The history and operation of the Internet: The development of its infrastructures

Stéphane Bortzmeyer,
AFNIC

Abstract:
The Internet is sometimes presented as a "cloud", something smooth, calm, free — something floating that pays no attention to borders and drifts where it wills. Nothing is farther from the truth. On the contrary, the Internet is an onerous set of physical, geographically located infrastructures. It also has a less materialistic infrastructure (but with a structural effect) made up of software, the protocols used by computers to communicate, relations between human beings and between organizations, and the naming systems and standards for the addresses of all these entities. What are these infrastructures and their functions?

A user, Anyone, from his home in Bobigny (Île-de-France), wants to browse the website of the government of Burkina Faso. Let us suppose that Anyone subscribes to Free, an Internet service-provider (ISP). Between Bobigny and Ouagadougou are 5140 kilometers and, above all, many different operators. Free does not know the Burkinan government, which, in turn, does not know Free. Nevertheless, the connection is made, and Anyone sees the webpage on his screen less than a second later. How is this possible given that no one is in charge of coordinating Free's connection with the government's website?[1]

The emergence of the Internet

The Internet's most important attribute is that it is EMERGENT. Each operator is busy with its own network, for which it purchases computers, hires technicians, and then configures everything and hooks up with a few nearby operators. Messages then pass from operator to operator until they reach their destinations. No one is globally responsible, but everything works.

In effect, the Internet is a NETWORK OF NETWORKS. Although no one is responsible for the connection between Bobigny and Ouagadougou, each of these networks is managed by an administration, firm or association that sees to making things work. On the Internet, these networks are called autonomous systems, (AS). Free is autonomous system 12322, and the Burkinan government website is hosted by ANPTIC, autonomous system 327871. There are approximately 69,000 active autonomous systems. They form the Internet, or rather their interconnections do.

Although an autonomous system is a somewhat abstract concept, it is not a purely nonphysical entity floating far above earthly pursuits. It has a nationality, is subject to local laws, has a banking account, etc.

[1] This article has been translated from French by Noal Mellott (Omaha Beach, France).

The material infrastructure

The Internet is material. Far from the sales pitch offering the dream of "clouds" in an "ethereality", the Internet depends on a material infrastructure. Most intercontinental traffic passes over undersea cables, which are very expensive to lay and maintain. Few societies or states are, by themselves, able to finance these cables. These cables are not public works like highways; they are often managed by consortiums or big firms. The critical nature of this material infrastructure becomes evident whenever a cable breaks down (*e.g.*, Sea-Me-We 3 in March 2020, when each part of the Internet's infrastructure was crucial given the restrictions related to the COVID-19 epidemic).

Next are the routers, the computers specialized in passing messages from one network to another. A router, which has interfaces with various networks, has to transmit data packets very fast. Your small "home box" is a router, one that is much less efficient and expensive, and consumes less electricity, than the big equipment used by Internet operators. The market of "core routers" is highly concentrated in the hands of three firms (Cisco, Juniper and Huawei). Given the high tech concentrated in routers, it is hard to break into this market.

Nor should we overlook the servers, the computers that, connected to the Internet, are waiting to receive queries from clients (for obtaining a message, downloading a film, or retrieving a document). Unlike cables and routers, servers are not an indispensable part of the Internet's infrastructure. There are many "peer-to-peer" applications that work via direct communications between end users' computers. In practice however, many uses of the Internet depend on servers, unfortunately so since this leads to greater concentration in this market and enables the firms that manage servers to exercise more control over uses.

Finally, these routers and servers are installed in data centers, specialized buildings with electricity, air-conditioning and security systems. Once again, all these facilities are managed by private firms; they are not a public infrastructure.

Software

The Internet's infrastructure is not just tangible machines and cables. It also includes a wide range of software. For example, a router with its many interfaces passes data from one interface to the other. It is material (a processor, memory, Ethernet cards) and, too, software. It has to route data packets in conformity with both local AS policies and the general rules for transmitting them (the Internet Protocol: IP). The software on high-end routers is in the router's specialized circuits. It is not executed by a general processor, but the principle is the same. A router has to be able to "talk" to other routers (routing algorithms) and to the system administrator who configures it.

At the top level of the Internet, the Border Gateway Protocol routes information between autonomous systems. This BGP is the same for the whole Internet, since it has to work for communications between different autonomous systems managed by different organizations. On the contrary, network administrators within an AS may freely decide on the protocol they want to use and may even choose a nonstandard protocol. The principle underlying the BGP is that each AS announces to its "peers" the routes that it has chosen and is able to join. The BGP implies a policy, since an AS does not announce routes that it does not want to use. For instance, AS administrators may, to a degree, decide to apply criteria such as "I want to exchange with all operators in my country (if they consent)" or, on the contrary, "I do not want exchanges with a certain party". "Schengen routing" refers to the criterion that traffic between two countries in the Schengen Area or the European Union is to remain within the Area or Union. BGP is not always easy to oversee. In any case, one party always hangs on the decisions made by other parties. If I want to route traffic through Orange but Orange does not want the traffic, I cannot force it to accept. This means that, in practice, policy-decisions are not always fully implemented.

Besides routing software, there are several other layers of software applications, including those used every day on our smart phones and computers. Many users think that these software programs are the Internet and soon forget that there are many other software programs in the Internet's infrastructure (*e.g.*, the programs used by the countless, indispensable servers.[2]

Protocols

I have used the word "protocol". What does it mean? A protocol is a set of rules that have to be followed by the machines that enter into communication. For two routers to be able to exchange routing information and thus offer connections on the Internet, their messages have to be formatted following precise standards that both machines can implement. A protocol might be informal, but it is better for it to be standardized, *i.e.*, formally stipulated in a written document that is validated and published by an organization that makes standards.

There are several, national or international, standard-making organizations. On account of the Internet's strategic value, it comes as no surprise that, within these organizations, heated debates sometimes flare up, as during discussions about a new version of the Transport Layer Security protocol, which uses encryption to protect the confidentiality of communications. The new version (1.3) of TLS, which patched some points of vulnerability in earlier versions, infuriated certain parties. After all, they had been using these points to snoop into the data being transmitted, an action that TLS should normally prevent.

Among standards organizations, mention should be made of the Internet Engineering Task Force, which sets the standards for the BGP and TLS protocols, as well as for several others. The IETF is open; there is no clear-cut barrier, apart from engineering know-how, to joining it. The standards it makes are published in RFC documents (requests for comments), the Internet's "holy book".

As seen with the example of TLS, a protocol is not neutral. Quite different parties are using these protocols but for varied purposes, negative or positive. Since protocols make some uses of the Internet easier, or harder, the choice of a protocol has consequences at the policy level. For instance, TLS 1.3 has made it harder to snoop into communications, this being an advantage for some parties but a disadvantage for others.

local.

Services

I have described the Internet as a network of networks. The user experience is much more complicated. When users say, "I found it on the Internet", they are, for sure, not thinking of optical fiber, routers and the BGP. What about the services offered on the Internet? There are two sorts of services: infrastructure services and visible services. Many of these services are more centralized than the infrastructure and are provided by foreign firms.

The phrase "infrastructure services" refers to the services that, though not part of the Internet in the strict sense of the word, are indispensable (or very important) for using the Internet. The best known are:

- the Domain Name System (DNS), which assigns domain names (such as www.annales.org). Since nearly all activities on the Internet require domain names, the DNS is critical. Its management is decentralized: the top-level national domain *.fr* is managed in France.

[2] According to a still current legend, a service can be "serverless", *i.e.*, built without servers. This is obviously false except for pure peer-to-peer systems. Servers are always present, even if they are "hidden" (*e.g.*, outsourced to AWS.)

● A content delivery network (CDN) comprises a set of computers that deliver contents (such as videos or webpages). CDNs are spread around the world in order to reduce access time. Content operators or specialized firms (American or Chinese) run them.

The "visible services" are those that Anyone knows (for example, websites). They are not indispensable for the operation of the Internet, but they are the very reason Anyone connects to the Internet. These are the services that, poorly designed, broke down during the COVID-19 pandemic, when the "confinement" of the population led to a more intensive use of the Internet. The Internet did not break down. These visible services are both very decentralized (since it is easy and cheap enough to set up websites or offer web-hosting services) but, in practice, highly centralized (since a few players, nearly all of them American,[3] dominate the market).

[3] American in the case of France. Elsewhere, the situation is quite different, as in China or even Russia.