Wladawsky-Berger on Linux and open standards

Irving Wladawsky-Berger, vice president of technology strategy, IBM Server Group, addressed the IBM Technical Developer Conference in San Francisco on Linux and open standards. Below is a transcript of his remarks.

Before giving an overview of our Linux strategy, I’d like to spend a few minutes telling you why it is that we in IBM have so aggressively embraced Linux and the whole concept of open-source software.

It is not some kind of New Age business model left over from the “dot.com” bubble that we fell prey to. We believe very, very strongly that open standards and open-source software are absolutely critical foundations for the IT business going into the future. And as a result, they are a very critical foundation for IBM.

Now why do we say that? Well, in the end, this is an information technology industry. And everything in the industry starts and stops with the march of technology.

Let’s take a look at what’s going on in technology. No matter where you look, whether it’s microprocessors, memory, storage, bandwidth, or displays, the rate of advancement is nothing short of spectacular. Not only is Moore’s Law alive and well, continuing the very boring flat lines you all see when you plot price/performance or capacity against time on semi-log paper, but in a number of cases, it is turning super-exponential, meaning that in areas like storage or bandwidth, every year the growth is a little faster than the year before and therefore the curve starts trending upward.

Now, the implications of this continuing improvement -- fifty, sixty, seventy, even one-hundred percent in some cases -- are really quite staggering. It means that the price/performance of microprocessors keeps dropping incredibly to the point where they are permeating all kinds of new appliances. You can easily foresee a future in which there will be large numbers of new appliances all taking advantage of these microprocessors, all this sophisticated software, all connected and integrated together.

But we also see at the other end some incredibly powerful supercomputers and storage and very big transactional systems, so the net result is this incredible advancement and growth in the technology in just about every possible dimension.

And if you look at the infrastructure absorbing all that technology, it is growing by leaps and bounds. In fact, the technology only has to advance sixty percent a year, well within the bounds of what’s going on, for us to say that five years from now, the infrastructure will have ten times more technology. And as a result of that, the infrastructure will be far more powerful, far more scalable, far more diverse, far richer than anything we have today.

That is incredibly exciting. It is also incredibly scary, because there are some very serious challenges that we need to address as this infrastructure grows by leaps and bounds.
One of them is, how do you manage it all? And that's why, earlier this year, in IBM, we launched our Project eLiza to develop self-managing capabilities in our servers. And while I’m not going to talk about that today, let me just say that one of the reasons we feel very comfortable in pursuing a strategy of self-management is because of the technology that we have and its declining cost.

We can leverage that technology for automatic intrusion detection, lots of caching work, lots of availability and redundancy and build that into everything we have. So not only is it now feasible, it is absolutely necessary if we're to move into the future.

But the second major grand challenge is what I most want to talk about -- the challenge of integrating, in the most flexible manner, this very rich and diverse infrastructure. How do you make sure that, as the Internet and the business infrastructure keep growing, all the pieces not only continue to connect to each other, they integrate even more tightly than they do today -- something all our customers tell us is necessary.

Well, the only way to make sure that these diverse pieces connect and integrate with each other is to base your software and all your protocols on standards, period. It’s no more complicated than that. Unless these diverse pieces observe standards, whether it’s TCP/IP or XML or WSDL, they won’t be able to connect, and therefore it is critical that all of the software we build, all of the products we build, be based on standards.

But, that’s not enough. In a number of very important cases, we need not just software with standards, but open-source software. The whole world of open-source software is actually open to quite a bit of discussion. There are some who think that it’s very good, and others who think it is very bad and have talked about it in somewhat disparaging terms.

There are many different points of view. Some people think of it as an ideological movement. That’s a small minority. The vast majority think of it as something very pragmatic, as what has to be done in order to move forward and make progress.

Let me talk about two major reasons among quite a number why I believe open-source software is so important in all the work that we’re doing to connect the infrastructure.

First the open source model is how research communities work together for innovation.

If you look at professional communities, whether it is high-energy physicists, organic chemists, or any other scientific discipline, there is a very long tradition of researchers writing papers and publishing them openly. Their peers then read the paper, write additional papers on the topic, and everybody builds on everybody else’s ideas.

The sharing of this information, whether you’re in a university, a research lab, or the private sector, is what advances research and innovation for the benefit of your community.

Now, when you write papers, it’s not enough to just be able to write them in English. In fact, as we all know, if you’re writing a paper in physics, you need additional means of expression, which is where mathematics comes in. Mathematics is a very good example of a “language” that enables physicists to share information in a way that every other physicist can understand and manipulate.

Likewise, more and more of the work we do can only be expressed in software.
In fact, if you try to write down complex software in English expecting others to know what you mean, it's just not going to work. So just as math is another language for expressing a lot of important notions in physics, software is another language for expressing a whole different and very important class of problems across many disciplines, especially computer sciences.

It would be ludicrous for me to try to cooperate with physicists and tell them, "Look, you can only read the English. You're not allowed to see the mathematical equations or software, which is the result of our research." Nobody would be able to work that way.

So linking communities that are producing innovation is one major reason why open-source software is so important, and not surprisingly, a lot of major open source advancements have come from that community.

There is a second major reason I'd like to talk about, which is that as we need to integrate this infrastructure, we need a layer of integration, with standard protocols and with ways for those protocols to communicate with each other, ways that all systems share. As our protocols get more and more sophisticated, whether they are the evolution of TCP/IP, whether they are the Globus protocols in Grid computing, the various protocols in web services, and on and on, it is very important to write those protocols down as software that's available to all. That way, when you put that reference software in your system, you have precisely what I have in my system, even though we may have totally different architectures and totally different vendors.

The need to keep advancing that horizontal layer so that it works across this huge diverse infrastructure is the second major driver of open-source software. And as the protocols get more and more sophisticated, not surprisingly, there are more and more communities that form to collaborate on those protocols, be they the Apache community or the Gnome or KDE communities, or the Globus community in grid computing, or all the other communities.

Does that then mean that all software has to be open-source? Absolutely not. Only that software that is really critical to your objective, whether it's research or whether it's horizontal integration, has to be open-source.

It's totally reasonable for everything else to be proprietary and to then compete on that basis in the marketplace. It's how IBM's Software Group makes billions in revenue. It's how companies and application developers build their infrastructure.

In fact, the better we integrate all the pieces -- open source and proprietary -- the more the industry advances and the better everybody can compete with their proprietary software, applications, services, and all the things that drive the IT economy.

The two coexist very well. Nothing probably says it better than what we have done with WebSphere. WebSphere, as you all know, is based on Apache and also integrates a number of other major standard and open-source software components, from Java servlets to XML parsers to the SOAP and UDDI components.

With open source software, we implement those components that are needed for the integration, we build proprietary software for those pieces that are not, and everything works very well.

The operating system -- the Linux kernel -- is open-source, so that it can work on everybody's architecture. Then on top of Linux, you build system management software, file systems, you build WebSphere, databases, all kinds of applications, which of course don't have to be open-source at all. It's a world
in which pragmaticsm rules: You use one for what it does best and the other for what it does best.

Now let’s talk about Linux. Linux continues to grow very, very fast, and in fact it continues to be the fastest-growing operating system out there. There was recently a bit of controversy about how fast Linux really is growing. But when the answer finally came back from IDC, they said, when you count all the preloads, you count all the downloads, you count all the Linux that works on new systems as well as in older systems, all-in-all Linux is far and away the fastest-growing operating system. Linux continues to advance, and there is no question that it is already the second high-volume platform in the industry behind only Windows.

Now in IBM, we are committed to embracing Linux across everything that we do. Linux runs on our Intel servers, on our Power-based servers, on our iSeries. It runs on our mainframes, on our OEM technology, on our storage servers.

Linux, as we know, is the only operating system of which you can safely say: It will run on architectures that have not yet been invented. You couldn’t say that about any other system, but Linux can be adapted to just about anything out there.

In addition, our software runs on Linux, and all of our services have been adapted to Linux. In other words, it permeates just about everything we do, and that is one of its main appeals.

Now, let me talk about the key marketplaces that we see with Linux, the four major segments where we see our customers using Linux today.

Number one is workload consolidation. One of the capabilities of Linux is that it makes it very easy to take work that has been distributed and consolidate it in larger servers. In fact, one of the big appeals of Linux on our mainframe, the zSeries, is the fact that our mainframe supports hundreds of thousands of virtual machines, each one running its own copy of Linux. That makes it a very good integrating platform for lots of distributed workloads -- for one thing, because it provides a much lower total cost of computing.

Linux also is superb in clusters, like the big supercomputing clusters that a number of universities and research labs are running. The rack-and-stack kind of clusters is ideal for Linux, because it scales horizontally very, very well.

The third style is the world of distributed enterprise. Not only can you consolidate distributed servers using Linux, you also can distribute servers with Linux. The distributed enterprise, where enterprises run large numbers of distributed servers, all centrally managed, is among the chief growth areas for Linux.

And finally, and perhaps most important for the future, is the world of appliances. In this world, you find Linux as an embedded operating system in all kinds of new applications, many of them major network servers, many of them file and print servers, and quite a number of them new kinds of information appliances.

We have customers in production now in all those categories, as well as others doing quite a few pilots.

Let me talk about some of them in each category. Winnebago Industries is a very good example of consolidation. They found it much more convenient to run their e-mail servers in partitions in a zSeries using a Bynari software package that’s totally compatible with exchange than they did in distributed e-mail servers. They found they saved quite a bit of money by that level of
Korean Airlines made a decision to consolidate some of their major distributed applications onto a zSeries. For example, the applications that schedule flight crews and run the Web site that flight crews access to see when they are flying all now run on Linux.

In Linux clusters we have a growing number of customers. The seismic industries -- Shell, Western Geophysical and a number of others -- are already doing very powerful seismic analyses with Linux clusters. Just last week the National Science Foundation announced the award of their Terascale Distributed Facility to four major centers: the National Center for Supercomputing Applications at the University of Illinois, the University of San Diego, Cal Tech and the Argonne National Lab.

The combined power of all the supercomputers, which are Linux clusters, is thirteen and a half teraflops, roughly eight in Illinois, four in San Diego, one at Argonne and about half a teraflop in Cal Tech. And these Linux clusters are all brought together as a grid.

In the world of distributed enterprise, Thrifty Car in Australia put Linux-based servers in each of their car rental offices, all centrally managed to cut costs and make sure that those distributed servers integrate very tightly with their central UNIX servers that are running mission-critical applications.

And finally, a growing number of companies are doing Linux appliance servers, from Akamai, which uses it in their caching servers, to Turning Stone Casino, which is actually creating a number of Linux gaming servers.

The number of customers keeps growing and more and more applications are coming alive.

We've also done quite a bit of market research on these and other customers. For example, we recently surveyed twenty-seven-hundred businesses, some very large, some very small, some classic businesses, and some new kinds of businesses. We asked what they found appealing about Linux, and this chart shows four key attributes that they like.

First, it's very cost effective. Second, it's very, very reliable. Third its performance is superb. And fourth, open standards facilitates their operations.

We also looked at what applications they've implemented and where they heading in the future, and you can see that the core of the Linux applications are in the Internet space: Web application servers, network servers, and so on.

But over the years we expect Linux to start also moving to more enterprise applications. And to facilitate that, a lot of work is going on in IBM with the Linux community and with other companies to help Linux move into the enterprise.

We have the Linux Technology Center, which is an organization in IBM with over 200 people that actively open-source code and provide skills for the Linux community to help it scale. We are working actively to make Linux 2.4 better in four-way SMPs and to help push it into eight-ways next year. We are doing journaling file systems, scheduling, and lots of other technologies.

Likewise, there is the Open-Source Development Lab, where IBM, Intel, HP, Hitachi, Fujitsu, and a number of other companies collaborate to help test Linux. We have a major center in Beaverton, Oregon and there is a satellite center in Japan.
There are quite a number of activities underway in the community to help scale Linux so that it becomes a better operating system in the enterprise.

Applications are critical here. And I'm very glad to say that the number of applications supported in Linux keeps growing.

Certainly, it's not enough. One of the signs of maturity in an operating system, is its portfolio of applications and with time and an even more robust Linux, the portfolio will continue to grow. But we have a superb base to build on, and that base keeps growing.

Finally, we have quite a number of programs to support Linux across all of our developers. Later this morning, Rich Michos will have a session on Linux, and he will go into a lot of detail on a number of these programs. And then on Thursday, Scott Handy will talk about the Linux programs we have for the software community.

To wrap up, we really see a very positive picture for the whole IT industry. We think technology will keep advancing at a rapid pace. We see the integration of that technology bringing us all kinds of new capabilities, and we very much see a very strong role for software based on standards, for Linux, and for open-source software in general.

Thank you very much.