Application Disaggregation

Optimising data centre infrastructure for smarter business outcomes

APRIL 2017

A whitepaper researched by

BROADGROUP

Consulting in practice

Contents

Executive Summary 3
Section 1 Challenge and opportunity 4
Section 2 Application models used in enterprises 5
Models used 5
Section 3 Solutions for smarter business outcomes 7
Geographical disaggregation 7
Power's role in geographical disaggregation 7
Multi-tier operations within a single data centre campus 8
The hybrid solution 9
Making the business case 10
Section 4 Industry sectors which could benefit from disaggregation 11
Section 5 Next steps 12

©BroadGroup 2017

Executive Summary

BroadGroup estimates that at least 85% of enterprises are missing out on potential cost-savings by applying a 'one-size-fits-all' approach to data processing and infrastructure.

To be successful in the long-term, CIOs need to develop cost optimisation strategies that can help disaggregate applications according to availability, latency and security needs and enable them to control IT operational costs effectively for the future.

This whitepaper reassesses existing approaches to data storage and highlights how modern operators can re-engineer the future of power intensive data centres. It demonstrates how enterprise users can optimise data costs by thinking differently about their data applications and signposts the potential for application disaggregation across a range of sectors. It also considers the pivotal role operators, such as Sentinel in North America or Verne Global in Europe, can play within global digital transformation.

> *"For companies to compete successfully, skills and assets need to be reconfigured to go beyond the limits of the current economic engine and provide alternative ways of serving customers."*
>
> Gary Hamel and C.K. Prahalad in 'Competing for the Future'

Read on to understand more about inevitable pressures impacting business leaders today and how CIOs can implement practical, application disaggregation solutions to drive more cost effective and sustainable business outcomes.

Section 1 Challenge and opportunity

Today's business leaders are challenged by the impact of digital globalisation on their operations and the need for ongoing transformation to compete effectively and secure the right business outcomes. In addition to this, there is constant pressure to reduce the bottom line and stretch budgets to achieve more with less. There is an opportunity for CIOs to play a central role in business transformation by devising smart strategies to improve data handling and data architecture, whilst still securing and protecting company information.

The opportunity to disaggregate applications is one that ALL companies involved with large amounts of data can do. It is not just about specialised applications, run by niche firms and verticals. There is good business sense in not paying more than necessary to handle and protect your data. Every company has specific data and applications that could be disaggregated and colocated in cheaper, more efficient and flexible data centres. An example of this is how the research community approach the deployment of their high-performance computing. They are happy to power the compute with a single direct mains supply. The storage would normally be supported by a traditional Tier 3 environment for researchers to access 24/7. As the compute volume is generally 70% of the total power consumption, this results in a significant reduction in their data centre costs.

A variety of options are available to CIOs in determining how to handle, locate and protect company data. Two possible solutions to achieve more cost-effective and sustainable data handling, processing and storing are:

- **Geographical Colocation** – European power prices are continuing to rise and the increasing volume of big data is putting a strain on existing power resources. This is a key factor when evaluating colocation. Naturally cold climate locations in the Nordics or Canada that offer low-cost, renewable energy solutions are becoming attractive outsourcing locations.

- **Multi-tier operations within a single data centre campus** – Many companies are typically dealing with big data in the same way as they do 'small data'. Contrary to popular belief that it is best to process applications in a singular environment; the reality is that many applications do not require the level of power and/or resiliency of a Tier 3 environment, and it is in fact more expensive and inefficient to apply a singular approach to data processing. When it comes to data, one-size doesn't fit-all.

Section 2 Application models used in enterprises

Large organisations run hundreds, often thousands, of applications. Understanding how specific end user applications can be handled is key for optimising costs when processing these applications. BroadGroup's discussions with CIOs revealed that simple 3-fold models are typically used to characterise applications and to direct how data is handled. Companies characterise applications in a variety of ways, for example: by organisational structure or data attributes such as latency, security and resiliency levels. This then guides decisions about geographical location - whether data remains in-house, deployed to cloud or located on a third-party data centre campus.

Models used

Enterprise 1

Differentiation of data applications at three levels

- **Person**
- **Group**
- **Enterprise**

Which applications are on each level?
What kind of quality attribute do you want for each?

- Availability
- Continuity
- Security

Enterprise 2

Three buckets of data

- **High spend, availability and resiliency levels** - business critical
- **Average latency needs, not time-sensitive** - cost-efficiency improvement
- **Reduced latency requirements** - enterprise back-office

Enterprise 3

Matching compute need to geographical location

- **Compute intensive -** specified data centre
- **Middle level compute -** variable, by application
- **Data storage -** anywhere

Some companies develop detailed matrices or dashboards to correlate applications with supporting elements and data centre requirements. The matrix shows an approach used in financial services:

Banking data centre application dashboard – an example

	Application 1	**Application 2**
Statutory requirements		
Basle II, MiFID, SEC, Sarbanes Oxley …	Yes	No
Disaster recovery limitations	Yes	No
Business requirements		
Network/Latency	High	Medium
Compute/hardware	High	Medium
Storage/archiving	Medium	Medium
Collaboration/access	Medium	Low
Criticality	High	Medium
Data centre requirements		
Power density	4-6 KW per rack	2-3 KW per rack
Security	High	Medium
Location	Within 100km	Within Europe
Cloud option	No	Yes
Outsourcing option	Yes	Yes

Source: BroadGroup/Financial Services sector

Section 3 Solutions for smarter business outcomes

One way business leaders can make more out of less is to disaggregate their applications, and this is applicable to any company with volume data. Being smart with data is not just placing everything in Tier 3. Why pay more to protect data that does not need protecting?

Two disaggregation methods that can improve costs and business outcomes are:

- Geographical disaggregation
- Multi-tier operations within a single data centre campus

Geographical disaggregation

Decisions based on geography are not always straightforward. In terms of latency-sensitive applications, a CIO will assess the type of connections going into a data centre and evaluate whether a solution is regional or country-based.

If data protection and data sovereignty are issues, they ultimately define geographical placement. Otherwise, for most commercial and Internet-based traffic, the decision on where to locate data is an economic one. Even financial companies who have traditionally required data centres close to users and exchanges can store significant amounts of data in remote locations with no discernible degradation to their performance. Research in 2016 by BroadGroup found that in the financial sector only 5-10% of applications require proximity to the London Stock Exchange.

Solutions can be different per country, contingent on how fibre-rich the networks are. There are, of course, other types of geographic disaggregation that require proximity to end users. Edge data centres, for example, are designed to support video streaming, retail and other bandwidth-hungry consumer applications.

Global networks and improvements in geographical latency enable companies to store and process data in alternative locations. Consequently, some data centre operators are beginning to offer regional hubs and host specific sector applications at a distance from major conurbations.

Enterprises can colocate non latency sensitive applications and compute to specialised data centres on the geographical rim, e.g., to the Nordic countries, Canada and the Pacific North-West. Non latency sensitive applications can include disaster recovery, back-up, storage, archiving, ERP systems, research and development, data mining, deep learning and machine learning.

Power's role in geographical disaggregation

Power accounts for up to 60% of total data centre operator costs. The cost of electricity is a key factor when evaluating colocation and other outsourced options – increasingly so as big data volumes spiral upward. New areas of power density have emerged with the growth of supercomputing; high performance computing (HPC) operates at densities up to 30kW. Ultra-high

densities of 40 kW – 60kW per rack are also now used, e.g., ASIC chip manufacturers in block chain computing for Bitcoin.

The growth in data is putting huge pressure on power grids. According to Northbridge Energy Partners, UK data centres are estimated to consume 3,000 to 4,000 MW, equivalent to 1.5% - 2% of total UK power consumption. This figure is forecast to grow at a rate of 3% - 6% annually.

Power-intensive applications, such as those using HPC, big data analytics and storage workflows can off-shore to low-cost-energy countries and regions, like Iceland, Scandinavia, Canada and areas of the US (Pacific North-West) and remote data centre locations in climates that support operational cost reductions and green energy.

The cost of power and advantage of cold climate location come into play as the ability to cool high density hardware is a key issue in supercomputing. Nordic countries with advantageous, low cost, green power profiles and cooler climates are proving ideal locations for HPC as well as non latency sensitive loads. Iceland, with its abundance of hydro-electric and geothermal electricity sources, is one such example of an attractive location.

Global players already benefitting from colocating in Iceland include:

- Automotive manufacturer BMW, which has been able to cut 3,750 metric tons of $C0_2$ emissions and 82% of its operating costs by processing computational fluid dynamics and other modelling and crash-test simulation applications within the Verne Global data centre

- Earlham Institute, one of Europe's leading bioscience and bioinformatics research centres use the Verne Global data centre for large-scale plant genomics and biology data analysis.

Multi-tier operations within a single data centre campus

The typical legacy model of single tier, single cost, single service in-house data centre is an inefficient way of operating. However, the majority of enterprise data centre users continue to place all applications within Tier 3 'mission critical' hosting environments, be these via colocation or in-house.

Not all applications in the enterprise portfolio require the same power profile, location or resiliency levels. The power versus resiliency paradigm is changing and traditional data centre infrastructure is becoming out-dated. There is no reason why applications such as archive, storage, back-up or research and development can function optimally at other levels of acceptable latency and this can save considerable expenditure.

Forward thinking companies are beginning to look at new ways of handling data applications to control costs and be more efficient. These include both in-house and third party solutions:

- Private clouds for auto-provisioning and cost allocation to business units
- Multi-tiers in a single data centre facility (in-house and third party)

In house, multi-tiering may prove a short-term solution as business dynamics change. E.g., whilst it owned PayPal, eBay built and used a multi-tier data centre with critical data located in the most resilient area and other applications located in areas of lower resiliency.

Multi-tier deployment depends on how companies have built their applications and how technically resilient they are. Issues of what happens to service if a company loses a facility and whether that is an acceptable business risk will play a role in the final decision.

Given the over-burdened or highly priced state of some power grids in Europe, it makes sense for CIOs and data centre operators to evaluate all options and to update their thinking by considering multi-tiered options. Today, this area of data centre design is gaining traction.

Two data centre operators already offering multi-tiered options are Sentinel in Canada and Verne Global in Iceland. Verne Global is leading the field in Europe as well as enabling users to access Iceland's advantageous power profile and grid reliability.

The hybrid solution

Whilst Nordic and Canadian data centres offer geographical disaggregation, Verne Global in Iceland is unique in that it offers both types due to its innovative Campus design. This hybrid solution provides enterprise users with the opportunity to combine both geographic and multi-tiered campus disaggregation at the same time in the same location.

The hybrid approach to computing can benefit the enterprise in two ways:

1. *Cost savings via optimisation*
   Not over-paying for proximity and resiliency that is not needed or warranted

2. *Reducing risk*
   Customers are no longer relying on power infrastructure that is outdated, unsustainable, in transition to renewables (such as in Germany) or suffering demand/supply issues and working to the limit as, e.g., in the UK

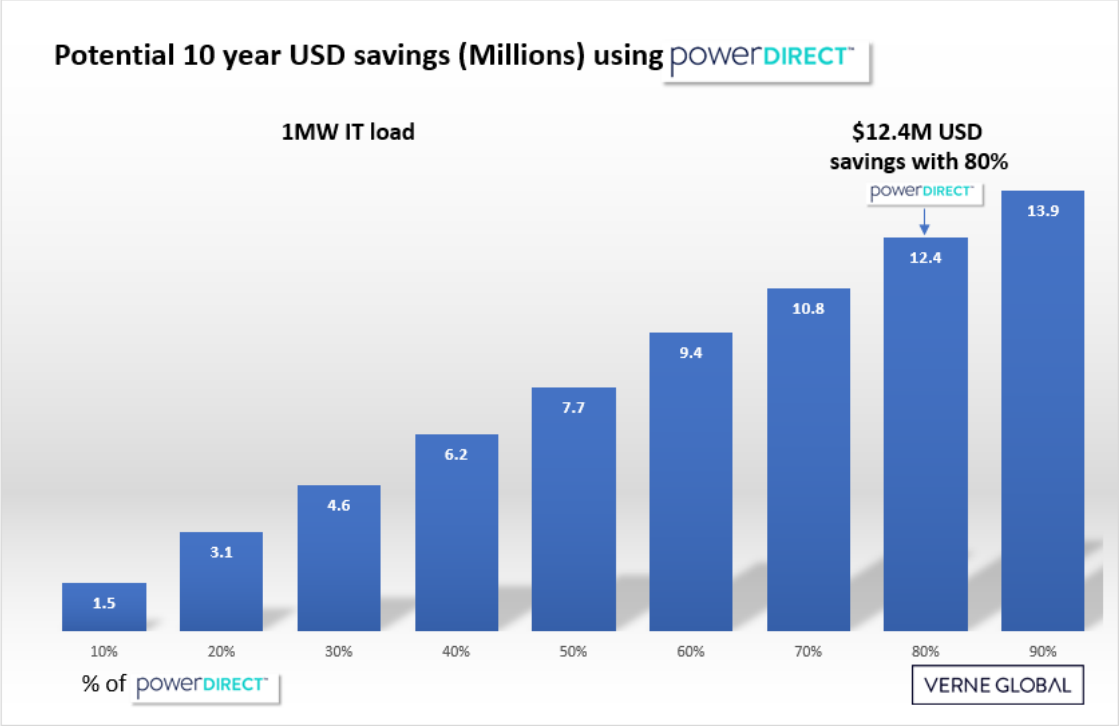
Verne Global's data centre in Iceland is designed as a multi-tier environment within the same campus. This design allows customers to run low resiliency workloads in one data hall and mission critical racks in another, so that costs match workloads efficiently.

This makes it possible to split applications across different profiles. For example, users can split their applications across Tier 3 (branded powerADVANCE), and Tier 1 (branded powerDIRECT) which offers high power density but with racks connected directly to the grid - with no UPS back-up. No UPS back-up delivers a reduction of 80% in power costs. At the same time, users are assured of 4 x 9's availability provided by Iceland's modern, industrialised power grid and abundant energy supply.

Making the business case

Savings from disaggregating data applications within the data centre can be substantial, particularly when the hybrid approach is used. In terms of power costs, savings of 60-80% can be achieved over 5-10 years if the applications are smartly placed in the correct hosting environment. Verne Global has calculated that, based on a 1MW load across a 10-year deployment, a saving of $12.4M USD is achievable. This is in addition to the 70-80% cost savings on power that is made when initially colocating to Iceland.

The graphic below illustrates the potential USD 12.4M USD savings that can be obtained by moving 80% of applications to the lower resiliency profile of the Tier 1 powerDIRECT™ environment.



As optical technology advances, connectivity to more remote locations is much less of a challenge. For instance, Verne Global's data centre campus in Iceland can reach customers in London or Paris in less than 20 milliseconds and New York in 41 milliseconds. This level of latency is suitable for a range of operations including database hosting, cloud computing, research projects, general business applications and disaster recovery. The downward trend in bandwidth costs will continue to facilitate colocation with data centre operators who can offer both innovative solutions and significant cost savings.

Section 4 Industry sectors which could benefit from disaggregation

Companies that have virtualised their applications can all potentially benefit from disaggregating applications – either geographically or by colocating in a multi-tiered data centre or by combining both. CIOs in all industry sectors will have applications in their portfolios that could be handled at lower resiliency. Research by Verne Global indicates that 25% of financial sector applications could run in a Tier 1 data centre. The table below gives examples of potential opportunities for cost reduction and smarter handling of data applications across multiple industry sectors:

Examples of industry sectors that could benefit from multi-tiering

Sector	Potential Lower Resiliency Applications
Academic	Research & Development projects
Automotive, Aerospace & engineering	Research & Development Manufacturing data Sales records Machine Learning/Deep Learning Artificial Intelligence Computational fluid dynamics Crash-test simulations Data analytics
Chemicals & Pharmaceutical	Research & Development Computational fluid dynamics
Financial Services	Quantitative Hedge-fund Analysis Machine Learning/Deep Learning Artificial Intelligence Data analytics Tax Accounting Insurance Pensions Any data sets accessed on an annual or infrequent basis 7-year compliance data
Healthcare	Research & Development Genome sequencing Lifestyle management data Bio-technical research Computational fluid dynamics Data analytics
Business Support Services	Email Business Continuity Back office administrative data Data storage Disaster Recovery Research & Development
Other	Film Rendering Post-production work Data Mining

Section 5 Next steps

Power is already the key challenge in the data centre world – not just the cost and availability, but reliability, environmental impact and sustainability. This issue is clearly getting more challenging as demonstrated by impending legislation in The Netherlands on carbon-neutral buildings or the desire for power created by electric cars.

In recent winters, many European power grids have run alarmingly close to their full capacity, as analysed in Northbridge Energy Partner's 'Mind the Gap' thought piece.

Data centre users interviewed in this report have typically been among the leading players in thinking about these issues. CIOs have gone back to basics on their compute usage and data centre requirements, and asked themselves key questions about what they need rather than what they are currently doing.

This transition has often been challenging. As well as the technical challenges in changing IT architectures and migrating data, there have also been the internal and cultural obstacles. Users often inherently want to think their data is residing locally, and run and managed internally.

Yet, the CIOs leading the way in application disaggregation are seeing clear and quantifiable benefits. They are seeing significant cost benefits in optimising their architectures, not just in locations but also 'right-sizing' to ensure that they are only paying for capacity they actually need.

Business are also benefitting from better performance and future-proofing their business against impending demands to restrict power usage and CO2 emission as CIOs align data handling strategies to support broader business financial targets.

Application disaggregation is changing from a 'nice to have' to something that all CIOs now need to implement, and this paper shows some of the clear benefits and practical considerations in achieving strategies that will lower IT operational costs and create more sustainable data management solutions for long term business success.

BroadGroup

Established in 2002, BroadGroup has achieved rapid recognition and growth through delivering quality research and insight in a number of niche and emerging areas of the telecommunications and technology sector. The Data Centre practice has been the main focus of the company since 2002.

BroadGroup has become the **leading European research and consultancy group** in this area, through three divisions; events, reports and consultancy.

The **events** business was launched at the end of 2004, both extending its value chain and market reach. Annually, BroadGroup holds events around the world, across Europe, Middle East and Asia. The flagship event, now known as Datacloud Europe, attracted over 1,500 delegates in Monaco in June 2016. The 13th Datacloud Europe will take place on 6th – 8th June 2017 at the Grimaldi Forum, Monaco. Other events look at different aspects of the data centre market from investment to power and cooling to demand in different regions of the world.

The **reports** business produces a series of in-depth reports covering data centre markets in the UK, Western Europe and Eastern Europe, Asia and Latin America, the rise of managed services and cloud services, the migration of outsourcing to third parties, and needs of particular vertical markets such as financial services.

The **Colocation Market Quarterly** report, launched in mid-2014 is a quarterly tracking service which offers high value *and* data integrity for operators and investors in Europe and internationally.

The **consulting** business has expanded rapidly over the last eight years, providing bespoke and customised services to users, operators, investors and development and government agencies, specifically in the data centre area.

BroadGroup Consulting is headquartered in Leatherhead, Surrey. It is incorporated as a limited company in the United Kingdom, and its registration number is 4894635.
Further details can be supplied on request.

Email contact: marion.howard-healy@broad-group.com
www.broad-group.com

www.data centres.com
http://www.datacloudcongress.com/