Orchestrating the computing and power needs of edge data centers

ELTEK
A brand of Delta

Executive summary

As the digital transformation of business and society progresses, data volumes are growing exponentially, putting pressure on Cloud computing capacity and power consumption. This particularly applies to highly time-sensitive processes at the network edge, close to the end user.

One solution is moving more time-sensitive computing processes to the network edge and edge data centers (EDCs). Processing data closer to the point of use not only reduces latency, it also offloads backhaul networks and reduces pressure on centralized data centers.

However, distributed processing requires a tight orchestration of computing and power infrastructure. This white paper explores the more effective approaches for aligning the two, and the technologies required to both manage a network of distributed EDCs and tap into new revenue streams such as ancillary grid services.

Contents

Section 1 - Introduction

1.1. The need for EDCs

In the wake of the pandemic, digital transformation has been accelerated and left virtually no part of society and the economy untouched.

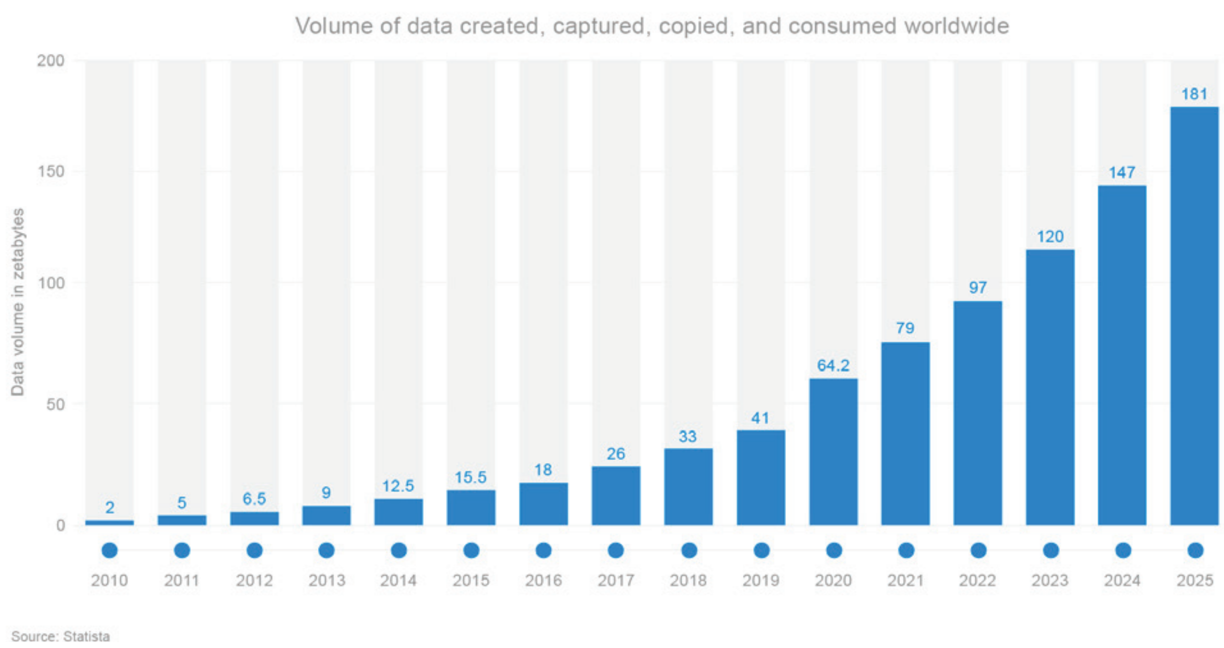
Businesses have moved to the cloud at pace to facilitate remote operations and integrate their supply chain management. By the same token, industrial automation and the Internet of Things (IoT) have seen an uplift. Once regarded as futuristic technologies, artificial intelligence (AI), machine learning, and augmented and virtual reality have all become part of our everyday lives.

Add to that the growth of mobile connectivity, culminating in the rollout of 5G mobile networks, and all these trends amount to ever-higher data processing needs.

Among all the drivers behind this expansion, AI is emerging as the most dominant. Training a good AI model alone will require substantial data acquisition and lots of processing power for training.

As the rollout continues, bandwidth will be stressed at every point in the network. One consequence of this is increased latency, which will only worsen as data volumes continue to expand.

Statista predicts the volume of data created, captured, copied and consumed across the globe will nearly double between 2022 and 2025, reaching 181 zettabytes.[1]



[1] Statista

However, only a small percentage of this data is typically retained - in 2020, it was only 2%.[2] This suggests that a lot of data is transient and mainly transactional.

Given the growing volume of data, it is becoming too complex and lengthy for applications with high-speed processing needs to send all data back to centralized, hyperscale data centers.

Not only would latency get in the way, but funneling all data back to the cloud would lead to expanding bandwidth requirements, not to mention the risks of losing access to data when the connections go down.

1.2. What is an EDC?

One way of avoiding these issues is a hub-and-spoke or client-server approach, where time-sensitive, latency-critical data is processed on the cloud edge. This is where EDCs come in.

They process and cache data locally, ensuring time-sensitive content can be delivered to an end device - a mobile phone, a PC, or an IoT-enabled piece of equipment - with minimal delays.

Less time-sensitive or low-latency data can still be sent to centralized data centers for analysis, analytics and long-term storage.

EDCs could form part of telecom networks, be located in or close to offices and factories, or even near homes.

Netflix, for example, has installed streaming cache servers in some ISP data centers so its customers can stream content quickly from local servers rather than from a server further afield.[3] This also reduces the need for and cost of international bandwidth, as the company doesn't have to cater for multiple downloads of the same content for each streaming customer.

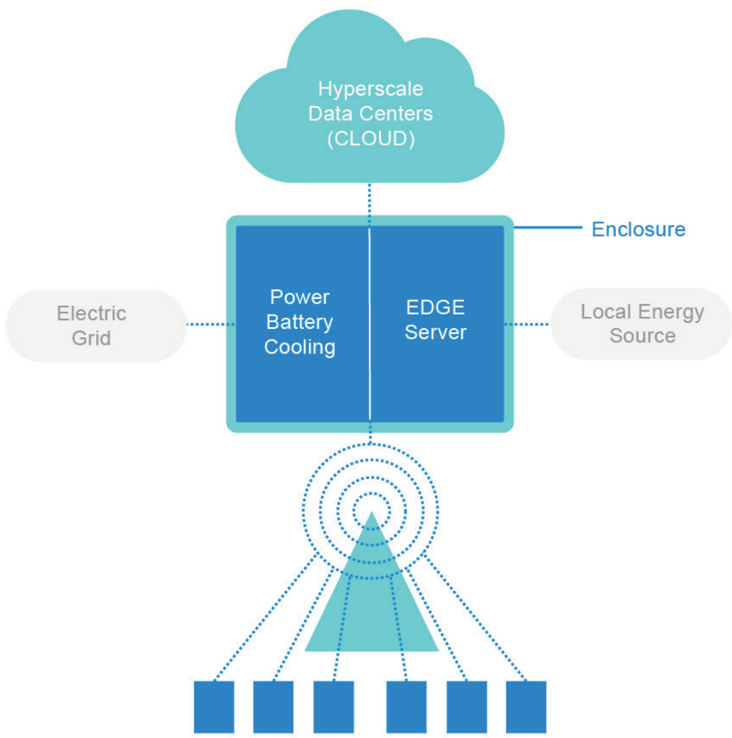
If applications are designed to use EDC capacity, this not only results in faster reaction speeds but also exerts much less pressure on the backhaul network and power consumption.

[2] Statista, [3] Voyager

With cybersecurity a perennial concern for businesses, another benefit of EDCs is that they allow for sensitive data to be processed “on-premise” rather than exposed to the risk of transferring to a hyperscale data center for computing. This means organizations can retain full control of critical services and maintain privacy and security.

EDCs vary in size, all the way down to micro data centers. These can be used for smaller businesses or at a radio site in a mobile network, for example.

Schematic of an Edge Data Center



1.3. EDC applications

EDCs are set to proliferate as computing needs continue to grow exponentially.

Between 2023 and 2032, analysts expect the market to realize a compound annual growth rate of more than 15%: from $9 billion to $50 billion.[4]

Their applications are manifold.

5G mobile networks will use EDCs collocated with mobile cell towers or network nodes to cache content locally. This will enable real-time applications and content delivery in areas with a high density of devices — such as city centers or office parks.

Time-sensitive IoT applications for infrastructure like smart factories or utility companies can use EDCs to speed up processing in areas such as predictive maintenance and quality management.

The EDC market is set to grow from **$9 billion** to **$50 billion** between 2023 and 2032

Similarly, autonomous vehicles will benefit from low-latency data sharing, which will increase their ability to react to changing road or traffic conditions in real time. The same applies to robotic applications, for example, in remote healthcare or smart manufacturing.

[4] Global Market Insights

Section 2 - Managing the power demands of EDCs

2.1. Optimizing power demand and availability

EDCs must be resilient and reliable, especially those positioned remotely. This also applies to power management, an area likely to see some unique challenges given the expected proliferation of EDCs.

One of the challenges in edge computing is balancing the dynamic relationship between power availability and processing needs. Both have to be managed dynamically, orchestrated and optimized relative to each other.

Many EDCs, especially those in telecom access networks, are likely to be located in areas where access to electricity is limited - and like any other data center, EDCs require significant amounts of electricity for both computing and cooling.

EDCs can help stabilize the grid

They may also be colocated with power-hungry users nearby, leading to competition for both energy and bandwidth, particularly in dense urban areas.[5]

Whatever the scenario, when the available power resources and the EDCs' power needs don't match, it is vital that the EDC can use a mix of electricity sources in addition to the public electricity grid.

Renewable power will be of particular importance, both to meet emission targets and avoid penalties, but also to establish grid independence and save costs. Batteries, solar PV and fuel cells may all be used to maintain a reliable energy supply. There are increasing numbers of EDCs in the field that use renewables, especially in telecom networks.

EDCs can also play a part in helping to stabilize the grid at times of peak demand.This includes the ability to draw down less energy from the grid in a demand-response scenario, allowing the grid operator to redeploy electricity where it's needed. EDCs could also be harnessed to feed power back to the grid (see section 3).

[5] Financial Times

Project Aniara - Designing the future for EDC power management in Europe

Project Aniara (part of AI-NET), a European research project within the Celtic-Next Eureka cluster, was set up to explore how digital transformation could be accelerated by network and service automation at the edge.[6]

This public-private partnership initiative is looking at the power and infrastructure needed to bolster Europe's 5G infrastructure, and find alternative approaches to the centralized cloud infrastructure considered inadequate for supporting Europe's digital transformation. It brings together companies like Eltek; several enterprise partners; universities in Sweden and the UK; and research institutes like The Fraunhofer Society in Germany and RISE Research Institutes of Sweden.

For the power management area of the project, a modular AC-UPS based on rectiverter modules from Eltek enables a seamless transition between power sources (grid, battery, PV), depending on the demand profile of the EDC at any given time.

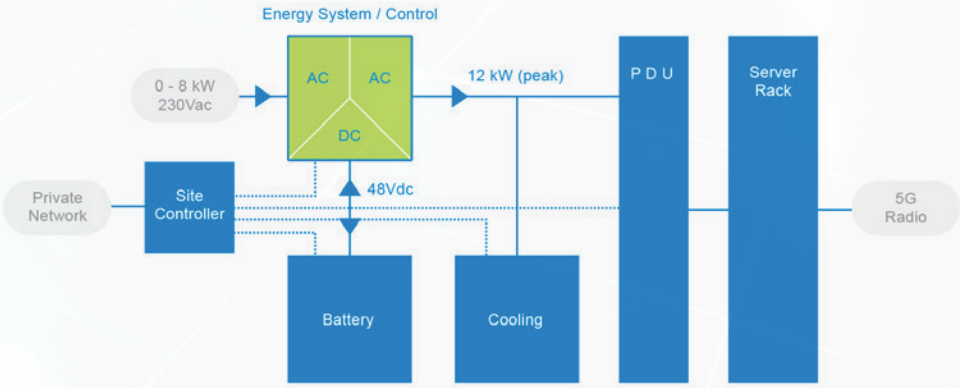
The rectiverter also lets EDC operators dynamically limit the power consumption from a specific source, for instance, which can be offloaded to the grid during peak demand.

A rectiverter module.

The underlying rectiverter technology helps optimize the orchestration of computing resources and energy utilization, maintaining the right balance. The server load can be supplied from both AC (e.g., grid) and DC (e.g., solar PV) input and the ratio of both can be controlled.

The rectiverter modules integrate all the functionalities of a rectifier, inverter and a static transfer switch in one compact box not much bigger than a typical rectifier. This helps reduce the footprint compared to other, less-integrated solutions.

Schematic of a 5G EDC with a rectiverter



[6] Aniara

2.2. Remote maintenance and operations

The design of EDCs will vary significantly from typical data centers. Not only will there be a lot of them, but most will be located remotely. Robustness and easy maintenance are prerequisites for achieving the highest levels of reliability, availability and resilience in all environments.

A central DC benefits from backup power generators, which many EDCs won't have due to their small size and location. In case of a failure or power supply disruption, EDCs therefore need to be designed to either switch to a battery backup system for an extended period of time - typically up to two hours - or reroute processing to another EDC.

While some EDCs will be located indoors, others are expected to be housed outside. Especially in 5G radio access networks (RANs), many EDCs will need to be located in close proximity to cell sites, including city rooftops or transmission towers, where floorspace is limited and comes at a premium. Therefore, a small overall footprint, as well as sturdy equipment ensuring protection from dust, water and other liquids will be critical. For example, Project Aniara aims to design an IP65-classified edge node as a solution, providing above-average protection from dust and liquids.

With large numbers of EDCs expected to be distributed around the country, operation and maintenance will also need to become more autonomous.

Unlike centralized data centers, EDCs will have no on-site maintenance staff and sending out maintenance crews will be unviable, if not impossible, in many cases. Remote maintenance and operation, as well as a high level of automation, will therefore be a vital criterion for EDC operators choosing their equipment.

Section 3 - Providing ancillary grid services

3.1. Backing up the grid

EDCs may be small individually, but taken together, they have the potential to play a significant role in future distributed energy systems.

In addition to demand-response services - lowering the power they draw from the grid during demand peaks (see section 2) - the combined distributed energy reserve in EDCs
will be an important asset for grid stabilization.

When a disturbance - such as a power plant outage - occurs and the grid frequency drops, this would be covered by grid inertia. This is stored kinetic energy traditionally provided by large rotating turbines in traditional power plants.[7]

However, the diversification of energy sources into renewables means there will be fewer of those plants, so other failover measures will be needed when inertia on the grid is low.

EDCs could fulfill some of these backup functions as a service provider to the grid.

They include fast frequency reserve (FFR) services, which kick in within 0.7 seconds and provide support for up to 30 seconds.

Similarly, EDCs could provide frequency containment reserve (FCR) services, which are slower to respond. Activation occurs around 60 seconds after a disturbance, and they can contain fluctuating frequencies for up to one hour.

EDCs are ideally suited to provide these services as they can be orchestrated to respond quickly.

[7] National Renewable Energy Laboratory

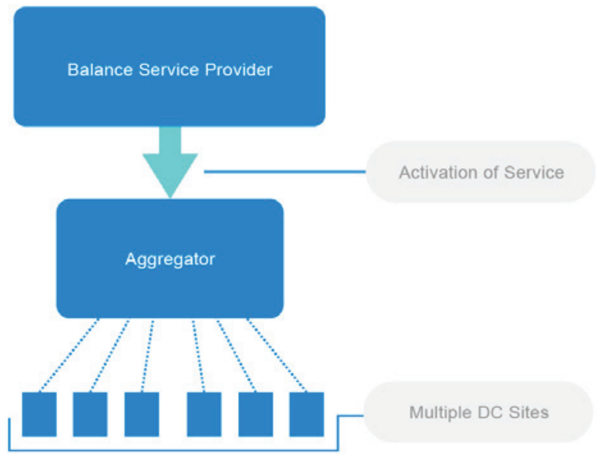
3.2. New revenue streams for EDC operators

Another scenario is a lack of transmission capacity.

It takes time for electricity generation to respond to increased consumption. This gap could be filled by the EDC power infrastructure - battery storage, on-site renewable sources - stepping in for a short amount of time.

Providing these services to the grid could represent a new revenue stream for EDC operators, by striking a deal with balance responsible parties (BRPs) charged with delivering reserves to transmission system operators (TSOs). This would be based
either on a dedicated pricing scale or a bid system.

Aggregation of edge data centers for electric grid support



To participate, power consumption would need to be above 100KW. As each edge node may only use around 20KW, this means many of them will need to be linked together and orchestrated to work in unison.

We may, therefore, see multiple, small EDCs being aggregated to act as a virtual power plant (VPP) to provide services for this kind of large-scale energy management.

For EDC operators and VPP aggregators to exploit these new revenue streams, choosing the right power and orchestration technologies will be vital. For example, to allow nodes to be powered up and down as flexibly as the grid requires, multimodal equipment such as rectiverters will be important technological enablers. Similarly, power systems need to be modular, so they can be tailored to meet diverse sites' needs both for their own operation and for supporting the grid.

Delta Xubus Edge - Why an "edge data center in a box" makes sense

Given the speed of expansion anticipated, the availability of integrated, prefabricated data center solutions could prove to be a trump card for prospective EDC operators.

Xubus addresses all the main challenges for customers rolling out EDCs:

Small footprint:
With space at a premium in many of the settings EDCs will be installed in, Delta's Xubus Edge 3.5KW data center[8] offers a rack-level solution and the smallest of five ready-for-rollout "data centers in a box". For example, its power distribution is integrated within the UPS, allowing customers to save valuable data center space.

Power conditioning:
Integrated within a cabinet, Xubus Edge is environmentally controlled with hardware and software for power conditioning and backup, power distribution, environmental management (cooling), security monitoring, data computing, data switching and data storage.

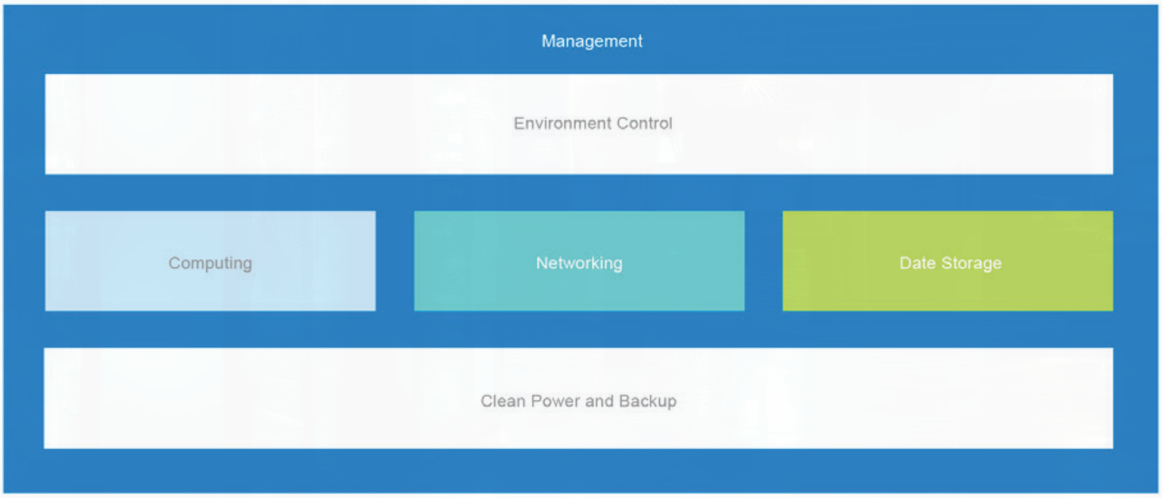
Integrated, optimized cooling:
Xubus Edge offers integrated environmental management that is optimized for each individual EDC's needs. Split cooling and backup free cooling come pre-integrated, so there is no requirement for additional on-site installation.

[8] Xubus Edge is not part of the Aniara research project mentioned earlier in the document.

Battery backup:
Xubus' battery backup systems are dimensioned to ensure a substantial failover period so that EDCs can continue to operate until repairs can be done. Where this is not possible due to the remoteness or inaccessibility of a site, traffic can also be rerouted to another EDC to avoid a drop-off in service.

Renewable energy:
Xubus Edge can draw energy from a range of sources, including renewables such as wind and solar energy. The EDC can either run on renewables with grid or battery backup, or draw on renewables and battery power when the grid fails. This is particularly important in applications such as telecoms networks, where EDCs may be located in areas that are not within easy reach of the grid, or where grid power is limited or heavily contested.

Design of a rack-level data center solution



Because they are compact, pre-tested and standardized, Xubus Edge DCs can be deployed rapidly, as well as save CAPEX and provide flexibility as the market grows.

Setup only needs to be configured once and all future DCs can then be set up in the same way, saving time and cost.

Conclusion

Rolling out EDCs will become vital for a wide range of applications as the data volumes we generate increase.

Beyond their immediate benefits for anything from telecom networks to self-driving cars and from streaming media to industrial IoT, they will also play an important role in the smart electricity grids of the future.

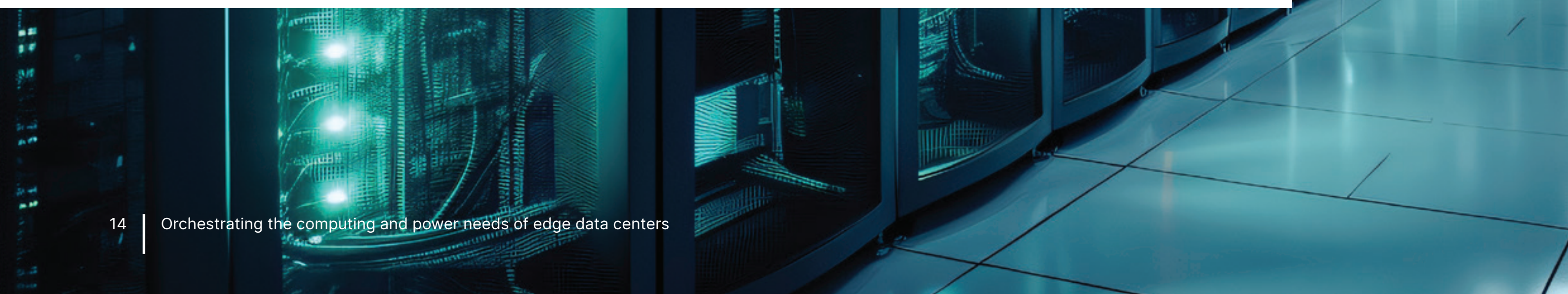
EDCs can potentially contribute to our net-zero future by providing important grid balancing services in distributed energy systems that are predominantly powered by renewables.

To make the most of both opportunities, the computing and power needs of EDCs must be carefully orchestrated to account for power competition and rollout in remote and other locations that are not prepared for edge computing. This includes the ability to remotely manage and maintain a network of distributed EDCs.

They have to be ready to use renewable energy and energy storage in order to be independent, save costs and support the grid as an additional source of income.

In selecting the right technology solutions, EDC operators also need to ensure their distributed data center infrastructure can be scaled up easily and incrementally as computing requirements evolve. In this way, they can apply a "pay as you go" approach to the investments they make.

Vital to enabling this flexibility will be smart power solutions like rectiverter modules that can juggle a variety of sources and handle a range of ancillary services, and EDC packages that can be rolled out rapidly as demand grows.



Delta EMEA Headquarters

Netherlands B.V.

Zandsteen 15, 2132MZ, Hoofddorp
The Netherlands
Tel: +31 (0)20 800 3900

www.delta-emea.com

Delta Telecom Power - EMEA Locations

Delta Electronics (Germany) GmbH

Ferdinand-Porsche-Str. 45, 60386
Frankfurt am Main, Germany
+49-69-42002-0

Delta Electronics (Switzerland) AG

Freiburgstrasse 251, 3018 Bern-Bümpliz, Switzerland
+41-31-998-53-11

Delta Solutions (Finland) Oy

Rajatorpantie 8, FI-01600 Vantaa,
Finland
+358-9-849-660

Delta Electronics (France) SAS co.

2 Rue du 19 Mars 1962, Zi Bastillac
Nord, 65000 Tarbes, France
+33 562 34 09 30

Delta Electronics (Italy) S.R.L

Building Spaces Eur Arte, unit 508 - 5F, Viale dell'Arte, 25, 00144 Rome,
Italy
+39 06 9931 0867

Delta Electronics (Poland) Sp. z.o.o.

23 Poleczki Str. 02-822 Warsaw
Poland
+48-22-335-2600

Delta Electronics (Czech Republic), spol. s r.o

Průmyslová 1306 /7, 102 00 Praha 10,
Czech Republic
+420 272 019 330

Delta Electronics Solutions (Spain) SLU

Ctra. De Villaverde a Vallecas, 265
1º Dcha Ed. Hormigueras – P.I. de
Vallecas 28031 Madrid
+34 91 223 74 20

Delta Electronics (UK) Ltd.

Hemel Hempstead, Hertfordshire
HP2 7EY,
United Kingdom
+44 (0)1442 219355

Delta Energy Systems MEA (South Africa)

Tuinhof Office Park, Unit C401, Karee
Building, 265 West Avenue, Centurion,
0157 South Africa

Delta Energy Systems (Sweden) AB

P.O. Box 3096 SE-350 33 Växjö
Sweden
+46 470 706 800

Delta Energy Systems AG - Dubai BR

P.O. Box 185668, Gate 7, 3rd Floor,
Hamarain Centre, Dubai, United Arab
Emirates
+971 4 2690148

Delta Electronics (Norway) AS

Terminalen 12-16, 3414 Lierstranda
P.O Box: 2340 Stromso, 3003
Drammen,
Norway
+47 3220 3200