MACHINE LEARNING THE RULES OF THE GAME

It may come as a surprise to many readers that Shell Ventures' latest investment is in Bluware, a company offering "alien technology" (its own description) based heavily on principles borrowed from the computer-gaming industry. No, Shell employees will not be playing *Fortnite* on their laptops (at least, not at work). Rather, it reflects Shell Ventures' interest in an apparently unique open-computing platform designed specifically to process, visualise and help to interpret colossal seismic datasets faster than ever before.

Houston-based Bluware, Inc. claims that its hardware and data architecture enable geologists and geophysicists to interact with seismic images as if they were playing a modern video game. Further, its built-in artificial intelligence enables seismic targets to be identified in a fraction of the time taken previously. So, how much of this is marketing hype and how much is reality? To find out, *TechXplorer* brought together Paul Endresen, Bluware's chief technology officer, and Tim Roden, manager of Shell's GeoSigns subsurface data interpretation platform, and put a series of questions to them.

What was behind Shell's decision to invest? This was not out of the blue, surely?
Tim: There have been links between the two companies for a long time. Bluware has provided Shell with software developers to support our proprietary subsurface workflows for more than 20 years, so the company has a pretty good idea of how Shell approaches seismic data interpretation.

In 2016, Bluware moved up a level when it joined forces with the two companies, Hue AS and Headwave, Inc., that were, together, developing the technology that is the basis of the current Bluware geology and geophysics offerings. (The combined company moved forward under the Bluware name.)

In fact, my team had been watching Hue and Headwave since well before the merger. I remember first seeing them in 2014 at a Society of Exploration Geophysicists conference in Denver and being immediately struck by the speed of some subsurface interpretation workflows being demonstrated on a laptop. Straight away, I wanted some of that performance!

You have to realise that seismic datasets are often very large – tens to hundreds of gigabytes. Now, consider that Netflix streams high-definition movies at about 3 Gb/h and we watch them over about two hours. On this basis, dealing with a seismic dataset is like streaming 10–20 high-definition movies simultaneously. Moreover, we want to be able to manipulate them on the fly. We struggle to do this on even an advanced technical workstation with 256 Gb of memory. By contrast, Bluware is now showing that this is entirely possible on a standard laptop. This opens all kinds of opportunities and is why we eventually took out a technology licence with Bluware towards the end of 2017.

In parallel with all of this, Shell Ventures was looking closely at all three companies. Once they merged, that interest increased. Arguably, our licensing the technology finally sealed the case for Shell taking a stake in Bluware. It has to be a good thing for both sides. We get access to a nimble partner with cool technology and some interesting friends. Bluware gets the benefit of our extensive industry experience, an insight into the real world of exploration and production. I think Paul will agree with that.

In simple terms, how does Bluware manage to achieve such performance?
Paul: Staying with the Netflix movie analogy for a moment, Netflix streams at about 3 Gb/h but a movie stored in the cloud amounts to more than 1 Tb of data. It would be hugely impractical to expect customers to download files of this size. Instead, Netflix uses data compression, marked datasets and adaptive signal quality (it focuses on what is changing in the scene from frame to frame) to screen movies on customers' TVs.

We use the same approach for seismic data processing. We compress the data in a special way and stream them at the quality the user needs for his or her workflow. The key to this is being able to randomly access the data and query it in the compressed form before pulling down the minimum amount of data at the optimum signal quality necessary for a chosen element of the workflow or a particular visualisation. The technology mimics the seismic data interpreter, who rarely wants to examine all the available data in great detail at once. He or she (or maybe a computer or an algorithm) will jump back and forth in the data and pick the necessary bits.

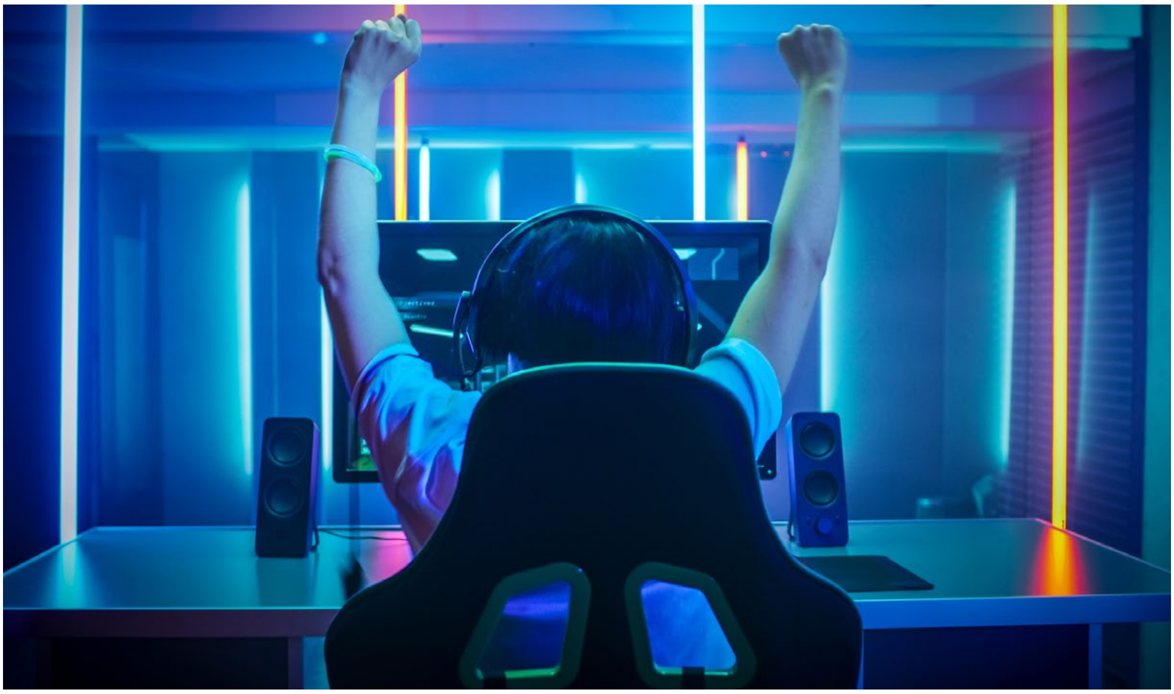
I must stress that it is much harder to do this than it sounds. It has taken us 12 years to perfect the necessary data compression and random-access capabilities. No other company has done it to our knowledge, at least for this sort of oil and gas application. Our preference is that we perform all this in the cloud, but we can easily install the capability at a user's premises.

Tim: The standard way of coping with large seismic datasets, without taking hours and hours, has been to throw hardware at it. Shell certainly did; we just bought bigger and bigger computers. Bluware's approach is much smarter. It truly is impressive how Bluware has made the data compression and random access work together and how it has built them into a usable platform on which we can build our own seismic workflows.

So, what is the link with the gaming industry?
Paul: Well, I started my career in the 1990s in the gaming industry, which is where the technology emerged. Game developers use data compression together with the latest graphical processing units (technology that Bluware also leverages) to enable players to take part in, or indeed to create, the action on the screen. Any time lag would spoil everything. Again, it involves processing only the data that matter to the user. It is the only way for so-called first-person shooter games to work on relatively simple game stations.

Another feature of the gaming industry is that the product development process has split. Instead of individual companies developing all the aspects of a game, there are now specialist gaming engine developers and separate content developers, or game designers, that use those engines to power their games. The gaming engine does all the computer science: the data handling, memory stuff, graphics rendering, visualisation, sound production and so forth. And, because the engine companies are very focused on what they are trying to do, the products are as good as they get.

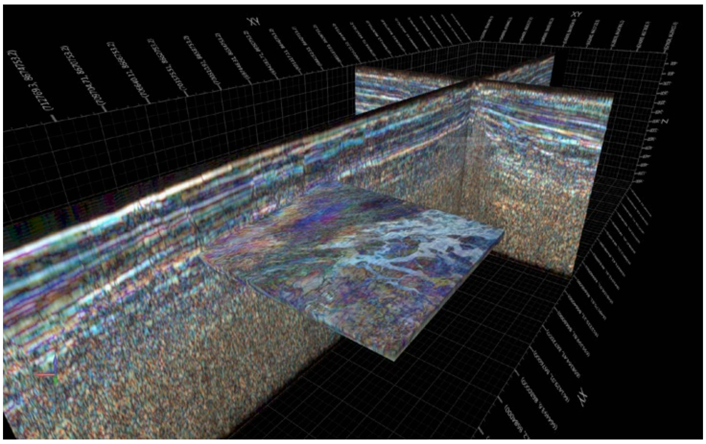
Bluware is applying that same approach in the seismic space. We have concentrated on providing a high-performance "seismic engine" – though we prefer to call it a platform – on top of which customers can design and build their own workflows for the geology and geophysics applications that matter most to them without worrying about all the underlying computer ▸▸▸



science. The platform fully supports Python, the computer language favoured by advanced data scientists for its productivity and flexibility. And we also guarantee that any developments and updates to the platform will not entail users having to change what they have built above it.

Tim: This is definitely a move in the right direction. Shell is not alone in having tried to do everything itself in the past: the data processing, the computing, the algorithms, etc. Actually, we should be focusing on the seismic interpretation "games" themselves and be leveraging the computing skills and, especially, the performance companies like Bluware provide.

Bluware also claims that machine learning is part of its platform. What role does this play?
Paul: We are actually referring to a subset of machine learning called deep learning. A good example is teaching a machine to pick out and differentiate between cats and dogs in various scenes. The challenge here is to ensure the images that you use to train the machine are completely random and that they reflect reality (one would expect the number of cats and dogs to be roughly equal). At the risk of oversimplifying the problem, if you inadvertently train the machine by showing it 90 dogs followed by 10 cats, it will then identify most pets as dogs. You will have introduced a bias into the machine.



Extremely fast data access enables complex processing algorithms to run in real time; here, they are turning seismic amplitude data into a picture of rivers that existed millions of years ago.

These days, most deep learning applications involve spending an inordinate amount of time massaging and validating the input data before machine training can begin. We are saving that time by using our random-access expertise to create valid training input. Compression also comes into it: that is, our ability to handle vast amounts of data very quickly. We can label data in real time and learn as we go along. Nobody else can do this. We call it interactive deep learning.

Tim: Machine learning is moving faster in the subsurface arena than I ever thought it would, principally for the identification of key geological features in the data such as faults, horizons and salt domes. But, following on from what Paul said, we spend a massive amount of time teaching a machine the patterns to look for that indicate, for example, the presence of a fault. We then send the tool out to users in Shell's various assets. The results we get back are often mixed because the performance of the machine depends on how closely the patterns in the seismic data for a particular asset resemble the training set. The interpreter often has to finish the process manually by adding features the machine missed and correcting any mispicks.

Paul: Bluware's approach makes that initial, time-consuming and, sometimes, erroneous training unnecessary. The machine learns from "watching" the seismic interpreter pick out the first few faults. It then attempts to find the next ones. The interpreter corrects the work the machine has done and sets it off again; whereupon, it performs better. Once the interpreter feels the computer is up to speed, he or she can just let it go.

Tim: An interpreter may spend weeks or months looking for all the faults in a dataset. Even then, he or she will have to interpolate between some of the seismic lines. The machine can do the same thing in hours. It is amazing, if not a bit scary, really.

Paul: As an example of this, we recently took a 12,000-km^2 survey with 5,000 lines in the dataset and asked the interpreter to label eight lines (12.5 km apart in this case) where there was a high

Paul Endresen & Tim Roden

likelihood of a salt dome. We then derived the rest of the salt across the entire survey in three hours on a laptop at an estimated accuracy of 99.9% compared with the human interpreter.

Where is Bluware technology currently being used in Shell?
Tim: We are climbing a steep learning curve. It has taken some time to get to grips with all the components of the Bluware platform, to get to know all its capabilities. That said, the October 2018 release of Shell's in-house-developed seismic-interpretation software package leverages Bluware's data compression and handling capabilities.

And we also have two proof-of-concept projects about to start. The first relates to a task called interactive stacking, which is used to generate optimum subsurface images from seismic data. We are hoping that, by using the Bluware platform for this, we will be able to access more data and still run the process on a standard laptop, thus extending an enhanced capability to a wider range of users.

The second project concerns interactive deep learning that Paul talked about. We will be trying it out for automated fault interpretation.

There is a sense that this technology is still in its infancy – that there is more to come.
Paul: We think there is. We are already seeking to extend the range of data the Bluware platform can support in the oil and gas business to well logs, drilling and production data, and rock core images. With Shell, we are doing some different work with fibre-optic sensor data. Distributed acoustic sensing downhole can generate 1 Tb of data every day that is very difficult to deal with using conventional processing technology. Processing drone image data pose similar challenges. Another possible use we have already examined involves data acquisition and transfer, even before it reaches the processing stage. We are certain we could transfer Bluware-compressed seismic data from the acquisition vessel directly to the cloud as soon as they are recorded. Currently, bandwidth limitations and the huge expense of the only alternative, satellite transfer, result in companies sending their processing people out to sea or transferring storage media by helicopter or even fast boats.

Tim: Bluware has set the tone for the future by focusing on one particular key aspect of dealing with large datasets. It is not trying to provide an end-to-end product; instead, it is selling high-performance data handling, computing and visualisation that users can leverage in a variety of ways in their own workflows.

I think that in the geology and geophysics area we will see an increasing move away from monolithic interpretation packages. For some time at Shell, we have been putting together our own seismic-interpretation software and working collaboratively with a range of outside companies. The increasing use of the cloud for not just data storage but also all aspects of the compute process is going to accelerate this trend. Over the next couple of years, I can see all the software pieces being up there in the cloud and us being able to grab what we need to construct the workflows that best suit our assets.

As the manager for subsurface interpretation at Shell, my job is to make available differentiated workflows to exploration technologists throughout the company. There is no single source for that and no off-the-shelf software package that can provide all our needs. We have to be able to easily integrate new and exciting technologies from wherever they emerge.

Paul: That is how Bluware sees it, too. Our clear preference is to see this develop in the cloud. If I am to be allowed the last word, I think the future is going to happen faster than we think, brought forward not just by the established players but also by small, very smart start-up companies with something genuinely different to offer. ■