THE LATENCY PROBLEM IN VERY LONG-DISTANCE OPTICAL GIGABIT LINKS

Abelardo Podcameni, Marbey M. Mosso, Cláudia B. Marcondes,
Joseuda Lopes and Gustavo B. Leal

Optical Systems and Microwave Group GSOM – CETUC / PUC-Rio
Center of Telecommunications Studies of the Catholic University of Rio de Janeiro
Rua Marquês de São Vicente 225, Rio de Janeiro, 22453-900 Brazil.
FAX: +(021) 294-5748; abelardo@cetuc.puc-rio.br

Abstract - The latency effect in long-distance gigabit links and networks may lead to some transmission penalties, as here discussed and illustrated through various examples. When transcontinental or overseas connections are considered, these penalties may become remarkably severe. By evaluating and discussing realistic cases, some countermeasures are suggested. These last may include an important conceptual change in several well-established protocols and a substantial increase in the data packet sizes together with some other modifications concerning the physical layer management.

Resumo - O efeito de latência em enlaces e redes de alta velocidade e longa distância podem levar a algumas penalidades de transmissão, como é aqui discutido e ilustrado por meio de vários exemplos. Quando conexões transcontinentais são consideradas, estas penas podem se tornar particularmente severas. Pela avaliação e discussão de casos realísticos, algumas medidas preventivas são sugeridas. Estas medidas podem incluir uma importante alteração conceitual em vários protocolos bem estabelecidos e um aumento substancial nos comprimentos de pacotes, assim como algumas modificações relativas ao gerenciamento de camada física.

Keywords: Latency, optical transmission penalties, gigabit networks, protocols, optical networks, optical transoceanic data links, FEC, TMN.

1. INTRODUCTION

The evolution in Optical Communications Network is such that bit rates of 2.48 and 10 Gbit/s are already in use. Higher rates as 40 and 100 Gbit/s are next in line [1]. Additionally, experts are forecasting that the future demand for Internet and multimedia services will dramatically increase. In this approaching scenario, the traffic, running through the infoways, will present a high degree of granularity and diversity. A huge concentration of traffic will arise from an enormous number of lower-speed traffic sources. These last sources will be mostly those operating through electric domain, using, for instance, technologies as ATM and IP. Quite often, SONET, unprotected Packet-over-Sonet [2-4], or still other arrangements to come, will be used. One may expect that the act of downloading, from a *virtual video rental shop*, a two-hour compressed movie (which shall be a 21-Gbit file) will take only a few minutes and will become a routine.

The shifting to the multi-gigabit rate domain will require a quite different operational strategy, with respect to the present ones. The reason is that the propagation delay effects dramatically change the traffic and information flux control [5,6].

The *total transmission delay*, t_{tot}, a message suffers, while progressing through a network, may be said to be composed by three parcels. The first one arises when the low speed traffic channels are routed into the multi-gigabit channel. There, a data package must compete with other ones, and a *buffering waiting time*, t_{wait}, will be present. The second parcel is the time, t_{ch}, for introducing all the bits of a packet into the high-speed channel. For a packet consisting of P bits and a channel operating at C bit/s, one has $t_{ch} = P/C$. Finally, due to the finite value of light speed, a propagation time delay or *latency*, t_{lat}, throughout the channel, will be present. The total delay, in a simple way, is:

$$t_{tot} = (t_{wait} + t_{ch}) + t_{lat} \qquad (1)$$

If the link distance were very short, or if the light speed would be infinite, then $t_{lat} \to 0$. On the other hand, as the channel capacity increases, both t_{ch} and t_{wait} decrease: so t_{lat} progressively becomes more relevant.

The present work aims at discussing the role that latency plays in a multi-gigabit network. Realistic parameters and accurate computer simulations are initially used, leading to a number of results, which are subsequently discussed. Finally, some solutions and techniques, borrowed from the process and transport management areas, will be used to peacefully live together with latency – at least in the short term.

2. FOCUSING LATENCY

The first type of latency problem, concerning multi-gigabit links, derives just from its own high pulse rate [6]. The pulse sequence may be so rapid that a certain pulse may enter the circuit while the previous one is still travelling by. This phenomenon is called pulse latency and is a severe suggestion against transmitting back and forth acknowledgement – or handshake – signals. By supposing that fiber pulse propagation occurs at 200,000 km/s speed,

Fig. 1 is furnishing a pictorial view of the maximum covered distance – for several bit rates - without the presence of pulse latency. Observe that a multi-gigabit link is not even able to reach a meter range, without pulse latency.

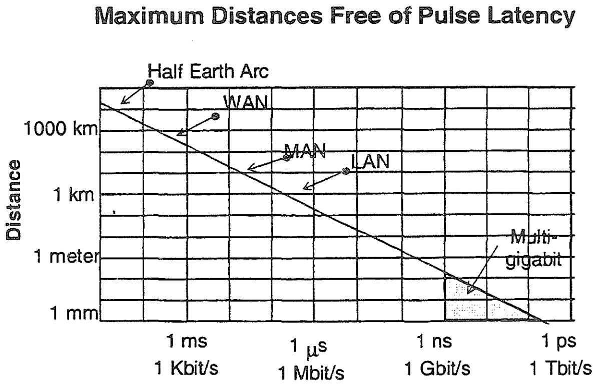


Figure 1. Example of a figure in single-column format. Maximum achieved transmission distances without pulse latency occurrence

A judicious approach concerning latency must be exercised. However, the finite speed of light is not the single latency problem in a multi-gigabit link. In Data Communications, the control protocols are rather packet-oriented than pulse-oriented. The traveling packets usually come from a multitude of low-speed sources and compete among them for being transmitted. Observe that it is very unlike that a multi-gigabit facility will serve exclusively a single user: rather it will be shared by many ones. Let then F be the fraction of the entire facility capacity, C, which is being used by the focused group of clients.

The F parameter, for each set of users, depends upon the level of activity practiced by this group. A classical queuing model of a Poisson stream of arriving messages requesting transmission may be assumed. In this model, the length of each message is exponentially distributed with a mean of P bits: the so-called M/M/1 system [7-9]. The F parameter may be written as:

$$F = \lambda P / C . \qquad (2)$$

where λ is the message arrival rate.

In this situation, it is possible to write t_{tot} [5], as below:

$$t_{\text{tot}} = P / [C(1-F)] + t_{\text{lat}} . \qquad (3)$$

In (3), the first parcel is recognized as $(t_{wait} + t_{ch})$, as already presented in (1).

The above expression of t_{tot} regards the mean time elapsed from the instant when the package first bit arrives at the tail of the transmitting queue, up to the moment that the last bit leaves the channel, including any propagation delay. Still in (3), it is not possible to change the second parcel, t_{lat}, provided the distance and the propagation media are given. However, the first parcel, $(t_{wait} + t_{ch})$, may be changed, if needed – meaning less delay – through proper choices of λ, P and C, as seen in (2). To decrease the delay associated with λ, P and C, when necessary – without increasing prohibitively the service costs – is the task expected from the network management staff.

The above issue may use the strategy of setting an acceptable lower limit for the $(t_{wait} + t_{ch})$ delay. A reasonable suggestion is not to try any further effort to reduce the $(t_{wait} + t_{ch})$ delay, beyond the so called critical value, defined as:

$$t_{\text{wait}} + t_{\text{ch}} \cong t_{\text{lat}} = t_{\text{crit}} . \qquad (4)$$

where, in the above equation, a very light load condition $(F \to 0)$ is assumed for evaluating the parcel $(t_{wait} + t_{ch})$.

Within the above-described scope, some realistic cases will be further discussed for enlightening latency mechanisms.

2.1. A GIGABIT OFFICE CASE

Let a set of users be considered, all of them in the same office, where a number of B-ISDN lines converge to a router leading to a gigabit link. The average data packet size may be assumed as 5 kbits, which is a typical average size for IP. Consider also that the main outgoing traffic destination is 2 km apart. The interconnection with the destination is made by sharing a 2.488 Gbit/s network. The office activity level may be modeled by (3) and the above share is F = 10%, within the 2.488 Gbit/s link.

Fig. 2 is describing this case. The delays are presented as a function of F, to cope with any future change in the office activity level. Three sets of results are there presented. The first is the $(t_{wait} + t_{ch})$ delay. The second is t_{lat}, assuming a 2-km long connection, with standard monomode fiber. The third is the t_{tot} set.

For F = 10%, one has $(t_{wait} + t_{ch}) = 2.2$ μs, and, $t_{lat} = 10$ μs, yielding a $t_{tot} = 12.2$ μs. The network is dominated by latency, i.e.: $t_{lat} > (t_{wait} + t_{ch})$. Two observations may be done.

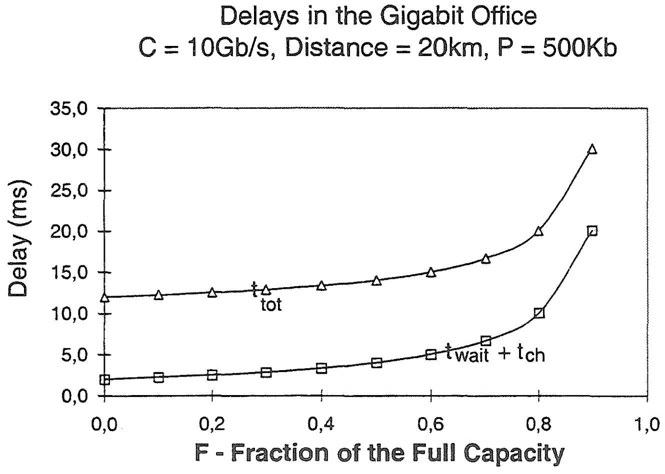


Figure 2. Latency effects in the gigabit office as a function of the facility loading capacity fraction utilization

(i) It is not quite effective to increase the gigabit channel capacity. This procedure would only reduce the parcel concerning the $(t_{wait} + t_{ch})$ delay; keeping the network still dominated by latency. For example, if the above channel capacity is increased to 10 Gbit/s, a lower value $(t_{wait} + t_{ch}) = 0.5$ μs results, while the latency value does not change. The new value of $t_{tot} = 10.5$ μs (from 12.2 μs) does not pay the cost of the increase in the bit rate.

(ii) The previous latency delay, 12.2 μs, is still very small when compared with usual values recognized as limits for acceptable quality of service. Only when the latency values would reach the range of several tens of millisecond, is that the quality of service (for voice or life video) may be depreciated. Nevertheless, the present example describes a short 2-km range link. When transoceanic distances together with gigabit capacities are focused, the latency delays may reach a range of such high values that might degrade the quality of service. Specifically for the above example, no change concerning the channel capacity is recommended for improving its effectiveness or quality of service.

2.2. GIGABIT NETWORK DESIGN GUIDELINES

The influence of the channel capacity interacting with the packet size will now be presented. For all subsequent cases, it will be assumed that the user level of activity is such that one always have $F = 0.2$. The concept of critical time, as given in (4), will be used, leading to similar concepts as critical capacity and critical package size.

The next presented results may be used as a yardstick to avoid sensible differences between $(t_{wait} + t_{ch})$ and t_{lat}. This balance is important as far as network managing is concerned.

When $(t_{wait} + t_{ch}) >> t_{lat}$, it means that network electronics is too slow. Some sort of improvement – if possible – must be performed within the network, in order to reduce the delay associated to $(t_{wait} + t_{ch})$.

When $(t_{wait} + t_{ch}) << t_{lat}$, there are two possibilities. First, the electronics is too fast, and maybe, money is being wasted, provided t_{lat} is less than a few milliseconds. Second, when t_{lat} is several milliseconds, it means that .the network is a very long distance one and it is heavily dominated by latency. In this last case, perhaps no economic benefit will be achieved in increasing electronics speed.

Fig. 3 presents the critical package sizes with respect to the transmission distance and channel capacity, by using the simple expression shown in (4). Given a package size, one must observe whether an operating point (corresponding to a given capacity and distance) is positioned above the pertinent package size locus. If so, the network delay behavior is dominated by latency. If not, it is dominated by queuing plus the channel delay.

In Fig. 3, a 500-bit package may be initially focused, which is a typical size for ATM applications. Regarding its transmission through a 10-km distance, the associated critical capacity is 15 Mbit/s. If higher capacities are considered, latency dominates, meaning (in theory) that a bit rate with more than 15 Mbit/s is too fast for this application. However, within this range of frequency, electronics cost is such that the use of 15, 30, or 50 Mbit/s will not make any sensible difference. Thus, small packets are better suited for short distances and/or low channel speeds.

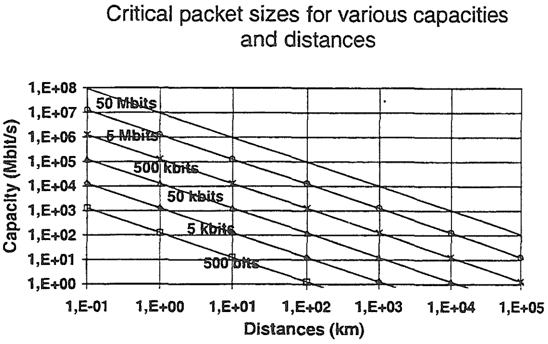


Figure 3. Capacity may be obtained by choosing the critical packet size for a given distance

Using a packet size of 5 Mbit for a 1000-km distance, the critical capacity would be 1.5 Gbit/s. In this case, latency time would be 5 ms, which added to (critical) queuing and transmission delay time, will result in ~9 ms. Instead, if a 40 Gbit/s channel were used, latency would dominate. With this last bit rate, the queuing and transmission time would almost vanish. The relevant point here is whether the cost of shifting from a 1.5 to a 40 Gbit/s channel is worth paying. Note that the total delay has only diminished from 9 ms to 5 ms. Thus, large packages are better suited for greater distances and/or higher channel speeds.

In Fig. 4, it is more clearly emphasized how large packages are best suited for transcontinental, or overseas, multi-gigabit links. Consider a 10,000-km distance, where a respectable 50 ms latency delay exists. Suppose a network where the high-speed channel operates at a 10 Gbit/s rate. Fig. 4 shows that the critical package size would be near the gigabit range: 800 Mbit. Presently, this packet length does not exist, neither its pertinent correspondent protocol. Consequently, packets of smaller size must be used alternatively, leading the network to work deeply dominated - again - by latency.

2.3. THE CONCEPT OF EFFECTIVE TRANSMISSION RATE

The effective transmission rate, R_{eff}, may be defined as the ratio between the amount of transmitted bits, B, and the time, T, taken to perform the complete task. It is apparent that the R_{eff} values decrease if one uses protocols calling for handshaking signaling. Handshakes imply in forth and back confirmation checks, and these add two extra latency periods for every check cycle. Furthermore, if errors are considered, R_{eff} further decreases, as detect-and-correct error cycle costs two latency delays per each error.

For bi-directional communications such as conversations (human or machine) and (video) conferences, the handshaking cyclic delays will provide for a further R_{eff} reduction. This happens even supposing that no errors should exist, that the remote part takes no time for answering, and that the back acknowledgement message is very short. Within this scope, it is easy to demonstrate that the R_{eff} may be put as:

$$R_{\text{eff}} = C/\{1+2[(t_{\text{wait}}+t_{\text{lat}})/t_{\text{ch}}]\}. \qquad (5)$$

here t_{lat} is expressed in seconds. If errors are present, this te is additionally degraded. It is apparent that $R_{eff} < C$.

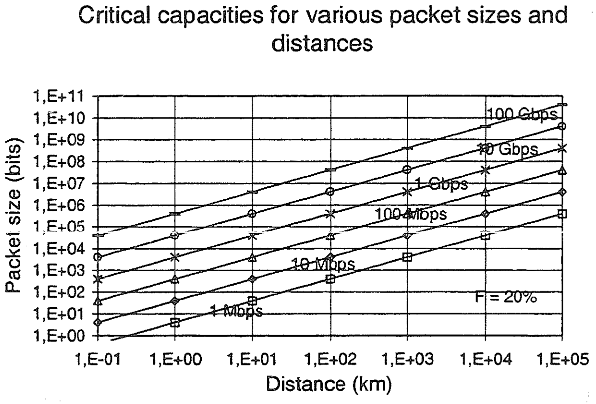
In the above expression, while dealing with anscontinental multi-gigabit sectors, it is quite common to ave: $t_{wait} << t_{lat}$. Consequently, (5) gets a simpler form:

$$R_{eff} = C/\{1+[2t_{lat}/(P/C)]\}. \qquad (6)$$

xample 1. Now, the above paragraph is illustrated. Let the ffective transmission rate for a prospective video global onference be evaluated, assuming a 10-Gbit/s capacity hannel, and that a 20,000-km distance transmission fiber ngth exists.

(a) Consider that the specific value of the critical acket size, as given by Fig. 4, will be initially used. This ize is 800 Mbit. Due to well established present electronic ircuitry speed, let be supposed that $t_{wait} << t_{lat}$, and (6) will en be used.

By using (6), first evaluate $t_{ch} = P/C = 0.08$ s. Secondly, valuate the latency time for 20,000 km: $t_{lat} = 0.1$ s. Next, bserve the value of the ratio $[(t_{lat})/(P/C)]$ is 0.1/0.08 = 1.25 inally, obtain: $R_{eff} = 10^{10}/3.5 = 2.86$ Gbit/s.


Critical capacities for various packet sizes and distances

igure 4. Packet size may be obtained by choosing the ritical transmission bit rate for a given distance

(b) Unfortunately, a packet as large as 800 Mbit is not a ecognized standard. Switch then to a 10-kbit packet, a ypical value for the SONET cell.

In this last case, $[(t_{lat})/(P/C)] = 10^{-2}/10^{-6} = 10^{4}$. The alue of the effective bit rate is then obtained as: $_{eff} = 10^{10}/(2\times10^{4}) = 500$ kbit/s.

The amazing difference between the two cases is xactly due to the delay concerning the back and forth onfirmation signals, which are more frequent when using mall packets. Thence, the less answers the destination oes, the less time it will take to end the conversation, creasing therefore R_{eff}.

It is possible to reduce the destination's answers by everal ways. The first is to transmit larger packets, equiring fewer acknowledgements. The second way is to ecrease the system intrinsic error rate. A third option is to pply error corrections at the destination node, by using a ode that allows for forward error correction, FEC. Finally, ne may employ special protocols, like *sliding-window*, for stance. This last, which is a part of TCP protocol, is used r satellite communications where a number (although limited) of packets are allowed to go before the first acknowledgement is received.

(c) For avoiding these back and forth delays and or special protocols, the use of a non-acknowledgeable-FEC would lead to the full transmission channel capacity utilization. In both (a) and (b) above cases, an $R_{eff} = 10$ Gbit/s would be obtained.

3. DISCUSSION

In modern networks, a very large diversity, concerning design parameter, arises, with respect to distance, channel capacity and packet size. With distances typically ranging from 20 m to 20,000 km, the resulting latency delays span over six orders of magnitude. With the possible choices for channel capacities ranging from 64 kbit/s up to 100 Gbit/s, the resulting transmission rates span over seven orders of magnitude. In contrast, the third parameter - the packet size - allows only for an extremely modest range of choice.

For a Global World, where one may live and work in Asia, while doing a job for an U.S. software house, the above paragraph argumentation has already reached a dramatic level. By continuously dealing with large distances and gigabit capacities, the need of extremely large packet formats, associated to pertinent protocols, may become even more eloquent.

The very near future is pointing towards the launch of the all-optical networks. Incidentally, these last solutions are indeed knocking at our doors. To cope with the forthcoming high-speed all-optical networks, a new whole family of protocols supporting FEC-based codes or very large transmission cells – in the Mbit and lower-Gbit range – ought to be developed.

Anyhow, for the time being, palliative measures – as the simultaneous use of extra connections for handling return signaling signals and/or known forms of parallel processing and pipeline [10] – may be used.

Large packets, on the other hand, may present some disadvantages. Perhaps the most significant one concerns with the excess of delay, associated with error retransmission.

4. CONCLUSIONS

In this work, the multi-gigabit links and networks were focused. Specifically, the latency problem has been discussed, together with queuing and channel delays.

A model of an efficient gigabit office, using a relatively short connection of 2-km, has been described. It has been shown that, when latency is dominant, increasing channel capacity produces no sensible benefit.

For transoceanic links and networks, the latency time delay may reach a value of several tens of milliseconds. Within this scope, it is not economical to substantially decrease the delay concerning queuing and channel delays below the millisecond range. On the other hand, the use of FEC-based codes may be recommended.

Still in very long gigabit links, alternative to FEC, it has been recommended to increase the packet size – at least – to the Mbit range, even knowing that in most cases, the near-Gbit length would yield better results. This last recommendation is - probably - the most relevant contribution concerning this paper. Figs. 3 and 4 are furnishing suitable packet sizes for any combination of distance and bit rate.

Incidentally, the use of non-fixed-size packets has already been suggested for the ACTS program, within the KEOPS project [11]. There, it has been recognized that the ATM fixed 53-bytes length is too short for most applications. A flexible different format is there presented for coping with very high-speed optical transport technologies. However, as a substantial increase in the packet sizes has not been suggested there, the use of long-distance links will still be jeopardized, by an effective drop in the transmission rate, as it has here been described.

In any case, it would be quite rewarding if the network management could be adaptive, in the sense of being able to equalize packet or cell sizes, according to each specific network delay. This last, would require intelligent protocols, heavily supported by sophisticated TMN.

In conclusion, this paper is reporting that long-distance optical links working in the multi-gigabit range must be regarded differently than those that cover short-distances. Some possible countermeasures for combating latency penalties, as the use of large data packets, FEC-based codes, or still new protocols combined with an adaptive network management, were here related. It is not expected that this issue will be solved with a single closed solution. Rather, more discussions are still needed in order to bring transoceanic multi-gigabit networks to an efficient degree of maturity.

REFERENCES

[1] S. H. Hersey and M. R. Wilson, "High Capacity Transport/Bandwidth Management Solutions: Realizing the Terabit Office", in: Proceedings of the 1998 Annual Multiplexed Telephony Conference, Vol. 1 (Atlanta), Aug 98, pp. 190-195.

[2] A. Tanenbaum, Computer Networks, Prentice Hall, 2nd-Edition, New Jersey, 1988.

[3] R. Handel, "Evolution of ISDN towards Broadband ISDN", IEEE Network Magazine, Jan 89, pp. 7-13.

[4] R. Ballart, and Y.C. Ching, "SONET: Now It's the Standard Optical Network", IEEE Communications Magazine, March 89, pp. 8-15.

[5] T. Miki, "The Potential of Photonics Networks", IEEE Communications Magazine, Dec. 94, pp. 23-27

[6] L. Kleinrock, "The Latency/Bandwidth Tradeoff in Gigabit Networks", IEEE Communications Magazine, April 92, pp. 36-40.

[7] L. Kleinrock, Queuing Systems, Vol. I: Theory, John Wiley & Sons, NY, 1975.

[8] M. Schwartz, Telecommunication Networks (Protocols, Modeling and Analysis), Adisson-Wesley Publishing Company, Massachusetts, 1988.

[9] A. Kershenbaum, Telecommunications Network Design Algorithms, Mc.Graw-Hill, Singapore, 1993.

[10] A. S. Tanenbaum, Modern Operating Systems, Prentice-Hall, New Jersey, 1992.

[11] F. Callegati, "Which Packet Length for a Transparent Optical Network?" SPIE - Symposium on Broadband Networking Technologies, (Dallas), Nov 97.

Abelardo (Belex) Podcameni received the BSEE and the MSEE by Catholic University of Rio de Janeiro in 66 and 69 respectively. His PhD status has been achieved in 1979, together with a promotion to the Associate Professor rank, at the same University. In 1977, Belex has spent four months in France, working for CNET on 11 GHz GaAs Fet Oscillators. In 1982, again in France, he worked for Thomson-CSF in the design of Dielectric Resonators for Stable Microwave Generation. From 88 to 91, he joined Intelsat Staff, at the Intelsat Spacecraft Program, in El Segundo, California, monitoring the five Intelsat-6 spacecraft construction, tests and launching. Presently, he is with CETUC, Telecommunications Studies Center at the Catholic University of Rio de Janeiro, as Associate Professor, doing research and academic activities on Optical Systems for Communications. Belex is a Senior Member of the IEEE, since 82, his biography is listed in "Who's Who in the World" and he has more than forty papers published in various International Journals and Symposia, both in US and Europe.

Marbey Manhães Mosso received BSEE and MSEE by Catholic University of Rio de Janeiro in 72 and 75, respectively. Presently, he is Head of the Optical Systems and Microwave Group and Assistant Professor at CETUC - Telecommunications Studies Center at the Catholic University of Rio de Janeiro - doing research and academic activities. In 73, Marbey worked on miniaturized microwave circuits in France for CNET, in 80 on dielectric-resonator-stable microwave oscillators for the Allen Clark Research Center – Plessey/UK and in 84 on microwave mixers for Anaren in the US. In 89, he was a Visiting Professor in the Maryland University, US. Again in the US in 93, he developed at Miteq TV-oriented optical analog systems. Since 87, Professor Marbey has authored several papers and has conducted a number of projects dealing with optical systems and networks.

Claudia Barucke Marcondes was born in Volta Redonda (Rio de Janeiro State) and has firstly concluded a Professional Electronics Course. Next, she was admitted to the Catholic University of Rio de Janeiro, where she graduated in 1997 on Electrical Engineering, with emphasis on the Telecommunications Engineering Area. In 1998, she was accepted for a postgraduate course in the same Catholic University of Rio de Janeiro, again in the Telecommunications Area. In 1999 Claudia has co-authored a paper presented in the AMTC99 Conference, dealing with Latency in Optical Transoceanic Gigabit Links. She succeeded finishing her MSEE in April 2000 (with honors), and her Thesis subject is related with Optical Network Management, Protection & Restoration implemented on the Physical Layer. Presently, Claudia is with a long-distance carrier telecommunications company, INTELIG, where she is a Transmission Engineer, dealing with optical network management.

Joseuda Lopes received the Electrical Engineering and the MSEE in Applied Electromagnetism – Optical Networks – degrees from PUC-Rio University, Brazil, in 1997 and

1999, respectively. She is presently with EMBRATEL working in QoS for the International Telephony Switching and Backbone Division. Simultaneously, she started the studies toward her doctorate in the beginning of 2000, comprising the introduction of intelligence in the physical layer of optical networks for optimal wavelength routing. Her interests include all-optical networks, WDMA, the wavelength routing and assignment problem and QoS.

Gustavo Bevilacqua Leal was born in Salvador (Bahia/Brazil) in 1970. He has graduated BSEE, 1994, in the Federal University of Rio de Janeiro (UFRJ) with emphasis in Electronics. He also carries a Master Degree in Mathematics, obtained in 1998, from the Pure and Applied Mathematics Institute (IMPA – Rio). During 18 months, he worked in a software-house as a software engineer. Next, he has enrolled in the PhD program of the Telecommunications Studies Center (CETUC/PUC-Rio), for pursuing a PhD degree in the Antennas field. Presently, he is working on the final stage of his thesis.