Banks and financial institutions are one of the main markets for the newest High Performance Computing products. The need for detailed risk analysis has led financial institutions to invest vast sums in data centres to calculate the best possible choices in their markets - and the ‘credit crunch’ has only increased the need for technology.

So - for a picture of where the rest of the market may follow, it’s always worth looking at what the banks are up to. After a period of steady use of in-house ‘grids’, churning through Monte Carlo calculations overnight, there are some interesting changes afoot in the market.

John Barr is an analyst with The 451 Group, focussing on High Performance Computing. “The main focus remains on risk - risk portfolio analysis, pricing of exotic options, a whole bundle of different risk areas that all come down to Monte Carlo analysis,” Barr says.

However, that analysis has speeded up in recent years, from a weekend process to overnight and now some banks are running calculations in real time, as decisions are being made.

“One model is to do the full risk calculation overnight and then modify the conclusions from that based on the real time data that’s streaming in during the day, using Complex Event Processing (CEP),” he says. CEP involves applying queries to the live data as it arrives, to pick up on any triggers or patterns and making decisions based on that.

“Combining that analysis of the real time data with continually updating your risk position enables you to make trading decisions within your trading strategy,” Barr says.

Adam Vile, head of technical consulting at Excelian, has been watching the same developments.

“Currently, in many banks, traders find out the risk on their trading book the day after they finish trading. They’re 24 hours behind in the calculation of risk. So there’s been a push towards intraday risk measurement and providing current, useful information.

“When you are receiving over a million messages a second, as the New York Stock Exchange does, that needs filtering and the key information has to be pulled out and processed, using something like CEP. Then that offers opportunities - if you can act programmatically on the millisecond timescale, quicker than anybody else, you can execute a trade or move your limit before they do, and so you have an opportunity to make money. So there are highly technical, proprietary trading capabilities emerging in banks for what’s called high frequency trading, using CEP.

“That’s connected to, but shouldn’t be
confused with, low latency trading - though you’ll have better chances of winning if you have both,” he says.

The big players in the CEP market have been shifting rapidly in recent years, Barr says, as acquisitions and mergers take their toll. Sybase, one of the principal providers, recently bought Aleri, which had previously purchased Coral8.

Microsoft and IBM, along with other large players, also have their own offerings for this market, recognising it as a growth area. Likewise, there have been changes in the grid and HPC market generally: the previous market leader, Data Synapse, was bought by Tibco in 2009 and has begun losing sales to Platform Computing’s Symphony product. Microsoft is also making significant in roads, Barr says, as its HPC stack becomes more mature.

It’s hard to keep track of who is using what, Barr says, because the financial world tends to be secretive about what it’s up to and what latest technologies are being used. The credit crunch, however, has certainly changed the way in which risk is analysed, he says.

Before the credit crunch, there was little focus on liquidity risk, for example - the risk that when you want to sell your shares, no one wants to buy them. Likewise, counter party risk - what if the organisation you’re trading with goes bust?

“One of the companies that I spoke to back then said they’d tried building counter party risk into their pricing models - but it made them uncompetitive, so they took it out again. Shortly after that, the credit crunch hit and a number of banks went out of business. So, the banks are far more concerned about doing the right thing now - both because there’s more awareness of risk and because the regulations are forcing them. The appetite for risk calculation has increased significantly,” Barr says.

Vile agrees, saying that the “regulatory requirements and, for example, the incremental Risk Charge, are driving the compute requirements by a factor of ten, just to continue to do business.”

The other big change in the financial analysis world has been a move towards the use of GPGPUs (General Purpose Graphics Processing Units) for risk analysis and the pricing of complex options.

“There’s quite a number of people using GPUs, particularly Nvidia GPUs, to accelerate a workstation. It lets them tackle jobs that previously would have required a large cluster. You can get speeds of up 30 to 50 times the standard workstation speed, and if you use an off the shelf package that’s been optimised to use Nvidia GPUs you get great performance without the complexity,” Barr says.

Power consumption is also much lower using a workstation with one or two GPUs than using a standard server rack, he says.

GPUs, of course, were designed to handle graphics and have been co-opted into general purpose computing. They are very useful for data parallel problems, Barr says.

“If you’re doing the same processing to large amounts of data, GPUs are very good at that.”

Likewise, FPGAs (Field Programmable Gate Arrays) are also being used. Effectively, each can be turned into many simple processors for handling data parallel calculations.

“They tend to run with slow clock speeds, so each calculation takes a relatively long time, but because you have a thousand of them working at one time, it’s actually much faster than an ordinary processor,” Barr says.

Banks have been experimenting with both GPUs and FPGAs for some time, Barr says. The biggest problem with them is the complexity of programming. Nvidia has done a good job of encouraging developers to create applications that companies can use, and making things as
simple as possible but to develop a new application is still challenging.

Likewise, using FPGAs has its problems. “You’ve got two problems there - you need to turn it into the special device that you need, which is complicated, and then you need to program this device that you’ve created, which is also complicated. So the complexity is quite significant - but when it works, it can be very, very good,” he says.

“They’re absolutely terrible to program,” agrees Vile. “There have been some success stories, but they’re just not going to be the magic button or panacea that solves the problem of needing ten times more compute power.”

On top of the programming challenge, software has to be written to suit the GPU architecture and that can often affect the results you get back, he says.

“There is an open standard emerging called OpenCL, for programming multicore and GPUs, but in fact, by using the abstraction layer, you lose a lot of performance. So people just aren’t using it.”

The principal difficulty is really the volatility of the financial services market, Vile says. While GPUs and FPGAs can be useful in areas like signal processing, or control systems where things don’t change regularly, they’re not suited to models that change frequently and require reprogramming.

The other big movement that Vile sees is towards cloud computing.

“That’s definitely one of the newer areas of interest, and banks are very keen to exploit it. Though there’s a degree of naivity there, and I think everyone thinks they’ll be able to buy whatever they need at low cost. But definitely cloud is under serious consideration because most investment banks now realise that they can’t maintain the size of estate that they’ve been doing up till now.”

This is likely to involve a mixture of internal and external cloud infrastructure, with banks using outside providers to provide scope for ‘burst’ capacity, rather than maintaining the overhead themselves.

“Oh obviously there are security concerns to overcome, and there are concerns around service levels, because the cloud providers right now aren’t very good in terms of the service they provide. But I think with Microsoft Azure coming up, that can make a big difference.”

Barr agrees: “Most firms have moved to some sort of enterprise grid, or private cloud, or whatever you want to call it, and there is some work happening in public clouds, although there are security concerns. I do find it ironic that the banks telling me they can’t do this, for security reasons, are the same banks telling me to do internet banking.”

It seems there’s plenty happening in the financial services technology market, and lots of business available for providers with the right products.

“It’s been surprising,” says Vile. “We all thought the credit crunch would mean a slowdown in IT development, but in banking there’s been a real increase.”

This article is produced by ICT KTN. Its publication does not imply any endorsement by ICT KTN of the products or services referenced within it. Any use of this article must include the author’s byline plus a link to the original material on the Web site.