David G Post, "In Search of Jefferson's Moose"

Back in 1950, Russian psychiatrist Immanuel Velikovsky published a best-seller called *Worlds in Collision* which argued that cultural myths of great catastrophes were based on other planets leaving their orbits and nearly hitting the earth. Years later an astronomer told me that, at the time, although he and his colleagues knew the astronomy was nonsense, they'd thought the anthropology was really interesting−until he'd talked to an anthropologist colleague who said that that although the anthropology was nonsense, the astronomy was really interesting.

In Search of Jefferson's Moose, published last year by law and policy professor David Post, draws parallels between the historical structure and growth of the U.S. in Jefferson's time with the recent structure and growth of the Internet. The book got good reviews, but every review I could find was by policy specialists, not historians or computer network experts who could evaluate his historical or technical claims.

I offer this review as a reasonably knowledgable computer network expert to provide some needed balance. In one sentence: I wanted to like this book but−despite extensive research−it is so riddled with errors of fact and interpretation that the errors discredit the author's conclusions about managed ("Hamiltonian") versus organic ("Jeffersonian") network organization and growth.

The first three chapters draw an analogy between Jefferson and the growth of the US in the 1700s and the growth of the early Internet. They make the unexceptionable point that exponential growth can be really fast, and anything that grows exponentially, such as the US population in the 1700s and 1800s, or the Internet in the 1990s and early 2000s, will become big enough to become important. But in chapter 4, the author runs into trouble when he attempts to explain what has allowed the Internet to scale up from a handful of academic networks to a globe spanning behemoth. He attempts to contrast the way the Internet routes traffic with the way that the Post Office routes mail, and manages to get both wrong. He calls the PO a ''centralized network,'' where a single authority has to know the route to every address. (There are such networks−notably Ethernet LANs and FedEx, which routes everything through Memphis−but the PO isn't one of them.)

Then he offers a bizarrely wrong description of Internet routing, claiming that messages just bounce around from node to node until they happen to land at the right one. In reality Post Office and Internet routing are more similar than different: addresses are hierarchical, with addresses in a single locality/network having the same non-local part. Using that knowledge both local POs and networks can recognize and deliver to their local addresses, and know how to send other messages to higher level sorting centers or routers where the same process happens at the next higher scale, with a few regional sorting centers or backbone routers that know routes to every PO or network. That is, the routing is distributed, but with overall coordination to manage the addresses. There are some important differences−notably that Internet routers manage route information automatically, so routes to newly added networks propagate across the net in a few minutes. Multilevel routing designs are not new or unusual; the phone system and package delivery compa-

nies use them. They're also not the important part of what allowed the Internet to scale, a topic we'll return to in a moment.

Chapters 5 through 7 continue the discussion of scaling, building on the misconceptions in Chapter 4, with a lengthy but largely irrelevant discussion of power law distributions which apply to the Internet just as they apply to all large distribution networks, computer and otherwise. Post mentions the oft-cited end-to-end principle which makes it easier to run new and unexpected applications, some of which become popular (the Web) and most of which don't.

Unfortunately, he completely fails to identify the key factor that allowed the Internet to grow so fast−its business model. There are a few technical enablers−notably the large IP address space that let the Internet run a decade longer than any of its competitors before running out of addresses, and the separation between TCP and IP that lets network routers scale up without having to worry about how many different conversations' packets they're handling−but the key to its growth is the no-promises business model.

Every Internet connection, from cheap consumer dial-up to gigabit backbones, is made on a "best effort" basis−a euphemism for making no promises. They'll deliver your packets if they can, but if it's unduly onerous or inconvenient to do so, they'll throw them away. In contrast, the common carrier model for telephone systems or package delivery promises to account for and deliver everything, which means that they need complex technical and contractual relationships to ensure that every bit of traffic is accounted for and that nobody deliberately or inadvertently overloads their neighbors with traffic. On the Internet, if you're overloaded, you just throw the traffic away, leading to a Mutual Assured Destruction model in which everyone is very polite to their neighbors, since impolite neighbors tend to get disconnected on a moment's notice. This makes the negotiation to connect networks a lot easier than it is for phone networks, and indeed people set up new connections between Internet networks every day.

It turns out that this sort of unreliable MAD network does a great job of delivering files from one point to another, and of sending short (a few hundred characters) messages, and those are the facilities on which all of the Internet's popular applications are built. File delivery underlies everything from e-mail to the web to Napster to Youtube, and short messages underlie the Domain Name System and Instant Messaging. Unreliable MAD does a lousy job of delivering a stream of data reliably at a given rate−that is, streaming audio and video−but it turns out that the modern Net is fast enough that audio sounds OK, and the market for actual streaming video, as opposed to Youtube style pseudo-streaming, is not very big. It's possible to build networks using Internet technology that stream reliably, which is the way that many phone networks now work internally, but their costs and management are those of the phone network, not the Internet. (This is the point where the end-to-end principle fails, since e-2-e gets reliability at the cost of delays when lost traffic is resent, and you can't change that without changing every router through which your traffic passes.) Post doesn't seem to be aware of any of this, which is unfortunate because the MAD model of interconection answers a lot of questions in the second half of the book that baffle him.

In the middle of the book is an Interlude in which he introduces caricatures of Jefferson, the advocate of decentralized unmanaged bottom-up rural Arcadia, and Hamilton, the advocate of top down command and control urban agglomerations. (You can probably guess from the title which camp Post is in.) In the second half of the book he looks at some of the issues involved in running the Internet−unfortunately with misconceptions just as bad as in the first half.

Chapter 9 tries to present the "code is law" idea popularized by Larry Lessig, but again Post botches his facts. The example he picks is the "referrer" field in the web's HTTP protocol, which a browser can use to tell a web server what page linked to the current one. Post thinks it is the key to pay-per-click ads like Google Adsense, telling the site who to pay. This is just wrong. PPC ads link to the ad network's server which records the click and then links to the target site; they don't depend on the referrer at all, which is just as well, since its use is entirely optional.

It is true that above the basic file transfer/short message level the Internet enables deployment of new and different versions of file transfer or short messages, but that is more of an optimization than a breakthrough. The first version of e-mail on the Arpanet, the predecessor to the Internet, was implemented on top of the existing file transfer protocol (FTP), and later reimplemented as a separate protocol to make it more flexible and reliable. Similarly, the Web can run on top of FTP rather than the usual HTTP, and occasionally does, with its HTTP adding flexibility and performance. A few parts of the code really are "law" that can't realistically be changed (the parts that make file transfer and short messages work well, and the traffic discarding which makes MAD possible and necessary) but Post doesn't distinguish between what's essential and what's just a detail.

On the next page, Post reports that the 4 billion addresses in the IPv4 addressing now used on the internet is likely to run out soon, and in one of his better insights observes that opinions about the effects of running out range from nothing to disaster. The solution to running out is to switch to IPv6, which provides so many addresses that every appliance in your house can have one. From this promising start, he stumbles again, claiming that with so many addresses we'll each have our own permanent addresses to take with us, removing any dependence on ISPs, but again, that's just wrong. IPv6 addresses are allocated in network groups just like IPv4, and neither lets individual users take their addresses with them. As Post is apparently unaware, permanent addresses are nothing new, since all Ethernet networks, including WiFi, assign one to each device. That works fine on a LAN where the routers or hubs can remember all of the addresses, but not on a large network, because remembering all the Ethernet addresses doesn't scale beyond a few hundred devices. The Internet handles Ethernet devices by also assigning them hierarchically allocated IP addresses, either manually using a system called DHCP on the local network. The success of the Internet was due to careful design that would scale up smoothly, which IPv6 preserves, but Post misses the point of IPv6 (to make a larger set of hierarchical addresses) and fails to grasp why permanent personal IP addresses wouldn't work.

Post rounds out the chapter with a look at the chaotic rough consensus governance of the

IETF (the Internet's standards group) comparing it with the ITU (the government and phone company standards group founded in 1865 and later made part of the UN) and the well known failure of the OSI network protocols in the 1980s and 90s. Although OSI was certainly a failure, Post appears not to know the difference between the ISO and the ITU. The ISO is a different Geneva international standards organization, not part of the UN, which developed OSI. The ITU later endorsed the OSI work but wasn't involved in its development. Post makes the absurd statement that "the UN [by which he means the ITU] couldn't build a network that could grow as fast as the TCP/IP network could grow." Disregarding the detail that he's insulting the wrong organization, he needed only to look in his pocket to see how silly that is; cellular mobile phones became available in 1983, about the same time as TCP/IP was deployed. There are a lot more phones than there are Internet users, and every phone is built on ITU standards.

TCP/IP succeeded where OSI failed because TCP/IP built incrementally on systems that were known to work (in particularly the Arpanet and Cyclades) with limited goals, while OSI tried build a much more expansive system to satisfy multiple politically hostile camps, with too many complex compromises and too much duct tape. Most important Internet standards precede the modern IETF, from back when the Internet was run by a handful of smart guys with government contracts who knew they had to adapt designs that worked because they didn't have the resources to reinvent from scratch, or, in the case of the Web, were developed outside the IETF, which then wrote its standards based on the running code. Standards processes of any sort work well when they codify and tidy up existing practice, and usually fail when they try to invent something new. The IETF has had its share of failures, and both the ISO the ITU, which learned a lot from the OSI fiasco, their share of successes. There's certainly a lesson here, but it's about the process, not about who appoints the participants.

Chapter 10 is about the Domain Name System (DNS) and the mess that is ICANN, the quasi-public non-profit that oversees the DNS. Post is astonished that back in 1998 it wasn't clear who was in charge of the DNS root, and apparently doesn't realize that it's still not clear, or at least doesn't realize that there are a dozen independent root server operators, none of whom (other than one root run by ICANN itself) have to do what ICANN says. Again, the key is the business model−or in this case the lack of one−leading to MAD. ICANN and the root server operators are very polite to each other, and are very conservative about what they ask of each other, since nobody wants to find out what would result from a serious disagreement. If you understand the (non-)business model, it's not hard to see how DNS governance has worked, but Post doesn't.

The penultimate chapter, Chapter 11, is about law on the Internet. Post cites the well known France vs. Yahoo case: Yahoo's auction site, located in the US, was selling Nazi memorabilia, which are legal in the US but illegal to sell in France, with rather difficult issues of jurisdiction. Post makes a reasonable (if somewhat simplistic) division of the opinions on this situation into Exceptionalists, who think that the Internet is so different that it will need a different legal system, and the Unexceptionalists, who note that jurisdictional disputes are not new and there's nothing that can't be handled. Post is a strong Exceptionalist, arguing, as best as I can tell, that solving the problem by harmonizing

laws doesn't work. Since the current legal framework doesn't work, he concludes it will have to change−a non-sequitur if ever there were one.

Anyone familiar with British libel tourism should know that simply because a law chronically leads to unfair and unreasonable results doesn't mean it's going to change any time soon, and the Great Firewall of China shows that a country−even a very large one−that wants to keep firm control over its part of the Internet can do so. He also greatly underestimates the amount of (and benefits of) harmonization happening in areas like online bank fraud, efforts that allow countries to cooperate on enforcement. I also see no evidence that France or any other countries outside the US see a problem in the Yahoo result; it's their country and their law, and they're not going to change it just because it makes life harder for companies in California. I can believe that we'll see some adjustments to make jurisdiction clearer and enforcement more predictable, but Post's prediction that Second Life will have its own legal system? Aw, come on.

Chapter 12 wraps up with a discussion of intellectual property law in cyberspace. Post appears to be in favor of liberty, opposed to tyranny, and, like Jefferson, wants enough IP protection to provide an incentive to create without being stifling, which is considerably less protection than we have now. For once I agree with him, although he doesn't know the history. If you look back a few decades, the 1940s through 1960s were the most furiously creative era in computing, during which nearly all of the hardware and software ideas underlying modern computing were invented. Perhaps by coincidence, software during that era enjoyed no IP protection whatsoever, no copyrights, no patents, no nothing other than the rare trade secret. The arguments about online IP are driven by the gatekeepers in the music and movie industries, not the creators or users of the material.

So what to make of this mess, in which every chapter that attempts to discuss the Internet is seriously undermined by substantial factual errors? Actually it's worse than that: the book is also full of little sloppy errors, like placing CERN, home of the Web, in the mountains of western Switzerland rather than down by the lake in Geneva on the French border. The overall result is a book in which nobody−from the author to the editor to the publisher−has made even the most cursory effort to ensure that this work gets its facts right. I don't know my American history well enough to comment on his presentation of it, but, as with Velikovsky's book, if he so badly botches the parts I do know about, how can I have any confidence in the rest?

There are some interesting lessons to be drawn from the history and development of the Internet, some of which I've tried to suggest as I've gone along, but to do so the writer needs to keep his or her facts straight. Post's opinion that successful systems need to allow the flexibility to innovate is of course true, but it's also trivial, and his belief that the Internet has organized itself is, as we have seen, mistaken. The real question is how to provide the maximum flexibility and still come up with something that works, since the coordination needed to make a billion node network work is both subtle and complex. I hope someone writes a book on that subject. Regrettably, this isn't it.

John Levine, johnl@taugh.com, Feb 2010

Revision date: 2010/02/21 03:51:34 (1.4)