Workshop topics

- Is there a lack of environmental policy measures and publicly available data on data centres?
- Are existing measures capturing the key elements of the methodological framework appropriately? Is there a need of adaptation of their methodological basis?
- Is there a need of new metrics / harmonisation of existing ones on the environmental impact of data centres?
- Should the Data Centre, in the context of environmental policy measures/methodologies/metrics, be approached as a system or should it be approached as its individual components?

Introduction to this paper/presentation

This paper and the presentation to the Workshop addresses the four topics above from the perspective of European and international standardization bodies before summarising the points made.

Digital data

Digital data is the foundation of today’s business, societal and social interaction.

The growth of the amount of digital data stored and transported across networks is fed by its susceptibility to damage or loss, the need for its accessibility and a rise in the number of ways in which that data is viewed (often leading to multiple identical copies being stored on multiple devices).

However, the main factor in the volume of data is the non-business and non-societal related applications – i.e. consumer activity such as video distribution from established television operators, Netflix, YouTube and social networking applications such as Facebook. In fact the primary result of increased broadband bandwidth and access has been this increase in consumer applications.

The main reason for the growth of data centres is to support this demand. There is a vast amount of survey data to support this position.

Data centre standardization

There has already been a significant standardization activity addressing energy and other resource consumption in data centres.

This began in Europe following Mandate M/462 and was initiated in ETSI via their Special Task Force 362 of which I was leader. This resulted (2009) in the production of ETSI TS 105174-2-2 “Broadband Deployment - Energy Efficiency and Key Performance Indicators; Part 2: Network sites; Sub-part 2: Data centres” which contains best practices information and basic information about the difficulty of defining universal KPI limits for data centres.
This was subsequently followed up by their Special Task Force 439 of which I was also leader. This resulted (2013) in the production of ETSI ES 205200-2-1 “Energy management: Global KPIs; Operational infrastructures Part 2: Sub-part 1: Data centres” which contains detailed specifications of Objective and Global KPIs (see below) for application in all data centres. ETSI have also formed an Industrial Specialist Group to review energy management in data centres (ISG OEU). This is a relatively closed group populated by major data centre users (both network operators and enterprise users).

CLC has also been involved in developing data centre standards via TC215. TC215 is producing a series of standards for data centre facilities and infrastructures (not IT equipment). EN 50600-1, EN 50600-2-1 and EN 50600-2-2 have been approved and/or published and other parts in the EN 50600-2-x series are underway.

ITU-T have released L.1300 which is, in large part, based on the Best Practices documents developed under the DG JRC activity on the Code of Conduct. Elsewhere in the international arena, ISO/IEC JTC1 SC39 WG1 is producing specifications for data centre KPIs but these are at an early stage.

In order to coordinate activity across all ESO bodies and to monitor activity elsewhere, the CEN/CLC/ETSI Coordination Group on Green Data Centres maintains a live list of all relevant standardisation activity entitled “Standardisation landscape for the energy management and environmental viability of data centres” and the current reports are available at:

- CEN/CLC/ETSI EXECUTIVE SUMMARY
- CEN/CLC/ETSI REPORT

What is a data centre?

All the main standardization activities in ETSI, CENELEC and ISO/IEC have a common definition of a data centre which is:

“a structure, or group of structures, dedicated to the centralised accommodation, interconnection and operation of information technology and network telecommunications equipment providing data storage, processing and transport services together with all the facilities and infrastructures for power distribution and environmental control together with the necessary levels of resilience and security required to provide the desired service availability”

It is important for such a general definition to be used because many commentators on data centres have a fixed view based on their experience - but that view does always match the true function of a data centre. With this working definition, standards have to support data centres ranging from the simplest composition, comprising a few cabinets in a space within an existing building, to something the size of a sports field and beyond.

The smaller data centres may not have separate infrastructures for power distribution and environmental control and may not even separately account for their resource usage. Just because these are not monolithic data centres in dedicated buildings with huge power supplies and environmental control systems does not make them any less a data centre. No one would notice the impact of such data centres in environmental terms as they would be within the “noise floor” of general building operation.

The size of a data centre is not directly linked to the importance of the data managed within them. Moreover, the reliability of the “data centre” may not be linked to their size or complexity since some organizations may chose to construct multiple centres, each of which has low level of infrastructural resilience, as a means of protecting their overall service availability. Resilience in such a situation may comprise other small data centres in other buildings interconnected by core and access networks.

This realization comes as a shock to those who conceive of a data centre as a custom-built building or part of a building using significant amounts of energy to function and to maintain function.

What is the role of a data centre?

The standards being written for data centre facilities, infrastructures and resource efficiency metrics describe at least four different types of data centres which support different business models - and the type and purpose of the data accommodated within them may vary dramatically.

How we determine efficiency?

Discussions regarding energy or other resource efficiency usually centre on how much resource is used for a given amount of work done - but that is not a universally applicable concept for data centres.
The different types of data centre and their customer’s profiles means that the relative importance of data storage, data processing and data transport varies widely. Some data centres will communicate little but store larger amounts of data while others may do the reverse. Therefore it is difficult to define the “work done” by a data centre. One option is consider the energy efficiency of the equipment within the data centre but this is futile since actual efficiency is dependent upon operational conditions rather than design Technical KPIs (see below).

Instead, most activity has focussed on minimising the resource required to support the IT processes (storage, processing and transport) and the KPIs reviewed so far have addressed the effectiveness of those resources in supporting the data centre function itself - this is termed “Supporting resource consumption, or supporting resource effectiveness”.

It should be pointed out however that resource consumption is proportional to the reliability demands of the data centre infrastructures. So high availability data centres with multiply redundant power and environmental control systems will use resources less effectively when compared to the consumption of those resources by the IT load itself - but this is intentional and vital for business continuity and cannot be penalised.

Metrics

The report of the CEN/CLC/ETSI Coordination Group on Green Data Centres describes KPIs in the same away as the ETSI ES 2050200-2-1 as follows:

- **Technical KPIs**: KPIs which provide information regarding resource efficiency of components, sub-assemblies and systems under defined operating conditions.
  
  It is tempting to assume that to create a data centre using such products with the best Technical KPIs will render it most efficient - and suppliers would like to promote that idea - but in reality the actual operating conditions may be completely different that those employed in the determination of the Technical KPIs.
  
  It is this message that data centre operators continue to deliver and they would prefer to consider a range of Objective KPIs.

- **Objective KPIs**: KPIs which address a particular operational objective

- **Global KPIs**: which combine Objective KPIs in particular ways - and which should be employed with caution.

Critically, both Objective and Global KPIs are overwhelmingly agreed to be trend indicators and not comparators between different data centres for the reasons explained above.

Equally importantly, we must distinguish between the specification of a metric and any definition of limits. Limits may be set by regulation or local agreement but are not the role of standards in this field.

Summary

In order to apply environmental policy measures to data centres, one has to determine the objective - remembering the potential conflict between those policies and others demanding wider availability of high bandwidth broadband delivery.

In order to discuss whether there is lack of available data on data centres and the policy use that such data may be put, one has to define the boundaries of the data centres for that information is sought.

The smallest enterprise data centres are not going to provide useful data (despite them being the largest cohort of any potential survey) whereas the largest may provide data but subject to their wide range of functions and business models will not always provide universally applicable information.

Standardisation has already considered in some depth the problems associated with the highly complex systems-based nature of data centre and has clearly differentiated the types of metrics that can be universally applied.

In addition, standardisation groups have considered the value of operational guides (indeed, the EU CoC Best Practices documents should have been developed and maintained as European Technical Reports many years ago) which can assist in improving the resource effectiveness of all data centres - independent of type, business model and size.

If policy is to be developed, having considered all the above issues, then ESOs stand ready to respond accordingly.