

[**COLUMNS**](https://adtmag.com/Articles/List/Columns.aspx)

**Q&A: Akamai takes grid computing to the Edge**

* By Will Kilburn
* 09/29/2003

**Q&A with Bill Weihl, Akamai Technologies**

|  |
| --- |
| http://www.101com.com/images/pub/adtmag/billWeihl_1261.jpg*Whether you call it autonomic, self-healing, distributed or utility computing, more and more companies are trying to find a way to harness the collective and ever-growing power and speed of computers and the networks that connect them. ADT’s Will Kilburn recently spoke with Bill Weihl (above left), CTO at Akamai Technologies, about EdgeComputing, his firm’s foray into the new space.* |

**Q: How does EdgeComputing distinguish itself from your core business of delivering static content?**

A: [Initially] what we did was to provide a giant ‘shock absorber’ for handling those variations in load on static content. What we’re doing with EdgeComputing is providing the same kind of shock absorber for handling load on applications. So if the load goes up, we will, in real time within a very small number of seconds, bring more instances of that application up on more machines and map the user requests to that other set of machines. We’re doing that today for a number of customers, enabling them to essentially pay for what they use, pay ‘by the drink’ in terms of the infrastructure that they’re using.

**Q: What is on the hardware end?**

A: This is joint work with IBM. We’ve evolved a very close technical partnership, and also a sales and marketing partnership with them over the last couple of years to put WebSphere Application Server on our Edge Servers.

From the customer’s point of view, this looks like a virtual extension of their enterprise data center. Normally, they would have a cluster of machines running WebSphere; today, using us, they would have a significantly smaller cluster of machines running WebSphere in their enterprise data center, and then a virtual extension of that onto our network where the application will run on some set of machines, and that set of machines will vary over time depending on load.

**Q: How do you predict, and not get overwhelmed by, those peak loads?**

A: We deal with that the same way we did for static content delivery. We have a lot of experience from the last five years in understanding customer load patterns and, in fact, not everyone is at peak load at the same time.

One of the things to understand about on-demand is that it’s not so much about the daily or weekly variation in load. That’s fairly regular. The customer response to marketing promotions, or world events like war or the blackout in the eastern half of the U.S., do not drive traffic cyclically. The large, unusual peaks, while they may be very large for an individual company, are typically relatively small for us as a percentage of our total load.

**Q: So the idea is to be adaptive and reactive here, rather than harnessing that power on a planned basis?**

A: If you look at what people have done with grid computing to date, most of the work has been focused around large, technical computations of some sort. An oil company with seismic data needs to take terabytes of data and crunch it on thousands of processors for several days to figure out where to drill their next oil well. They don’t want to have those processors sitting all the time, to use them three times a year for these enormous computations. They want to be able to essentially rent that capacity -- tap into a utility grid to use that when they need it.

What we’re doing is something that other companies in this space-- IBM, HP, Sun -- are beginning to move toward: real-time management of resources in response to load that’s driven by the real-time business processes.