the helicon
where great minds come together!

A Practical Approach

The iOS Protocol

Contents

Part 1 – Introduction

ICT staff roles and workflows aren't the only things changing as ICT evolves and budgets tighten – amongst others, data centre power- and cooling system designs may also need an update.

ICT managers should re-evaluate how their data centres run. Lower energy use doesn't have to mean sacrificing hardware's reliability and performance. Replacing old hardware and updating data centre cooling systems provide great benefits.

Consolidate server hardware

One of the most effective ways to improve energy efficiency is to compute with fewer servers, virtualising multiple applications to run simultaneously on each server. Virtualised servers host as many Virtual Machines (VMs) as the physical server hardware - CPU cores, memory, network I/O and so on, allows.

The potential impact of server consolidation is stunning. A traditional physical server running a single enterprise application might use 5% to 15% of the server's total computing resources. That same server could host 10 VMs, each using an average of 8% of the server's total computing capacity, replacing 10 physical servers while still leaving 20% headroom in its total computing capacity.

Server consolidation is not an all-or-nothing proposition, and varies depending on the type of workloads and the technology and team supporting them. Organizations that are new to virtualization may start with limited consolidation using noncritical applications and then gradually increase efficiency consolidation levels and virtualise more important workloads. Virtualization raises the importance of systems management tools and practices to track, monitor and control VMs.

Adopt energy-efficient servers

Server upgrades and consolidation are often approached as independent projects, yet the two initiatives both increase overall energy conservation. Businesses can easily virtualise the existing server farm, then systematically upgrade servers during subsequent technology refresh cycles. Server upgrades give IT teams the chance to optimise consolidation and balance the distribution of VMs across servers. Taken together, virtualization on more energy-efficient server platforms will make significant improvements in energy conservation.

New server designs provide greater computing capacity while reducing data centre energy consumption. A new Intel Xeon processor dissipates 65 W of heat, compared to 150 W just a few years ago, even with many more cores and memory. Part of the energy savings is in using slower processor clock speeds and relying instead on processor performance enhancements. These enhancements include methods like hyper threading, which basically allows a microprocessor to do the work of two processors, and processor throttling, which adjusts clock and voltage settings based on computing demands.

New server designs with energy-efficient processors and advanced power- management and conservation capabilities can in turn increase virtualization and consolidation, further reducing energy demand.

Improve data centre cooling systems

Consolidating enterprise workloads into fewer, more energy-efficient systems has a welcome side effect: The servers produce less heat, and that means the data centre requires less cooling.

Emerging standards from the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) also support server operation at elevated temperature and humidity levels. For example, class A4 servers carry an allowable upper limit of 45 degrees Celsius and 90% relative humidity.

The biggest challenge for most organizations is using the existing cooling infrastructure in the most efficient way and introducing alternative cooling methods where appropriate.

Take advantage of elevated operating temperatures and explore smaller computer room air conditioning (CRAC) designs and other efficient cooling systems, especially because alternative cooling technologies continue to evolve, and reap large savings.

One avenue to better cooling efficiency is the aggressive use of containment, such as a hot aisle/cold aisle strategy. This approach directs cooling airflow to the server rack spaces rather than to an entire room, so the CRAC system does not need to work as long to achieve the desired operating temperature.

When ASHRAE thermal guidelines are followed, containment can also limit IT staff exposure to uncomfortable temperatures and humidity.

Unfortunately, more efficient cooling can actually harm existing mechanical refrigeration systems. As cooling demands fall, CRAC cycle times get shorter, and frequent compressor starts and stops will wear down the large, building-grade cooling systems.

Moving cooling closer to the IT equipment is a more efficient cooling alternative for data centres. One option is to move the mechanical refrigeration from the building's roof to the server area, using smaller, high efficiency, in-row air coolers.

Part 2 – Overview of the iOS Programme

The iOS Programme outlines the concepts, frameworks and standards as they pertain to the data centre purpose, the data centre definition, its components, layouts, topologies, methodologies, designs, plans, documentation, availability, security, safety, resilience, capacity, efficiency, operation, telecommunication, power and electricals, cooling and mechanicals, cabling, civil, structure, architecture, lighting, earthing, signage, connectivity, information technology, software, application, monitoring, network and operations centre, servers, storage, backup, recovery, mirroring, archival, and any method, technology, product or service that relates to the hosting, delivery and facilitation of the organisation and its business applications.

We have identified six distinct 'layers' that make up the complete DC environment, as follows:

1. Facilities

A set of components, at the facility level, utilised for applications delivery, housed at the data centre node.

It is an independently operated space to house data centre infrastructure. The requirement specifications of Applications, Platforms, Compute requirements and Topology dictate the qualifications and provisions of the Data Centre Site.

It specifies the power, cooling, etc. infrastructure and related services that a site is outfitted in order to meet the specifications of the layers it supports.

2. Hardware package

The logical arrangement of computing resources. Compute is a layer in which processing requirements of the application are defined in an abstract form. The compute cycle cloud would subsequently map to actual virtual or physical cores in the IT Infrastructure (ITI) layer.

The ITI layer is a set of physical components at the information technology level, utilised for the application delivery, housed at the data centre node. The ITI specifies the network, the servers, the storage, etc. infrastructure and related services that a given site is outfitted with in order to meet the specifications of its supporting layers.

ITI specifications require explicit levels of: Availability, Performance, Capacity, Economics, Security, Presence, etc.

3. Applications

The methodology by which the applications are delivered (Application Delivery Mode).

The Application requirements specification lead to an application Hardware package selection.

Common Cloud based delivery platforms are:

- Business as a Service (BaaS)
- Software as a Service (Saas)
- Platform as a Service (PaaS)
- Infrastructure as a Service (IaaS)
- Etc.

The Application topology specifies the connectedness, continuity and boundaries signifying a set of correlations, connections, dependencies, and relations between various data centre nodes (Hardware packages) for the delivery of the organisation's applications. Topology specifies the physical location, interconnectivity and interrelation of the data centre's Hardware packages.

Application specifications require explicit levels of: Capacity, Flexibility, Compliance, Performance, Availability, Security, Economics, Presence, Operations, etc.

4. Operating systems

The collection of software that directs a server's operations and management, controlling and scheduling the execution of other programs, and managing storage, input/output, and communication resources.

Operating system specifications require explicit levels of: Availability, Performance, Safety, Capacity, Economics, Security, Efficiency, etc.

5. Silicon

Silicon refers to the server microprocessor(s) that accepts digital data as input, processes it according to instructions stored in its memory, and provides results as output.

Silicon specifications require explicit levels of: Reliability, Connectivity, Safety, Security, Efficiency, Availability, Economics, Presence, Capacity, Compliance etc.

As per the above, the authors are of the opinion that the Uptime Institute's Data Centre Tiering model portray a narrow vision of all the data centre disciplines and cover little ground as to what the real purpose of a Data Centre is, it does not consider all the layers that make up a data centre and their effectiveness, that it literally serves no practical purpose at all.

The iOS Programme redefines the way data centres are defined, looked at and modelled, and hence from both conceptual and physical points of view, optimised towards the purpose they are intended to fulfil.

The iOS approach

Application Environment (AE)

Applications live, exist, and interact only in an environment that supports its purpose. The data centre caters to that need by being the infrastructure that supports the application environment (AE) and its well-being.

Comprehensive

It is founded on the basis of bridging the existing data centre industry gaps, in its predecessors' standards and guidelines by covering a comprehensive outlook of the data centre, its investment, its design (especially in terms of power- and cooling efficiency), its people, and its management.

It starts off by first establishing a baseline of where the DC is currently positioned in terms of the above.

The iOS Programme effectively utilises what is available and re-usable and makes the required short-term, medium-term and long-term change recommendations to address the issues of Redundancy, Reliability, Connectivity, Safety, Security, Efficiency, Availability, Economics, Presence, Capacity, Compliance etc. that is pertinent to the specific site.

Efficiency-driven

The iOS Programme takes great measures to ensure efficiency compliance of data centres in both operational as well as infrastructural sense as rigidly as other parameters such as availability, security, redundancy, etc.

Management infrastructure

Data Centre Infrastructure Management (DCIM) is a category of solutions which are created to extend the traditional data centre management function to include all of the physical assets and resources found in the Facilities and ICT domains. Since DCIM is a broadly used term which covers a wide range of data centre management values, each deployment may include a subset of the full DCIM needed and expected over time.

It may be supplemented by or integrated with other management software that includes computer and server management, virtual systems, proprietary systems, automation, and the management of the services and applications used for data processing.

Data centre management requires a number of tools, IT policies and strategies to create and maintain a secure and efficient data centre.

International

The iOS Programme is outfitted to bringing the world's best and greatest together and allow the data centre stakeholders benefit from the synergy.

Operation conducive

Our philosophy dictates that a data centre's availability, security, safety, compliance, etc. cannot be measured unless its operational parameters are fully analysed and understood. Merely having highly redundant components without measure of how such components are operated, by whom, in accordance to what policies, etc., by virtue of continuous measurement we fulfil the iOS Programme's requirement of compliance.

Effective

One of the iOS Programme's core missions is practicality and effectiveness.

Lack of understanding the real needs and addressing them to the point, to give all stakeholders in the data centre a fair approach towards handling their mission-critical applications sparked the emergence of the Programme.

Cloud-inclusive

Virtualisation, as in abstracting away from actual resources and establishing cloud-based infrastructure is the essence of this day and age.

The iOS Programme proudly takes the initiative of analysing and evaluating cloud-based infrastructure in both the physical and the logical aspects as a key basis of our next generation data centre framework.

Localised

The iOS Programme is built with the capacity to be localised to towards both logical and physical realms. Whether the requirement is localising to a specific industry, to a specific geographic area, to a specific site (or all of these) or transforming itself into local polices and norms, the Programme caters to that need and promotes such localisation for ease of access, enhanced effectiveness, and compliance, to say the least.

Application-centric

We acknowledge and promote that the purpose of a data centre is delivering applications effectively (performance and cost) and that the data centre is but the infrastructure supporting the application environment, it is the sole and prime entity that through the iOS Programme provides a guideline for the data centre from an application-centric approach.

This approach eliminates redundant and goalless efforts, lack of understanding data centre’s purpose and what is it supposed to do, which in turn enable the stakeholders to the set the right expectations and plan in alignment and towards ultimately satisfying the users of these applications needs.

Part 3 – The iOS methodology

Ever increasing business demands, brought about by an increasingly connected world, has put more and more pressure on organisations' data centre capabilities. The increase in processing demand has led to significant growth in the amount of technology being placed in data centre environments, the consequences of which are now being felt by many.

A large increase in power consumption has led to spiralling overheads, and with the increase in power comes an increase in heat. Legacy systems and infrastructures can rarely cope with these new influencing factors, resulting in poor performance and demanding constant attention.

Downtime means frustrated customers and lost revenue and business opportunities. It's therefore of utmost importance for companies to audit the data centre to discover (potential) risk factors which could lead to downtime before it is too late. By having an in-depth understanding of the (potential) risks in your data centre or telecommunication site, you are able to prioritise investments for improvement in line with the risks in different areas.

Consequences and outcomes

The inability of a data centre to provide the desired output means that applications will not run at the speed required to deliver the greatest financial return for the organisation. In some instances, the inability of the equipment to work under the heat conditions in which it is placed may mean total failure and therefore no access at all to business-critical operations.

The cost implications increase in the face of under-performance of the organisation. It becomes harder for an organisation to justify continued spend on the data centre environment, despite return on investment studies, when previous performance has indicated otherwise.

There are two solutions available – expansion or consolidation. However, in many cases neither is straightforward:

Expanding the space available for servers may solve the problem of heat affecting performance, but it is rarely an option and does not cater for a reduction in power. Cooling and power management is now paramount as more and more computing needs to be carried out in smaller footprints. Although some organisations are looking to move their physical presence out of major urban areas and into more spacious semi-rural areas, most do not have the luxury of being able to do so and are actually in need of reducing the overall physical space they use.

So more organisations are looking to consolidation, but how do you consolidate into space that you do not have? Floor space is a considerable problem with factors including available floor space, height, power supply limitations and cooling inefficiencies, posing limitations on further expansion.

Identifying the key issues

Until recently there has been an obvious lack of joined-up thinking between the owners of each element.

The purchase cost of server equipment has fallen over time to the point where the simplest solution for lack of capacity is to buy new equipment. This is also making it easier to solve the issue of equipment that has reached end of life or does not meet the new business requirements – replace it. However, in most instances, legacy equipment has been left in place, either because it is used to run less critical applications, or because the cost and effort of removing it is high. It is estimated that as much as one in three servers are comatose, i.e. it is consuming energy and requires cooling, but does not do any significant (or none at all) computing.

Aside from the heat consumption and power generation issues concerned with the server, there are several aspects that must be considered at the same time. There have been instances where companies have invested in new server capabilities without checking that they have the necessary physical space to cater for its integration, that is, the door isn't big enough!

The issue of size is not something that can be levelled at blade servers only. Many organisations have made massive investment in introducing blade computing to their environment in order to save space and reduce power consumption.

These savings are the traditional view, however it is true to say that blades can reduce operational costs through quicker set-up and update times as well as significantly reduce the level of cabling required.

This modular approach of blades seems to suit many companies; however it is not always the answer. Although blades do consume less power than traditional server set-ups, the increase in density of blades means that the issue of heat is still prevalent. In many instances, organisations are faced with using only two thirds of the rack as the blades cannot cope with the heat when full, and in order to meet the processing needs of the data centre, that means more racks.

The racks themselves play a large part in controlling heat dissipation and are subject to equal levels of technological advancement as the processor. As well as ensuring a free flow of air within the rack through structured and consistent cabling, the heat must be evacuated from the rack, but cannot be simply dumped out of the space for fear of overloading the cooling system of the environment.

More organisations than ever are creating 'server farms' in order to meet the needs of the online economy, created from scratch or by gathering existing equipment from across the organisation and placing them together. If this is done without an acute understanding what the effect will be, which is often the case, there is a danger of less attention being given to the physical environment in which they sit. The consequences of the heat distribution can be dire and is often only apparent once serious problems have occurred.

The setup of the environment is primarily driven by four core factors: power supply, the air-conditioning system, available space, and the security of the data centre. Additionally, the data centre has become part of the normal working environment, which means that health and safety concerns are more important than ever. In the past there has been good practice, whereas now it is subject to legislation.

The overall cooling of the room is driven by the Computer Room Air Conditioner (CRAC) set-up and the efficient use of the air supply plenum, the space below the raised floor. Organisations are finding that the positioning of the CRAC units with respect to the overall geometry of the room and the heat load distribution, rather than relying on what is perceived as the logical placement, results in a more efficiently cooled area. Equally, particular care and attention to the air supply plenum, ensuring that it is completely enclosed and therefore correctly pressurised, is necessary for correct airflow and to prevent moisture collecting in the data-centre.

The footprint or the area surrounding the racks that must be free of obstruction to allow for air circulation, and the setup of the rack aisles is another critical factor in managing the overall heat load of the data centre.

There is, however, a physical limit to the cooling which can be provided by the traditional approach of wall mounted CRAC units with a raised floor plenum. In any event, the very act of considering capacity for a whole room assumes that a uniform power/ heat density is found throughout the computer room, but data centres are heterogeneous environments containing a mixture of server topologies, plus storage and communications equipment.

Common practice used to be to simply spread out the higher powered equipment around the data centre. But companies are increasingly dependent on IT and higher density form factors like blade servers answers the need to achieve better utilisation of space. It is an easier and cheaper solution than increasing the size of the data centre. The problem that arises is that racks of high density equipment, i.e. more than 10kW per rack in legacy data centre space, which is typically built out only to between 2kW and 5kW per rack position, are going to cause the emergence of hot spots which in turn will cause equipment failure.

The security of a data centre is paramount. It is not hard to understand the potential damage caused, physically and operationally, to a business if the data centre is damaged by humidity, moisture or even fire.

Additionally, correctly sized Uninterruptible Power Supplies (UPS) are critical to support and protect the load following a power failure – something that is a real possibility considering the pressure being put on the supply.

Many of the elements within the data centre discussed have to be thought about when considering security in conjunction with efficient operation, as statistics show.

> *Concrete, brickwork and gypsum walls with fire protection do not help IT Hardware against humidity and moisture. 75% of cable/pipe work entry points are overlooked and can be the main source of entry for killer gases. 85% of building engineers, architects and fire-fighting experts have no idea of the critical values specified for protecting magnetic disk media and related computer hardware*
>
> *Source: Debus11*

It's impossible to remove all of the risks entirely, however steps should be taken to minimise them where at all possible. This requires a thorough risk assessment audit, ensuring that all the necessary aspects are covered.

Audit, audit, audit

There are different ways to approach the management requirements of a data centre, yet one consistent message is to make sure that the organisational and therefore data centre requirements are thoroughly understood prior to beginning any kind of deployment, and to provide a baseline against which these requirements can be measured can only be achieved through a thorough auditing process.

Thought through properly, the data centre becomes a success, done badly and the failure can be significant. There are several questions from the outset that organisations should be asking themselves:

- What are the business objectives?
- What is in place and how well or inefficiently is it currently running?
- How do we make the best use of what is already in place?
- What type of infrastructure do we think we need and what is our starting point?
- Are we embarking on a consolidation exercise, beginning a refresh or starting from scratch?

By undertaking a full audit, the logistical elements of building and running the data centre can then be tackled one by one, reducing the risk of significant problems to a minimum. Additionally, the involvement of the necessary parts of the business can be sought from the outset, reducing the chances of any political issues that may arise.

Historically, the data centre has been solely the responsibility of the ICT department whereas now the facilities resource of the organisation should be just as heavily involved. Both should take a lead in the development of the data centre, because when individual involvement is disjointed, so generally is the solution.

Equally, involving the commercial decision-makers from within the organisation, throughout the development process, will pay dividends. If they have a better understanding of what work is required to deliver the commercial objectives of the organisation from an IT perspective, the business case, finance and timescales can be worked on to achieve the best results.

The buy-in from the rest of the organisation will only stay in place if the proposed solutions actually work. More importantly, undertaking significant infrastructure changes involves risk which needs to be reduced to a minimum.

By identifying the conditions of the data centre and testing the solution thoroughly prior to deployment, the level of confidence in what is proposed will be high. Proving the concept will deliver long-term benefits by being able to test what effect unexpected and future changes to the environment may have.

The iOS approach

Using in-house developed tools and methodologies, as well as methodologies, tools and input documentation based upon work done by Lawrence Berkeley National Laboratory, the Green Grid, ASHREA, The Technology Innovation Agency (TIA) and the Building Industry Consulting Service International (BICSI), to name few, we comprehensively gather accurate data about all the elements of the data centre.

We fully customise our approach for each client in order to meet the client's specific requirement and needs, based on a Business Requirement Specification that ensues through discussions with the relevant element owners.

Amongst others, the following areas can (and should) be addressed:

- Management and Organisational Best Practices
- Technical Best Practices
  - Airflow Management
  - Air Handler Systems
  - Humidification
  - Plant Optimisation
- IT Equipment
  - Servers including Blades
  - Consolidation Maturity
  - Virtualization Maturity
  - Racks and Eco Racks
  - Monitoring and Control Systems
  - Access Control
- Electrical Infrastructure
  - UPS and Power Management Systems
  - Lighting
- Commissioning and Retro commissioning

The processes, guidelines, tools, methodologies used and recommendations are presented in a formal document at the completion of the process.

Phase 1: Inform

The first step in the iOS process is the Inform phase, in which we form a picture of the client's current situation and the goals and expectations of what we're about to do.

Because the needs and requirements of each client is unique, and each element of the data centre may be at a different level of maturity or sophistication, it is important that all the elements of the data centre are audited during this phase.

It is important to get an accurate *before* picture at this stage so that after the project is complete, we can use the phases' findings as a gauge to answer the questions, *did we add the value we hoped to add?* — or — *did we achieve what we set out to achieve?* We will also consult with the client about key metrics, how they might be improved, and how our processes will address those issues.

We will endeavour to get the client to articulate what they're really hoping to accomplish, from the word go. What we do — and encourage our clients to do — is get in a room and really clarify business goals, strategic goals, the technical goals, and how they're aligned.

At the same time, the objective is to determine what strategies and tactics will fulfil the client's goals.

For example, if the client says the goal is to rollout and implement a power management system, it's important to have the client representatives explicitly define what they're expecting it to accomplish.

Phase 2: Discover

Once the client has articulated the needs the project should fulfil, we move into the Discover phase. During this period, we prefer to meet with users to seek out and validate the similarities between what the business wants to achieve and what the users of the information really want and need. Our goal here is to reconcile the differences between the business goals and the users' expectations. At this point, we attempt to learn everything we can about the client's infrastructure:

- Air conditioning: does it have the relevant capacity?
- Airflow: is it distributed correctly?
- Monitoring: can temperature and humidity be efficiently monitored?
- Underfloor areas: are they clean?
- Cabling: is it structured, tied back and properly marked and labelled?
- Power: is there adequate redundancy and backup, and can it be monitored and controlled?
    - Does it include server energy efficiency policies and measurement?
- Security: is it adequate and controlled?
- Racks and enclosures: how are they deployed?
- Servers and processors: How and where are they deployed?
- Server deployment maturity: What level of consolidation has been achieved?
- What level of virtualization has been achieved?
- Etc.

At the end of both the Inform and the Discover phases, we meet again with the client to *calibrate* the views and information we've gathered.

We try to articulate it in a way that says*: Here are the business objectives and what the requirements are. Here are the similarities, here are the discrepancies, and here are a set of recommendations to reconcile those.*

By comparing these goals we can lay out a tactical plan of action in which the most important features are implemented first. In that way we give the client the next most important functionality that they need while working toward the ultimate goals of the business, and ultimately, over the duration of the project, meeting the objectives of both the users and the business.

Phase 3: Plan

After the users' and the business' objectives are thoroughly investigated and discussed, the next step is to formulate a plan. During the Plan phase, every detail of the solution will be specified, by asking questions like:

- How are we going to solve the problems that were identified?
- What tools, products or services do we use?
- How do we lay out the capabilities?
- What are the tactical steps?
- What is the architecture?
- What can be automated at this stage?
- What short-term fixes (low or no cost) can be implemented?
- Etc.

When creating a project schedule, every effort is made to specify all of the client's key objectives uncovered in the Discover phase and revisit those points after each milestone is met. The plan is continually re-defined if required, as the client, the users, and we review the deliverables that the iOS Programme comes up with.

During this phase it is important to create the technical prototype. At this point, we are looking for technical areas of risk or circumstances that will make this project and the processes perform poorly, and we create a controlled benchmark against which implementation can be measured.

Any issues that will affect the day-to-day functioning of the system should be investigated and alleviated during the technical prototype phase. In the case of an application rollout, issues include, but are not restricted to:

- How much physical intervention (time, resources, environment change(s), etc.) will be required?
- How will the intervention impact the availability of the data centre?
- How much of the processes and requirements can be automated?
- What is the intervention time frame for each audited device?
- How many, and of what calibre are the physical resources that will be required?

Phase 4: Build

The Build phase includes planning and building the solution, creating test plans, testing the solution, and creating maintenance plans to sustain the program.

Throughout this phase, we do lots of quality tests — even daily, if possible — to test each unit of deliverable. Then we tie it all together and do integration tests to ensure that we have all the pieces, and do they work well together? Ultimately, we would push that out in a staging environment and do acceptance tests.

As part of the Build phase we capture all the audit data into a discreet Configuration Management Data Base. The proposed system comprises modules which cover the full gambit of data centre management.

In brief

The need to manage the data centre environment, and in particular power consumption and heat load has become paramount. Although this is now firmly on the agenda of many manufacturers, the advances that is being made will take time to spread across the interested parties and become standard solutions. Even then, the issue of managing the environment will not be 'solved' as it is fair to argue that, for the foreseeable future, processor output requirements will also continue to rise and change.

There are several tactical solutions that organisations can consider deploying to help with the management of this situation, however the overriding message is that thorough and consistent planning is needed in the first place. Taking advice and understanding the start point, irrespective of the intended deployment, is paramount. Only then can an organisation build a clear picture of exactly which tactical solutions are suitable for them.

Much has been promoted about blades as a solution for lower power consumption and limited space. Equally, they do allow an organisation to act in a modular fashion when it comes to their computing needs, a step closer to utility computing. However, their very design means that more and more are packed into smaller spaces and so the low power consumption by each is offset by the total consumption of all, and so power- heat load issues still remain.

What we do know is that there are some fundamental, physical guidelines that organisations can adhere to which will deliver results irrespective of their overall requirements. These practices, such as ensuring the data centre is secure, cabling is structured and not obstructive, using hot and cold aisles, managing and controlling the power consumption, all add tangible benefits.

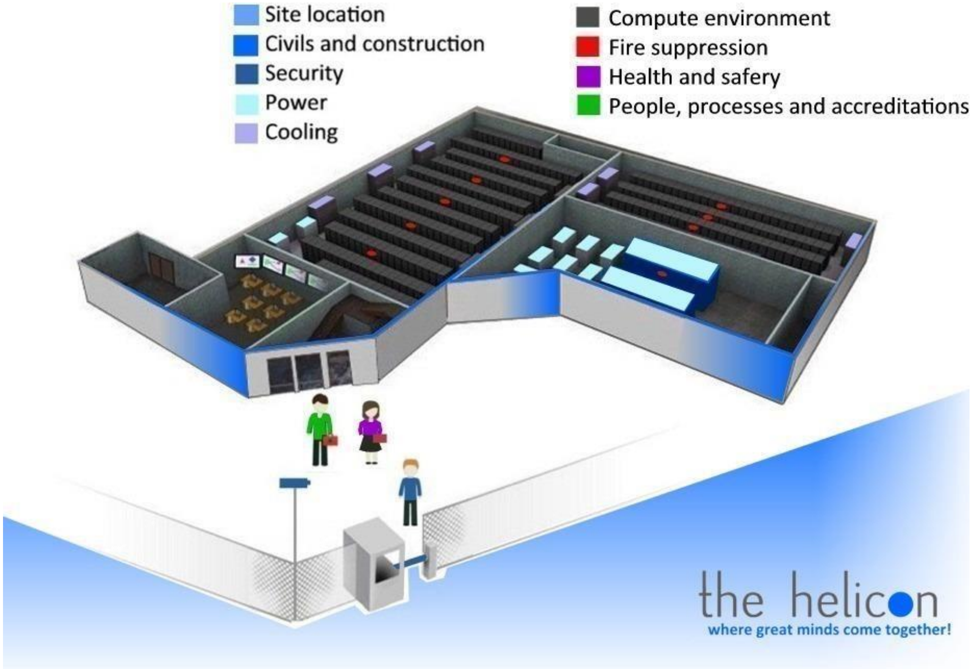
There are huge numbers of factors to consider and this should not be done in isolation of one or the other, but if planned well in the first instance, the risk of an ineffective data centre is dramatically reduced.

Part 4 – The iOS data centre assessment

The iOS Data Centre Assessment is a vendor-independent assessment of the physical facility and operating practices of your current or prospective data centre. It reviews the physical facilities: location, power, cooling, security, processes and procedures. It also does a high level review of the existing Application Environment.

At the end of the assessment we will provide you with a report based on our findings that:

- Details the current state of your data centre
- Outlines 'quick wins'—improvements you can achieve rapidly and with little investment
- Provides recommendations on how to align your data centre with your stated business and ICT objectives



Our team spend time at your site, interviewing business and ICT stakeholders to understand their objectives, and analysing the current condition of your data centre against those objectives to identify any gaps. A detailed examination of the physical data centre covers:

- The site itself—its location and any related risks or vulnerabilities
- Construction of the building
- Security arrangements for site access
- Power—infrastructure, redundancy and resilience
- Mechanical, electrical and cooling capacity and effectiveness, including scope for future growth
- Data Centre rooms—including environment, size and layout
- Hardware—high level observation of server technologies and racks, including:

- Server Energy Efficiency (SEE) - if servers are enabled to report
- Server's level of consolidation
- Server's level of virtualisation
- Rack Cooling Index (RCI) and Return Temperature Index (RTI) metrics

- Staff numbers, skills and knowledge, if appropriate
- Health and safety
- Accreditations – adherence to applicable standards such ISO, TIA-942, ASHRAE etc.
- Processes—benchmarking against iOS best practice processes that are based on industry standards

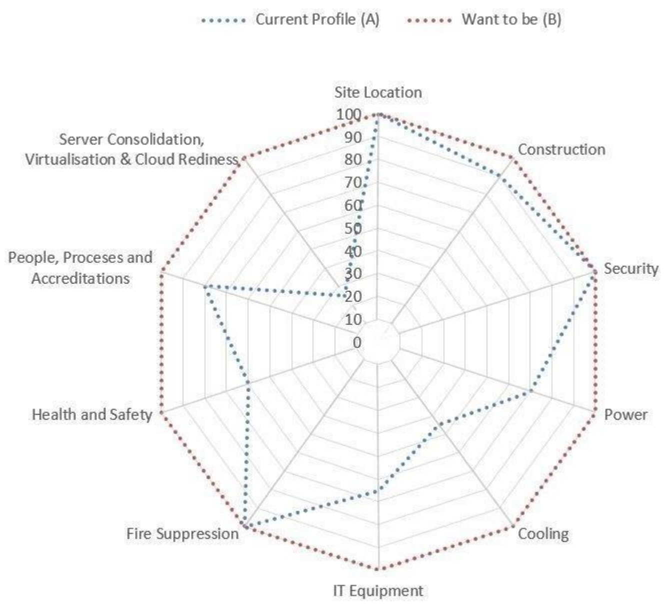
Benefits: Reduce operating costs, improve compliance and prepare for the future

Implementing our recommendations could:

- Raise your data centre's capability and performance in line with business requirements
- Contribute to operational cost savings and your Green ICT agenda
- Help you improve data centre resilience and reduce risk
- Support your data centre vision and strategy
- Help you prepare for future change or site relocation
- Assist in preparations for external audits for regulatory compliance

With the iOS Data Centre Assessment you get a clear understanding of your data centre's physical capacity, how effectively it's being used, and what improvements are needed to support current and future business needs.

We analyse various dimensions of your data centre and identify the gaps between where you are today (A) and where you need or want to be (B).



An iOS Data Centre Assessment study analyses 14 key data elements about your ICT environment and issues four reports:

- Site breakdown report
- Platform breakdown report
- Environmental report (Green report)
- ROI report

These reports are used to create an enterprise view of your current and future ICT states and contain data on server frame and operating system counts, square meterage, racks, power, cooling, carbon dioxide emissions, etc.

The ROI report helps you prioritise IT expenditures and focus on the optimisation paths that will have the most positive impact on your ROI and business.

Fast, customised and specific recommendations

Recommendations are based on the cost of implementing the technique relative to the overall savings estimated over five years. Recommendations are customised to your environment, unlike the one-sise-fits-all solutions common in the industry.

In the past, similar analyses have taken from three to six months. With an iOS Data Centre Assessment, a full assessment can be performed and results returned in only one to three weeks, depending on the location, size and complexity of your environment.

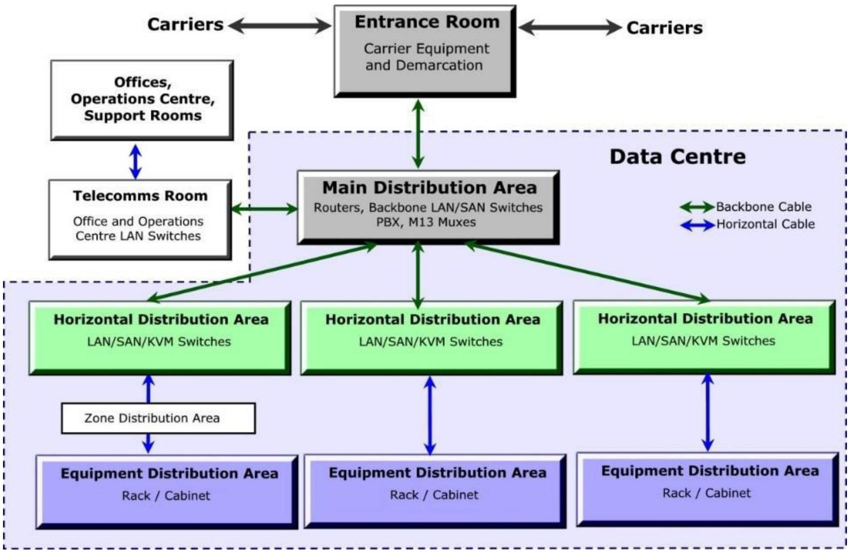
Quick view of assessment *complexity*

The differences between data centre and computer / server room design don't amount to a hill of beans for most people. The terms are often used interchangeably, but using them correctly makes a big difference if you're trying to communicate with a data centre design firm or an ICT expert.

Data centres are designed to provide a secure, power protected, environmentally controlled space used for housing server, network and computer equipment. As the ICT operating centre for an enterprise's ICT service delivery, a data centre site may utilise the entire site and building shell.

The design of computer / server rooms is more limited in scope. A computer room is merely a functional space within an organisation - it serves as a secure environment for the equipment and cabling directly related to the critical load.

The easiest way to tell the design of a data centre from that of a computer / server room is by looking at how the space's functional areas (as per TIA 942) are put together. A data centre is a larger space composed of smaller spaces, such as a computer room, network operations centre, staging area(s) and conference rooms.



In either case, data centre design and computer room design are both accomplished by identifying the key design criteria for the two main areas of the project focus – the technology infrastructure and services (ICT) and the support infrastructure and services (the facility).

The key design criteria are:

- Business Objects (Scoped)
- Availability Requirements
- Power and Cooling Densities

While site selection is also a criterion for data centre projects, a computer room design project can be as involved as a bigger base-building project or as simple as an upgrade of an existing room in a building.

The goal of any data centre is to provide continuous availability of all ICT services. Therefore, whether discussing the network design strategy or the facility design strategy the overall philosophy remains the same:

- Keep it as simple as possible
- Design for scalability
- Utilise modularity wherever possible
- Be flexible and adaptable to change

Best Practice' assessment criteria

5 ~ Enterprise Class data centre: +500m^2 / 1000+ servers; extensive external storage

4 ~ Mid-tier data centre: <500m^2 / 500+ servers; extensive external storage

3 ~ Localised data centre: <95m^2 / 26-500+ servers; moderate external storage

2 ~ Server room: <50m^2 / 2-25+ servers; no external data storage**

1 ~ Server cabinet/rack: <18m^2 / 1-2 servers; no external data storage**

** in the DC/SR size columns indicate suggested best practice*

*** Refers to external storage, within the facility but not on the application server(s), or external to the facility, i.e. DR site or Cloud Service*

Best Practice	1	2	3	4	5
Consolidate servers into common/logical platforms	*	*	*	*	*
Virtualization - get economies of scale using existing consolidated servers	*	*	*	*	*
Cloud Services readiness	*	*	*	*	*
Mechanical: Air Flow Management					
Hot aisle/Cold aisle layout			*	*	*
Blank unused rack u-space	*	*	*	*	*
Use appropriate air diffusers			*	*	*
Position air supply and returns to minimise mixing	*	*	*	*	*
Cocoon/POD to provide high-density spot cooling			*	*	*
Minimise leaks in raised floor			*	*	*
Construct under-floor plenum to optimise air-flow and delivery			*	*	*
Mechanical: Air Handler Systems					
Use redundant air handler capacity in normal operations			*	*	*
Configure redundancy to reduce fan power use in normal operation			*	*	*
Control volume and velocity with variable speed drives on fans based on space temperature			*	*	*
Mechanical: Humidification					
Use widest suitable humidity control band			*	*	*
Centralise humidity control			*	*	*
Use lower power humidification technology			*	*	*
Mechanical: Plant Optimization					
Use free cooling / adiabatic cooling / waterside economisation			*	*	*
Monitor system efficiency			*	*	*
Right size chilled water cooling plant			*	*	*
Electrical Infrastructure					
Load balance incoming phases	*	*	*	*	*
Right-size and distribute use across circuit breakers	*	*	*	*	*
Maximise UPS loading	*	*	*	*	*
Specify minimum UPS efficiency at expected load points	*	*	*	*	*

IT Equipment: Specify high efficiency power supplies	*	*	*	*	*
IT Equipment: Consider equipment power consumption in specification	*	*	*	*	*
Lighting: Occupancy Sensors			*	*	*
Lighting: Low voltage/heat technology, i.e. LED	*	*	*	*	*
Lighting: Bi-level			*	*	*
Lighting: Task orientated			*	*	*
Commissioning and Retro-commissioning					
Perform a peer review	*	*	*	*	*
Engage a commissioning agent			*	*	*
Document testing of all equipment and control sequences			*	*	*
Perform full operational testing of all equipment	*	*	*	*	*
Recalibrate all control sensors			*	*	*
Install efficiency monitoring and measurement equipment, i.e. power & environmental			*	*	*

In conclusion

Management of an ever-changing data centre and server room landscape is a challenge for both ICT and Facilities staff.

The need to manage the data centre environment, and in particular power consumption and heat load has become paramount. Although this is now firmly on the agenda, the advances that is being made will take time to spread across the interested parties and become standard solutions. Even then, the issue of managing the environment will not be 'solved' as it is fair to argue that, for the foreseeable future, processor input and output requirements will also continue to rise and change.

There are several tactical solutions that organisations can consider deploying to help with the management of this situation, however the overriding message is that thorough and consistent planning is needed in the first place. Taking advice and understanding the start point, irrespective of the intended deployment, is paramount. Only then can an organisation build a clear picture of exactly which tactical solutions are suitable for them.

Much has been promoted about blades as a solution for lower power consumption and limited space. Equally they do allow an organisation to act in a modular fashion when it comes to their computing needs, a step closer to utility computing. However, their very design means that more and more are packed into smaller spaces and so the low power consumption by each is offset by the total consumption of all, and so heat load issues still remain. On this basis, blades are not going to be a suitable solution for all.

What we do know is that there are some fundamental, physical guidelines that organisations can adhere to which will deliver results irrespective of their overall requirements. These practices, such as ensuring the data centre is secure, cabling is structured and not obstructive, using hot and cold aisles, managing and controlling the power consumption, all add tangible benefits.

There are huge numbers of factors to consider and this should not be done in isolation of one or the other, but if planned well in the first instance, the risk of an ineffective data centre is dramatically reduced.

We ease the resource burden required to gather, assess, and present the data in a usable format through our comprehensive offering that is affordable, professionally presented and maintainable to help make those tough decisions a little easier.

Please do visit our website www.thehelicon.co.za for more information, or drop us an email at info@thehelicon.co.za if we can help in any way.