

CLOUD HPC

INNOVATION

ACCELERATED



Ad

COME SEE US AT SC2018 IN DALLAS, TX AT BOOTH #2833

NOVEMBER 12-15<sup>TH</sup>

Explore Our Demo Pods featuring New Technology and Advanced Features!

HOME >

COMPUTE >

Oil Immersion Cooling Cranks Up Oil Simulation HPC

# OIL IMMERSION COOLING CRANKS UP OIL SIMULATION HPC

November 9, 2018

Dan Olds



Being at the forefront of high performance computing, as the oil and gas industry has been from the very beginning, also means always dealing with issues of power and cooling. Energy companies know better than most that there has to be some balance between what you are doing and the energy that is consumed to do it. That applies equally well to extracting barrels of oil from the ground as it does to running the seismic analysis and reservoir models that help make this a profitable endeavor.

We recently caught up with Down Under Geosolutions, which is one of the companies that wrestles with that balance every day. DUG offers geoscience services from the menu below, which is quite a wide range of options aimed at the oil and gas industry. Not surprisingly, performing these activities requires *a lot* of computing power. Oil and gas companies deliver huge amounts of data on tape to DUG, who then does the processing and delivers the output on tape back to the companies. This is all normal stuff so far, right?



Stuart Midgley, DUG’s System Architect, explained how DUG’s facilities are different from almost every other datacenter in the world in terms of its approach to building and cooling its datacenters. DUG has dubbed its approach “Extreme Data Center Cooling” and it is a response to trends its experts see in the HPC industry.

Innovative G&G Services and Software

Land/Marine/Transition Zone

2D/3D/4D

Hydrophone/Geophone/Node/Dual Sensor

Seismic & Well Data

Seismic processing

Full waveform inversion

Depth imaging

Petrophysics

Quantitative interpretation

Basin-wide velocity model building

DUG HW and SW solutions

Download/Interactive/Seismic

The first trend Midgley noted is that much of the processing today is performed on accelerators like Nvidia Tesla GPUs and Intel Xeon Phi parallel processors and sometimes co-processors. These massively parallel accelerators are power hungry – using about 250 watts per socket, and even more in some cases – thanks to their big compute motors and the high bandwidth memory that is closely attached to them to keep their many threads fed.

While many devices used in systems are using less power, the electricity burned by accelerators and thrown off as heat is much higher even as performance per watt is going down. Even the latest generations of CPUs from Intel burn progressively more juice because adding more and more cores does not come for free as Moore’s Law is slowing down. The point is, when you consider that an analytics system might use four or more accelerators, you are talking more than 1,000 watts before you factor in the CPUs, memory, storage, network, and so forth. And every vendor either has, or will soon offer, two-way servers that support up to eight GPUs – meaning that you will need 2,000+ watts just to run the GPU part of the configuration.

Moreover, with shrinking chip manufacturing processes (moving from 14 nanometer to 10 nanometer or 7 nanometer and lower), expect to see CPU power demands increase from today’s 125 watt to 145 watt level for middle-bin processors all the way up to 250 watts to 350 watts for high-end parts at the top of the bin stack in the next few years. (**The top-end “Skylake” Xeon SP-8180M**, with 28 cores, is already at 205 watts, and **the dual-chip “Cascade Lake AP” module** that Intel will launch soon is probably around 330 watts or so, we estimate.) So a highly configured GPU-enabled server of the future with dual processors could easily require 2,500 watts or more watts just to power the CPU and accelerators.

But we also must consider the power consumed by memory. As the number of memory channels increases, the number of DIMMs in the system will also increase as noted on the slide below.

Memory going up

- Power used by memory is considerable
- 10W/Dimm
- With multiple memory channels the number of Dimms goes up
- 80W in memory... as much as a lot of CPU’s
- Memory doesn’t really slow down, have to refresh every n clock cycles
- It is also a dominant cost

POPULAR ARTICLES



OPENACC: THINGS ARE GETTING SIMPLER MORE QUICKLY



THE WIDENING GYRE OF SUPERCOMPUTING



INTEL DOUBLES DOWN ON DOUBLED UP XEONS FOR HPC



WHAT TO SEE AT SUPERCOMPUTING 2018



OIL IMMERSION COOLING CRANKS UP OIL SIMULATION HPC



Large servers today offer up to 24 DIMM slots, requiring 240 watts of power, which, from a power consumption standpoint, costs more than having a third high performance processor in the configuration.

A heavily configured analytics system draws a lot of power now and will draw even more with the advent of new higher-power processors. Here’s how things look based on what we know today and our future estimates:

| Component    | Power Consumption Today | Estimated Future Power Consumption |
|--------------|-------------------------|------------------------------------|
| CPU          | 145 x 2 = 290 watts     | 235 x 2 = 470 watts                |
| Accelerators | 250 x 8 = 2,000 watts   | 300 x 8 = 2,400 watts              |
| Memory       | 24 x 10 = 240 watts     | 32 x 10 = 320                      |
| SSD drives   | 11 x 4 = 45 watts       | 11 x 4 = 45 watts                  |
| Hard drives  | 12 x 8 = 96 watts       | 12 x 8 = 96 watts                  |
| Total        | 2,671+ watts            | 3,331+ watts                       |

As Midgley points out, all of this processing power (and energy) isn’t just idling along in the datacenter. It is being used to crunch ever increasing amounts of data generated by IoT, tracking things and people, building more high precision models, and extracting more value from data using machine learning. Old data is often crunched again and again as it is used to build more sophisticated models and add greater precision to existing models.

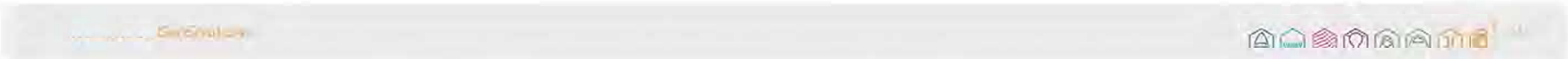
DUG needs serious hardware due to the complexity and compute intensity of the new algorithms it is using, for example:

Why the huge machines?

dug

New algorithms are very compute-intensive and very effective eg.

- 3D shallow-water demultiple
- High frequency FWI
- Elastic FWI



While most of us don’t know the details behind the algorithms listed above, Midgley said that just doubling the frequency of the High frequency FWI algorithm will result in a 10X increase in compute load.

High frequency FWI

dug

Double frequency = 10x compute

- Currently doing ~15Hz
- Want to do 125Hz
- Currently have 25PFlops
- 2.5EFlops needed!



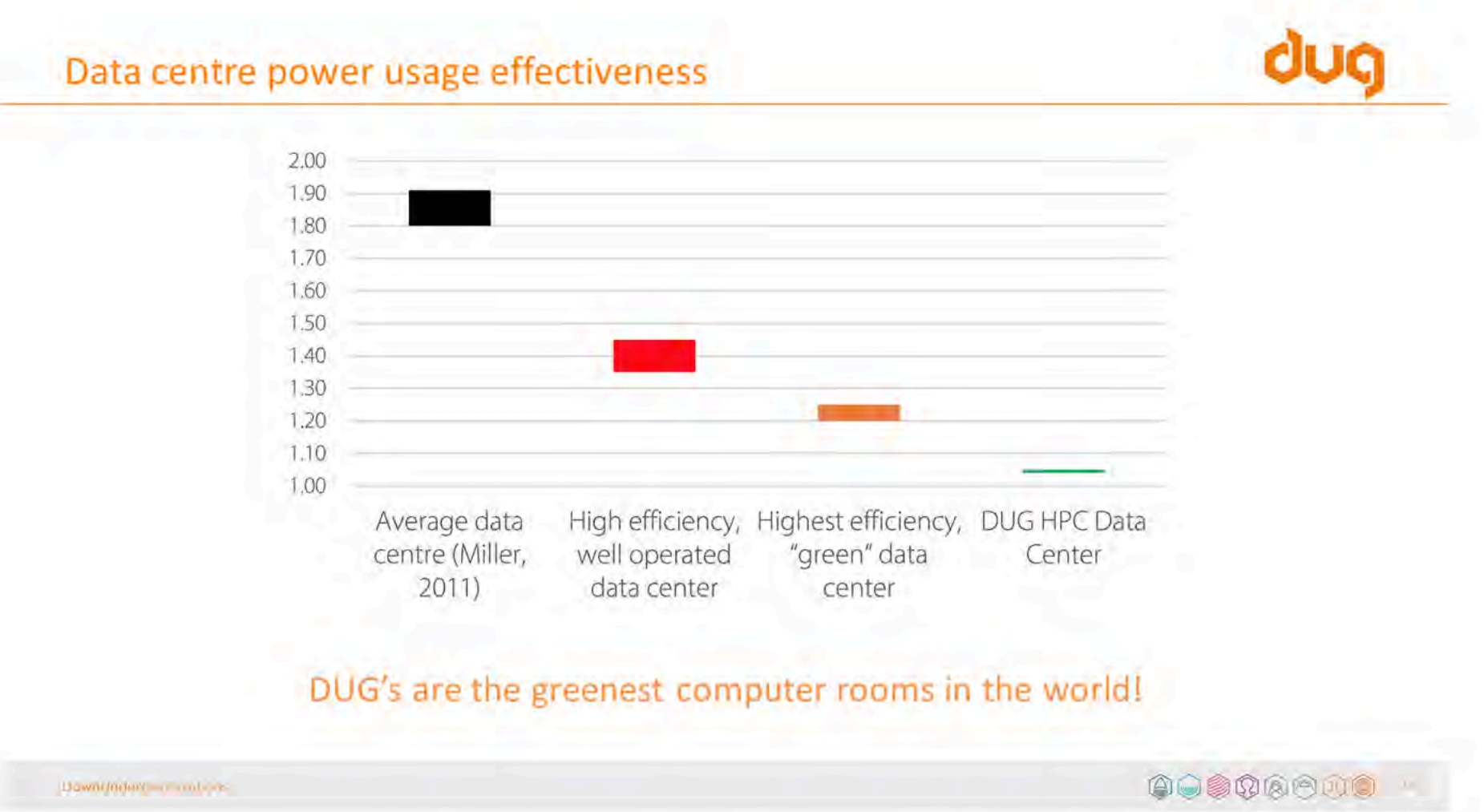


DUG is one private organization that absolutely has a need and a use for exascale systems.

## HOW MUCH POWER IS ACTUALLY BEING USED?

The answer to this question is one we can't print in a family newsletter, but it was an expletive ending in "load" and it is a term used in the IT industry plenty. As much as 3 percent of the world's energy (around 420 terawatts) is used by datacenters. This amount of power is far higher than the total consumption of the United Kingdom. That's a lot of power and data center usage is only rising.

There is a measure of datacenter energy effectiveness called the Power Usage Effectiveness, or PUE. The PUE of a datacenter is the ratio of the total amount of energy used by a datacenter facility to the energy delivered to the computing equipment (servers, storage, networks) that is actually doing useful work. The lower the PUE, the more efficiently your datacenter infrastructure handles delivering power and cooling to your gear.



Cooling is the main reason why we see such a wide range of PUEs on the chart above, Midgley explained. The PUE for DUG's HPC datacenter is about 1.04, which puts them in the winner's circle when it comes to power efficiency. This score is better than Google, whose best PUE was 1.08 according the search engine giant's last publicly released numbers.

DUG has made cooling the main target in its drive for highly energy efficient HPC computing. Why? Well, there is not a lot that organizations can do about CPU and memory energy usage, or the network or the storage for that matter. Sure, they can use lower power parts like CPUs, GPUs, and memory, but that won't get their compute workload done in time to meet their stringent deadlines. They can drive up utilization, and there are all kinds of ways to keep machines humming. Organizations are left with cooling and voltage transformation as the only components they can really control. And control them, DUG does.

The major goals behind DUG's extreme cooling initiative was to get more flops per watt of power while at the same time increasing density and reducing all types of waste.

Midgley pointed out some of the more unusual ways of reducing the cooling load on datacenters, first by pointing to Microsoft's deep sea datacenter – a column filled 864 servers that has been deployed to the ocean floor near Scotland's Orkney Islands. He also discussed the approach of putting datacenters in naturally cold environments, which would definitely decrease PUE. (Facebook is doing this in Sweden.) Midgley also talked about direct contact liquid cooling, which is pretty good for CPUs and GPUs, but is very cumbersome when you try to cool DIMMs.

One of the alternatives that he brought up was two phase cooling where you immerse servers into a vat of specially engineered liquid that has a low boiling point (say 65 degrees C). When the systems heat up the fluid, it boils, and the heat is conveyed to the top of the vat via steam. The steam is then re-condensed into a liquid and drops back into the vat. (It's a kind of weather.)

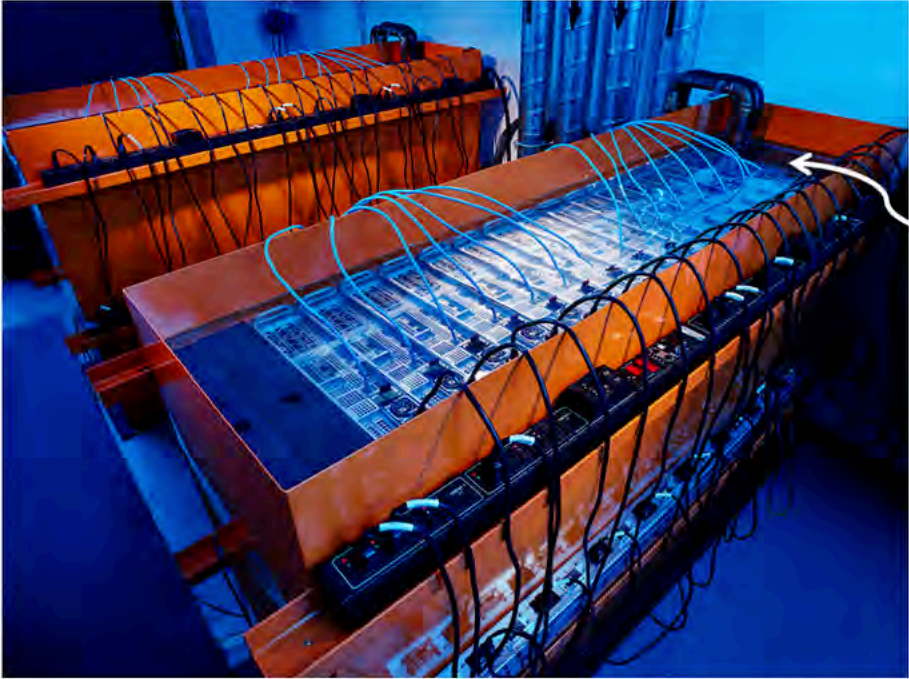


This is a highly efficient cooling methodology, but, according to Midgley “from a practical point of view, I wouldn’t touch it with a barge pole.” He went on to say that there are lots of issues with this type of cooling. The first is that some of the liquids become acidic when they absorb moisture, which can cause motherboards to delaminate over time – a bad thing. This is why these vats are typically sealed into containers, which could, theoretically, burst under pressure – another problem.

## SINGLE PHASE IS EXTREME ENOUGH

DUG’s liquid cooling solution is to dunk its servers in giant vats of oil – but not mineral oil, as others have used. The fluid DUG uses is very common – it’s a food grade oil that is the same substance used in Johnson’s Baby Oil. It has a very high boiling point, and won’t burn until at least 280C, although Midgley used a blow torch to heat it to over 400 degrees C and still couldn’t get it to flash. So from a safety standpoint, DUG’s oil is well suited to the liquid cooling task.

DUG Solution



Patent publication number: WO 2017/091862 A1

DownUnderGeoSolutions



The input water temperature with the DUG system is a warmish 30 degrees C (around 86 degrees F) and the output temperature is 38 degrees C (100 degrees F). As Midgley said, “It’s actually quite simple to get the return water at 38 degrees C water down to 30 degrees C. It’s very hard to get 50 degrees C (122 degrees F) air coming out of your server down to say, 18 degrees C (65 degrees F) input air in your datacenter.”

DUG uses water as the main heat exchanger from that fluid to the outside air. DUG has a significantly high water bill of \$3,000 (Australian) per month at \$.20 per kiloliter, meaning DUG is using something like 15,000 kiloliters monthly. However, that’s a small amount of money compared to the energy cost savings DUG are realizing from their system.

DUG has also paid a lot of attention to voltage transformation, which is an often overlooked place for finding increasing efficiency and lowering costs.

What is it really all about?



Thermodynamics

- Trying to increase the input temperature
- Don’t use air as the transport medium
- Have as few dT’s as possible
- Move the fluids as slow as possible
- Cool as much as possible with the advanced cooling system



DUG has reduced its power costs by an estimated 43 percent overall. DUG's newest facility will use 13 megawatts of energy. Let's deconstruct those numbers a bit. The lowest electricity rate in Australia is \$0.27 per kilowatt-hour. This means it will cost \$2,365 per year to purchase one kilowatt of Australian electricity. This yields a megawatt cost of \$2.37 million per year, meaning that DUG's 13 megawatt installation will cost the company \$30.75 million per year to keep the lights on and the systems running – if they're paying retail prices for their energy. Which they almost certainly are not. Since DUG is a high volume user of electricity, the company qualifies for wholesale rates. In Australia, this can vary from \$109 per megawatt-hour to \$75 per megawatt-hour. Taking the average of these prices yields an average cost of \$92 per megawatt-hour. Over the course of a year, this will cost them \$10.47 million to run its 13 MW facility. So annual savings would equal 43 percent of this total, or \$4.51 million, which shows the savings power of saving power.

Other advantages of the DUG immersion cooling methodology include:

- Space saving
- Simplification of technology
- Increased performance of all components
- Longer lifespan
- Increased reliability
- Reduction of components (heat sinks, plastics, refrigeration oils, etc.)
- Simplification of design and manufacture
- Increased density

One of the most surprising points that Midgley talked about the utter simplicity of the setup of its system. It doesn't require sophisticated studies of air flow and air flow management. It doesn't require massive air conditioning units. It is just designing a vat, which is straightforward engineering work, then stripping the unnecessary fans, shrouds and so on the system and component boards, which increases density and also reduces power draw. According to Midgley, the DUG vats are "dirt cheap" to produce, costing less than a standard liquid cooled rack while holding many more systems.

Components fail less, so less maintenance is needed and, as a kicker, the great cooling capacity has given DUG the ability to run its accelerators considerably faster than stock, which gives DUG free compute capacity. What's not to like?