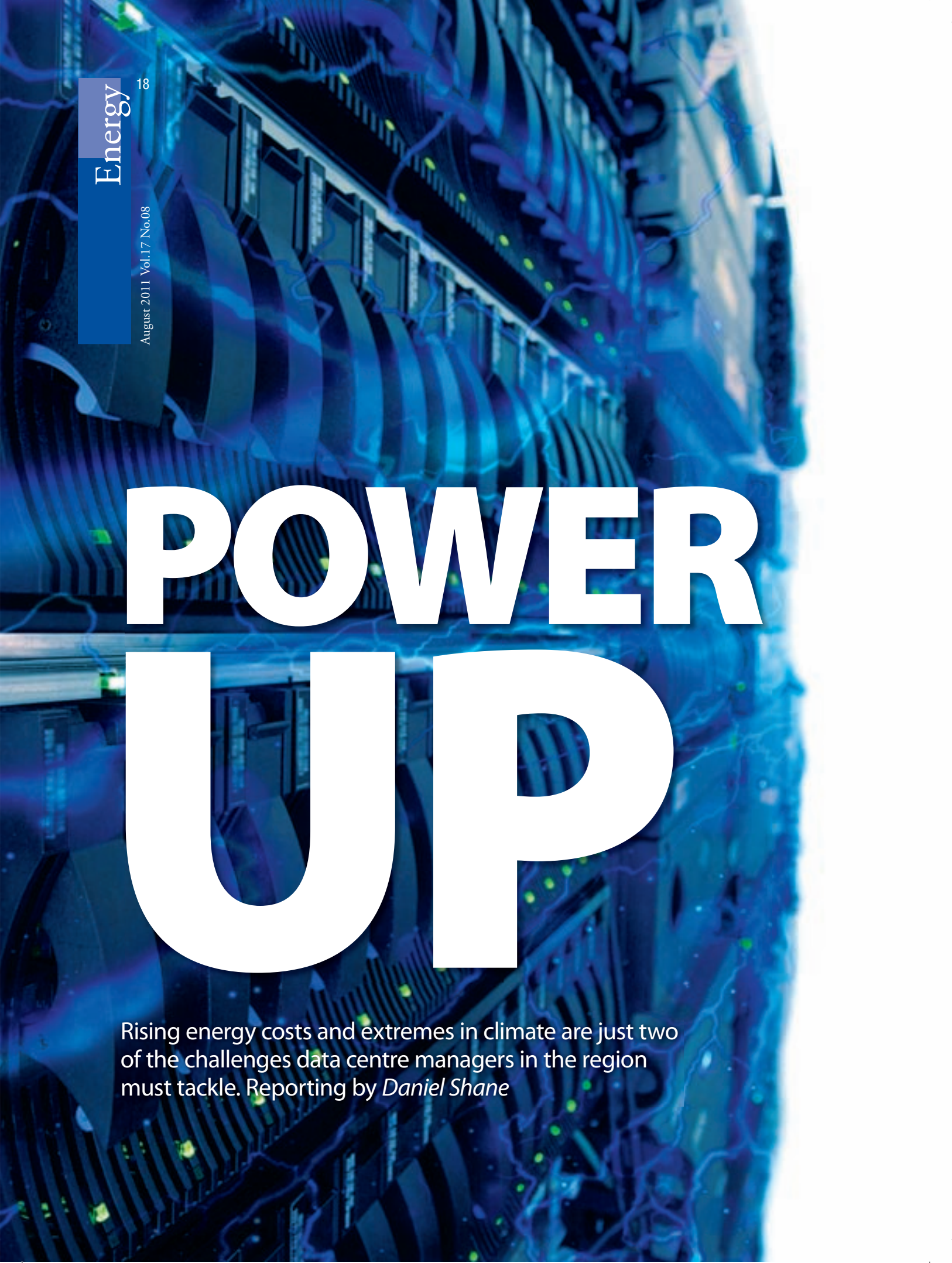
August 2011 Vol.17 No.08



POWER UP

Rising energy costs and extremes in climate are just two of the challenges data centre managers in the region must tackle. Reporting by *Daniel Shane*

Energy consumption in the Middle East has on average seen sharper growth in recent years than any other region on the planet.

Data published by the International Energy Agency earlier this year revealed that between 1997 and 2007, power usage in the UAE soared 71%, in Qatar 113%, Saudi Arabia 62%, and in Oman more than 135%.

By way of comparison, energy consumption in the US rose by just 9% during the same period. Granted, as developing market economies, it is no surprise that growth in Middle East energy consumption is higher – still, the IEA statistics give an idea of the power supply challenges the region must accommodate.

A part of this growth in power use can be attributed to the uptick in businesses setting up in the Middle East, and corresponding rise in the number of data centres – titanic facilities that house and run stacks of technology infrastructure, including servers, storage, and networking components, as well as a galaxy of associated applications.

Mammoth energy consumption has not come cheap either. Research published by IT market watcher Gartner late

"Companies are facing the challenge of transforming their technology infrastructure environments into agile, energy-efficient and cost-effective assets to drive business growth,"
says eHosting Datafort's Abdulrehman Ubare

last year estimated that the costs associated with powering and cooling these facilities accounted for 12% of all expenditure on data centres. The study forecast that this percentage is highly likely to increase in the coming years.

According to one expert, data centre managers and CIOs in the Middle East are beginning to feel the pressure in terms of balancing energy efficiency with innovative, technology-driven corporate strategies. "More and more companies are facing the challenge of transforming their technology infrastructure environments into agile, energy-efficient and cost-effective assets to drive business growth," explains Abdulrehman Ubare, head of technical operations, eHosting DataFort. "While increased pressure is being placed on CIOs to deliver more business services at a reduced cost, data centres are approaching the limits of their energy, cooling and space resources."

Ubare believes that in the region, typically, managing power consumption and costs in terms of IT has only been given lip-service by business leaders. "Historically, the cost of energy, and the cost of the data centre power and cooling infrastructure has not been on the radar for most CFOs and CIOs, and have not been considered in total cost of ownership (TCO) models," he claims. "This was a reasonable assumption during the 1990s when server power and energy costs were substantially lower."

However, as established, rising power usage and unit costs mean that organisations have in recent years started to seriously reconsider how they use energy.

Another local industry expert believes that power and energy efficiency considerations must be forethought during the planning and design phases of a new data centre. Vinayak Burungale, senior technical architect at Tech Mahindra, an India-based systems integrator, says that businesses must perform an in-depth audit of both their present and future requirements in terms of IT. "A detailed analysis and proper capacity planning plays a very important role in managing the challenges related to power, cooling and space," he reckons. "It is very important in the new age for organisations to thoroughly study the specification of the equipment, predict future business accurately and design the data centre infrastructure accordingly."

One technology that appears to have helped organisations manage their IT energy consumption more efficiently is virtualisation. By abstracting operating systems and applications from their hardware layer, virtualisation allows businesses to run several or more applications on one piece of tin. The theory is that if organisations can run hardware at higher rates of utilisation, they will have less of an infrastructural requirement. With a smaller, consolidated IT estate, the energy cost burden is obviously lower.

"Virtualisation means you can consolidate some servers, and get rid of others,"
says Gartner's Rakesh Kumar


Hot climates in the Middle East increase cooling challenges, says Tripp Lite's Vipin Sharma

Rakesh Kumar, vice president of Gartner Research, explains further: "Virtualisation makes a big difference. It allows better utilisation of the underlying hardware, both servers and storage, and even networks. So for every given dollar that you're spending on a server, you're using that server more effectively by having more virtual machines on there and running more applications. That means for every dollar of energy you're spending on running that box, you're now able to get more applications running on there."

To give an example, Kumar says that if you line up two servers side-by-side; one running operating systems and applications at a 70-80% utilisation rate, and one switched on, but running nothing. He claims that the one running nothing will consume about three-quarters the quantity of energy as the one at almost full utilisation. "You don't need as many servers, and you can consolidate others down and get rid of them," he adds. Kumar advises that the best time to virtualise infrastructure is during a wider hardware refresh, and the process should be taken in stages.

One of the most significant developments in data centre energy over recent years has been in measurement. The Power Usage Effectiveness (PUE) scale was developed by The Green Grid consortium, and aims to measure how efficiently a data centre uses its power. By dividing the total amount of power used by a data centre by how much power it consumes for IT equipment, PUE gives a single digit score on the facility's IT efficiency. The ideal score on PUE is generally somewhere between 1.0 and 2.0.

However, Gartner's Rakesh Kumar believes that while PUE is a helpful gauge, it is not the all-purpose indicator that the IT industry requires. "One of the problems we've had in the industry, especially a few years ago, was there was no uniform way of comparing the energy distribution of one data centre to another, or the efficiency of one data centre to another," he observes.

"The problem with PUE is that it's a point in time snapshot. In six months time, it may change because the [organisation] may put more IT equipment in there or various things may change. It's semi-useful but it's also semi-dangerous, as it can give a misleading impression," Kumar adds.

Others, however, argue that getting sufficient power to run a data centre at all in the Middle East can be difficult. "Inadequate power supply from [the] grid is also seen as a major problem. [Some parts of] the region [are] just coming out of turmoil and development work and projects have been stalled for quite some time," claims Vipin Sharma, head of Middle East at

power management firm Tripp Lite. "Power generation and the quality of power distribution needs to be developed so that more data centres can come up in this region."

According to Sharma, power supply problems can be greater in the Middle East than in other parts of the world. This, he claims, is due to the monopolies on supplies that often exist in the region, meaning that only one energy provider is available in a given territory. Because of this, data centres can be left without a backup power source in any event of an outage.

If there is one advantage the Middle East does have in terms of designing data centres for power efficiency, it is that the region contains less legacy infrastructure, and many deployments are Greenfield ones. This gives IT and data centre managers something less of a headache in terms of integration and maintaining performance during any IT refreshes.

However, for those sites that are already in existence, keeping existing operations going during any kind of switchover can be an intimidating obstacle. These challenges also apply to the power and cooling aspects of a facility, as one expert explains.

"Adding new equipment to the data centre while continuing to run the old equipment in parallel increases the power and cooling requirement significantly. Managing power and cooling during such cases becomes very challenging," claims Tech Mahindra's Vinayak Burungale. "Equipment like [server] blades and storage require huge power and high density cooling. Provisioning the same in legacy data centres becomes challenging, due to the amount of effort and flexibility required for re-designing."

Burungale adds that making alterations to power and cooling systems in an existing site requires meticulous preparation. He advises IT managers to install heat monitoring tools, so any systems that overheat during any major work, can be quickly pinpointed and remedied.

STAYING COOL

It is not only securing a reliable power supply, and then managing its usage and costs, that data centre managers in the region must concern themselves. With stacks upon stacks of infrastructure hosted in any one facility, the question of how to stop it from overheating becomes an urgent one.

Tripp Lite's Middle East head of sales, Vipin Sharma, says that issues of data centre cooling are exacerbated in the Middle East. "The most important challenge in the region at the moment is the need for cooling, as the ambient temperature is too hot, which in turn drives a higher cooling cost compared to other data centres based in cooler climates," he explains. Considering outside temperatures in the Middle East frequently top 50 degrees Celsius, it is easy to see why cooling is a hot topic for CIOs in this part of the world.

Chatsworth Products International is a specialist in the area of data centre cooling. According to the company's regional sales director, Sundeep Raina, says that there are several design considerations that should be taken into account to ensure that the data centre is maintained at an adequate temperature, such as passive cooling and aisle level containment. "The ultimate level of isolation is aisle level containment," he advises. "Cold aisle and hot aisle containment allow for isolation by creating enclosed rows. This type of containment increases efficiency, allowing you to utilise 90% of the airflow in the containment space with greater cold air volume."

He suggests that while the capital investment to implement energy efficient power and cooling systems in the data centre can in some cases be a considerable one, the investment is not necessarily just about saving cash, but also improving business agility and performance. "One might hear the argument that energy costs aren't as high in the Middle East compared to the rest of the world. Even though this may be the case, it still makes plenty of sense to choose a power solution that will optimise all facets of data centre performance, whether you're designing a new facility or upgrading an existing one," Raina believes.

"However, getting the necessary budget for power management is becoming far more complex," he concludes.

The flaw in Moore's Law

Moore's Law is one of the longest-running and best-known trends in computer hardware. Coined in 1970 by the founder of Intel, Gordon Moore, the theory states that the number of transistors that can be economically attached to an integrated circuit doubles approximately two years.

Since then, Moore's Law has held up phenomenally well, proving itself to be true time after time.

However, some believe that the persistence of Moore's Law is one of primary reasons why IT energy consumption has spiralled so drastically over the years.

"If you look at Intel's design of processors and Moore's Law, all they have been doing is putting more and more processing power on to a smaller and smaller bit of silicon, and not really worrying about the fact that this generates more heat and requires more energy," explains Rakesh Kumar, vice president at Gartner Research.

"From one generation of processor to the [next], the power needed was going up by something like 50% to 60%. In a single laptop or a computer that's not going to cause a problem, but the blade servers that were coming out had a very small form factor, and server manufacturers were putting multiple chips on the same board," he continues. Kumar says that this issue has been severely exacerbated in data centres, where up to 50 blade servers can be stored in the same rack and run within close proximity of one another, generating "a concentration effect" in terms of heat.

"These are offshoots of Moore's Law that no-one has really thought about until the last three or four years," Kumar reckons.

However, he does see progress in terms of chip design with Intel's recent Tri Gate '3D' semiconductor. By adding an extra gate to each transistor that is vertically above the other two, Tri Gate pledges to leak fewer electrons, and subsequently consume less power.