



RUNAS RADIO



<http://www.runasradio.com>



Richard  
Campbell

RunAs Radio is a weekly Internet Audio Talk Show for IT Professionals working with Microsoft products. The full range of IT topics is covered from a Microsoft-centric viewpoint.



Greg  
Hughes

*Text Transcript of Show #0108*  
(Transcription services provided by [PWOP Productions](#))



**Pedro Manfredi Does High Performance Computing!**  
**May 6, 2009**



## Pedro Manfredi Does High Performance Computing! May 6, 2009

[Music]

**Brandon Wenn:** From [runasradio.com](http://runasradio.com), you're listening to RunAs Radio, the Internet audio talk show for IT professionals with Richard Campbell and Greg Hughes. This is Brandon Wenn, announcing show #108, with guest Pedro Manfredi, recorded Thursday, April 16, 2009. RunAs Radio is produced each week by PWOP Productions, providing professional media and podcasting services online at [pwop.com](http://pwop.com). You can follow the boys on Twitter at [twitter.com/runasradio](http://twitter.com/runasradio).

**Richard Campbell:** You're listening to RunAs Radio. I'm your host, Richard Campbell. With me as always, my co-host Greg Hughes.

**Greg Hughes:** Hey everybody. Richard, how are you today?

**Richard Campbell:** I'm doing just fine, sir. Busy, busy, busy. You know, my house is almost finished.

**Greg Hughes:** It's amazing.

**Richard Campbell:** It's been over a year.

**Greg Hughes:** Yeah. How long has it been? It's been I think like 11 years? No, not that. It has been more than a year though, hasn't it?

**Richard Campbell:** Yeah.

**Greg Hughes:** It's been quite the saga actually.

**Richard Campbell:** Well, we're getting what we wanted, man. I'm building my dream home but it's taken -- we moved out January of last year.

**Greg Hughes:** Yeah, that's awesome.

**Richard Campbell:** So hopefully by the end of May it will all be over but we'll see, it's still going on.

**Greg Hughes:** Well, congratulations on that. I know that's a big deal for you guys.

**Richard Campbell:** Yeah. We're trying to finally get back to normal life. We are getting some emails once in a while. Folks are actually asking if we were going to be at TechEd and I had to tell them that unfortunately you're not, you've got other commitments. We'll be doing Speaker Idol and all the usual things at TechEd so we are going to be there. Well, Greg is going to be there in spirit, I'll be there in person.

**Greg Hughes:** Yeah, that's a good way to put it.

**Richard Campbell:** If you want to find me, you'll probably find me at TechEd online area, that's usually where I hang out most of the time so come by and say hi. If you got any questions, concerns, ideas for shows, criticisms, things you'd like to see, send us an email, [info@runasradio.com](mailto:info@runasradio.com). All right, Greg, let's introduce our guest. Pedro Manfredi is a Microsoft veteran, currently senior technical evangelist and managing Global ISVs with the Microsoft Developer & Platform Evangelism team. Before joining DPE, he held several positions at Microsoft: Architect Evangelist, Principal Consultant, Technical Account Specialist, and Support Engineer. Welcome, Pedro.

**Greg Hughes:** Hey, Pedro.

**Pedro Manfredi:** Thank you very much for having me.

**Richard Campbell:** So you know, looking at that list of titles, you've sort of grown through the company. Started on the support engineer side and sort of moved into evangelism.

**Pedro Manfredi:** Yeah. Actually I started providing support for some arcane Microsoft products like Microsoft C, COBOL, Microsoft Fortran.

**Richard Campbell:** Wow.

**Greg Hughes:** Wow.

**Pedro Manfredi:** Believe it or not, Microsoft used to offer a Fortran compiler.

**Richard Campbell:** I remember. That's back in the pre-Visual Basic and the whole Windows side of things when there was Microsoft Professional Basic, Professional Fortran, Professional COBOL, and Professional C.

**Pedro Manfredi:** Yeah, that says a lot about our age.

**Richard Campbell:** Yeah, it does, doesn't it?

**Greg Hughes:** When did you start working at Microsoft, Pedro?

**Pedro Manfredi:** I started back in Brazil in 1992 as a support engineer, and then after 12 years in Brazil I got a promotion or whatever you call it, but moved to Florida to manage the architecture evangelism team there back then, and two years ago I got this position here at the headquarters.

**Greg Hughes:** Up in Redmond, yeah.



## Pedro Manfredi Does High Performance Computing!

May 6, 2009

**Richard Campbell:** And now these days you're focused on High Performance Computing?

**Pedro Manfredi:** Yes, yeah.

**Richard Campbell:** Isn't all computing high performance computing?

**Pedro Manfredi:** That's a good question. Usually we hear from our partners and even customers that they want the products to run on high performance computing and not on the usual let's say low performance of Windows. So there's a very particular approach to doing this type of computer station and that's where the HPC product from Microsoft tries to fit in.

**Greg Hughes:** So what are we talking about then? It sounds like we have normal performance computing and then high performance computing, but what exactly do we mean when we say high performance computing?

**Pedro Manfredi:** This is the type of computing that's usually being used in engineering, automotive, aerospace, or biochemical, those are academics in government. This is a type of computation that usually you measure response time in hours or even days instead of sub-seconds or seconds.

**Richard Campbell:** So doing stuff like simulating airflows over wings and cars, that kind of thing?

**Pedro Manfredi:** That's pretty much it, yeah.

**Richard Campbell:** So I mean, we use to call it supercomputing, didn't we?

**Pedro Manfredi:** Yeah. Sometimes we -- that's interchangeable, supercomputing and high performance computing is pretty much the same thing.

**Richard Campbell:** Okay, but it used to be supercomputers were built by Cray and they lived in their own buildings with their own power supplies and...

**Greg Hughes:** Those were cool computers.

**Richard Campbell:** They were very, very cool but they were hugely expensive.

**Greg Hughes:** They look like big round couches that weren't very comfortable.

**Richard Campbell:** Yeah.

**Pedro Manfredi:** Yeah, that's the history of the supercomputing or high performance computing.

Nowadays, you can have a cluster of computers working for high performance load built out of market servers that you can pretty much buy from any vendor.

**Richard Campbell:** Right. So I ganged together a bunch of Dell 1U pizza boxes and make it my own supercomputer out of it.

**Pedro Manfredi:** Yes, you can.

**Richard Campbell:** So then this really is just a software thing, is it?

**Pedro Manfredi:** It is pretty much software but when it comes to really the top tier high performance software that runs on top of those machines, you have additional ingredient which is the Interconnect. The network interface that powered those high performance computers is usually InfiniBand or Merinet, those are high-end network cards that provide you with very low latency and lots of true-put.

**Richard Campbell:** I've heard of InfiniBand before. This is a card essentially that has a really, really high speed networking?

**Pedro Manfredi:** Yeah, it's a network card with specific implementation and usually you paired the network card with proper switch, InfiniBand switch, and that's going to provide you with a very high bandwidth and very low latency.

**Richard Campbell:** Okay. I think it's all about speed, speed, speed. This is not necessarily like fiber optic cable or anything like that. It's just a particular kind of networking?

**Pedro Manfredi:** Yeah, actually you can -- usually you have the physical medium that you use to connect those cards to the switch, it can of course fiber optic but the magic is actually on the card and the switch.

**Richard Campbell:** So how much faster is something like InfiniBand over regular good old Ethernet?

**Pedro Manfredi:** It can have, for example, latencies in the realm of two or three seconds and true-puts of 100 megabytes per second depending on the configuration so it's really, really fast.

**Richard Campbell:** Wow, okay. So obviously for ganging a whole bunch of machines together and connecting them for high speed communications like that, we have to have some kind of workload that parcels itself out nicely across many machines.



## Pedro Manfredi Does High Performance Computing! May 6, 2009

**Pedro Manfredi:** Yeah. Usually a high performance cluster, we resort to some sort of iScheduler or an application that's going to help administer the cluster and provides the administrator with all the support in order to maintain the clusters up and running. It starts with the deployment of those clusters. You can have a few nodes to thousands of nodes. So you got to have some sort of infrastructure that allows you to easily deploy from bare metal to an up and running operating system in a very short time.

**Richard Campbell:** Okay and I guess, yeah, the big thing here is this is still the same sort of work we used to do with the Cray's but it's less expensive so more people can use it. Is the workload really changing much like what do people need to do with these things?

**Pedro Manfredi:** The workload is pretty much the same. Like I said, we are talking about engineering automotive, aerospace, or even finance sector doing all those pricings of derivatives and stuff. The good thing with these new offering from Microsoft, we were talking about the interconnect that are very powerful but you can also view the under the desk cluster based on standard hardware and you can achieve very good performance and the goals of lowering the time tool solutions using Windows HPC.

**Richard Campbell:** Isn't multi-coring of computers in general just driving us this way anyway? I mean, we're already buying four cores, quad-core laptops.

**Pedro Manfredi:** Yeah, this is a good question. Of course you can have parallelism on the same chip as when you're buying a multi-core, but usually when we talk about high performance computing we are talking about distributed memory. In that case, you have many different computers put together with the distributed memory of those computers and with the software solution you can manage all those computers at a single advantage.

**Richard Campbell:** So literally shared memory, that's what the InfiniBand stuff is about, it's just being able to share memory efficiently.

**Pedro Manfredi:** Yeah. There is a distinction. When we say shared memory, you have a single computer, a bunch of memory, and you can have multiple processors or multiple CPUs and cores sharing the same address space, the same memory. When we talk about distributed memory, we consider different CPUs, different computers but the solution is going to take advantage of that distributed memory. Another set of CPUs and memory and you put together many tools and you work like it were a single solution.

**Greg Hughes:** And distribute out the units of work across those nodes.

**Richard Campbell:** Yeah. I'm still trying to get my head around the difference between this and stuff like SETI@home which was just massively distributed as well.

**Greg Hughes:** Maybe that's a good comparison to how, if you take it to SETI@home, which a lot of the geeks, those of us out there at least know about that and a little bit of how that works, how does that compare to what you're talking about?

**Pedro Manfredi:** This set of work is much more great computing type of solution in which you just offload the processing of a set of data to a remote computer, but you don't have a single vision of all the computers, you just get the result out of the computation but you don't have the vision of what's currently being done and you cannot administer, or for example if you want to put together different machines on a single set of nodes, for example, and you want to administer that set of nodes as a single entity, you don't have that on great computing as it is right now.

**Greg Hughes:** Right. You're chunking up the data into small chunks, sending them out to each of your nodes that have agreed to do work but they're doing that work individually and then shipping the results back...

**Pedro Manfredi:** Yes, correct.

**Greg Hughes:** As oppose to they're not working in concert with each other.

**Pedro Manfredi:** No, no, they don't know about each other.

**Richard Campbell:** And so there's a separation of work and there's a separation of work. This is the idea of not being that separated, that every node has to interact with the other node very tightly.

**Pedro Manfredi:** Yeah, they are working on the same computation and coordinating to get the solution.

**Richard Campbell:** This still sounds like the challenge of HPC is really finding the right project and how to craft that project to harness the hardware.

**Pedro Manfredi:** Yeah. The hardware piece of the equation is pretty much solved because like you said we don't need those specific built computers to create the cluster. You can do that from a standard hardware that you buy. For example, there's the logical disk that has a reference. You can build a



## Pedro Manfredi Does High Performance Computing! May 6, 2009

cluster out of the Mac mini. You can build a Windows cluster, Windows HPC cluster out of the Mac mini.

**Richard Campbell:** That's hilarious.

**Pedro Manfredi:** I can later provide the link to this, there is a good article in the Microsoft site that explains that. So you pretty much can have this cluster out of standard hardware. That's the big difference when you compare with the previous high performance computing, those supercomputing approach in the past.

**Richard Campbell:** Okay.

**Greg Hughes:** How big can we go with this? How many actual individual pieces of hardware and how big can it get? Is there really a limit to it?

**Pedro Manfredi:** This is a good question. We have right now, there's a list of top 500 of the most powerful computers in the world and Microsoft is the number 10 on that list and this is a cluster of 1300 nodes when we are talking about 5600 CPUs on a single cluster.

**Greg Hughes:** Wow.

**Pedro Manfredi:** So it can grow, there is no hard limitation, logical limitation on the number of nodes. Usually you find on corporation, you find from number of hundreds to thousands of nodes. The challenge is how to manage that level of infrastructure.

**Greg Hughes:** Sure. Just to help paint a picture, can you maybe give just one good solid example of where this type of computing environment really pays off? Do you have an example, like a real world example?

**Pedro Manfredi:** Yes. For example, let's take Boeing for example. They have big clusters, they are doing these new depths and they need to evaluate the friction of the air on the geometry of the plane so this is the type of computation that requires probably hundreds of nodes working in conjunction in order to provide a solution to the problem.

**Greg Hughes:** Because of the complex geometry of the plane and the complex or the possibilities of what the air could do as a fluid. So you're doing the fluid analysis and what-not.

**Pedro Manfredi:** Yeah, we are talking about the molecules on the air interacting with the different molecules on the plane...

**Greg Hughes:** Right.

**Pedro Manfredi:** The materials that the plane is based on like a steel or so.

**Greg Hughes:** That's a lot of molecules.

**Pedro Manfredi:** It's unbelievable, the size of the problem.

**Greg Hughes:** Well, I like airplane examples so thanks for that one, that's really cool.

**Richard Campbell:** My friend recently got his pilot's license so he's a little biased.

**Greg Hughes:** And an airplane. Not that there's many molecules on my airplane but, I mean we're counting molecules and I assume there's got to be work. I haven't done that stuff but there's got to be some assumptions that are made about how things work because certainly we have to do a lot of modeling of the whole fluid dynamics analysis, it could just be incredible.

**Pedro Manfredi:** Yeah.

**Richard Campbell:** I just feel like that's a very specialized area of work. I'm just trying to get a sense for why there are applications of HPC style of apps, that these under the desk styles, if you forgive me, are not that huge but maybe there are data mining opportunities. I'm just trying to get a sense for not everybody needs to figure out the wind flow over an air foil.

**Pedro Manfredi:** Yeah but for example you can have a solution if you are building a stadium, an arena and you want to model the flow of air inside this arena when you are still in the design phase. So you can use this type of software to model the flow of air and see what is the proper air-conditioning solution that is going to be able to support such a huge open space.

**Greg Hughes:** Right. I guess you can also model and do -- is it full of people, is it partially full, and all those types of things too.

**Pedro Manfredi:** Yeah, yeah, this is a good example because we have one example of a company that was building a stadium in Europe and they needed to model not only the building, the stadium, but also model the flow of people and all the prospects of a huge mass of people in the same environment. So they created with a media company that moves, rendering special effects. This company builds a model of people on the stadium and in conjunction with completely fluid dynamic application, they were able to model the more close as possible scenario of the fill up of the stadium and the flow of air. So those are the types of solutions that we are talking.



## Pedro Manfredi Does High Performance Computing!

May 6, 2009

**Richard Campbell:** Crowd modeling is a fascinating concept and definitely something where I see you wouldn't go to a supercomputing center to do crowd modeling. Likely you would do it in smaller scenarios with those smaller systems so to me that seems to be a great example of just kind of new applications in this space for it to make sense to figure that sort of thing out. So what's the developmental environment like for HPC? What do folks need to prepare for if they're going to start working in this space?

**Pedro Manfredi:** Well, in the Windows environment. This is an important area and that's where Microsoft is trying to make the Windows HPC as broadly available as possible. You have Visual Studio integrated with the Windows HPC so you have debugging and you have the many languages that are available on the Visual Studio product, but usually the languages used for those solutions are C or C++, lots of Fortran code.

**Richard Campbell:** Still.

**Greg Hughes:** Good old Fortran, yeah.

**Pedro Manfredi:** It's still Fortran. It's very likely that the latest racing cars that we see and the latest gadgets, electronic gadgets that we use, those products they were built, designed using the Fortran piece of code very likely.

**Greg Hughes:** When it comes to doing math, Fortran really is a pretty, pretty lean mean and a pretty good way to do it.

**Pedro Manfredi:** It's still the language of choice.

**Greg Hughes:** Yeah.

**Richard Campbell:** Hey, I guess that sort of leaves out the .NET development environment then.

**Pedro Manfredi:** Yeah. There is an implementation of the underlying inter-process communication that used it on the HPC space which is the MPI, Message-Passing Interface. There is an implementation of MPI for .NET so it's possible to use .NET to develop such applications as well.

**Richard Campbell:** Okay.

**Greg Hughes:** Interesting.

**Richard Campbell:** Well, I think it's an important part of this is typically going to be an IT initiative where they're going to want to scale out an app for trying to these more complex tasks and the development environment is an important part of it.

You need to build apps that can do this. Is there much in a way of a third-party market for software in HPC? Can I buy like this crowd analysis software off the shelf?

**Pedro Manfredi:** Yeah, this is something that when we compare the Windows environment with the Linux, in Linux you have to resort to different vendors in order to put together a solution. Open source vendors provide different pieces of the HPC cluster. With Microsoft, we have pretty much everything that you need, at least the base infrastructure is available from Microsoft. So you have the pieces of deployment integrated with Windows, the Windows Deployment Services. You have the administration and job scheduling with monitoring execution and so on, so you have pretty much a complete solution stack based on the Windows. Of course you can acquire additional software. For example, if you need the parallel file system, there are some vendors providing parallel file systems on the Windows HPC product as well.

**Richard Campbell:** Interesting. Okay, I guess that's part of it, it's not everybody has to be innovative here. Building something new, it's taking a look at the infrastructures out there today and saying I can buy an off the shelf solution to working this stuff out.

**Pedro Manfredi:** The idea that the administrator is not going to concentrate on the underlying infrastructure but pretty much concentrate on the problem itself, the engineering problem, the biochemical problem at hand. So the solution is integrated and available without any major burden put on the administrator.

**Richard Campbell:** Right. So they just buy it, fire it up, run it on its own. How do you connect up to HPC? Is it still that regular Ethernet card, the same old Windows GUI we know well?

**Pedro Manfredi:** Yes, it is Windows, the Windows HPC product is based on a version of Windows. You can install the HPC pack on top of Windows Server Standard or Enterprise or any version of Windows.

**Richard Campbell:** Okay.

**Greg Hughes:** Okay, so it sits on top of it then.

**Pedro Manfredi:** It's on top and other familiar administrative tools and everything that you know and love about Windows, it's there and it's available to the administrator.

**Greg Hughes:** So how does one actually acquire the tools that are used to do this? Where do



## Pedro Manfredi Does High Performance Computing! May 6, 2009

you get the stuff that you add on top of your Windows Server?

**Pedro Manfredi:** There is a special SKU for Windows HPC server. You acquire, of course if you are a corporation and you have a Microsoft representative and a Microsoft contract, you can acquire the HPC pack to run on top of Windows, or you can acquire the product itself with everything, the license for the OS and the license for the HPC pack.

**Richard Campbell:** Typically this is just a standalone rig, not a lot of work in the areas of redundancy, or do you ever run stuff like SQL Server or any other software on it, or there's just a custom app made for HPC?

**Pedro Manfredi:** Considering that we are talking about the standard Windows, you can of course run anything on top of it, but license for the Windows HPC server limits the workload that you are running to just HPC workloads.

**Richard Campbell:** Okay.

**Pedro Manfredi:** If you want and you need to use the same infrastructure to run additional Microsoft or to debug a software, you have to get the specific license of Windows. You can have Windows Standard or Windows Enterprise and on top of that have the HPC pack plus the applications that you want to run. It's not a very common scenario because given the demands that those applications require, you're not going to see more than the HPC itself on the same box.

**Greg Hughes:** Sure.

**Richard Campbell:** Right. This is the ultimate example of dedicated hardware.

**Pedro Manfredi:** It is, it is, yeah.

**Richard Campbell:** And certainly I can see any application working with the SQL Server on a different box to get its workload and to write out its results but it's not going to be on the same box.

**Pedro Manfredi:** Yeah, that's probably the most fitful scenario.

**Greg Hughes:** How does an application, how is it made to be HPC aware so that it will run? Maybe on the IT side of things or for the non-developer, like me, side of things, to help understand, what's the core differences between say you're a standard, runs on one machine or enter architecture, so the standard application versus this high performance one?

**Pedro Manfredi:** You can have different types of applications. You can have the parallel parlance. You have for example what they call a embarrassing parallel type of application. In that case, you can run the application on multiple nodes without the need of the nodes to be communicating with each other. That's what they call embarrassing parallel. The applications are going to work on the specific set of data and there is no coordination between the applications. Yeah, of course when the cluster to manage the launch of those applications on different nodes and monitor, the execution and the status of those applications but they don't demand any communication between each other.

**Greg Hughes:** Okay.

**Pedro Manfredi:** The more typical type of HPC application which is the distributed memory in which you have a number of nodes with probably the same piece of code working on a set of data that's shared between those different nodes trying to solve the specific problem. So you have an application that is built to take advantage of this inter-process communication and then as I said you are going to utilize the methods best in interface specification and you have on the Windows platform an implementation of the MPI stack.

**Richard Campbell:** Okay.

**Pedro Manfredi:** Does that answer your question?

**Greg Hughes:** Yeah, I think so. It just kind of paint a picture of this is the high level kind of how it all fits together. It sounds like we're talking about a pretty specialized platform for doing very specialized work for the most part.

**Pedro Manfredi:** It is specialized, you are correct but you can also run any type of application on top of it. Discoverer doesn't care about you, it just tries to run. If it decides a test that can be executable then it's going to distribute this executable on all the nodes that you specify and it's going to monitor the execution of those applications so you can run pretty much anything in parallel.

**Greg Hughes:** Okay. So if I wanted to like get really weird about it and I wanted to say I want to install this and then I wanted to try and run Exchange Server on this platform, is that completely out there and totally unrealistic or are you saying that that's something possible to do?

**Pedro Manfredi:** I would say that's unrealistic.

**Greg Hughes:** That's what I thought.



**Pedro Manfredi Does High Performance Computing!**  
**May 6, 2009**

**Pedro Manfredi:** You can try to do that. There is no any impediment or anything that's going to prevent you from doing that but it's not something that's going to give you any benefit right now the way it's built.

**Greg Hughes:** Got you, okay. So it might work but it would be using massive resources for something that maybe you will not be able to take advantage of it.

**Pedro Manfredi:** Yeah. Those are different types of parallelism. When we talk about Exchange or when we talk about SQL Server and BI for example, those are different types of parallelism.

**Greg Hughes:** Pretty interesting stuff.

**Richard Campbell:** Absolutely. Pedro, we're down to the last couple of minutes of the show. Any shout outs, places people should go if they want to learn more about high performance computing?

**Pedro Manfredi:** Of course. You can have lots of information about the Windows HPC product at the [microsoft.com/hpc](http://microsoft.com/hpc) and I just want to say that Microsoft has a very competitive product on the HPC space and this competitiveness can be compared to Linux on the deployment administration side and secure to all the integration of Microsoft tools so it's an effective solution today for medium and small or even large high performance computing blaster.

**Richard Campbell:** Excellent. Pedro, thanks so much for coming on the show.

**Greg Hughes:** Thanks Pedro.

**Pedro Manfredi:** Thanks for having me.

**Richard Campbell:** And we'll talk to you next week on RunAs Radio.