Internet Interaction Pinged and Mapped

John S. Quarterman[1]
Smoot Carl-Mitchell[2]
Gretchen Phillips[3]

Abstract

We ping approximately 4,500 domains, 1,428 of them in Europe, six times daily, seven days a week. This paper gives an overview of both the techniques of data collection and interpretation of this data with regard to network latencies to selected domains in Europe. It includes a geographic representation of the data that gives new insight into IP networks as they are actually used in real time across time zones, providers, and national boundaries. The heart of the paper is the accompanying maps, which show usage patterns on the Internet in Europe as they cross the continent from east to west.

I . Introduction

Geographical plotting of latency data obtained by use of ping (ICMP ECHO) permits a perspective view of the relative distance of selected major domains in Europe according to one of the metrics that matter most on the Internet: round trip time.

II . The Maps

The maps included with this paper show round trip times from our hosts (tic.com) to selected hosts in Europe. They use averages for data collected at six different times of day on every Wednesday from December 1993 through March 1994.

Unlike many previous maps from MIDS, these maps show just the Internet, not any other network. The Internet for our purposes is defined as all hosts pingable with ICMP ECHO from tic.com, i.e., it explicitly does not include hosts behind firewalls, and it explicitly does not include hosts connected by protocols other than IP: not OSI, not UUCP, not FidoNet, not NJE (EARN or BITNET) [3].

Also unlike many previous maps from MIDS, these maps do not show total host populations, not numbers of users, and not traffic. They show selected hosts, and for those hosts they show latencies, or round trip times.

The maps use an intuitive technique for displaying latency data. The size of a circle indicates the round trip time to a host at that location. Bigger circles indicate closer hosts (lower latency) and smaller circles indicate more distant hosts (higher latency), providing a sort of three dimensional perspective view. The upper left legend gives the latency scale, which is logarithmic. The unit is the millisecond, so 1000 indicates one second.

Colors indicate the number of hosts at a given latency and location. The upper right legend gives the host count scale, which is also logarithmic: red for 1, orange for 2, and so on through yellow, green, and blue to violet. The use of colors gives four dimensions of data on each two dimensional map: latitude and longitude (icon placement), latency (icon size), and number of hosts at each latency and location (icon color). For some of the more densely surveyed cities, it is possible to see Bell curves in the size and colors of the icons plotted, with red (few hosts) on the outside and inside of the icon, and indigo or violet (many hosts) between. Even on a black and white display, the different latencies are visible, and some of the differences in numbers of hosts can be seen by differences in grey scale, but the maps are best seen in color.

The geographical projections were made with the GMT package from the University of Hawaii [4, 5], slightly modified by MIDS. Host localization, latency computation and aggregation, and icon location, size, and color plotting were done with custom software written by MIDS.

1 John Quarterman is Senior Technical Partner in Texas Internet Consulting. He can be reached as jsq@tic.com

2 Smoot Carl-Mitchell is Managing Partner in Texas Internet Consulting. He can be reached as smoot@tic.com

3 Gretchen Phillips is the Manager of the UNIX Support Group at the University at Buffalo. She can be reached as gretchen@acsu.buffalo.edu

III . The Data

We have collected the data by using the ping program to send ICMP ECHO requests to a list of approximately 4,500 hosts on the Internet around the world. 1,428 of them are in Europe, loosely defined to be the area that appears on the accompanying maps, which includes the entire continent of Europe from Iceland to the Urals, plus some of the Middle East and North Africa. We ping the complete set every four hours, or six times a day. The worldwide list was derived from Mark Lottor's domain survey [2] by omitting all domains with no subdomains, and then using DNS to find MX records, and then following those to the IP forwarder, in order to derive the main mail host for each domain.

IIIA . Privacy and Security

The letter we sent CERT (Computer Emergency Response Team) with an overview of the description of the methodology for their information is available by anonymous FTP, Gopher, or WWW from ftp.tic.com, URL gopher://ftp.tic.com/matrix/mmq/latency/overview.

With FTP,

log in as user anonymous with password guest and retrieve the file found in: matrix/mmq/latency/overview

With Gopher,

look under MIDS, then MMQ, then latency, then overview.

This latency survey follows all the guidelines for Internet measurement activities set forth in RFC 1262 [1].

We do not believe there are any privacy or security concerns, since ICMP ECHO does not permit viewing any actual data or addresses traversing any network link. We are not trying to ping any hosts on the other side of firewalls, and most firewalls would block such requests anyway. The load imposed by six pings a day is negligible, both to the host being pinged and to any intervening links.

We do not and will not identify individual users, hosts, or institutions. Of course, if you are receiving our pings, you can tell that easily by looking at your router logs.

In addition, anyone who objects to being part of this ongoing latency survey can send us a message requesting their host be deleted, and we will send no more pings to that host.

IV . Interpretation

Different latencies to different hosts in the same time period may be due to the number of intervening hops and the speeds of the links involved. Different latencies to the same hosts at different times of the day are more likely to be because of differing loads.

The maps included with this paper show latencies from TIC to Europe at midnight, 4AM, 8AM, noon, 4PM, and 8PM Central Standard Time, which are 6AM, 10AM, 2PM, 6PM, 10PM, and 2AM GMT. To avoid loading the Internet with even the slight traffic of these pings, we are careful not to do them too rapidly in succession, so the times shown are not precise; the actual ping time could be up to two hours later. We ping every day of the week, but for these maps we used only data from the 18 Wednesdays in the four months from December 1993 through March 1994. We omit data after March to avoid the complication of daylight savings time (summer time).

Two of these maps, in particular, are conveniently during business hours for both continents (8AM CST and 2PM GMT) and off hours for both continents (10PM CST and 4AM GMT). The business hour chosen is busier in Europe than in North America, so most of the difference in latencies should be due to traffic generated in Europe. This does appear to be the case, since the latency circles are smaller during busy European hours on all the maps, showing greater latency at those times.

It's clear that Tromsø, Norway, and Ankara, Turkey are closer to TIC than Lisbon at any time of day. They used to be closer than Moscow, until a faster link was connected to that city through St. Petersburg from Helsinki earlier this year. Moscow and Lisbon become farther away during the busy European hour shown, and their latencies also spread more. You can clearly see two distinct latency circles around Moscow, presumably corresponding to two different routes into the city, with different bandwidths. These two circles are most clearly visible on the 6 and 10 AM GMT maps, in the morning in Moscow. The same effect is even more pronounced with Budapest and Warsaw. As the load picks up in the afternoon, both circles shrink, with the bigger one collapsing in the 14 GMT map to almost the same size as the smaller one. The size of the latency circles increases, and they separate again, in the maps for later hours. Except the map for 2AM GMT or

5AM Moscow time shows higher latencies than the ones immediately before and after. This could be due to traffic to North America, where it is only 6PM on the west coast.

Helsinki shows a similar pattern, with substantial load showing by 10AM GMT or noon local time. Outside of Helsinki, latencies to most of the rest of Finland do not vary much with the hour of the day. The same appears to be true in Norway and Sweden outside their capital cities. Apparently there is plenty of bandwidth to carry all the traffic for the rest of those Nordic cities.

Across the Gulf of Bothnia from Helsinki, St. Petersburg, Russia, Talinn, Estonia, and Riga, Latvia are visible. All three of these show marked variation during the day, with Talinn showing both the best latencies and the greatest variation.

A wider spread of latencies with less distinct divisions between latency circles is common for much of Europe during busy hours. Considering that the maps show averages of data per host for the same set of hosts over several weeks, these spreads of latency do not indicate variance of round trip times to individual hosts. A wider spread of latency per city probably indicates that there are different routes to different hosts in that city, and that those different routes perform differently under load. In other words, carriers with faster links or simply excess capacity at any link speed will probably diminish less in latency with load. A slower link will probably display a higher latency (smaller icon) at any load, but the differences among link speeds should be more visible with greater load, and this does appear to be the case.

For example, from looking at these maps, it is likely that Bratislava is connected through a slow link, since it becomes much farther away during peak hours. The data apparently includes pings of only one host there, but the differences in latencies to that host at different times of the day are very evident.

Looking at France, it is rather interesting that the best latencies (shortest round trip times shown by largest circles) are not in Paris. Latencies from Texas to Paris run a second or two, while some other nodes in France have latencies closer to half a second. Presumably this is because the links into Paris have high bandwidth but are heavily used.

Locations of the domains surveyed were determined by a variety of techniques, involving WHOIS, X.500, netfind, WWW, FTP, finger, and gazeteers. The key localization and aggregation technique actually involves telephone numbers. Separating the Slovakia and the Czech Republic turned out to be relatively easy, since all former Czechoslovak area codes greater than 42-7 for Bratislava are now Slovakian, and the rest are Czech.

We are suspicious that we have placed too many Spanish nodes in Madrid. It is likely that some university subdomains are for campuses in different cities throughout the country, for example. We also wonder about the remarkably low latencies (large circles) around Madrid at 2PM GMT or 3PM local time. Could this be lunch hour?

The perceptive reader will wonder where are Tunis and Cairo. We did not find in the data we used to build the list of sites to ping. Cairo was connected somewhat after we started this project. We will add both of those cities and others to later versions of our ping list.

In all six maps from measurements over the whole day, it is possible to see peak hours moving across the four time zones of Europe, from Sverdlovsk and Moscow in Russia in the east to Portugal, Ireland, and Iceland in the west.

The point of this paper and these maps is not, however in the details that we the authors can deduce from the data or the maps; we suspect that European network administrators will see patterns that we do not. The point is that these maps represent a new method of seeing the Internet. By remote sensing using ping (ICMP ECHO), we can determine an important internal characteristic (latency) of the Internet in some detail, and by use of geographical maps we can make important features of the data directly visible. These maps show the Internet in Europe as a whole, yet with differing characteristics of its component parts still visible.

V . See Also

Further results will be published in geographical map form in the publication *Matrix Maps Quarterly*, and later in other publications, perhaps in other forms.

VI . Acknowledgements

We are most grateful to those individuals who have helped in the collection of the data. Especially, we would like to thank Eric McKinney.

VII . References

References

1. Cerf/CNRI, Vinton G. Editor, ‘‘Guidelines for Internet Measurement Activities; RFC 1262,’’ *Request for Comments*, (1262)Network Working Group, Internet Activities Board, (October 1991).
2. Lottor, Mark, ‘‘Internet Growth (1981-1991); RFC1296,’’ *Network Working Group Request for Comments* (RFC1296)Network Information Systems Center, SRI International, (January 1992).
3. Schwartz, Michael F. and Quarterman, John S., ‘‘The Changing Global Internet Service Infrastructure,’’ *IRENAP* **3**(3) pp. 8–25 Meckler, (Fall 1993). ISSN 1066-2243.
4. Wessel, P. and Smith, W. H. F., ‘‘Free software helps map and display data,’’ *EOS Trans. AGU* **72** p. 441 (1991).
5. Wessel, P. and Smith, W. H. F., ‘‘The GMT-SYSTEM v. 2.1.4 Technical Reference & Cookbook,’’ SOEST/NOAA (1993).

Author Information

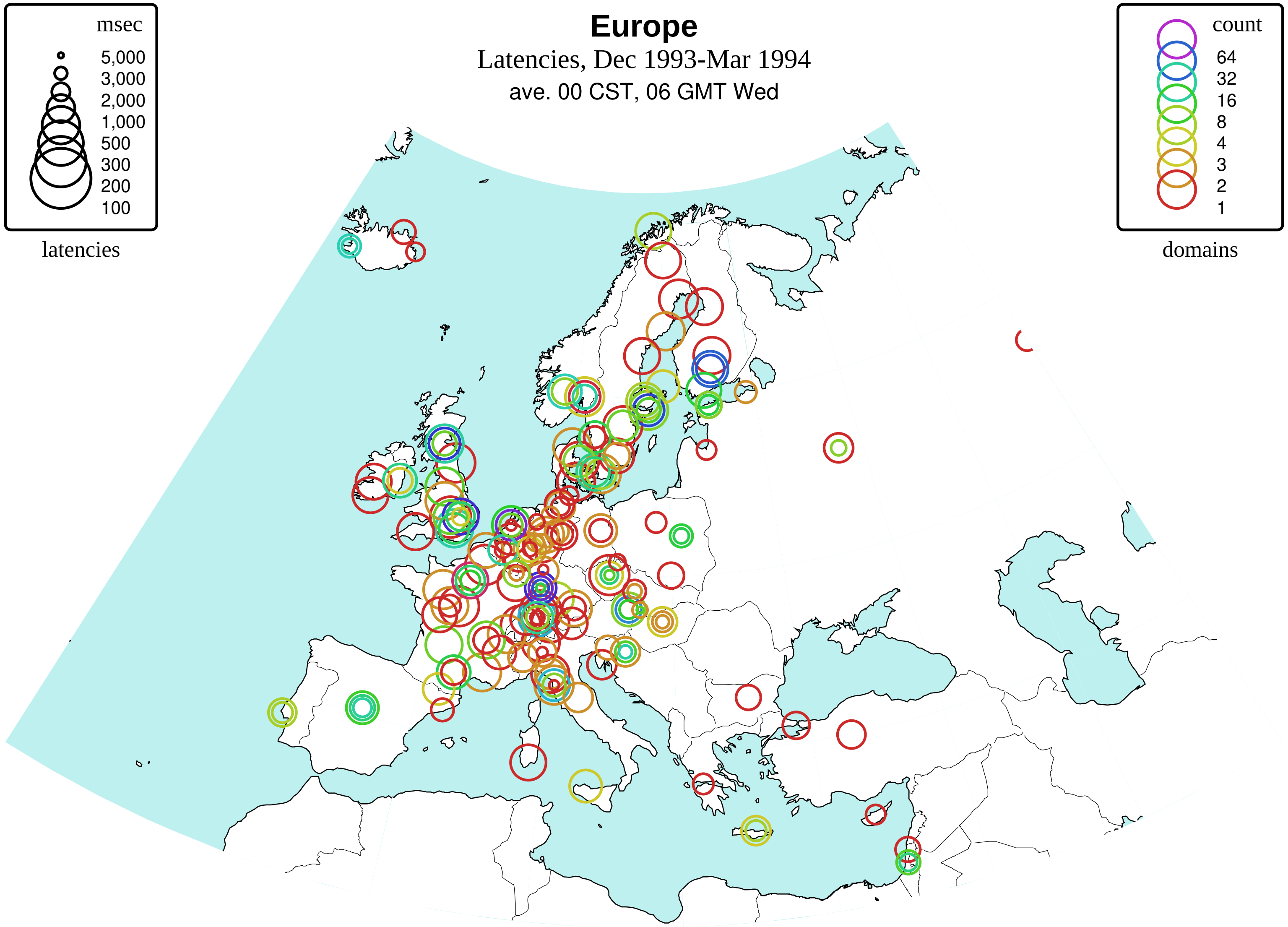
Smoot Carl-Mitchell and John S. Quarterman are partners in Texas Internet Consulting (TIC), which consults in networks and open system with particular emphasis on TCP/IP networks, UNIX systems and standards. They also write articles and books. The books include *Practical Internetworking with TCP/IP and UNIX*, 1993, and *The Internet Connection: System Connectivity and Configuration*, 1993, both from Addison-Wesley. Quarterman wrote *The Matrix: Computer Networks and Conferencing Systems Worldwide* (Digital Press, 1990), and co-authored *The Design and Implementation of the 4.3BSD UNIX Operating System* (Addison-Wesley, 1989), and *UNIX, POSIX, and Open Systems: The Open Standards Puzzle* (Addison-Wesley, 1993). Carl-Mitchell and Quarterman, as Matrix Information and Directory Services, Inc., publish *Matrix News*, a monthly newsletter, on paper and online, about contextual issues crossing network, geographic, and political boundaries, and *Matrix Maps Quarterly*, a color publication of maps of networks, including the Internet.

Gretchen Phillips is currently the manager of the UNIX Support Group at the University at Buffalo. As such, she is responsible for directing the group in all areas of UNIX support, including system configuration and installation, networking, and daily operations, as well as documentation and end user support. This group provides support for several hundred UNIX workstations, primarily Sun and SGI, networked across two campuses and centralized UNIX timesharing for over 4000 faculty, researchers and students. She has been involved in UNIX system administration since 1983.

Europe

Latencies, Dec 1993-Mar 1994

ave. 00 CST, 06 GMT Wed



+1-512-451-7602

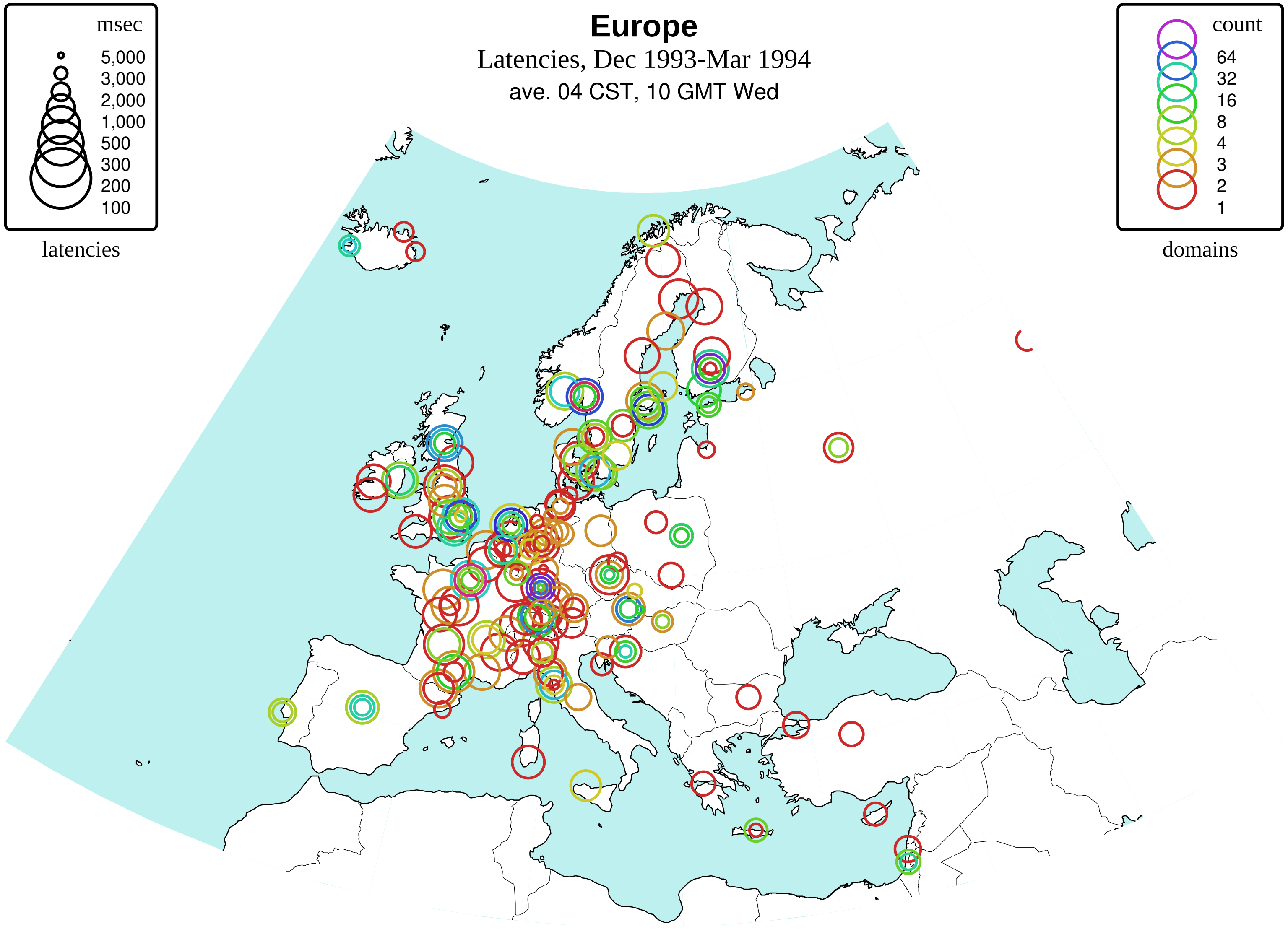
mids@tic.com

1:40,000,000 Albers Conic Equal-Area projection 94.04.22_00:46 (0 42 0)

Europe

Latencies, Dec 1993-Mar 1994

ave. 04 CST, 10 GMT Wed



+1-512-451-7602

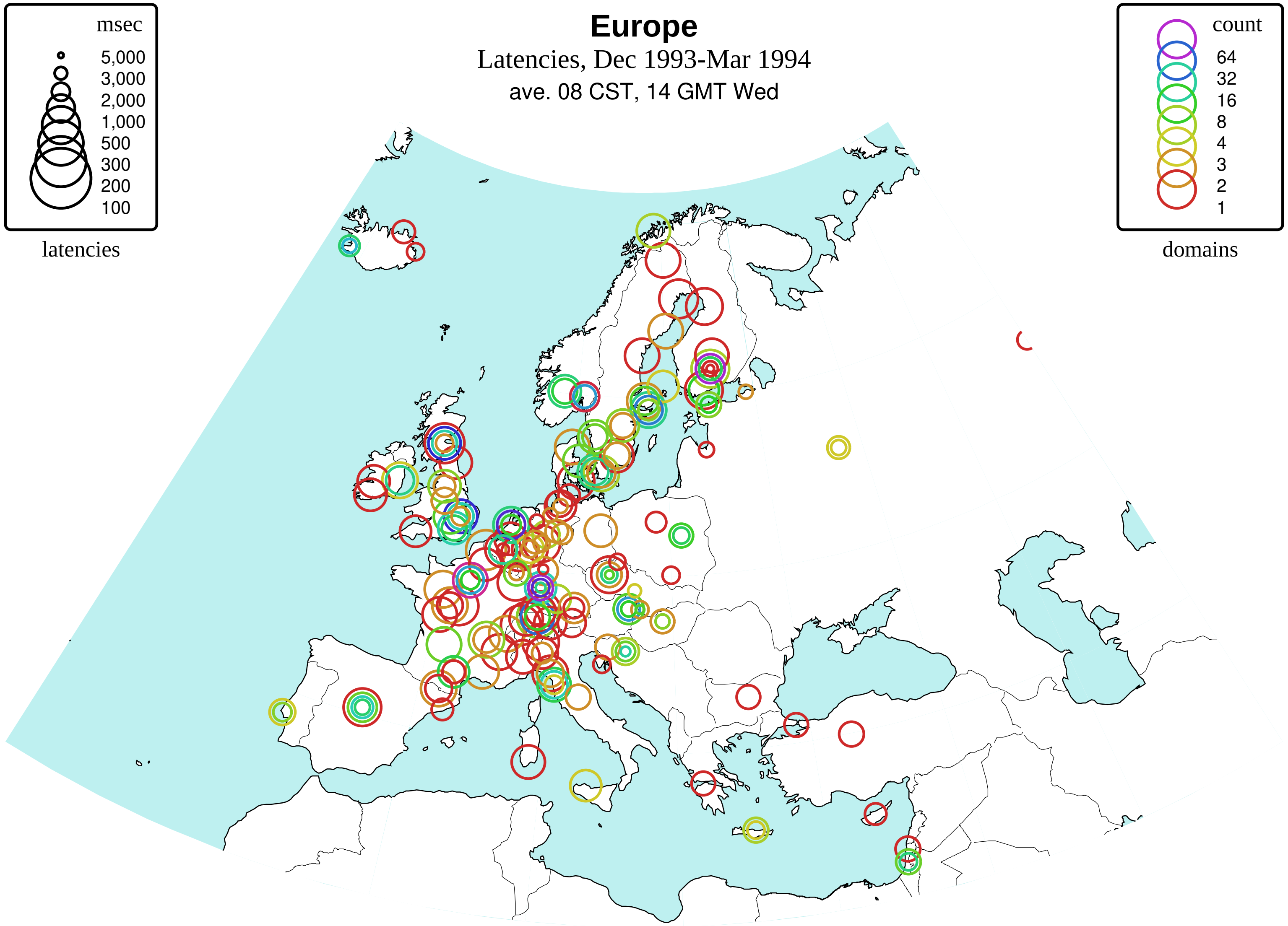
1:40,000,000 Albers Conic Equal-Area projection 94.04.22_00:48 (0 42 0)

mids@tic.com

Europe

Latencies, Dec 1993-Mar 1994

ave. 08 CST, 14 GMT Wed



+1-512-451-7602

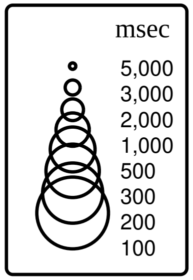
1:40,000,000 Albers Conic Equal-Area projection 94.04.22_01:02 (0 42 0)

mids@tic.com

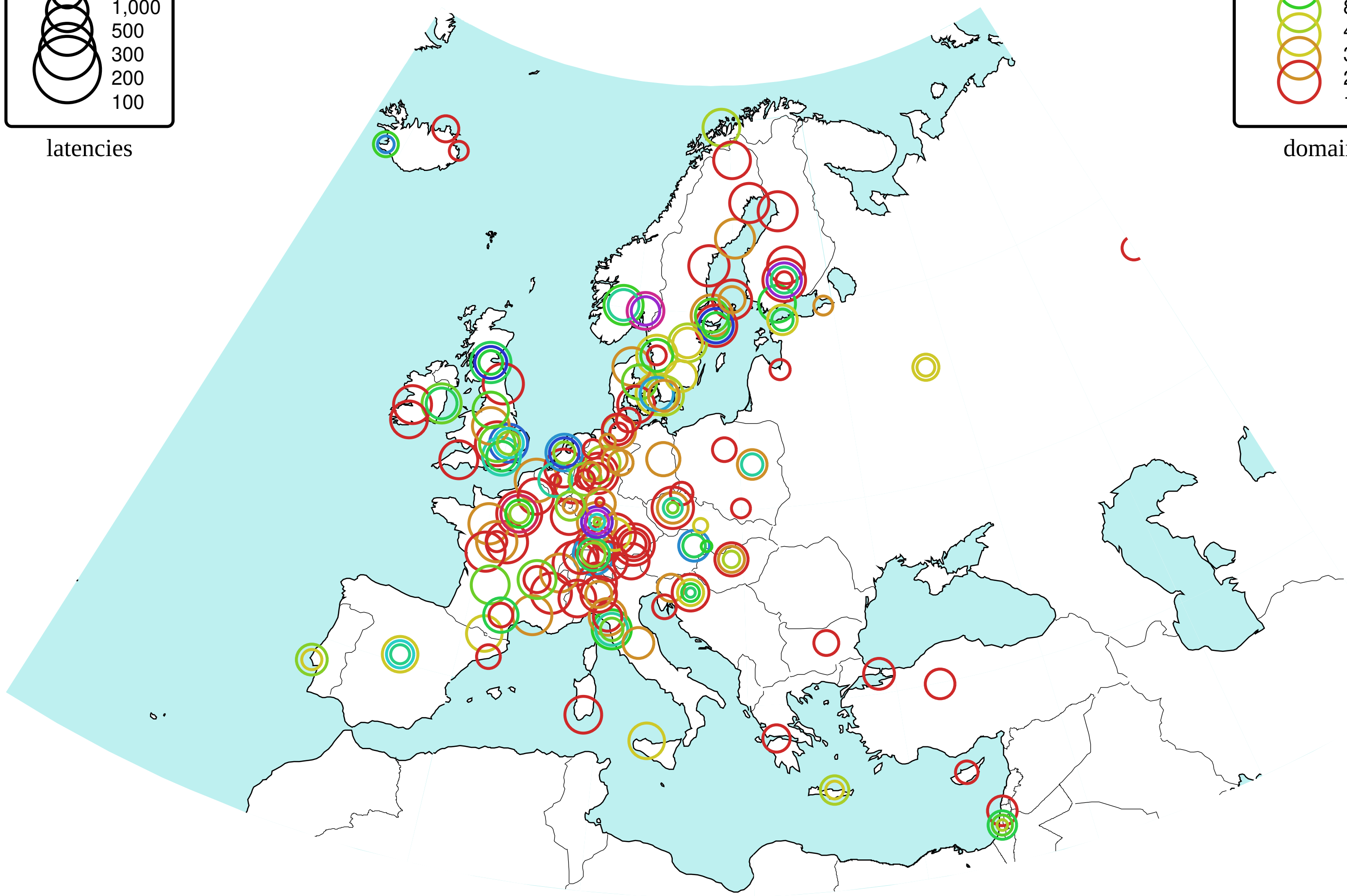
Europe

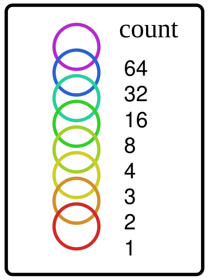
Latencies, Dec 1993-Mar 1994

ave. 12 CST, 18 GMT Wed



latencies





domains

+1-512-451-7602

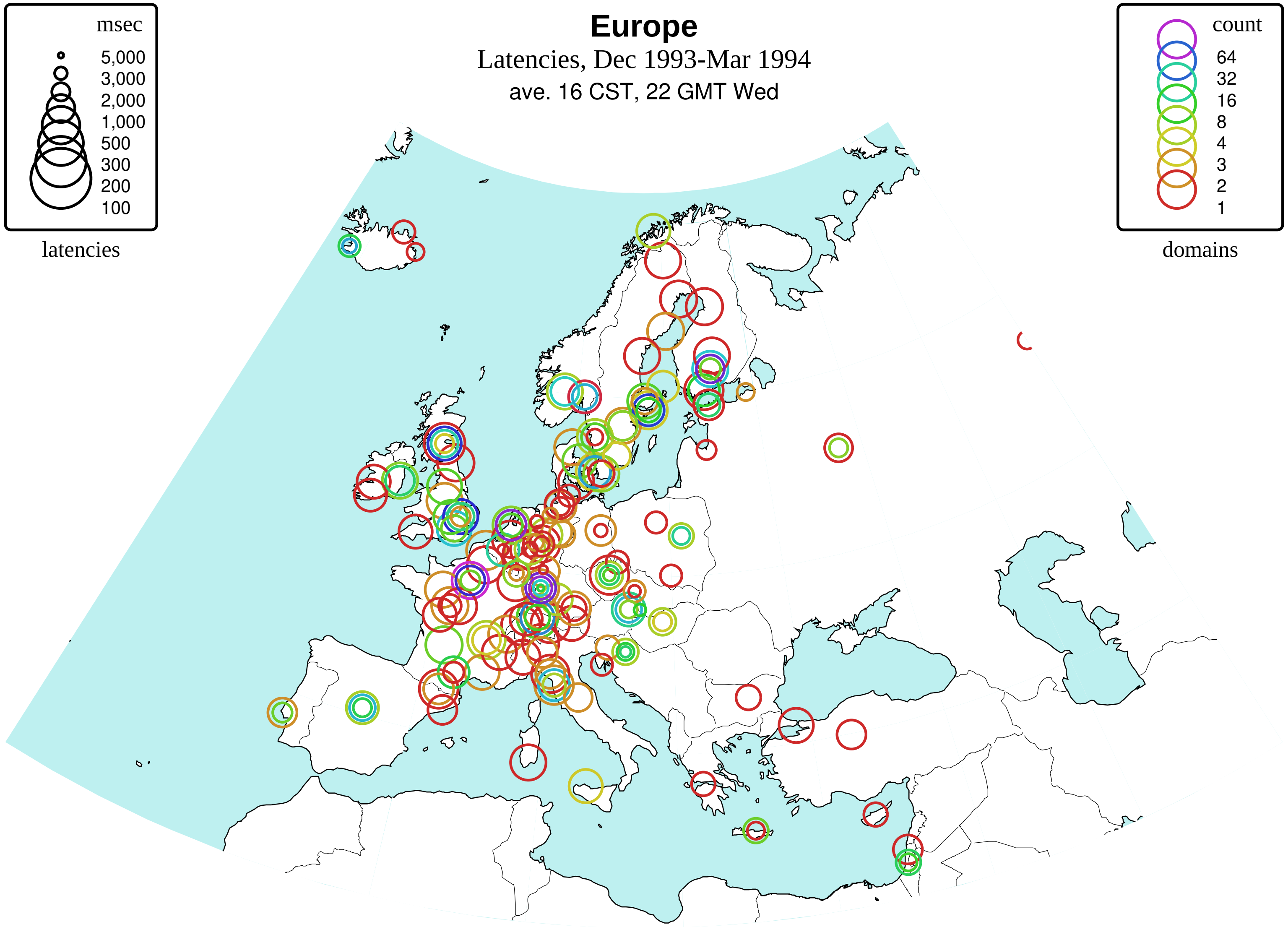
1:40,000,000 Albers Conic Equal-Area projection 94.04.22_01:05 (0 42 0)

mids@tic.com

Europe

Latencies, Dec 1993-Mar 1994

ave. 16 CST, 22 GMT Wed



+1-512-451-7602

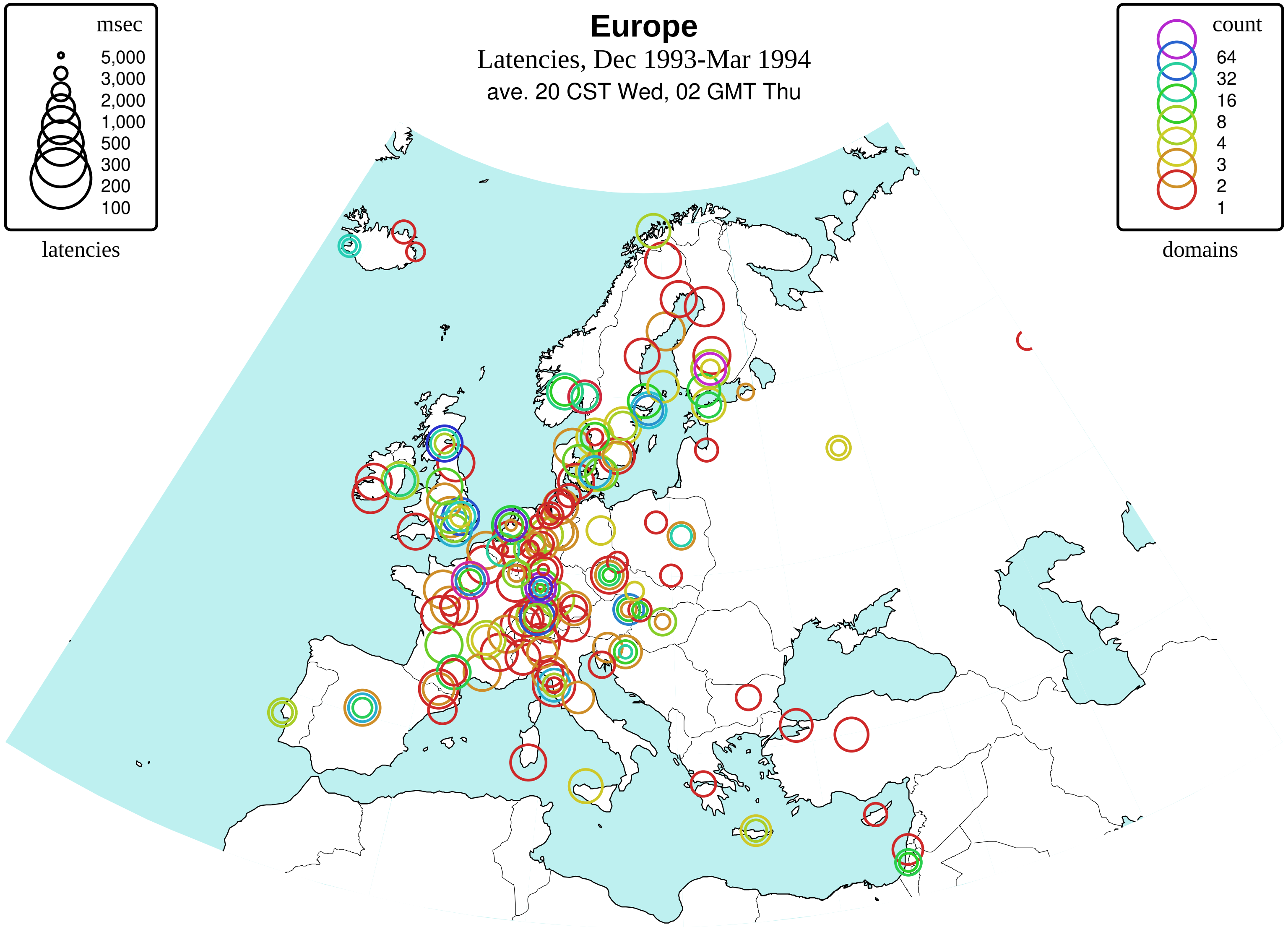
mids@tic.com

1:40,000,000 Albers Conic Equal-Area projection 94.04.22_01:07 (0 42 0)

Europe

Latencies, Dec 1993-Mar 1994

ave. 20 CST Wed, 02 GMT Thu



+1-512-451-7602

1:40,000,000 Albers Conic Equal-Area projection 94.04.22_01:10 (0 42 0)

mids@tic.com