These days it is all about the data.

Don MacVittie, 2007-25-07

The other night I made a fateful error on one of my hobby sites. I cut from Microsoft Excel and pasted into an HTML window.

Now I know better than that. Office products create some of the most bloated and convoluted XML/HTML known to man. Their attempts to be generic *and* control presentation make the HTML output of any MS Office product unwieldy at best. That's not likely news to you. Other products are bad, MS Office products are legendary.

But I had been working on this article forever, it was 14 pages long before the spreadsheet was inserted, and I didn't feel like writing the table HTML myself. It was also 2am and I had to work the next day - and this was the last thing I had to do. So I tried the shortcut.

No surprise, there was a ton of white space between where I inserted the table and the text that was supposed to be right after it. If that's the worst I got out of it, I was lucky. But the white space wouldn't delete. So I went and looked at the HTML source that had been inserted.

I got a firm reminder to *never* paste from an MS Office product into HTML documents. There was bloat a-plenty, and I had to wade through it to find the cause of the blank lines. It turns out that they filled a bunch of empty cells with font formatting information and a carriage return. I removed the carriage returns, and the white-space went away.

That didn't relieve the bloat, something I will be going in and fixing tonight. This article is already huge, I don't need fonts defined for empty cells, or color references that define the color black for text that would have been black anyway.

F5 (and competitors) have products that can help with this type of problem. Our Web Accelerator product could cut down delivery times in a variety of ways, and our WANJet also offers options.

But that's not the solution. What we need is to hold those vendors who auto-generate HTML and XML to a strict standard of non-bloat. We need to force them to remove the waste, and give us just plain old HTML or simplified XML, thus saving our Ethernet cables from melting.

Unfortunately, I cannot think of a single group that could enforce such standards. OASIS could for Open Source, but for commercial software? They're going to do what makes the best business sense to them, and until we are sending so much of a given applications' data over the wire that we begin making buying decisions on bloat, they're simply not going to put man-hours that could be used for product enhancements into bloat reduction efforts.

So what do you do? You look into products like our Web Accelerator and WANJet - or their competition - or you don't allow users to do things like "Save as HTML" from an Office product and then put it on the Intranet (or Internet, depending upon usage). That's tough to enforce, so looking at the products is probably your best bet.

Slow web-sites should be a thing of the past... But they're not, and bloat is part of the reason. Lose some weight, put yourself on an HTML/XML diet, or get an exercise machine like WANJet or WebAccelerator - at least that kind of exercise machine you'll use once it's in place.

Or we can start a consortium to ridicule vendors and open source projects with overbloated applications. That might be more fun, but it's a slower solution... And might make a few enemies. Not that I've ever worried too much about angering people who were wrong.

The addendum to this story? Today I noticed that the entire first column is missing from that table in the published version of the document. So much for the easy answer, tonight I get to fix it. This time, I'm writing the table by hand.

Don.

Reading: *Panzer Aces* by Franz Kurowski

Imbibing: *Water*

F5 Networks, Inc. | 401 Elliot Avenue West, Seattle, WA 98119 | 888-882-4447 | f5.com

F5 Networks, Inc.
Corporate Headquarters
info@f5.com

F5 Networks
Asia-Pacific
apacinfo@f5.com

F5 Networks Ltd.
Europe/Middle-East/Africa
emeainfo@f5.com

F5 Networks
Japan K.K.
f5j-info@f5.com

©2016 F5 Networks, Inc. All rights reserved. F5, F5 Networks, and the F5 logo are trademarks of F5 Networks, Inc. in the U.S. and in certain other countries. Other F5 trademarks are identified at f5.com. Any other products, services, or company names referenced herein may be trademarks of their respective owners with no endorsement or affiliation, express or implied, claimed by F5. CS04-00015 0113