Report outlining scenarios for the evolution of the pan-European infrastructure for European research and education networking

IST-2001-34925

Study into European Research and Education Networking As Targeted by *e*Europe

SERENATE

Deliverable no. D18

Delivery type: Report
Security: Public document

Authors: Dai Davies, DANTE
Karel Vietsch, TERENA

Date: 30 December 2003

CONTENTS

Executive Summary

In Europe, the provision of network services to research and education is organised at three levels: the Local Area Network to which the end-user is connected, the national infrastructure provided by the National Research and Education Network (NREN), and the pan-European level provided by GÉANT. An alternative to this structure could be the creation of regional networks covering a number of countries, or the (partial) replacement of the pan-European GÉANT network by bilateral connections between bordering countries. The latter option could in theory reduce costs as international connections could be established between neighbouring nodes close to the border. However, this model would bring considerable problems with regard to organisation and management. The current organisational model, with one NREN per country and close collaboration at the European level, has been a success factor for the development of research and education networking in Europe. It is expected that this model will remain in place and will continue to be a key to success for at least the next 5-10 years.

There are large differences between countries in the way that the national and international levels of research networking are funded, varying from total central funding to substantial funding via the connected institutions. As to the European level of research networking, GÉANT is partly funded by the European Commission, but the larger part of the GÉANT costs is covered from national contributions. Those costs are shared between the countries on the basis of an algorithm that has the effect to smooth the very large differences between the costs of communication services to individual countries. This approach is positive from the point of view of European coherence, but it has some negative side effects. As a consequence of the emergence of new applications that require very high network capacities and put heavy demands on network availability and end-to-end performance, it will become increasingly difficult to sustain the current cost-sharing approach.

In many parts of Europe, researchers have a reasonable environment of research and education networking. However, in several regions in Europe the situation is far from satisfactory and arguably getting worse in relative terms. When comparing the three levels of infrastructure provision, on the campus, nationally and internationally, for many European researchers the major source of limited network performance is primarily at the campus. Networking requirements will grow dramatically over the next 5-10 years, in all areas of research and in all countries. There are opportunities to start completely new research activities that were prohibited until now by lack of very high performance network facilities. There is a remarkable interest in Grid computing. User expectations have evolved beyond the provision of pure bandwidth towards the supply of more complex services. There will be a large demand for authentication and authorisation services, as well as for mobility – the ability to access networks wherever the researcher happens to be.

Research networking in Europe has developed rapidly in the past five years. This has been true particularly for pan-European co-operation. For most locations in Europe, the liberalisation of the telecommunications markets has made access to leading-edge communications technology possible. Prior to 1996, the monopolistic structure of the telecommunications markets meant that, internationally within Europe, telecommunications was rationed with capacities being limited by the monopoly suppliers. As a result, it was not easily possible to extend national research initiatives to a pan-European basis. The removal of this barrier has provided the opportunity for much enhanced European collaboration. Increasingly, it is becoming possible for researchers to disregard national frontiers in Europe and to plan research co-operation on a pan-European basis, knowing that the network infrastructure will be available that is necessary to support such co-operation.

It is the fundamental objective of the European research networking community to develop and extend the network infrastructure to provide optimal support to researchers across Europe, thereby acting as a key enabler for the European Research Area. The combination of the innovative spirit of research networking in Europe with the ability to offer efficient pan-European services is an important asset. The enormously increased capacity that is available has enabled technical developments, which make it possible for research networks in Europe to offer an advanced technical platform to their users. This platform provides the basis for application development, and for the deployment of new applications that are capable of operating in a distributed manner across Europe.

The emergence of scientific applications in which individual sites are capable of generating as much traffic today as the historic aggregate demand from a whole country, will lead to significant new challenges in meeting user requirements. A number of marketing, technical and commercial issues arise. It will be an important activity for the European research networking community to develop a market segmentation of the users and to use it to target and improve support to users. The GÉANT cost-sharing model will need to be revisited, as a direct translation of pan-European costs into charges to users may deter demanding users from taking advantage of shared network solutions.

Historically, it has always been assumed that it will be a logical and cost-effective network development to deploy higher transmission speeds and to continue using a router-based network. Now, it is not at all certain whether this will be the obvious next step. In relation to equipment, there are two important questions.

The first question relates to the building blocks for transmission. Current networks are based on 10 Gb/s building blocks, and there are serious doubts whether 40 Gb/s is an appropriate building block. The lack of router interfaces operating at this speed, the cost of such interfaces and the lack of interest from suppliers of DWDM hardware to make 40 Gb/s available all strengthen these doubts.

The other question relates to the cost of interfacing bit streams. As the cost of providing transmission has dropped, the relative amount of money spent on routing equipment has risen. Alternative, hybrid architectures, using a combination of switches and routers, are likely to offer a more cost-effective solution. This architectural approach will be better suited to deal with the emerging traffic patterns where individual applications generate very large point-to-point flows. However, this approach raises new technical and managerial challenges.

The ability to meet the growing network demands from the European research community depends on a number of factors, including in particular the way in which the market for telecommunications capacity will develop. There are also issues surrounding the development of user demands, in particular from those groups of users who are organised on a European basis and who might potentially implement their own solutions. In order to give some insight into these issues, three scenarios have been developed; they are described briefly in the final chapter of the current report.

1. Introduction

SERENATE is the name of a series of strategic studies into the future of research and education networking in Europe. The SERENATE (Study into European Research and Education Networking As Targeted by *e*Europe) project aims to contribute to European policies, social objectives and economic development by providing inputs on initiatives that could help to keep European research networking at the forefront of worldwide development. The objective is to provide important inputs to the development of policies by the European Commission, but also to national governments and funding bodies, the management of universities and research institutions, and the National Research and Education Networks (NRENs).

The current situation is that European NRENs and the wider European research networking community are at the forefront of developments. While much of the history of European research networking over the past two decades was characterised by the need to keep up with developments in North America, Europe currently has a leading position in many aspects of research networking. Gigabit networks are being implemented by a number of NRENs, and in other countries plans for such networks are being elaborated. At the European level, GÉANT, the network interconnecting the national research networks of the various European countries, has been a significant step forward, introducing 10 Gb/s in the core of the network and offering a wide coverage of 2.5 Gb/s capacity. More ambitious longer-term numerical targets may now be appropriate. Similar developments are to be expected at the national and local levels of research networking.

SERENATE contributes to achieving these networking goals by investigating the strategic aspects of the development of such "superfast" networks, looking into the technical, organisational and financial aspects, the market conditions and the regulatory environment. As a result, by the end of the project, the relevant policy makers, funders and managers of research networks in Europe will have at their disposal a set of recommendations and background materials that will enable them to set their policies for the further development of European research networking.

SERENATE is funded by the European Commission as an Accompanying Measure in the Information Society Technologies programme of the Fifth Framework Programme for Research and Technological Development. The project consists of fourteen interrelated work items, each looking into an aspect that is of strategic importance for the evolution of European research networking. The findings and conclusions from each of the work items are presented in a publicly available SERENATE report.

In the final phase of the SERENATE project, the SERENATE Steering Committee has developed consolidated views on the various developments that have been investigated in the SERENATE studies. These views are described in three reports.

The first of these reports is SERENATE deliverable D14 "Report discussing future scenarios for the funding of network infrastructure in the European research networking community, and of related costs". That report is based on the work items whose results have been described in the six reports listed below, and gives their synthesis:

- Deliverable D3 "Report on the experience of various communities that have experimented with "alternative" models of infrastructures" presents a number of case studies of "customer-controlled" networks in various places around the world.

- Deliverable D4 "Report on workshop on operators' views on infrastructure and likely evolution" presents the views of traditional and alternative providers of telecommunications and network services on the status of the pan-European telecommunications and network infrastructures and their likely evolution.

- Deliverable D6 "Report on present status of international connectivity in Europe and to other continents" provides an overview of the current market for international connectivity in Europe.

- Deliverable D7 "Report on the expected development of the regulatory situation in European countries relevant for the SERENATE project" describes the regulatory framework under which the research and education networks in Europe will operate in the future.

- Deliverable D8 "Report on workshop on National Research and Education Network models" reports on a workshop in February 2003, where European NREN managers discussed the progress of the SERENATE studies and in particular SERENATE's findings about the opportunities for research network organisations to obtain some form of "ownership" of the infrastructures that they use.

- Deliverable D13 "Report on the expected evolution of international connectivity in Europe and to other continents over the next five years" provides an integration of the work that was reported in deliverables D4, D6 and D7, and presents forecasts of the market development and the price dynamics of the transport and infrastructure market.

The current report, which is the second SERENATE report with consolidated views, is based on deliverable D14 (and hence on the other six reports mentioned above), in combination with the results of two other SERENATE studies, which are described in deliverables D9 and D10:

- Deliverable D9 "Report on the availability and characteristics of equipment for next-generation networks" gives an overview of the technical developments that can be expected in the next five years, dealing with the transmission components, the switching/routing components and the changes that might be expected in network architecture as a consequence.

- Deliverable D10 "Report on the networking needs of users in the European research community" presents the results of a study into the network service requirements of European researchers, which was conducted among others through a questionnaire and a workshop.

The current report integrates results from all the publications mentioned above, and describes a number of scenarios for the future evolution of the pan-European infrastructure for research and education networking.

It should be noted that the present report is still only an intermediate step in the SERENATE studies. The final report[1] of the SERENATE project, which is the third report with consolidated views, is based on the current report, the conclusions from a number of case studies[2] on the inclusion of user groups outside research and higher education, the results of a study[3] into the "digital divide" in European research networking, and the feedback received at SERENATE's final workshop[4].

[1] SERENATE deliverable D21 "Summary report on the SERENATE studies"
[2] see SERENATE deliverable D15 "Report on examples of extension of research networks to education and other user communities"
[3] see SERENATE deliverable D16 "Report identifying issues related to the geographic coverage of European research and education networking"
[4] see SERENATE deliverable D19 "Report on Final Workshop results"

2. Structure and Organisation of European Research Networking

2.1. The Organisational Structure

Network services that are provided to users in the European research and education community are organised at various levels with different geographical domains. These levels are managed by separate organisations.

The level closest to the researcher, teacher or student is the Local Area Network (LAN) at the site (for example, the university campus) where they work. This level is the responsibility of the organisation (i.e., the research institute, university, college, school etc.) that runs the campus.

The next level is the national research and education network that provides the connectivity between the local networks of research and higher-education institutions in a country. This level is the responsibility of the National Research and Education Network organisation of that country. In some countries campuses are not directly connected to the national network but via Metropolitan Area Networks (MANs) or regional networks.

The third level of the research and education network infrastructure provides international connectivity between researchers in Europe. It is essentially provided by the GÉANT network, which interconnects the national research and education networks. GÉANT is managed by DANTE on behalf of the NRENs.

Connectivity to research networks in other continents is either obtained directly by the NREN or provided by DANTE via GÉANT. The same holds for connectivity to the commercial Internet.

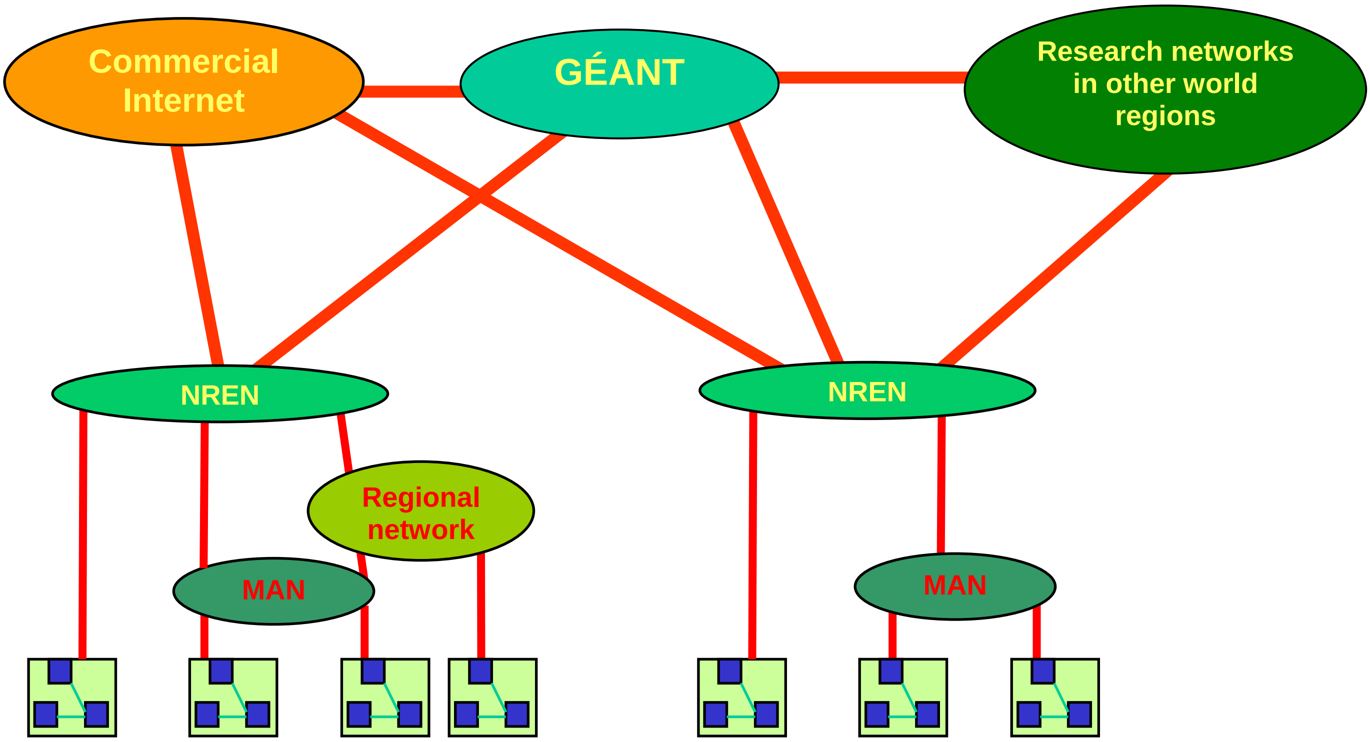

Figure 1. Structure of local, national and international research networking

Alternative structures could be envisaged. One of the options would be to create a number of regional networks, each encompassing a (small) number of countries, which when interconnected would serve as a simpler trans-European network, with less connection points than the current one. NORDUnet in the Nordic countries provides an example of such a regional network. Another alternative approach could be to replace the trans-European backbone network by direct interconnections between national research and education networks in neighbouring countries. The latter option could in theory reduce costs as international connections could be established between nodes in neighbouring countries close to the border. Border hopping, i.e., combining this border-crossing concept with some form of ownership of the international fibre/wavelength/capacity, could be attractive for countries where market prices for international connectivity are very high, and in situations where the distance between nodes connected to two different national research networks is relatively short.

Border hopping will not replace the three-level structure of European research networks. However, border hopping might be a suitable solution to establish direct interconnection between research institutions in neighbouring countries with special communication needs, as a complement to the trans-European network.

Because of the many managerial, technical and organisational problems associated with the alternative structures mentioned above, they can be expected to be implemented only in a small number of special cases. The general organisational model of research networks in Europe, with different levels related to geography, a single NREN per country and close collaboration at the European level, can be expected to remain essentially unchanged, at least for the next 5-10 years.

It should be noted that this structure has been key to the success of European research networking over the past years. It is a model that is increasingly being copied in other continents, for example in the Asia-Pacific region and in Latin America.

Whilst it is recognised that the future challenges for research networking in Europe will require increased technical and human resources, it is the case that the current organisational structure represents an optimum approach in terms of funding sources and matching national and European objectives. For this reason, the continuation of that structure is a basic assumption in the remainder of this report.

2.2. Funding of Research Networks

In most European countries, educational establishments are funded almost entirely from public resources and research institutions are funded to a large extent from public resources. Therefore, in the end, research and education networks are basically financed by taxpayers' money. However, the paths along which the funds are channelled from the government to the responsible organisations may vary.

In most countries, funding of the local network facilities within an institution is considered to be the responsibility of the institution itself, as the LAN is considered part of the environment provided by the relevant research or education organisation.

In practice, there are large differences between countries in the way the national and international levels of research networking are funded, varying from total central funding to substantial funding via the connected institutions. The optimal solution clearly depends on national circumstances. A large majority of countries have a mixed system, which can work well if expenditures of long-term benefit are centrally funded and some of the services whose costs can be directly related to individual institutions are funded through those institutions.

As to the European level of research networking, GÉANT is partly funded by the European Commission, but the larger part of the GÉANT costs is covered from national contributions. Those costs are shared between the countries on the basis of an algorithm that has been agreed by the participating NRENs and can be updated by them as needed. One of the effects of the chosen algorithm is to smooth the very large differences between the costs of communication services to individual countries. We will discuss those differences in chapter 4.

The current approach to sharing the costs for GÉANT is rather simple. All connected networks share the complete set of network costs, based on the assumption that any individual NREN has the possibility to make use of all the network resources. There is no attempt to allocate costs on the basis of actual usage. This simple approach encourages coherence and stability, and limits the viability of "opportunistic bypass", i.e., the situation where an NREN would organise its own connectivity on those routes where this can be done cheaply, but would rely on the pan-European network for more difficult and expensive locations.

The current approach to cost sharing assumes that an annual subscription will be made for a specific capacity. That subscription is based on two factors:

- the underlying cost of international connectivity to a country
- the capacity subscribed for by the NREN in that country.

The cost of provision of service to a country is heavily dependent upon the cost of international connectivity to that country. Figure 2 compares the variation in the connectivity costs in GÉANT with the variation in subscriptions to GÉANT. When the cost of subscription to the network is compared with the underlying costs of providing service, it can be seen that the GÉANT cost-sharing approach significantly reduces the differences. This is the smoothing effect of the cost-sharing algorithm that was mentioned above.

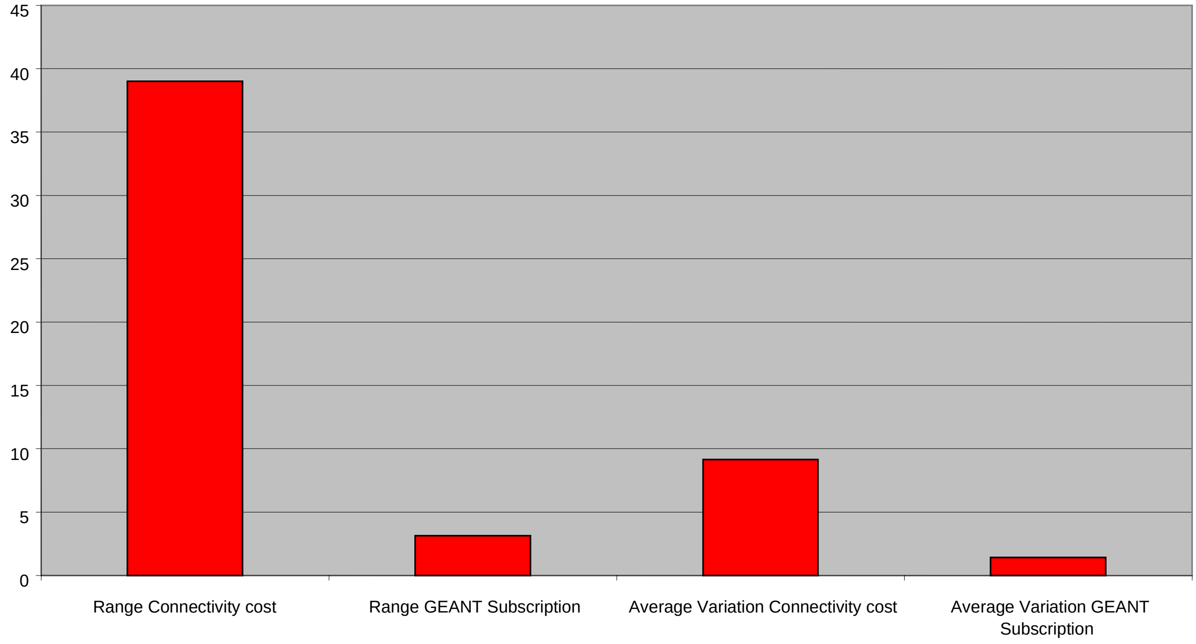

Figure 2. Variation of connectivity cost and of GÉANT subscription (average 1996-2003)

Figure 3 illustrates how this smoothing has developed in the period 1996-2002.

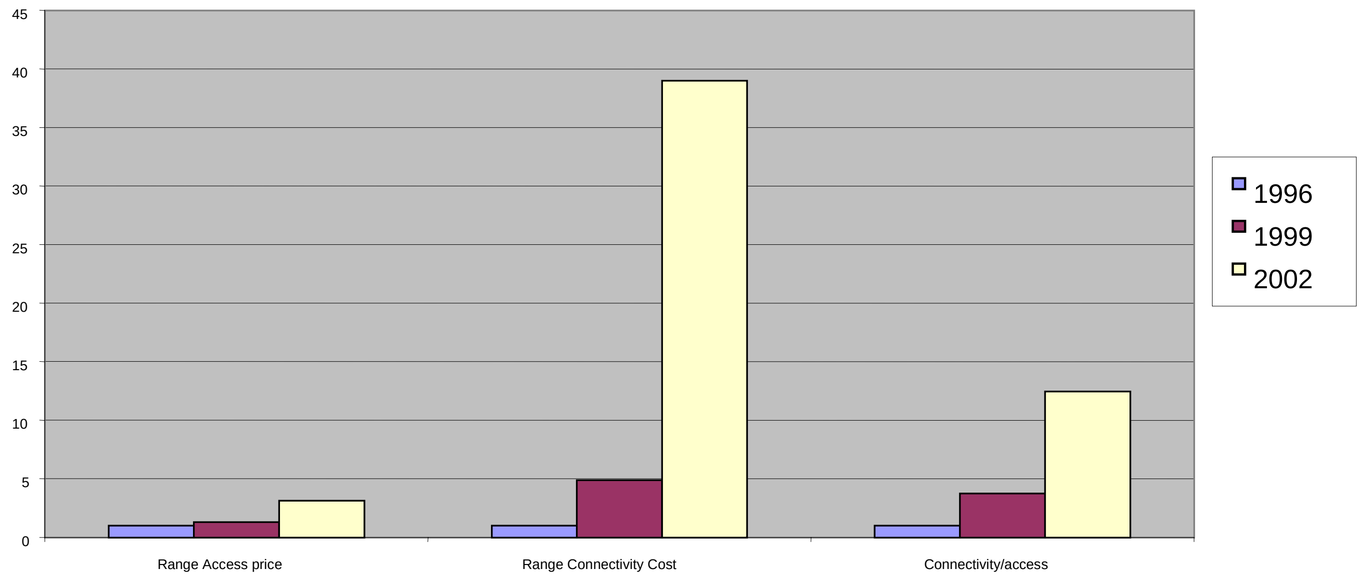

Figure 3. Development of cost sharing 1996-2002

The central set of bars in this figure shows how connectivity prices have diverged over this period, while the leftmost set of bars shows how prices for access to the pan-European research network have diverged. The right-hand set of bars illustrated the way in which cost sharing has increasingly smoothed this difference. While this approach is very positive from the point of view of European coherence, it has two negative side effects. Firstly, it reduces the political pressure to resolve the issue of market imperfections, because, to a very large extent, it reduces the effects and visibility of these imperfections. Secondly, and more importantly, it can make the network look expensive, particularly for those users who require only connectivity between a limited number of cheap locations. Historically, this has not been important, since most individual users consume small amounts of capacity when compared with the total network capacity. In chapter 3 we will describe the emergence of new applications that have significant capacity requirements compared with the network building blocks. As a consequence of those new applications, the current cost-sharing approach, which is based on a single cost for total access to all network resources, will become increasingly difficult to sustain.

3. Development of User Requirements

It has been an important part of the SERENATE studies to investigate the future networking needs of members of the European research community. There is impressive evidence of growing network requirements from all areas of research. These needs will grow dramatically over the next 5-10 years, in all disciplines and in all countries. At the same time as many of the natural sciences are pushing towards a very broad deployment of Grid computing, there are compelling examples of how research in the humanities could benefit greatly from advanced networking, while the aspirations of, for example, social scientists, ecologists, musicologists and geographers are also very challenging.

At the SERENATE workshops and in their replies to SERENATE questionnaires, researchers have given examples where their research would become much more efficient if network capacities could be increased by one or two orders of magnitude. Examples were also given where there is a possibility of starting completely new research activities that were prohibited until now by lack of very high performance network facilities.

The SERENATE studies found a remarkable interest and involvement of researchers in Grid computing. Such involvement was reported from astrophysics and astronomy, particle physics, computer science, earth sciences and oceanography, protein modelling, photonics, chemistry and many other disciplines.

The expectations of network users have evolved beyond the provision of pure bandwidth towards the supply of more complex services. There are concerns about security, privacy and confidentiality. There is likely to be a strong demand for authentication and authorisation services in the research and education area. There will be a growing demand for researchers to be able to access networks and their own usual set of network and information services from wherever they happen to be. There is general pressure from end-users that research and education network organisations should give more attention to the end-to-end aspects of communication, including issues related to quality of service.

Recently, there have been cases where a single instance of a new application, such as a Grid file transfer, a remote immersive virtual-reality session or the transmission of very high definition images, has exceeded the aggregate flow that one usually sees from a whole country with thousands of simultaneous users. Indeed, the recent availability of very large transmission capacities, at much reduced process, will enable co-operative European research projects that require significant bandwidth for their own use between a limited number of locations, to implement a networking solution to support their needs. In the past, the cost of transmission precluded the organisation of such distributed projects on a European basis. GÉANT and its predecessors only supported the aggregate demands of users connected to the national networks, with no individual user representing more than a small fraction of the total traffic.

The development of European projects in which individual sites are capable of generating as much traffic today as the historic aggregate demand from a whole country, will lead to significant new challenges in meeting user requirements. A number of marketing, technical and commercial issues arise. In an environment of a large population of individual users, where each user represents only a small element of network demand, it is not relevant to try and address individual user needs. However, when groups of users require significant network resources, it is necessary and appropriate to try and understand the group requirement and to seek to address it directly. Understanding user needs is not simply related to providing performant networking. To achieve better understanding, it is necessary to have a segmentation of the user base, reflecting their needs and requirements. It is apparent that certain users are highly demanding in terms of their networking needs. This is typically expressed in terms of performance, quality or security requirements. In the future, it will be necessary to have a more precise view of who these users are, and how they can be supported best if they are to be encouraged to fulfil their networking needs using a solution based on the shared facilities of the NRENs, as opposed to a separate solution dedicated exclusively to their own needs.

An important future activity for the European research networking community is to develop a market segmentation of the users and to use it to target and improve support to users. The precise form of segmentation is not defined yet, but it is likely to group users into three or four categories, ranging from those whose needs can be met without any special support, to those whose requirements are such that specific engineering may be needed to meet them. The latter group is recognisable today to a reasonable extent, but is not clearly defined. The networking community needs to obtain a much more precise view of the users that constitute this latter group. There are two reasons for this. Firstly, by definition, the user group requires technical solutions that are to some extent tailored. Secondly, the costs of providing them with connectivity are likely to be significant. These users will therefore carefully look at the cost of the service. User groups with a defined and limited topological requirement will look at the direct cost of implementing connectivity to meet their needs and compare it with a shared-cost solution based on NREN costs. If the NRENs directly translate the pan-European costs in their charges to users, then that may deter demanding users from taking advantage of shared network solutions.

4. Consequences of Developments in the Telecommunications Markets

The liberalisation of international telecommunications in Europe in 1999, as a result of an initiative of the European Commission, had a dramatic effect on the market for connectivity. Prior to liberalisation, the market was best characterised as one where rationing prevailed; prices were kept artificially high; access to high-speed connections was severely limited. This protected the lucrative, monopolistic, international switched-voice market from bypass, since it was very expensive to acquire the international connectivity. Capacities were kept artificially low to limit the bypass opportunities. More importantly, from the point of view of pan-European research networking, the market acted as a serious inhibitor to technical innovation and the provision of advanced services.

Liberalisation changed all this as far as pan-European research networking is concerned. Figure 4 illustrates the general development of prices of connections used to build the pan-European network, over a five-year period. By the year 2000, a new generation of competitive, trans-European operators had started to make state-of-the-art technology available. The prices charged by these operators for international connections between major centres of population fell dramatically across much of the geography of western and central Europe. As a result, the cost of capacity expressed in terms of euro per Mb/s per year[5] fell by a factor of 6,000 in a period of three years. At the same time, the connections made available jumped from 34 Mb/s circuits to 10 Gb/s circuits. Connections with larger capacity will generally be cheaper in relative terms than connections with smaller capacity, because there are some economies of scale in the provision of such connections. However, even allowing for this, the real reduction in cost of connectivity was a factor of 200. For research networks, this is an unprecedented change in underlying costs and it has facilitated major changes in the way that research networking can be exploited at a pan-European scale, notably by:

- allowing access to the most advanced transmission technology, which has enabled the provision of very high speed network connections to the most demanding scientific applications
- making it economically feasible to distribute co-operative research and development activities across Europe, relying upon cost-effective network communications to enable truly pan-European co-operation.

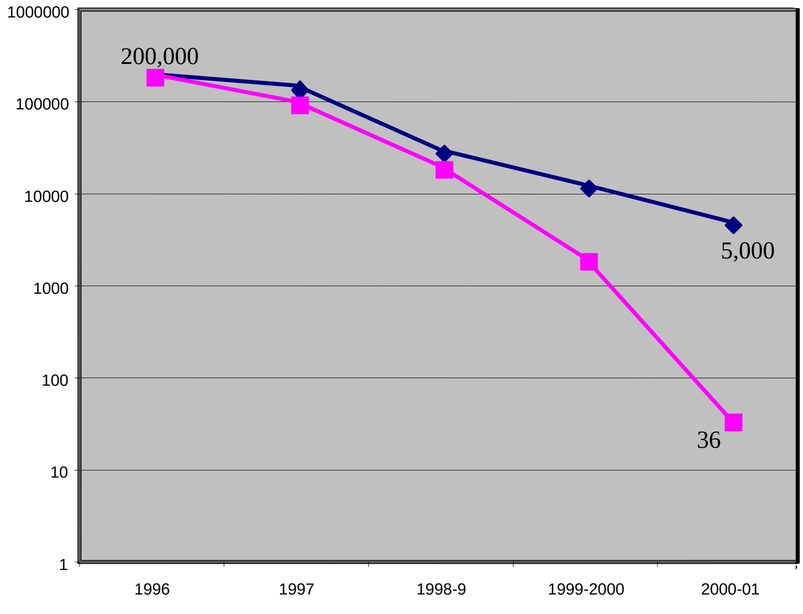

Figure 4. Evolution of market competitiveness: international intra-European connectivity

The graph shows the development of international connectivity prices in the period 1996-2001 as measured by the response to tenders for GÉANT and its predecessors. Prices are expressed in the simple measure of euro per Mb/s per year. The upper line shows the average offer prices and the lower line shows the lowest offer prices. Note that the graph is on a logarithmic scale!

[5] This is the annual lease cost of the connection in euro per year, divided by the capacity of the international connection expressed in terms of its speed of operation in Megabits per second.

These developments have been extremely positive for European research networking. Nevertheless, when looking several years ahead, there are a number of factors that will potentially constrain the very rapid progress that has been made in the last four years. These are discussed in some detail below. Although the analysis is based primarily on the intra-European international cost data, many of the conclusions are, in the context of European research networking, equally applicable at a national level and even at a local level.

Although liberalisation has had a dramatic effect on the European telecommunications market, the effects have not been uniform across Europe. As far as the prices for international connections are concerned, this is illustrated by the divergence of the two lines in Figure 4. The very large differences that exist between the most effective and least effective markets have created a significant variation in cost for provision of the same service. The geographic variation, in market efficiency and competitiveness, presents a real barrier against the development of homogeneous service in Europe. Although the regulatory environment across Europe is becoming increasingly harmonised and allows research network organisations to implement their own infrastructure where that is cost effective, there remain informal barriers to the creation of a fully liberalised market, both nationally and internationally, within Europe.

Historically the international component of networking in Europe has always been the weakest link, with slower connections than those available in national networks and on campuses. The change in those international telecommunications markets in Europe where liberalisation has taken place, has been extremely rapid. Where international connections have become much more cost effective, national, regional and campus networks have often not managed to keep up. These days, capacity limitations at campus and national level are more likely to create bottlenecks. There remains a need to ensure that the end-to-end service experienced by users is independent of their location in Europe. In addition, as described in chapter 3, as performance improves, there are emerging groups of users whose requirements for service are significant when compared with aggregate usage of international connections by a country today.

In the current report, a number of scenarios will be formulated that represent potential future developments of the technical and market environments, which will influence the evolution of research networking in Europe. In developing these scenarios, the following assumptions have been made:

1. The NRENs and their common pan-European and global interconnection facilities are the networking basis for international co-operation in the fields of science, the arts and the humanities.

2. These facilities shall support both the generality of users and those specific groups of users whose requirements are, in themselves, significant on a pan-European or global basis.

3. These facilities shall form a major element of the European Research Area.

4. European researchers shall have appropriate access, independent of the location in Europe where they are situated.

5. There is an expectation that groups of users, organised on a European basis, whose network usage is significant compared with general aggregate usage, will be requested at some point to recognise the real cost of the networking facilities that they use.

The main element of the costs for pan-European research networking today is that of connectivity. Therefore, the development of connectivity prices in the next five years will have a significant effect on the overall budget required to support the service provision. Although, as noted above, the general trend in the development of prices has been very positive, this overall trend masks a number of factors that need to be considered in scenarios for the future.

The substantial reduction in price has not been uniform across Europe. Figure 5 illustrates this. It plots the relative cost of international connections in relation to the number of suppliers offering connections to a country. It can be seen that a very large divide has opened up between the most competitive locations in Europe and other, more expensive routes. As a consequence, the cheapest European connection to a country is a factor of 40 cheaper than the most expensive one. The principal reason for the opening of this divide is that competition has developed more effectively in some areas of Europe than in others.

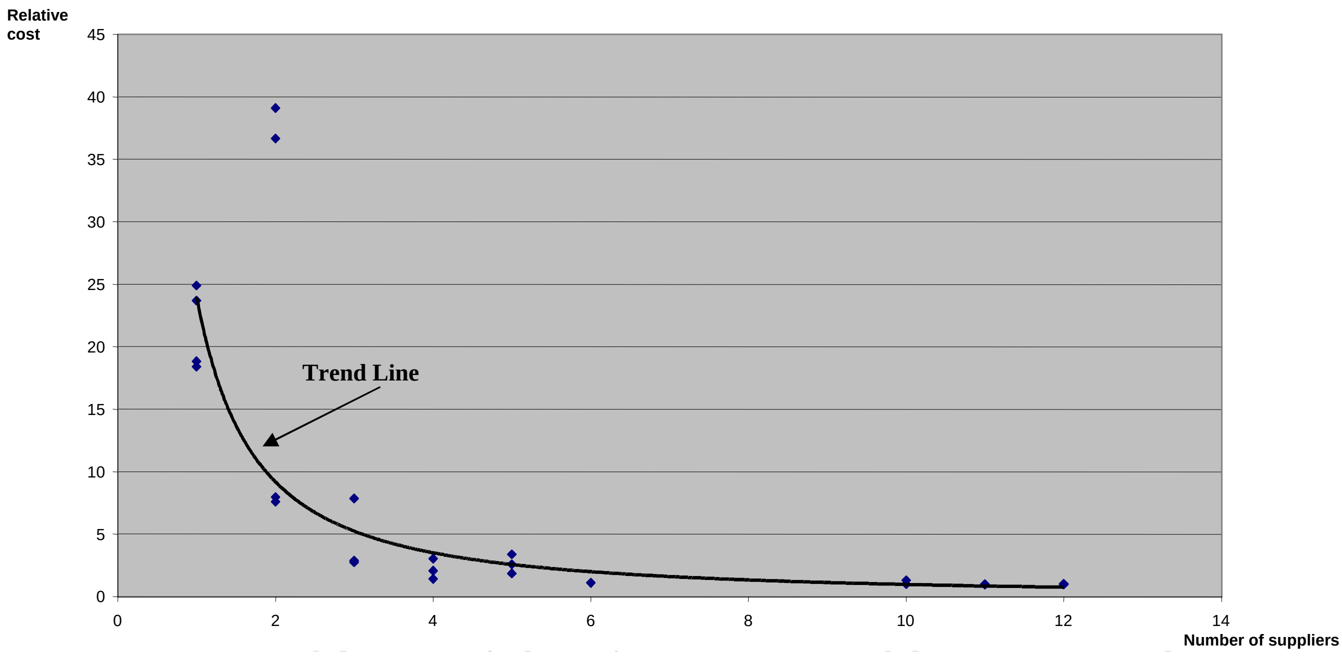

Figure 5. Relative cost of connectivity compared with number of suppliers – GÉANT tender data 2000-2001

The figure plots the relative cost of connectivity in response to the GÉANT tender when compared with the number of suppliers offering international connectivity. Each point represents the relative cost of international connectivity to a country, with the cheapest connectivity being normalised to 1. The relative costs have been normalised to take account of the fact that slower-speed connections are relatively more expensive than higher-speed connections. The trend line represents the best fit among these data points, excluding the two outlying data points at the top left-hand corner of the graph.

Another important point is that the very large reduction in costs represents something of a one-off effect. As the market became more competitive, there was a major re-adjustment from monopoly-based prices to prices that are more closely aligned with costs. Once this has happened, prices will decline more slowly, driven by increasing demand and technical developments.

There are three factors that define connectivity costs. The first is the speed of operation of the transmission systems. In general, Dense Wavelength Division Multiplexing (DWDM) systems today are based on 2.5 Gb/s or 10 Gb/s channels. Secondly, there is the distance between the end points. DWDM costs are heavily distance dependent. Historically, distance has not been an important factor in the prices charged for connectivity. But the distance-dependent character of DWDM will become a feature as competition increases and prices align more with underlying costs. Finally, there is the volume of capacity on a given route. The initial investment in a DWDM system is capable of supporting many wavelengths. The marginal cost of adding wavelengths is very small, and the average cost of a wavelength therefore varies, depending on the extent to which the system has already been partially filled. This is of considerable benefit for routes where there is a significant level of installed capacity (for example, London-Paris). However, those countries where market demand will be more limited, and international routes therefore have less installed capacity, will be disadvantaged.

Where the market is competitive, prices will be based on underlying costs. The two important factors that determine the cost are the route length and the level of demand for wavelengths. Figure 6 indicates the relation between route length, demand for wavelengths and connectivity costs for DWDM systems today. It can be seen from this figure that distance is an important factor. Therefore, that factor should have very significant influence on the topology of future networks. Even in a competitive market environment, it will mean that locations further away from the networking centre of Europe will have higher costs for their connections. And there will also be higher costs for countries where there is limited demand for international connectivity.

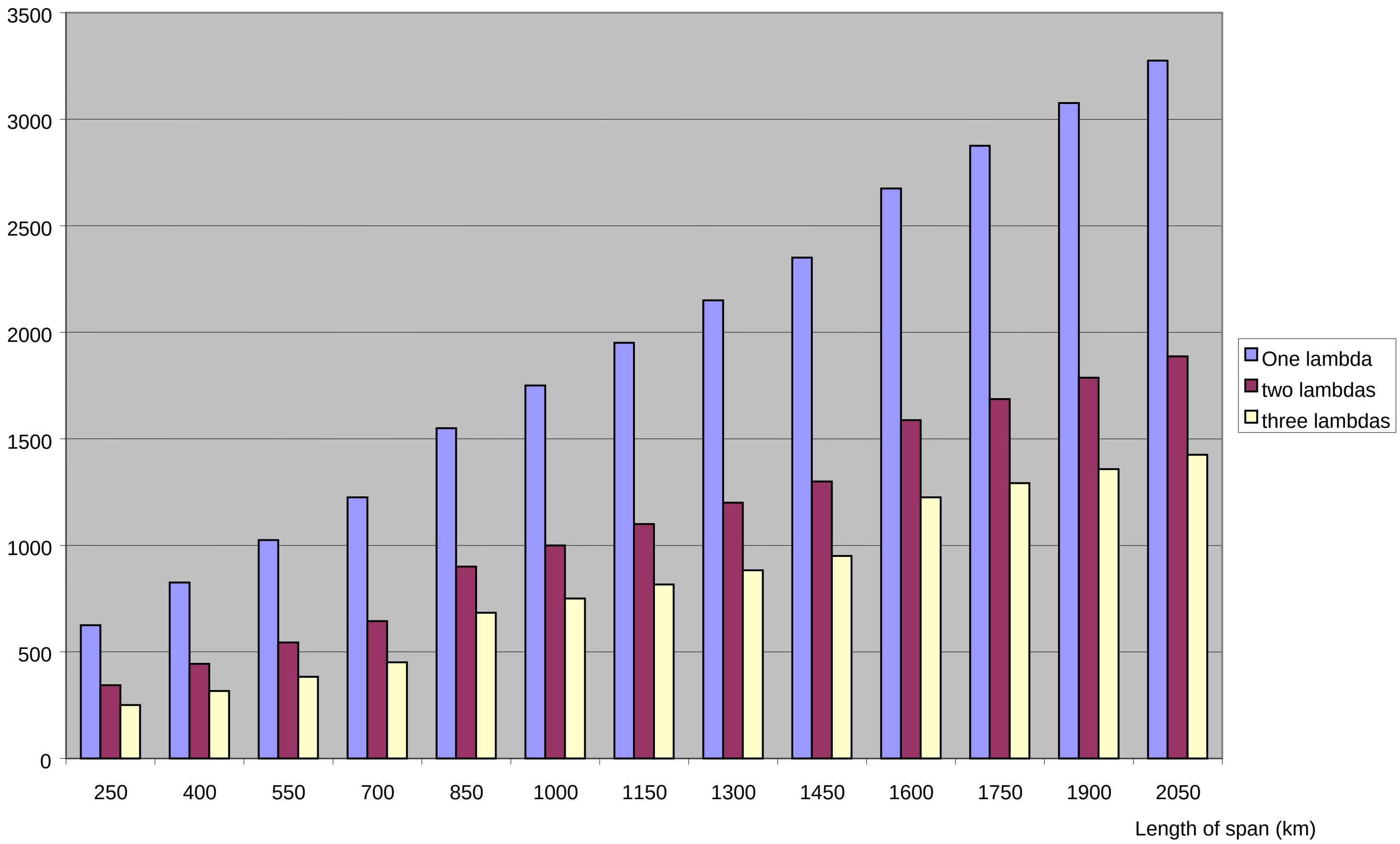

Figure 6. DWDM costs (investment cost in euro per lambda per year) versus route length (in km)

The SERENATE report on the present status of international connectivity[6] showed the current relative state of liberalisation of the European markets for international connectivity on the basis of an analysis of the GÉANT tenders. It indicated large variations in the competitiveness of these markets, with only ten countries having a really competitive market for international connectivity. Subsequent re-tendering activities for GÉANT have demonstrated some improvement in competitiveness, particularly in respect of Portugal and Slovenia. Although these locations remain relatively expensive, the new offers represent nevertheless a significant development of market competitiveness.

In order to understand how the cost of connectivity might develop in future, further analysis has been carried out.

A very basic indicator of competitiveness is the willingness of suppliers to offer DWDM connections. At the end of 2002, it was still not possible to acquire DWDM connections on 16 out of the 37 routes within the GÉANT network. These routes are very uncompetitive. Their suppliers are not prepared to allow access to the DWDM technology on which the connections themselves are based.

In order to determine the extent to which current GÉANT connectivity prices for DWDM reflect underlying costs, further analysis was done on the 21 routes where DWDM is available. This analysis was based on the distance matrix for GÉANT, which contains the route lengths of the circuits deployed in GÉANT today. Figures 7 and 8 show the relative costs, on a per-route basis, of DWDM connections in GÉANT, for those routes where DWDM is available. This is based on the assumption, illustrated by Figure 6, that connectivity costs are distance dependent. The annual price per route has been divided by the route length. The results are plotted for each route relative to the cheapest route, which has a relative cost of 1. The analysis is done separately for 2.5 Gb/s routes and for 10 Gb/s routes. It is interesting to see that the variation of wavelength

[6] SERENATE deliverable D6 "Report on present status of international connectivity in Europe and to other continents". That report is also the source of Figure 4 and Figure 5.

prices is significant and that the variation is smaller for 10 Gb/s routes than for 2.5 Gb/s routes. The latter point can be explained by the fact that the 10 Gb/s routes are generally between more competitive locations.

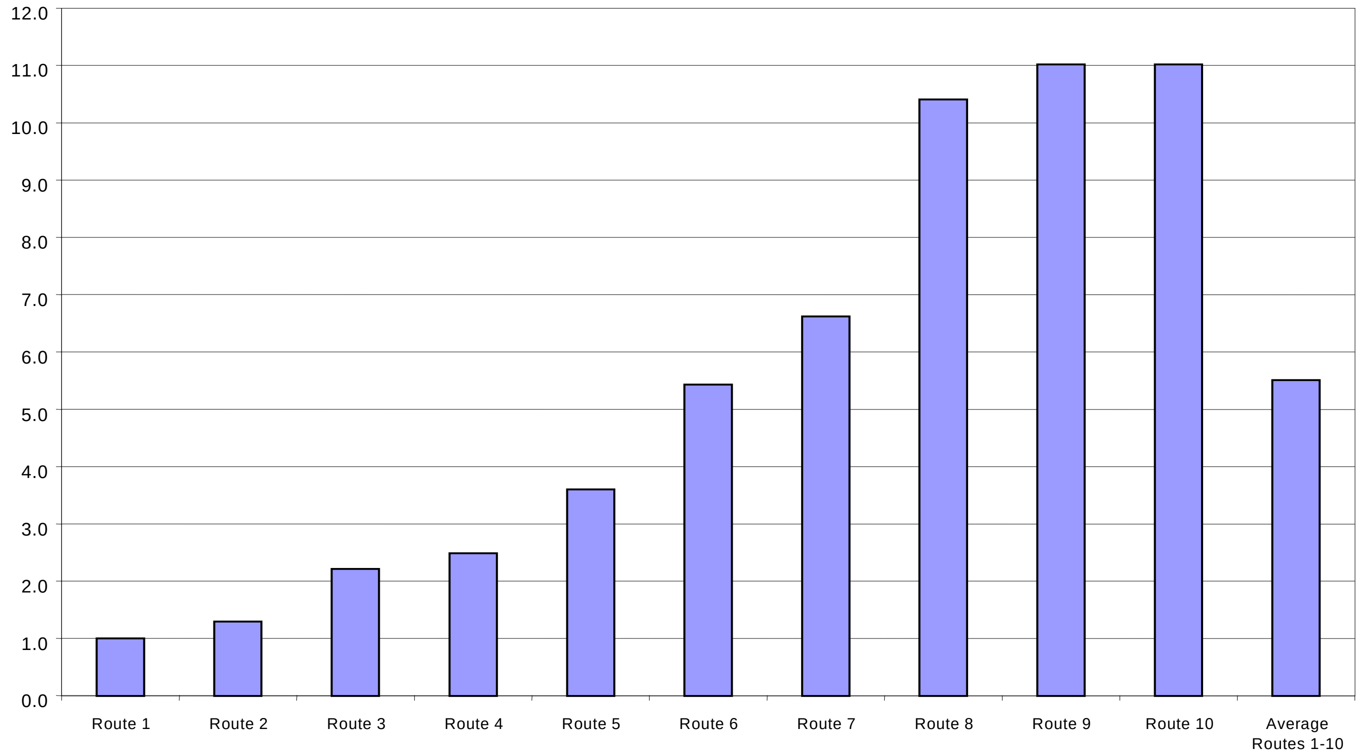

Figure 7. Relative costs of 2.5 Gb/s routes per km

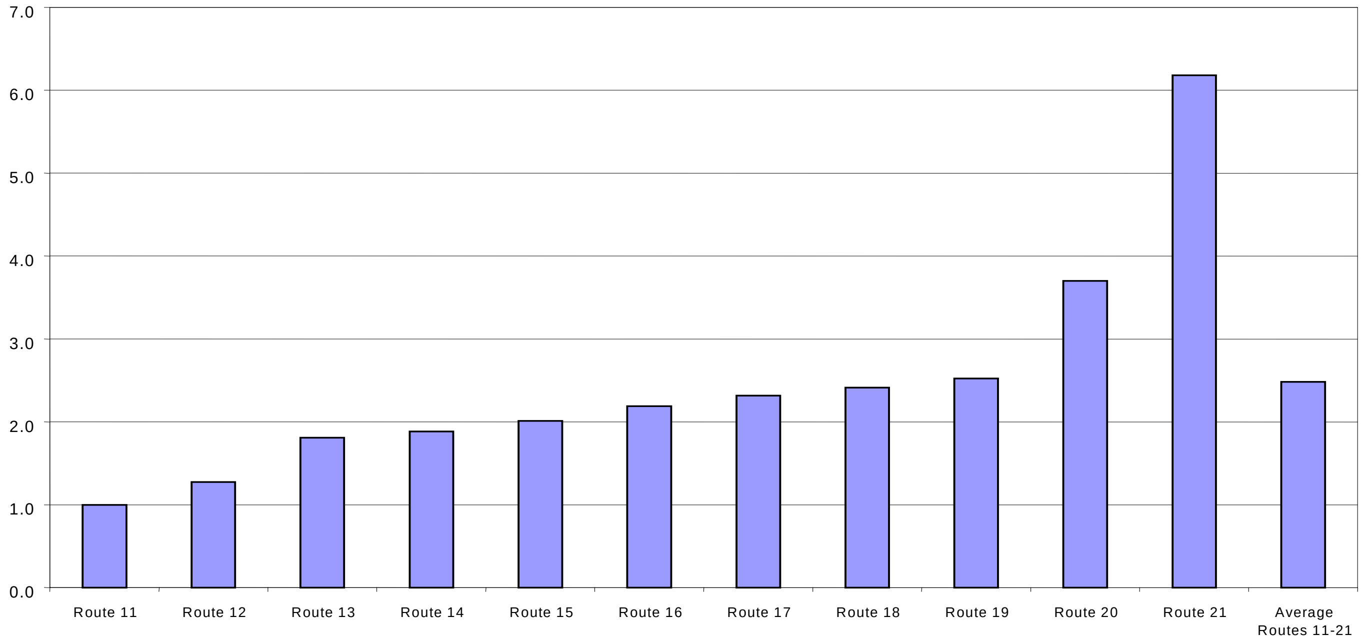

Figure 8. Relative costs of 10 Gb/s routes per km

The analysis of the relation between cost and distance for GÉANT shows very considerable variation in the underlying price of wavelength capacity. When allowing for the fact that longer spans will be more expensive, the variation in the underlying price of wavelengths is a factor of 11 between the cheapest and the most expensive wavelength when expressed as a cost per year, per kilometre (i.e., factoring out the distance element). The current total cost in GÉANT for wavelength capacity is approximately 14 million euro per year. Assuming that all GÉANT wavelengths were provided at that cheapest per kilometre cost, the total cost of this capacity would be approximately 4 million euro per year. This is a very significant difference and is a further indication of the relative level of competition, within the wavelength market, as far as GÉANT is concerned. Bearing in mind that the SDH[7] capacity in GÉANT, which consists of 3 Gb/s, costs 16 million euro, the potential opportunity for network growth, if all transmission capacity were available at the most cost-effective rates, is extremely significant.

There is likely to be a gradual improvement in competitiveness, which will be of benefit to locations in Europe that are currently more expensive. However, as the geography of GÉANT expands, it is probable that the newly connected countries will have less liberal markets, and therefore the price problem is likely to continue. Moreover, for some countries that are currently already connected, most notably Greece, price remains a very real problem in the current situation.

[7] Synchronous Digital Hierarchy

5. Consequences of Technical Developments

The SERENATE report on equipment for next-generation networks[8] has provided an extensive analysis of the likely technological developments in the next five years. Historically, it has always been assumed that it will be a logical and cost-effective network development to deploy higher transmission speeds and to continue using a router-based network. Now, it is not at all certain whether this will be the obvious next step as far as pan-European research networking is concerned. The SERENATE report on equipment raises two important questions.

The first question relates to the building blocks for transmission. Current networks are based on 10 Gb/s connectivity building blocks. There are serious doubts whether 40 Gb/s is an appropriate building block. The lack of router interfaces operating at this speed, the cost of such interfaces and the lack of interest from suppliers of DWDM hardware to make 40 Gb/s available all strengthen these doubts.

The other question relates to the cost of interfacing bit streams. As the cost of providing transmission has dropped, the relative amount of money spent on routing and switching equipment has risen.

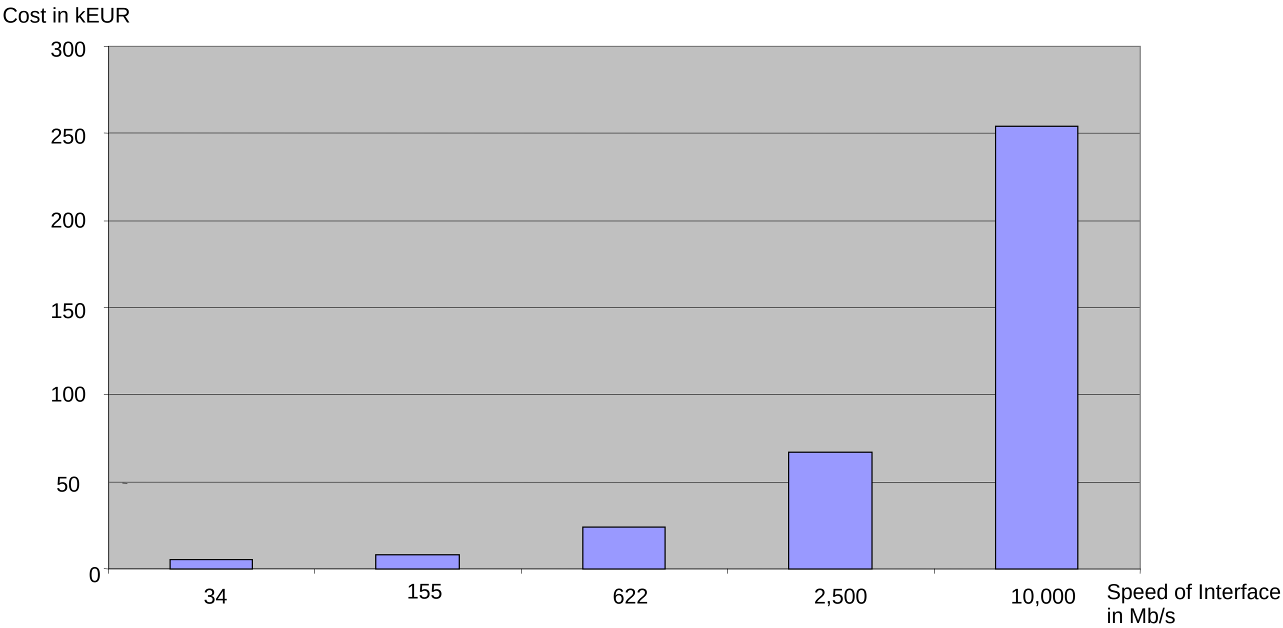

Figure 9. Interface costs

Figure 9 plots the relation between interface costs (based on list prices) for a router port and speed of operation of the port. It illustrates the very significant increase in the cost of interfacing circuits to a router as speeds become higher. Striking differences emerge when one compares the cost evolution of interfaces with the cost evolution of connectivity when speed of operation increases. These are shown in Figure 10. Whereas the relative cost of connectivity declines with increasing speed, thereby yielding economies of scale, this is not true for router interfaces. In general, the figure even understates the divergence, because it is based on data for SDH connectivity. In comparison, DWDM connectivity is cheaper per unit of capacity than SDH (i.e., 2.5 Gb/s DWDM will cost relatively less than 2.5 Gb/s SDH), so that for DWDM the economies of scale will actually be greater than illustrated by Figure 10.

As the percentage of total network costs that is spent on hardware increases, the impact of the development of hardware costs on the development of total network costs will grow. This will be true particularly in a pan-European context, as is demonstrated by the analysis in chapter 4. That analysis showed continuing scope for reduction in connectivity costs in GÉANT, which are likely to be translated into connections with higher speeds, thereby increasing the percentage of expenditure on hardware. In the cheapest part of the European network, the capital cost of interfacing wavelengths to routers is greater than the annual cost of operation of the wavelength.

[8] SERENATE deliverable D9 "Report on the availability and characteristics of equipment for next-generation networks"

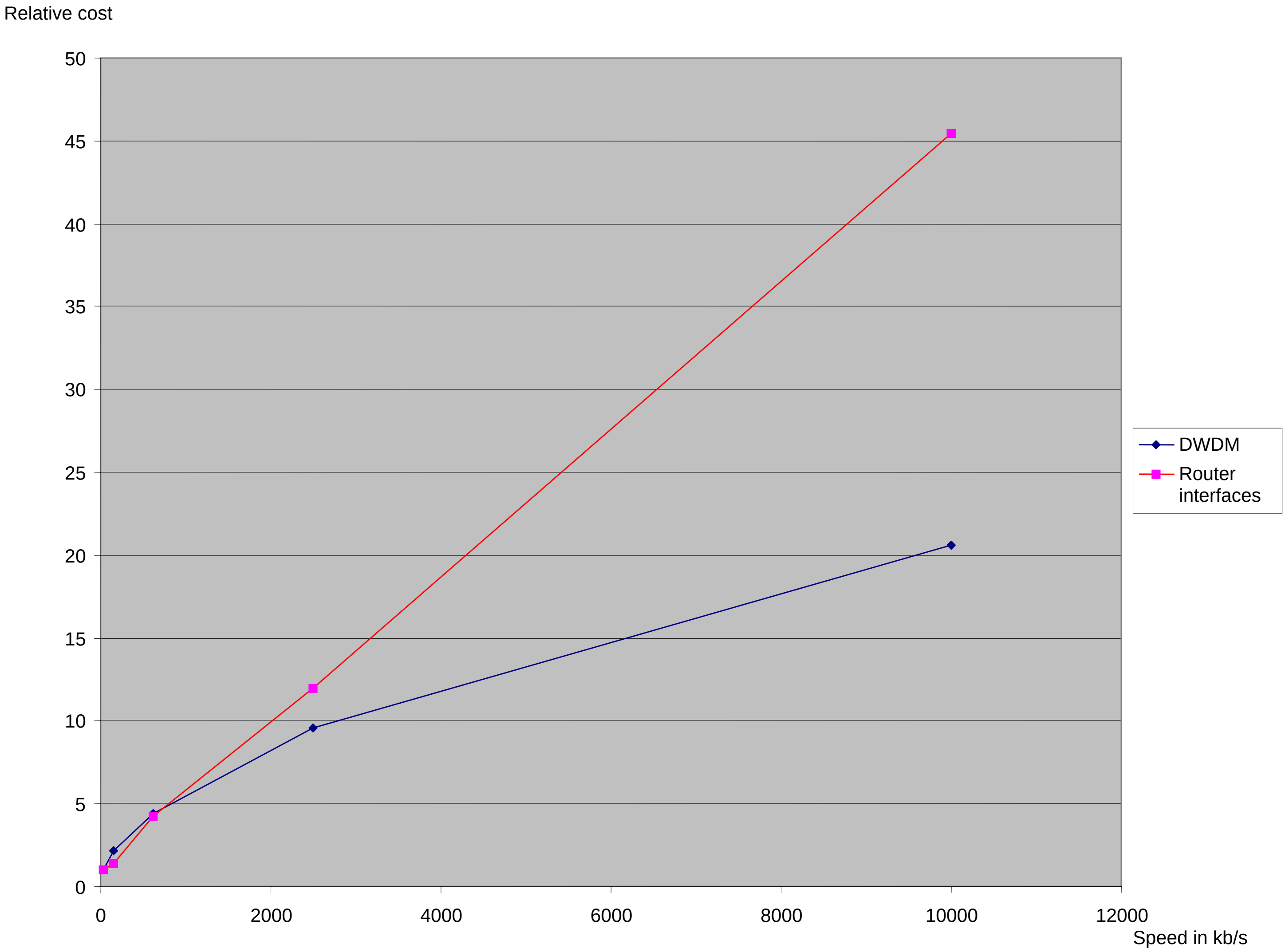

Figure 10. Relative cost of DWDM connectivity and router interfaces with speed of operations

Still higher speeds of operation (40 Gb/s or more) could lead to further significant increases in the cost of interfaces for routers.

A major issue to be resolved is the architecture and the structure of the next-generation networks. In particular, the balance between routing and switching is one that needs to be addressed. The current and previous generations of research networks have been based on simple IP services, using a routed architecture, and operating at increasing speeds. The emergence of a new class of applications that may have bandwidth requirements equivalent to the basic transmission building blocks of the networks mean that it is relevant to look at alternative architectures. These applications generate point-to-point bit streams. Therefore, they do not require per-packet routing decisions. The entire stream can be switched between end-points. The fact that interfacing bit streams to switches is significantly cheaper than interfacing them to routers re-emphasises the possible viability of alternative architectures. Figure 11, which compares the costs of optical cross-connect switch interfaces with those of router interfaces at different speeds, shows the savings that can potentially be made by using switches.

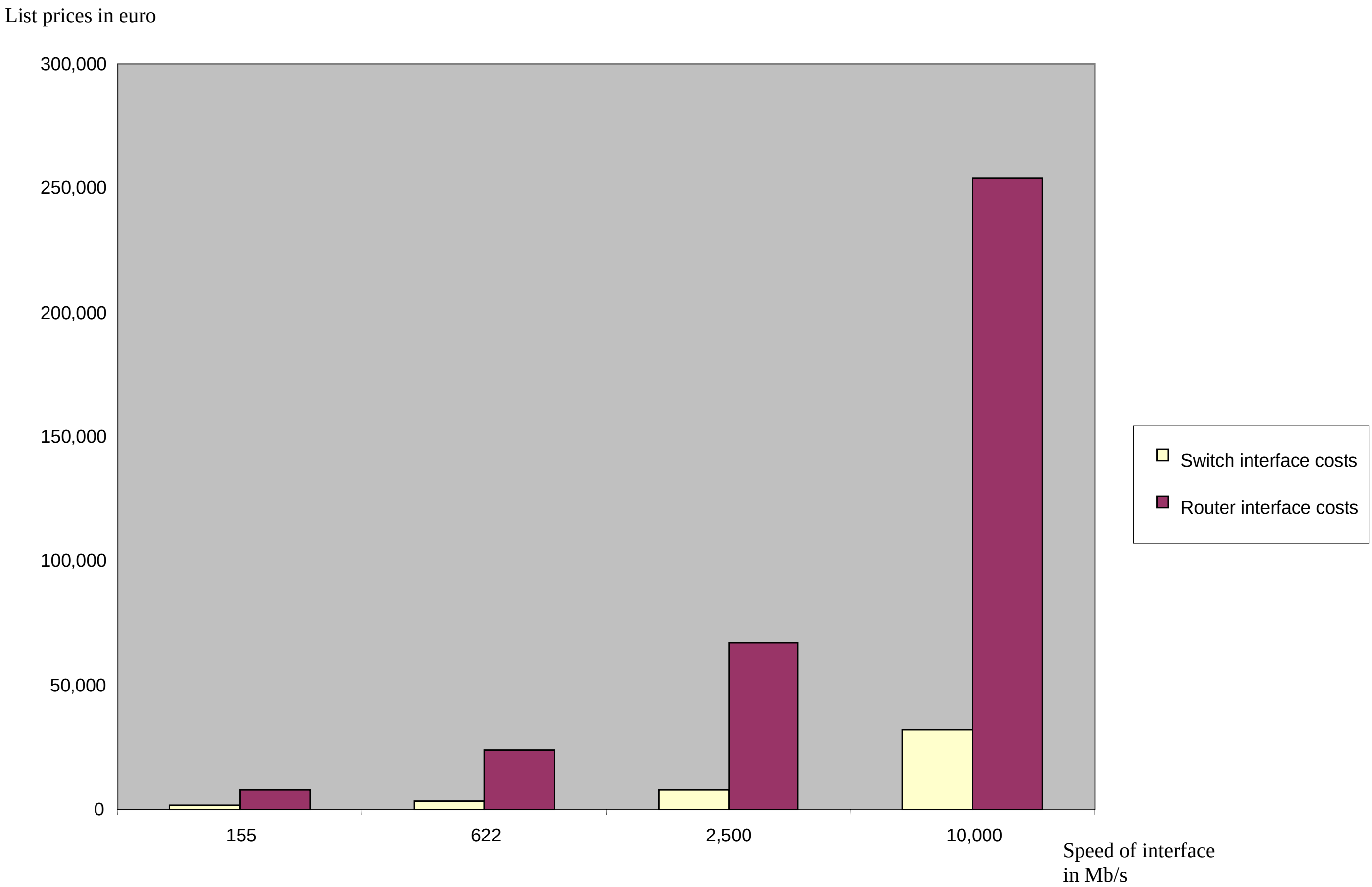

Figure 11. Relative cost of switch interfaces and router interfaces

However, progress towards new architectures is also heavily dependent on the ability to control the various kinds of equipment. It is this area of network management where the most significant technical questions arise. A switched architecture dedicates network resources to an individual user. This resource allocation is relatively easy to manage within a single management domain. However, it is much more complicated to manage this process across an environment of multiple management domains, which is typical of research networking in Europe. By contrast, a routed architecture does not dedicate resources and is, in this sense, simpler to manage across multiple management domains.

6. Conclusions

On the basis of the analysis in the previous chapters, a number of conclusions can be formulated:

1. The liberalisation of telecommunications markets in Europe has been very beneficial for European research networking. It has provided access to very high speed connectivity and to leading technology. The very significant reduction in prices in much of Europe has changed the underlying economics sufficiently to enable the deployment of applications that require very high speed connections. There, network capacity requirements of an individual application are significant when compared with the aggregate traffic that is exchanged currently between two European countries.

2. Although there has been significant reduction in telecommunications prices across Europe, progress has not been uniform. In a period of four years, we have moved from a position where prices were independent of location in Europe to one where the difference between the cheapest and the most expensive locations is a factor of nearly 40. In the more expensive countries, access to advanced transmission technology is also severely limited.

3. Even in the more competitive areas of Europe, there is still significant variation in the costs of connectivity. If all connections were available at the most cost-effective rates, then there would be considerable scope for increasing the pan-European research network capacity. In the absence of stronger liberalising forces this is unlikely to happen. A gradual improvement in competitiveness is a more probable development.

4. Historically, European research networking has been based on Internet technology making use of routers. The development path has been to increase the speed of network operation. The declining cost, and increasing speed, of transmission has meant that routers are becoming an increasing component of network expenditure. Alternative, hybrid architectures, using a combination of switches and routers, are likely to offer a more cost-effective solution. This architectural approach will be suited better to dealing with the emerging traffic patterns where individual applications generate very large point-to-point flows. However, the approach raises new technical and organisational challenges.

5. The current cost-sharing arrangements for GÉANT do much to reduce the variation in underlying costs between different European countries by "smoothing" the cost differences. It will be necessary to review how costs are shared, as new forms of network usage develop in which only a subset of network resources is used but that usage is very intensive.

6. Changes in network economics and technology are enabling new applications where demanding users require access to significant amounts of network resources. In order to ensure that these user needs are properly met, it will be necessary to get a better understanding of the requirements of these users and to tailor solutions to these requirements in order to consolidate such demand.

7. Scenarios

In general, the existing GÉANT network has more capacity, over much of its geography, than is currently needed to support the traffic that it carries. There are a number of trends in traffic demand - in particular relating to the requirements of scientific projects on a pan-European basis - that may well change this picture. In addition, the aggregate traffic growth from the existing user base in the research community is a factor of around 1.7 per year. (This figure is based on the period from September 2002 until September 2003.) The combination of organic growth, together with potential new demands, will lead to the need to increase capacity at some point in the next 1-2 years. The ability to meet this growth depends on a number of factors, including in particular the way in which the market for telecommunications capacity will develop. There are also issues surrounding the development of user demand, in particular from those groups of users who are organised on a pan-European basis and who might potentially implement their own solutions.

In order to give some insight into these issues, three scenarios have been developed. The intention of the scenarios is to look at various potential developments in the areas of cost and availability of infrastructure and support for users. This analysis only relates to the trans-European network element. In order to have a complete picture, it will be necessary to complement it with similar views of national and campus developments.

7.1. The Tailor-made Networking Scenario

The scenario envisages that there will be positive developments regarding the cost and availability of pan-European infrastructure, giving rise to enhanced demand from users, and in particular advanced users with specific and well-defined networking needs. As a consequence, in addition to supporting the "average user", the European research networking community will be faced with the challenge of organising specific support for groups of users. In this scenario, "tailor-made" networking, where the needs of a limited set of demanding users have a significant influence on the service portfolio and operational support, will cause organisational change and development in the NRENs. From the point of view of the supply of connectivity, the competitive market will continue to develop both geographically and in time as a result of effective implementation of the regulatory framework of the European Union. This means that for those countries where international connectivity is currently based on SDH, there will be a move to DWDM connections. For all locations, network operators will provide connectivity at prices that are closely related to their underlying costs of network provision.

Looking at the current GÉANT costs and applying the assumptions described above, the total network capacity could increase significantly at no extra cost. In addition, increasing volumes of DWDM will give rise to continuing reductions in unit costs. A further positive aspect of this scenario will be that the current very significant smoothing in the GÉANT cost sharing, which means that there is cross-subsidising between cheaper locations and more expensive locations, will be unnecessary. It will still be necessary to develop an equitable cost-sharing approach that mitigates the effects of distance. The current cost-sharing model is based on annual demand for global connectivity. In this scenario, a more flexible way of allocating costs to significant users will be developed, so that their usage will be more precisely reflected in the costs incurred.

7.2. The Networking-For-Many Scenario

In this scenario, there will be limited change as regards the competitiveness in the market place, and the current digital divide in research networking in Europe will remain an important factor. In order to overcome the problem of relatively expensive transmission routes, a self-provision approach to transmission is introduced. Because of the rather large investment costs required to achieve this, in general the effect of the approach will be to yield more capacity rather than to reduce overall costs. In order to justify the finances used, NRENs will concentrate on maximising the number of users so that overall traffic volumes will increase. This will be the case particularly in smaller countries, where NRENs will extend their services to new groups of users, such as schools, in order to increase usage volumes and thereby benefit from economies of scale.

It is questionable whether in this scenario it will be possible to preserve the current GÉANT-like infrastructure while at the same time providing cost-effective connections, at high capacities, to projects with more limited geographic requirements and specific technical needs. Whether or not that will be possible, will depend partly on the implementation of new and innovative approaches to cost sharing, whereby high-usage applications requiring a limited geography only pay the marginal costs of the connectivity that they consume. Nevertheless, there is a strong possibility that some groups of users will establish their own infrastructures, because they will judge that their special requirements are not met.

7.3. The Business-As-Usual Scenario

In this scenario, the market for connectivity will not become significantly more competitive. In some regions, additional "market failures" may even give rise to increased prices and reduced availability of connections. As a consequence, the widening digital divide will exclude the introduction of new countries with a history of a monopoly market and very high telecommunications prices. Developments in the market in the currently involved countries will imply that there will be even greater divergence of the underlying costs of research network provision. There will be no political initiative to address the issue of the digital divide. Therefore, NRENs in countries where prices for international connections are relatively cheap will be unwilling to pay the increased premium that will be needed to reduce the effects of the digital divide, and the cohesion of the GÉANT network will be jeopardised. The service that is provided on a pan-European basis will be limited to a best-efforts IP service that provides basically the lowest common denominator of the services required across Europe.

This scenario is likely to lead to the creation of independent high-performance networks that cater for the specific demands of those users who demand high and predictable performance. The geography of those networks will be limited to those regions in Europe where there is a high level of competition in the provision of telecommunications services. Some projects may be able to exploit this approach, but in general there will be no shared pan-European network of any significant level of performance.