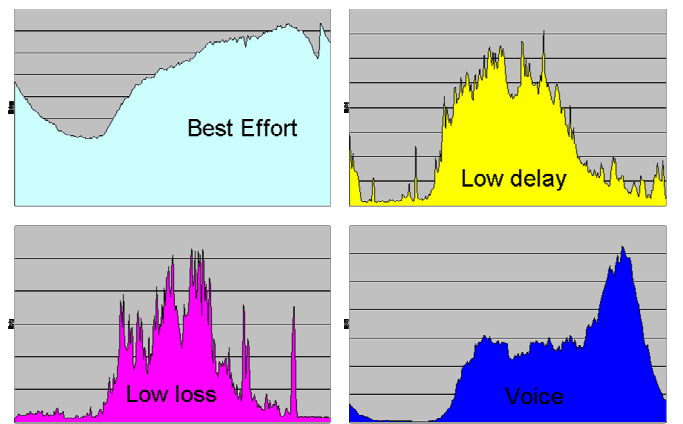

Figure 1: Daily profiles for different QoS classes for one backbone link (normalized diagrams)
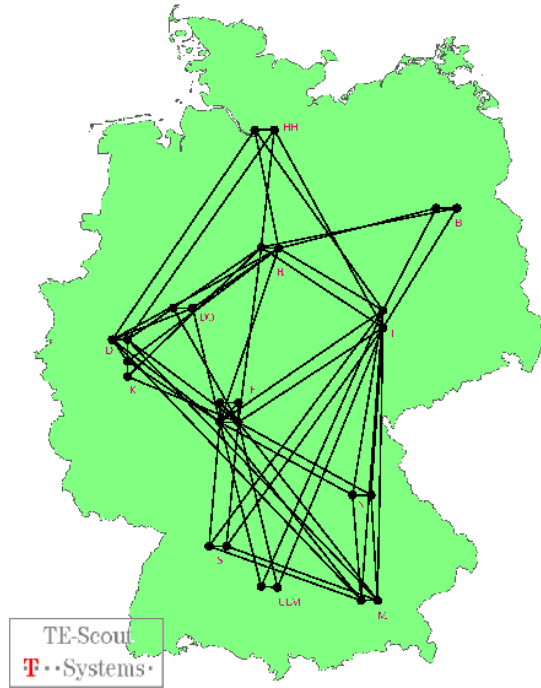
Figure 2: Example network topology.

3.1 Estimation convergence

Dependent on the initial estimate vector t_{EQ}, problem (TGQ) may result in negative values for individual traffic matrix entries. In this case we apply an iterative procedure where the negative elements per origin-destination (o-d) relation are set to 0 and the negative volume is added to the values of the positive entries. After this modification, the sum over the QoS classes per o-d relation is unchanged and all entries of the traffic matrix are non negative. This result is used as a new start vector t_{EQ} for the next iteration with equation (TGQ). The diagram in Figure 3 shows that the iteration scheme converges: The percentage of the negative traffic matrix volume is reduced from 2 % to 0.02 % after only 20 iterations.
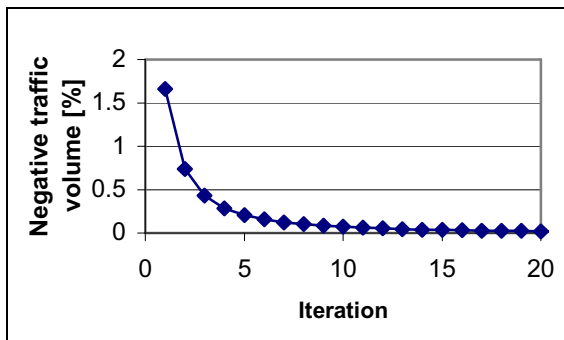

Figure 3: Estimation convergence

3.2 Estimation accuracy

For analyzing the estimation accuracy we assume given traffic matrix values for all QoS classes. This given vector of traffic matrices is denoted by t_Q^*. We use a random number generator to set the values in the range between 0 and 100000. In a simulation, the traffic of the

four QoS classes is routed according to the IGP metrics of our backbone network. From the routed traffic we can then calculate the link utilizations $\hat{x}_k^{(l)}$ for each link k and QoS class l. The traffic matrix t^* of the total traffic is calculated from the sum over the given QoS traffic matrices. Finally, t_Q is estimated from t^* and x_Q as a solution of (TGQ) and we calculate the relative error E as measure for the estimation accuracy.

$$ E = \frac{1}{qn^2} \sum_i \frac{\left| t_{Q_i} - t_{Q_i}^* \right|}{t_{Q_i}^*} $$

If we take into consideration all elements of the traffic matrices, the mean relative error is dominated by the small elements. As the contribution of these elements to the overall traffic volume is very small, we focus on the mean relative error E_α of those elements that are larger than the $(1-\alpha)$ quantile of the traffic matrix element distribution function.

Figure 4 we compare E_α from the (TGQ) method (normal lines) with results from a tomography model with a start vector from the gravity model (TG model: dashed lines, same color). The tomography model (TG) is applied to each QoS class separately. The results show that the mean relative error is reduced by a factor of 2.5 compared to results from the (TG) estimation. Nevertheless, a mean relative error in the range between 30 % and 100 % indicate that further improvements are needed.
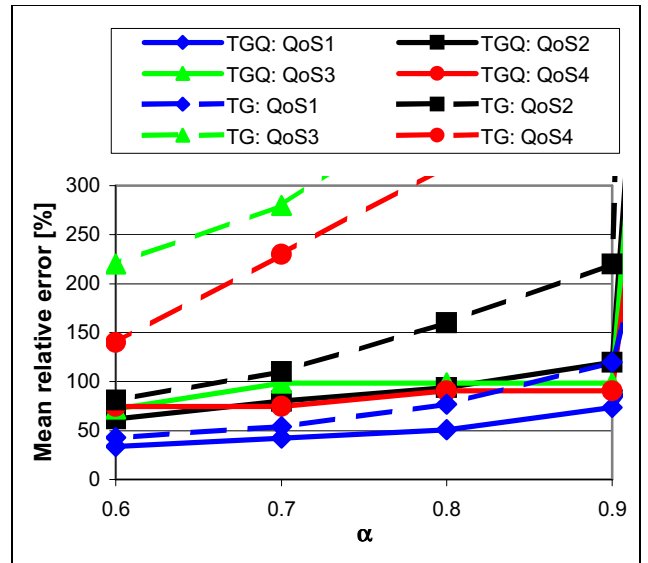

Figure 4: Estimation accuracy.

One application of QoS class specific traffic matrices is QoS dependent routing - especially if the structure of the QoS traffic matrices differs from class to class. The estimation result is useful only, if the structures are identified accurately by the estimation method. The following accuracy test analyzes this problem. We assume two different QoS classes. Class 1 traffic originates at nodes 1 to 12, class 2 traffic originates at nodes 13 to 26. Each source sends 100000 traffic units to each other node of our network. Nodes 1 to 12 and nodes 13 to 26 respectively are not in a close topological region. The nodes are sorted by name alphabetically and numbered from 1 to 26. The estimation result for the QoS class specific traffic matrices is shown in Figure 5 and Figure 6. The matrix elements are colored as

follows: green: less than 1000 traffic units, yellow: between 1000 and 50000 traffic units, red: more than 50000 traffic units. Therefore, a perfect match is achieved, if the first 12 rows of the QoS 1 matrix are red and the others are green and vice versa for QoS2. The structures of the QoS traffic matrices are accurately identified in the estimation.

The mean relative error E is 14.6 % for QoS class 1 and 11.2 % for class 2. If we calculate the E_α, only the positive elements of the traffic matrices are considered. In that case, the mean relative error E_α is 6.3 % for class 1 and 5.3 % for class 2.

classes (Figure 7). Obviously, the total traffic volume is dominated by the best effort traffic (QoS4) and the structure of the QoS4 matrix agrees with the structure of the total traffic matrix (Figure 8). The structures of the QoS1, QoS2 and QoS3 traffic matrices are different to the structure of the total traffic matrix. This is mainly caused by specific application architectures (e.g. position of voice gateway routers), or locations of specific customers with relatively high QoS demands. From the QoS1 traffic matrix in Figure 9 (voice traffic), we can see that a lot of traffic is destined to node 1 and node 2, the location of voice gateways. Main sources are the nodes 7 to node 14.
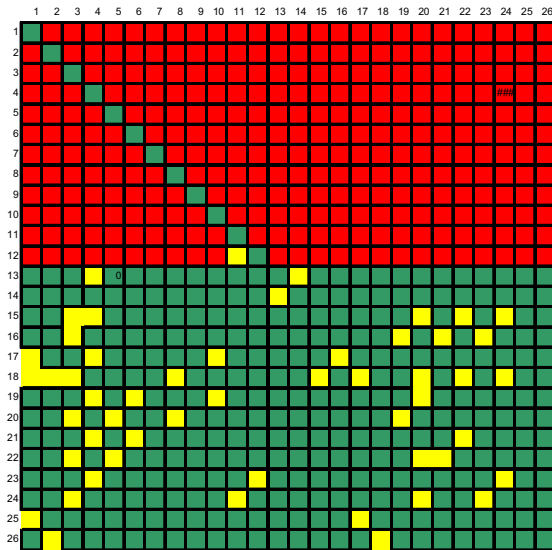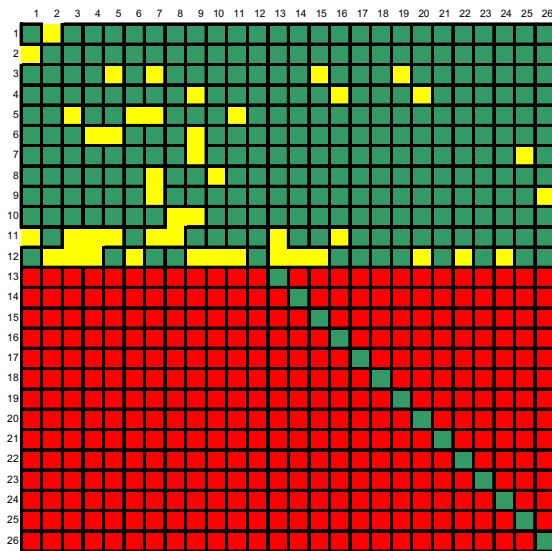

Figure 5: Estimation of QoS 1 traffic.

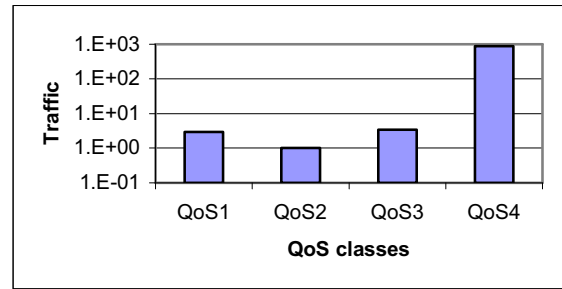**Figure 7: Distribution over QoS classes.**

Figure 8: Overall traffic matrix.

Figure 6: Estimation of QoS 2 traffic.
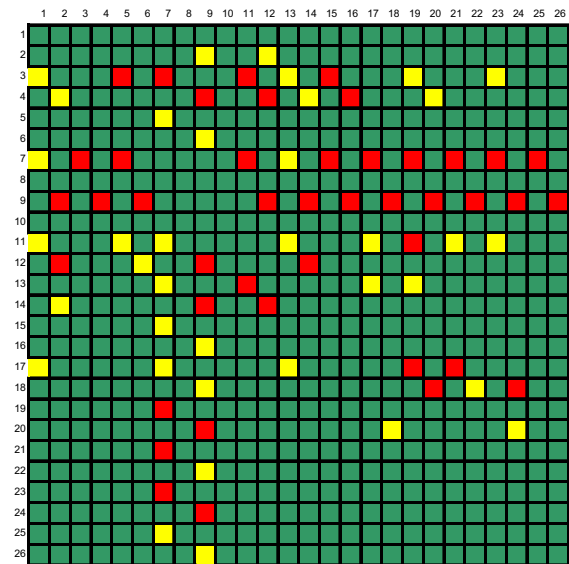
3.3 Estimation results with real network data

In the following section the QoS class specific traffic matrix estimation (TGQ) is applied to data from our backbone network. The total traffic matrix is calculated by means of the LDP method. In the considered network topology only 12 of the 26 backbone routers could measure the link utilization statistics per QoS class. Therefore, we expect more accurate results when all the routers are upgraded to an OS version that allows measurements per QoS class. First we show the relation of the amount of traffic between the four QoS
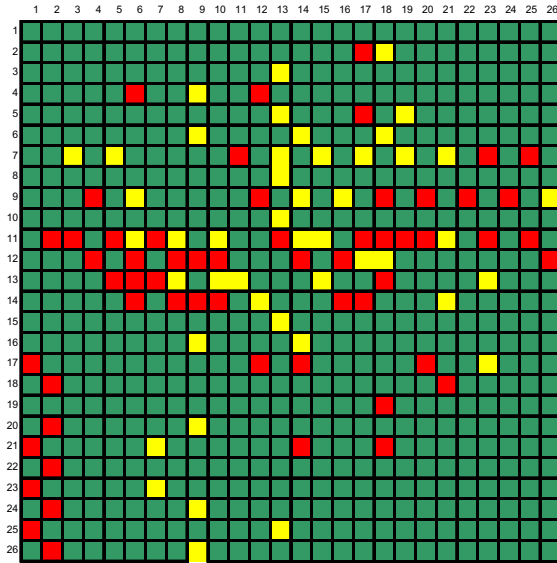
Figure 9: QoS 1 traffic matrix.

4. CONCLUSION

We have presented a new model and numerical results from a real world network for QoS class specific traffic matrices. We demonstrated that this method can significantly improve the estimation quality. The numerical results show the differences in the demand structure for different QoS classes. Thereby we motivate why the availability of good quality QoS class specific traffic matrices is important for efficient network planning and traffic engineering.

The absolute size of the estimation errors is still quite high for the intended use of the QoS class specific traffic matrices. On the other hand a good estimation of the demand structure of the traffic matrices could already be achieved and the need for QoS class specific traffic matrices will only grow with the traffic growth in non-best-effort traffic classes. Figure 7 shows that there is still some time to improve the estimation quality.

Future improvements of our method include the usage of state-of-the-art methods from numerical linear algebra (e.g. sparse matrix SVD). This would allow us to investigate larger network topologies. Furthermore, we aim to incorporate a larger number of QoS class specific link utilizations – not only from links within the backbone network but also from ingress links into the network - to improve the estimation quality.

5. REFERENCES

[1] *Netflow Aggregation*, Cisco IOS release 12.0(5)T.

[2] E. Rosen, A.Viswanathan, R. Callon: *Multiprotocol Label Switching Architecture*, IETF RFC 3031, Jan 2001

[3] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, G.Swallow: *RSVP-TE: Extensions to RSVP for LSP Tunnels*, IETF RFC 3209, Dec 2001

[4] L. Anderson, P. Doolan, N. Feldman, A. Fredette, B. Thomas: *LDP Specification*, IETF RFC 3036, Jan 2001

[5] S. Schnitter, M. Horneffer, *Traffic Matrices for MPLS Networks with LDP Statistics*, Proc Networks2004, VDE-Verlag, Vienna, 2004.

[6] S. Schnitter, T. Morstein and M. Horneffer, *Combining LDP Measurements and Estimation Methods for Traffic Matrices in IP/MPLS Networks*, Proc. Networks2006, VDE-Verlag, New Delhi, 2006.

[7] S. Schnitter, G. Haßlinger, Heuristic Solutions to the LSP-Design for MPLS Traffic Engineering, Proc. Networks2002, VDE-Verlag, Munich, 2002.

[8] M Horneffer, *IGP tuning in an MPLS network*, NANOG 33, Las Vegas, 2005.

[9] B. Fortz, M. Thorup, Internet Traffic Engineering by Optimizing OSPF Weights, Proc. IEEE INFOCOM'2000, 2000.

[10] J. Kowalski, B. Warfield, *Modeling traffic demand between nodes in a telecommunications network*, Proc. ATNAC'95, 1995.

[11] J. Cao, D. Drew, S. Wiel, B. Yu, *Time-Varying Network Tomography: Router Link Data,* Bell Labs Tech. Memo, 2000.

[12] Y. Zhang, M. Roughan, N. Duffield, A. Greenberg, *Fast Accurate Computations of Large-Scale IP Traffic Matrices from Link Loads*, Proc. SIGMETRICS'03, San Diego, 2003.

[13] A. Gunnar, M. Johansson, T. Telkamp, *Traffic Matrix Estimation on a Large IP Backbone   A Comparison on Real Data,* Proc. IMC'04, Taormina, 2004.

[14] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Croz, A. Greenbaum, S. Hammarling, A. McKenny, D. Sorense, *LAPACK User s Guide*, SIAM, 1999.

[15] M. Berry, *Large Scale Singular Value Computations*, International Journal of Supercomputer Applications, 6:1, pp. 13-49, 1992