# The HTTP Chasm

by Randy Charles Morin

For ten years, the World Wide Web revolutionized our world by turning the Internet into the primary means of communicating information. Some of us are using the World Wide Web to read our morning news. Others are researching the Web to find out why their goldfish is under the weather. The simple quantity of information now available on the Web is moving our already information rich world to a new level.

This revolution created a dot-COM bubble in the late 90s where any company that could process an HTTP request were valued at billions of dollars even with little revenue to back up these valuations. In the late 90s, I was working for a small dot-COM from Toronto. Our initial public offering garnered two hundred million dollars in cash for the company and soon after we were valued at ten billion dollars. Every employee in the company was a millionaire on paper.

On February 29th of 2000, Jim Cramer gave a keynote speech at the 6th Annual Internet and Electronic Commerce Conference and Exposition. The beginning of the speech is replicated here.

> You want winners? You want me to put my Cramer Berkowitz hedge fund hat on and just discuss what my fund is buying today to try to make money tomorrow and the next day and the next? You want my top 10 stocks for who is going to make it in the New World? You know what? I am going to give them to you. Right here. Right now.
>
> OK. Here goes. Write them down -- no handouts here!: 724 Solutions (SVNX:Nasdaq), Ariba (ARBA:Nasdaq), Digital Island (ISLD:Nasdaq), Exodus (EXDS:Nasdaq), InfoSpace.com (INSP:Nasdaq), Inktomi (INKT:Nasdaq), Mercury Interactive (MERQ:Nasdaq), Sonera (SNRA:Nasdaq), VeriSign (VRSN:Nasdaq) and Veritas Software (VRTS:Nasdaq).

His speech was intended to show investors value in these already very valued stocks. But his speech most serves as the checkpoint in time where the bubble was at its eclipse. After a few more weeks of bubbling, the bubble burst. Lockup periods and greed prevented nearly all from realizing any money and sent many investors into bankruptcy. The NASDAQ plummeted from an intra-day high on March 10th 2000 of $5132 to an intra-day low of $1387 on September 21st 2001, ten days into the aftermath of September 11. Seventy-three percent of the value of the NASDAQ had disappeared in less than a year and a half.

Hundreds of dot-COM companies like Excite@Home, Exodus, 360 Networks spent themselves into bankruptcy. Others like JDS Uniphase ($56 billion) and Nortel Networks ($19 billion) posted one-time losses greater than and near their market valuation. Most dot-COMs had very little revenue to begin with and still they posted decreasing revenues quarter-over-quarter for several quarters.

## 1 Dot-COM?

It was easy for analyst to name this The Dot-COM Bubble. But the bubble did not only affect dot-COM companies Yahoo!, InfoSpace, Digital Island and the like. The bubble

also had an affect on network equipment manufacturers (Nortel and Cisco), wireless equipment and software manufacturers (724 Solutions and Ericsson), silicon chip manufacturers (Intel and Motorola) and many more. More accurately, it was a bubble created by the World Wide Web. But the World Wide Web had also unveiled other sister industries like the Wireless Web (Openwave and Aether) and E-Commerce (Ariba and I2).

All of these technologies have one thing in common. They are based on a common protocol that was the root of World Wide Web. The protocol is known as HTTP, or Hyper Text Transfer Protocol. The World Wide Web simplified is HTML (Hyper Text Markup Language) documents being transferred between computers using the HTTP protocol.

Openwave's version of the Wireless Web is called WAP (Wireless Application Protocol), which simplified is WML (Wireless Markup Language) documents being transferred to mobile devices using the HTTP protocol. E-commerce is often associated with the transfer of XML document using the HTTP protocol. Each company that I have mentioned deal primarily with the processing of HTTP transactions.

Nortel, Ericsson and JDS make network equipment that serves as the heart of the Internet and 75% of the Internet traffic is HTTP transactions. Intel and Motorola create chips and Dell and Compaq create computers that serve often as HTTP intermediaries and end-points. Aether, 724 Solutions, Ariba and I2 create server software, which service HTTP transactions.

Let me conclude this section by more accurately naming what happened at the turn of the millennium as the HTTP Bubble. Yes, it's all Al Gore's fault. If he wants credit for the World Wide Web, then let him have it.

# 2 History of HTTP

Now that I've named what happened at the turn of the millennium, let me give you the history of what led up to this phenomenon. This is a brief history of the HTTP protocol.

In the beginning, the world of networked computers existed as a hodge-podge of mostly incompatible network protocols. The worlds of networks were in chaos.

In late 1981, ARPANET announced its intent to transfer its network from its legacy NCP (Network Control Protocol) based network, to a new network, called IP/TCP (Internet Protocol-Transmission Control Protocol). This announcement was documented as RFC801, http://www.ietf.org/rfc/rfc0801.txt.

Sidebar

> I particularly enjoy the way the acronym changed from IP/TCP to TCP/IP over the years.

In 1982 a few organizations began adopting TCP/IP as their transport and network protocol suite, including ARPANET and the DoD (US Department of Defense). Around the same time, people began calling the connected set of TPC/IP networks "The Internet".

In the meanwhile, the computing world realized that network standardization was necessary and a group was formed from ISO (International Standards Organization) called OSI (Open Systems Interconnection). This group created the OSI Reference Model and many organizations like the DoD and IBM announced full support for this initiative. The model described seven layers of network computing.

- Physical
- Data Link
- Network
- Transport
- Session
- Presentation
- Application

The OSI Reference Model was a great advancement in network computing. Unfortunately it suffered from three problems. First, the Internet and its protocol suites were growing fast and did not fit well into the model. I've seen some people try to jam the SMTP or FTP protocols into the presentation layer, but this is clearly incorrect. An Internet email application's presentation layer is MIME encoding and the SMTP protocol is a transport layer component, just as TCP is a transport layer component. So, the email application's protocol stack for sending emails is Email (application), MIME (presentation), SMTP (transport), TCP (transport) and IP (network). You'll also see two transport layer components in nearly every Internet application.

Second, security in the form of encryption, was accounted for, but was never accepted by the masses. Any transport layer can use SSL to provide for encryption. In this model, a small encryption layer exists as an appendage of the transport layer. This was never considered in the OSI model.

Third, scalability cried for stateless computing, which in turn made the session layer in the OSI model obsolete. This is a new phenomenon in computing that is only a few years old. I heard it best expressed by Don Box, one of the fathers of the SOAP protocol. Most programmers are still developing very stateful systems. This will change.

The Internet was declared an interim measure until OSI protocol suites could replace it. TCP/IP fit well into the model, but applications based on it (SMTP, FTP, etc) did not.

Then on August 6th 1991, Tim Berners-Lee, the grandfather of the World Wide Web wrote the following executive summary of his WorldWideWeb project.

From: Tim Berners-Lee
Subject: WorldWideWeb: Summary
Newsgroups: alt.hypertext
Date: 1991-08-06 13:37:40 PST

In article <6484@cernvax.cern.ch> I promised to post a short summary of the WorldWideWeb project. Mail me with any queries.

WorldWideWeb - Executive Summary

The WWW project merges the techniques of information retrieval and hypertext to make an easy but powerful global information system.

The project started with the philosophy that much academic information should be freely available to anyone. It aims to allow information sharing within internationally dispersed teams, and the dissemination of information by support groups.

Reader view

The WWW world consists of documents, and links. Indexes are special documents which, rather than being read, may be searched. The result of such a search is another ("virtual") document containing links to the documents found. A simple protocol ("HTTP") is used to allow a browser program to request a keyword search by a remote information server.

The web contains documents in many formats. Those documents which are hypertext, (real or virtual) contain links to other documents, or places within documents. All documents, whether real, virtual or indexes, look similar to the reader and are contained within the same addressing scheme.

To follow a link, a reader clicks with a mouse (or types in a number if he or she has no mouse). To search and index, a reader gives keywords (or other search criteria). These are the only operations necessary to access the entire world of data.

Information provider view

The WWW browsers can access many existing data systems via existing protocols (FTP, NNTP) or via HTTP and a gateway. In this way, the critical mass of data is quickly exceeded, and the increasing use of the system by readers and information suppliers encourage each other.

Making a web is as simple as writing a few SGML files which point to your existing data. Making it public involves running the FTP or HTTP daemon, and making at least one link into your web from another. In fact, any file available by anonymous FTP can be immediately linked into a web. The very small start-up effort is designed to allow small contributions. At the other end of the scale, large information providers may provide an HTTP server with full text or keyword indexing.

The WWW model gets over the frustrating incompatibilities of data format between suppliers and reader by allowing negotiation of format between a smart browser and a smart server. This should provide a basis for extension into multimedia, and allow those who share application standards to make full use of them across the web.

This summary does not describe the many exciting possibilities opened up by the WWW project, such as efficient document caching. the reduction of redundant out-of-date copies, and the use of knowledge daemons. There is more information in the online project documentation, including some background on hypertext and many technical notes.

Try it

A prototype (very alpha test) simple line mode browser is currently available in source form from node info.cern.ch [currently 128.141.201.74] as

/pub/WWW/WWWLineMode_0.9.tar.Z.

Also available is a hypertext editor for the NeXT using the NeXTStep graphical user interface, and a skeleton server daemon.

```
Documentation is readable using www (Plain text of the instalation instructions is
included in the tar file!). Document

        http://info.cern.ch/hypertext/WWW/TheProject.html

is as good a place to start as any. Note these coordinates may change with later
releases.
```

This alpha of what is today the World Wide Web, was based on the exchange of text and SGML documents over a new transfer protocol called HTTP/0.9. The HTTP protocol was eventually formalized in May 1996 as HTTP/1.0 in a document called RFC1945 (http://www.ietf.org/rfc/rfc1945.txt). Later versions of the same standard were RFC2068 (http://www.ietf.org/rfc/rfc2068.txt) and RFC2616 (http://www.ietf.org/rfc/rfc2616.txt).

RFC
: An RFC or Request For Comments is a formal document released by the Internet Engineering Task Force. Nearly all Internet standards are initially released as RFCs before they are adopted as standards.

Although HTTP/1.0 was only formalized in 1996, early releases of the protocol in 1992 had started the HTTP tornado. By mid-1992, the World Wide Web was beginning to take form and WWW sites started to grow in numbers by late 1992. In 1993, the number of WWW sites grew uncontrollably and the World Wide Web revolution was at full throttle. By the end of the decade, WWW traffic represented more than half of the total Internet traffic.

Early HTTP website successes like Yahoo! served almost exclusively HTML documents. By the late 90s, new markup languages like OFX (Open Finanical Exchange) were also being served by the HTTP servers. PDF, Word, Excel, Powerpoint, MPEG, MP3 and many other document types are now being transferred over the World Wide Web.

It eventually became obvious that the SGML, HTML and OFX languages required very complex parsers in order to read the data correctly. But at the same time, devices like cell phones and palm-sized computers were growing in number. In order to marry small devices with limited power and the growing number of SGML applications, a new subset of the SGML protocol was created. This protocol required a lighter-weight parser. The protocol was XML (eXtensible Markup Language). From XML, new markup languages (i.e. WML) were created to serve smaller device.

XML had another advantage in that the lighter-weight parser resulted in less parsing overhead. This made XML over HTTP a competitive RPC (Remote Procedural Call) protocol. New XML over HTTP RPC protocols became popular in early 1998, with SOAP (Simple Object Access Protocol) leading the charge.

Then shortly after the turn of the millennium, the HTTP bubble burst.

# 3 History of the PC

So what does the history of the PC have to do with the HTTP bubble? Let me explain.

Exactly fifteen years before the World Wide Web, another technology-based tornado was overtaking the world, the Personal Computer tornado. Just over fifteen years ago, another technology-based chasm was overtaking the world, the Personal Computer chasm.

Coincidence? In the words of Obi-wan Kenobi, “In my experience, there is no such thing as [coincidence]”.

Note

Obi-wan said luck, not coincidence.

In January 1975, Popular Electronics magazine published an article on the Altair 8800 computer. This article encourages Paul Allen and Bill Gates to join the PC tornado. The first Altair 8800 is shipped in April of 1975. This article is to the PC revolution as Tim Berners-Lee’s article is to the World Wide Web. It’s a marker in time, indicating the start of the tornado.

In March 1976, Steve Wozniak and Steven Jobs complete the Apple I computer. In July of the same year, Apple ships its first computers. In 1977, the PC market begins to grow with the Apple II, Commodore PET and TRS-80.

In 1980, IBM starts a secret project in Boca Raton Florida called “Accorn.” One year later, the IBM PC is introduced. The introduction of the IBM PC marked the point in time when the PC moved from the domain of the techno-hobbyist to a larger but still limited domain, the techno-savvy.

In 1983, Applet introduced the Mac, the first user-friendly computer that relied on intuitive user input instead of command-line programs and options. Microsoft shipped Windows in about 1985, ten years after the Altair 8800 was shipped.

Then the PC bubble burst. In 1985, PC shipments fell over 20% in the US. From 1984 to 1986 shipments of PCs did not grow at all. The 1989 recession exacerbated the problems in the industry adding a third year of negligible US growth in the history of the PC market. From 1984 to 1990, the PC industry grew on average 3% per year. In the eight years from 1976 to 1984, the industry grew on average greater than 100% per year. In the eight year from 1990 to 1998, the industry returned to higher growth rates averaging 18% per year.

What happened in the middle six years is called a chasm. This is where the industry can no longer grow by selling product to the techno-savvy. Growth can only continue when the industry compels the common folk to partake in the tornado. For the PC, it was the usability offered by the Mac and Windows Operating Systems that compelled the common folk to buy a PC.

# 4 Tornados and Chasms

The Personal Computer and the World Wide Web will remain the greatest tornados of the second half of the previous century. They will be remembered and documented in detail. But nobody will remember that ten years into the tornados, over supply and a ceiling on demand in the industry created a chasm, a drop off so severe that many of companies responsible for the initial tornado disappeared into bankruptcy.

In his books “Inside the Tornado” and “Crossing the Chasm,” Geoffrey Moore authored the idea of tornados and chasms in the growth of an industry. The current lag in the HTTP tornado may last for some more years, but eventually growth and the tornado will return.

The successful companies that defined the PC tornado are Intel and Microsoft. An investment of $1000 in either company during the chasm is now worth about $100,000. The same investment in Compaq, Apple, Radio Shack, Hewlet-Packard and IBM would be worth $20,000, $10,000, $5,000, $5,000 and $3,000 respectively. Never mind that some of these companies pay dividends.

In these first years of the World Wide Web, industry growth was about 100% per year. Of late, growth has turned into contraction. Some have suggested that the failure of dot-COM's resulted in about 20% contraction of the industry. This all sounds very familiar. By 2005, growth will return to the double digits. It's just history repeating itself.

## About the Author

Randy Charles Morin is the Lead Architect of SportMarkets Development from Toronto, Ontario, Canada and lives with his wife and two kids in Brampton, Ontario. He is the author of the www.kbcafe.com website, author of Wiley's Programming Windows Services book and co-author of many other programming books and articles.