This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 649972.

EURECA Training

Module 6: Data Centre Contracts and Risks

Target Audience: Procurement/DC ICT&FM/Operations/CSR/Environment

Version 2.0, 28 July 2017, Contact: https://www.dceureca.eu

Agenda

- Starting questions
- Structure of the EURECA courses
- Learning outcomes
- Course contents
    - What is the nature of the risks in data centres?
    - Why identifying these risks is important?
    - What risks exist and how do we identify them?
    - What is your uptime expectancy/SLA and/or resilience level?
    - How to avoid and mitigate risk in the data centre?
- Starting questions to be answered
- References and further reading

Agenda

- **Starting questions**
- Structure of the EURECA courses
- Learning outcomes
- Course contents
  - What is the nature of the risks in data centres?
  - Why identifying these risks is important?
  - What risks exist and how do we identify them?
  - What is your uptime expectancy/SLA and/or resilience level?
  - How to avoid and mitigate risk in the data centre?
- Starting questions to be answered
- References and further reading

Starting Questions

1. What are the risks to my DC?
   a) SLA Violations
   b) Service Outages for Planned Maintenance
   c) Hardware faults causing unexpected outages

2. How resilient is my DC?
   a) Tier I
   b) Tier II
   c) Tier III
   d) Tier IV

3. What can I do to reduce the risks associated with my DC?
   a) Ring-fence critical resources
   b) Build business cases for remedial action
   c) Implement fault tolerance to prevent single points of failure

Agenda

- Starting questions
- **Structure of the EURECA courses**
- Learning outcomes
- Course contents
  - What is the nature of the risks in data centres?
  - Why identifying these risks is important?
  - What risks exist and how do we identify them?
  - What is your uptime expectancy/SLA and/or resilience level?
  - How to avoid and mitigate risk in the data centre?
- Starting questions to be answered
- References and further reading

Overview EURECA training KIT

EURECA Training Toolkit	Procurement	Module 1: PPI for Public Sector Procurers and ICT Managers
		Module 2: Business Case Development
		Module 3: Legislation and Policies
		Module 4: Procurement Strategy
		Module 5: Tendering
		Module 6: Data Centre Contracts and Risks
	Technical	Module 7: Data Centre KPI’s and Standards
		Module 8: The EU Code of Conduct for Energy Efficiency in Data Centres
		Module 9: The Data Centre Maturity Model

Agenda

- Starting questions
- Structure of the EURECA courses
- **Learning outcomes**
- Course contents
    - What is the nature of the risks in data centres?
    - Why identifying these risks is important?
    - What risks exist and how do we identify them?
    - What is your uptime expectancy/SLA and/or resilience level?
    - How to avoid and mitigate risk in the data centre?
- Starting questions to be answered
- References and further reading

Learning Outcomes

- Learn to identify and understand the true risks involved in data centre operations
- Understand the importance of mitigating risk in operating data centres
- Better understand the relationship between contracts / SLAs and risk mitigation
- Identify how to avoid some of the obvious data centre risks

Agenda

- Starting questions
- Structure of the EURECA courses
- Learning outcomes
- **Course contents**
  - What is the nature of the risks in data centres?
  - Why identifying these risks is important?
  - What risks exist and how do we identify them?
  - What is your uptime expectancy/SLA and/or resilience level?
  - How to avoid and mitigate risk in the data centre?
- Starting questions to be answered
- References and further reading

Agenda

- Starting questions
- Structure of the EURECA courses
- Learning objectives
- **Course contents**
    - **What is the nature of the risks in data centres?**
    - Why identifying these risks is important?
    - What risks exist and how do we identify them?
    - What is your uptime expectancy/SLA and/or resilience level?
    - How to avoid and mitigate risk in the data centre?
- Starting questions to be answered
- References and further reading

What is a data centre?

A data centre is a building or a segregated area of a building whose sole purpose is to provide secure, highly available, and reliable IT services – Risk Avoidance and Mitigation is key to successful operations (Think of an Insurance Policy)

Agenda

- Starting questions
- Structure of the EURECA courses
- Learning outcomes
- **Course contents**
  - What is the nature of the risks in data centres?
  - **Why identifying these risks is important?**
  - What risks exist and how do we identify them?
  - What is your uptime expectancy/SLA and/or resilience level?
  - How to avoid and mitigate risk in the data centre?
- Starting questions to be answered
- References and further reading

Why is Risk Important for Data Centres?

What introduces risk into data centres?

- IT equipment requires very stable power (A UPS both conditions and provides autonomy)
- IT equipment is far more sensitive than humans
- Equipment warranties are based on narrow environmental operating envelopes (ASHRAE)
- Multi-disciplined environment with historically poor communication (IT + M&E + Security etc)
- High levels of expectation from the business (sometimes with little understanding!)
- 24x365 operation (Aircraft are arguably easier to maintain!)
- Potentially very high profile – Potential loss of Revenue and/or Reputation

Unfortunately Many Risks are not Obvious

Some areas of potential risk:

- Risks based on False Assumptions and Poor Communications
- Hidden Infrastructure Risks Inherent in the Site
- Reliance on Infrastructure rather than Operational capability
- Perceived Protection based on Contract or SLA
- Inappropriate or poorly understood Metrics
- A focus on cost or Energy Efficiency rather than Risk Avoidance
- Regulatory Risk

Energy Efficiency rather than Risk Avoidance

DATA CENTER eureca PROJECT

- The Sole Reason for a data centre to exist is to provide centralised, highly available, and reliable IT services, not to be “Green”
- To strive for green can introduce risk by reducing resilience e.g. Some energy efficient Delta Conversion UPS systems with limited frequency correction
- A “Green” DC requires the IT load to be “Green”...........
- There is no such thing as a green Data Centre – just greener.....

CARBON TRUST

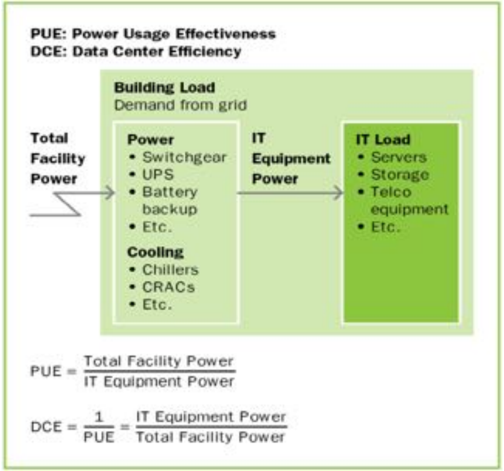
climate savers smart computing

Inappropriate Metrics - Example

Reliance on PUE – The Implications

- SLAs based on PUE are often not thought through. Often used without full understanding and can have unexpected consequences, particularly for those operators that do not control IT deployment.
- Can create perverse incentives such as increasing energy usage and cost as well as introducing risk to the business – NB PUE reduction can be achieved by INCREASING power consumption.......

the green grid™
get connected to efficient IT



PUE: Power Usage Effectiveness
DCE: Data Center Efficiency

$$PUE = \frac{\text{Total Facility Power}}{\text{IT Equipment Power}}$$

$$DCE = \frac{1}{PUE} = \frac{\text{IT Equipment Power}}{\text{Total Facility Power}}$$

IT Reliability vs. Infrastructure Reliability

Is the IT infrastructure resilient or redundant?

- E.G. True site level Active/Active mirroring or Synchronous replication between sites

Can this be regularly fully tested and has it been?

- If **YES** dependency on Mechanical and Electrical (M&E) infrastructure is reduced (and potentially also M&E costs)
- If **NO** dependency on infrastructure reliability and resilience is increased

Are IT and M&E contracts aligned?

- IT and site infrastructure contracts are frequently misaligned and do not support each other. This is of increasing concern in contract supporting “Cloud” services.

Availability and Reliability – Not The Same Thing!

Reliability

- The probability of failure-free system operation over a specified time in a given environment for a given purpose

Availability

- The probability that a system, at a point in time, will be operational and able to deliver the requested services
- Always measured in terms of percentage of uptime vs. downtime; the closer to 100% the better

Availability and reliability are related but distinct

- Availability takes into account the time that the system is out of service
- Unreliable systems can have a high availability if there is a short restart time (Not useful in data centres)
- IT systems will typically take far longer to restart than M&E systems

Availability and Reliability – Not The Same Thing!

Availability

- **Availability,** in its simplest form: **A = Uptime / (Uptime + Downtime)**
- Availability is the percentage of time when system is operational. Availability of a component can be obtained by the formula below:

$$A = \frac{MTBF}{MTBF + MTTR}$$

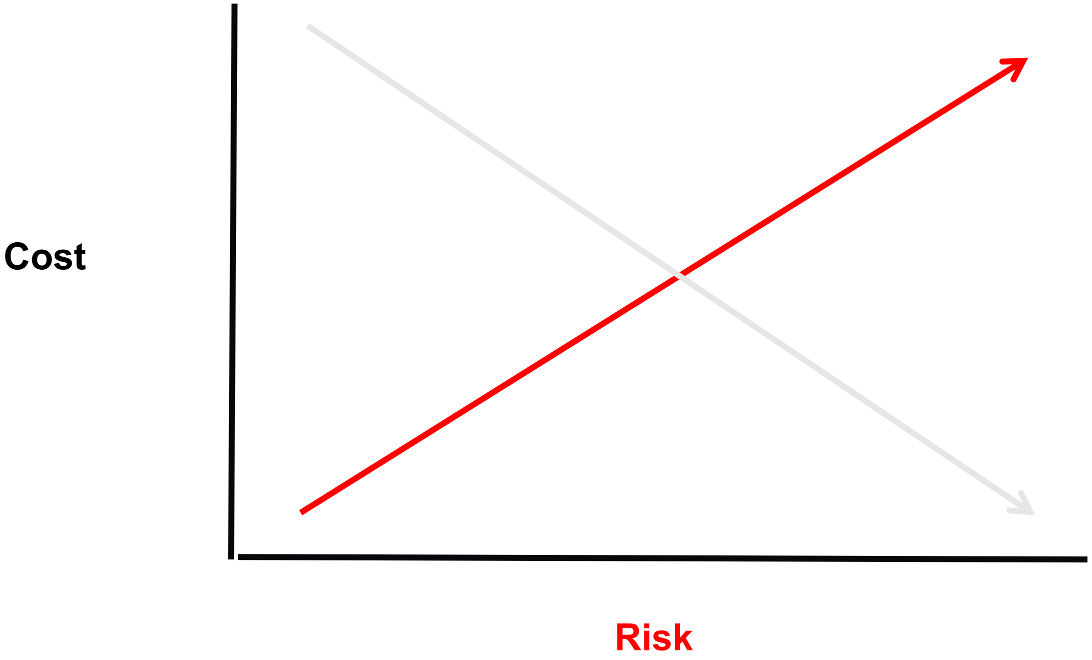
Downtime

- Downtime per year is a more intuitive way of understanding the availability. The table below compares the availability and the corresponding downtime.
  - 90% (1-nine) - 36.5 days/year
  - 99% (2-nines) - 3.65 days/year
  - 99.9% (3-nines) - 8.76 hours/year
  - 99.99% (4-nines) - 52 minutes/year (Tier III approximately)
  - 99.999% (5-nines) - 5 minutes/year (Tier IV approximately)
  - 99.9999% (6-nines) - 31 seconds/year

Availability gives information about use of time. **Reliability** gives information about the failure-free interval.

- Both are described in % values.

Why are Data Centres so Challenging?



Many data centres lack the budget to properly address Business Risk

Which is more important to your business – Cost Reduction or Risk Reduction?

Perceived Risk Protection or Real Protection?

- Contracts and SLAs are frequently driven by customer or management demand without an appreciation of either the infrastructure requirements or consequences
- Offering SLAs or expecting to be able to deliver services without fully understanding the underlying physical infrastructure and more importantly its limitations introduces risk
- 100% SLA is pointless if the supporting infrastructure is not capable of supporting 100% availability, regardless of what a contract might require
- Some customers may be unaware of a site's limitations yet force suppliers to agree a 100% SLA – Do not believe the Marketing Department!
- Financial penalties will not cover reputational damage and will only rarely cover lost revenue

Perceived Protection based on Contract or SLA

OK so just move it into “The Cloud”

- AKA “Someone Else’s Data Centre” - the risks are merely being moved elsewhere
- Moving to a shared infrastructure or becoming dependent on an infrastructure of unknown quality or appropriateness will introduce risk rather than remove it
- Financial penalties will not cover reputational damage
- Risk cannot be outsourced – Only the Management of Risk

Agenda

- Starting questions
- Structure of the EURECA courses
- Learning outcomes
- **Course contents**
  - What is the nature of the risks in data centres?
  - Why identifying these risks is important?
  - **What risks exist and how do we identify them?**
  - What is your uptime expectancy/SLA and/or resilience level?
  - How to avoid and mitigate risk in the data centre?
- Starting questions to be answered
- References and further reading

Avoid some Common Issues that Introduce Risk

- Lack accurate/complete inventory and layout of current environment
- Overprovision of power/cooling because of vague capacity requirements and management
- Facilities team and IT team have disjointed objectives
- Decisions often made with narrow, short term perspective
- Critical resources are not ring fenced
- No compelling evidence or business case to justify remedial action
- Business requirements and supporting infrastructure are not fully understood
- Fear of disrupting ongoing operations

The Importance of Operations vs. Infrastructure

DATA CENTER eureca PROJECT

- Uptime Institute analysis of Data Centre Failures shows that 73% of incidents are the result of Human Error and are therefore avoidable*

Uptime Institute™

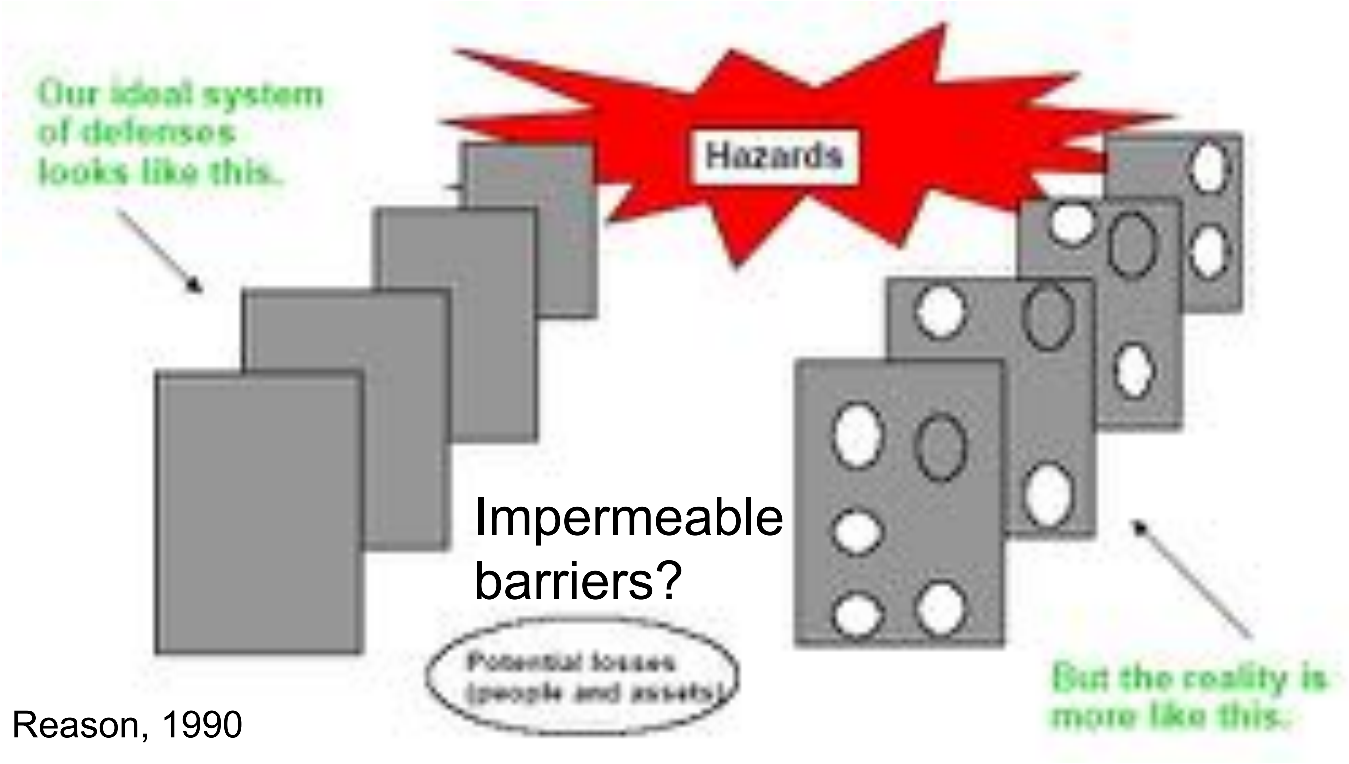
* Based on Uptime Institute™ Abnormal Incident Reports up to 1 July 2010

- This includes both Individual and Institutional issues
- It is essential to recognise the importance of Human Factors – many organisations do not
  - There is a 50/50 chance of doing something wrong when performing an unfamiliar task at speed and under pressure
  - This drops dramatically to 0.04% for a familiar, often performed task carried out by well motivated and highly trained staff

HEALTH & SAFETY LABORATORY

Jerry Williams, HEART Methodology

Why do Incidents Happen?


Our ideal system of defenses looks like this.
Hazards
Potential losses (people and assets)
But the reality is more like this.

Impermeable barriers?

James Reason, 1990

Operational process and procedures should be designed and implemented to “plug the holes”

Hidden Infrastructure Risks

- Uptime Institute Tiers – Frequently misunderstood and misapplied yet used ubiquitously
- Examples:
  - Self assessed sites probably do not reach the expected or desired UI Tier
  - Is the expectation to protect against a single event or multiple events?
  - Do IT staff understand the M&E Infrastructure and its constraints?
  - What is meant by an Outage or Failure?
- SLAs are pointless if the underlying infrastructure is inadequate. The danger is that businesses think they are protected when they are not
- The IT infrastructure requirements are rarely factored in and false assumptions are often made about IT resilience

Hidden Infrastructure Risks

Uptime Tiers - The Real Implications

- Tiers define “Opportunity for Maintenance” **NOT** redundancy (N+1 etc)
- Tier III is NOT **Fault Tolerant** (only protects against **PLANNED** events)
- Tier III does **NOT** remove Single Points of Failure
- **ONLY** Tier IV addresses unplanned events
- Tier II is likely to require service outages even for planned maintenance
- Tier I **WILL** impose some service downtime for planned maintenance (E.G. statutory 5 yearly testing)
- Frequently confused with TIA-942 requirements yet TIA do not certify data centres, nor do they certify others to do so.....
- 99.9999% availability is greater than Tier IV and yet still anticipates 32 seconds of downtime per year.

Agenda

- Starting questions
- Structure of the EURECA courses
- Learning outcomes
- **Course contents**
  - What is the nature of the risks in data centres?
  - Why identifying these risks is important?
  - What risks exist and how do we identify them?
  - **What is your uptime expectancy/SLA and/or resilience level?**
  - How to avoid and mitigate risk in the data centre?
- Starting questions to be answered
- References and further reading

Business Objectives and Infrastructure

It is essential to ensure site capability (Tiers) and business requirements are fully aligned

The Tier Topology references the ability to carry out scheduled maintenance and addresses fault tolerance so the following apply:

Maintenance:

Can the organisation (or service), *afford* to take the computer room down to perform pre planned infrastructure maintenance (***planned downtime***) ?

- If Yes require Tier I or Tier II only
- If No require Tier III or above
- If No consider moving applications to higher Tier DC
- **OR** redesign applications to be more resilient

Fault Tolerance:

Can your organisation (or application), accept unplanned incidents or events taking your computer room and IT services down (***unplanned downtime***) ?

- If Yes require Tier III or lower
- If No require Tier IV
- (Consider moving applications to higher Tier DC)
- **OR** redesign applications to be more resilient

Implications of % uptime

Reliance on % uptime or the ‘9s’......

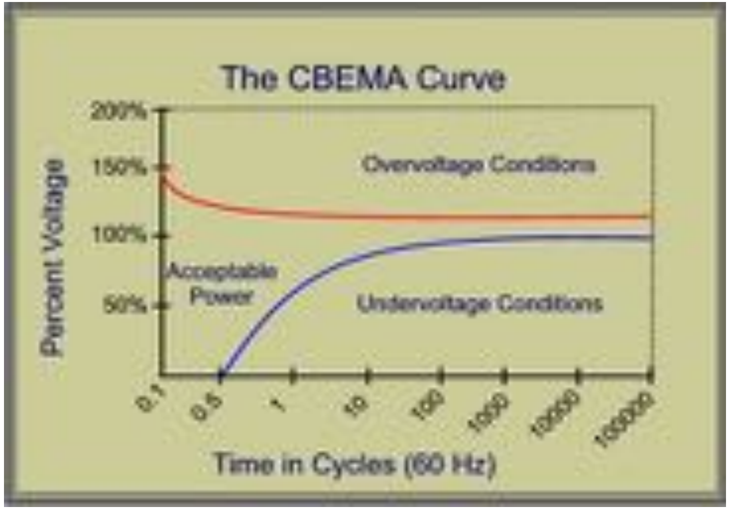
- The ‘9’s are merely based on statistical equipment failure rates (MTBF) and do not take into account operational practices
- 99.9999% Uptime = 32 seconds per year service outage. This would not be acceptable to most
- Or 32 X 1 second outages. Or 64 X 0.5 second outages etc.
- 99.9999% availability is significantly greater than Tier IV
- A single 25ms break will halt ALL services for a considerable period of time
- This is why percentage uptime has been dropped by The Uptime Institute from the latest Tier Topology update

Inappropriate Metrics

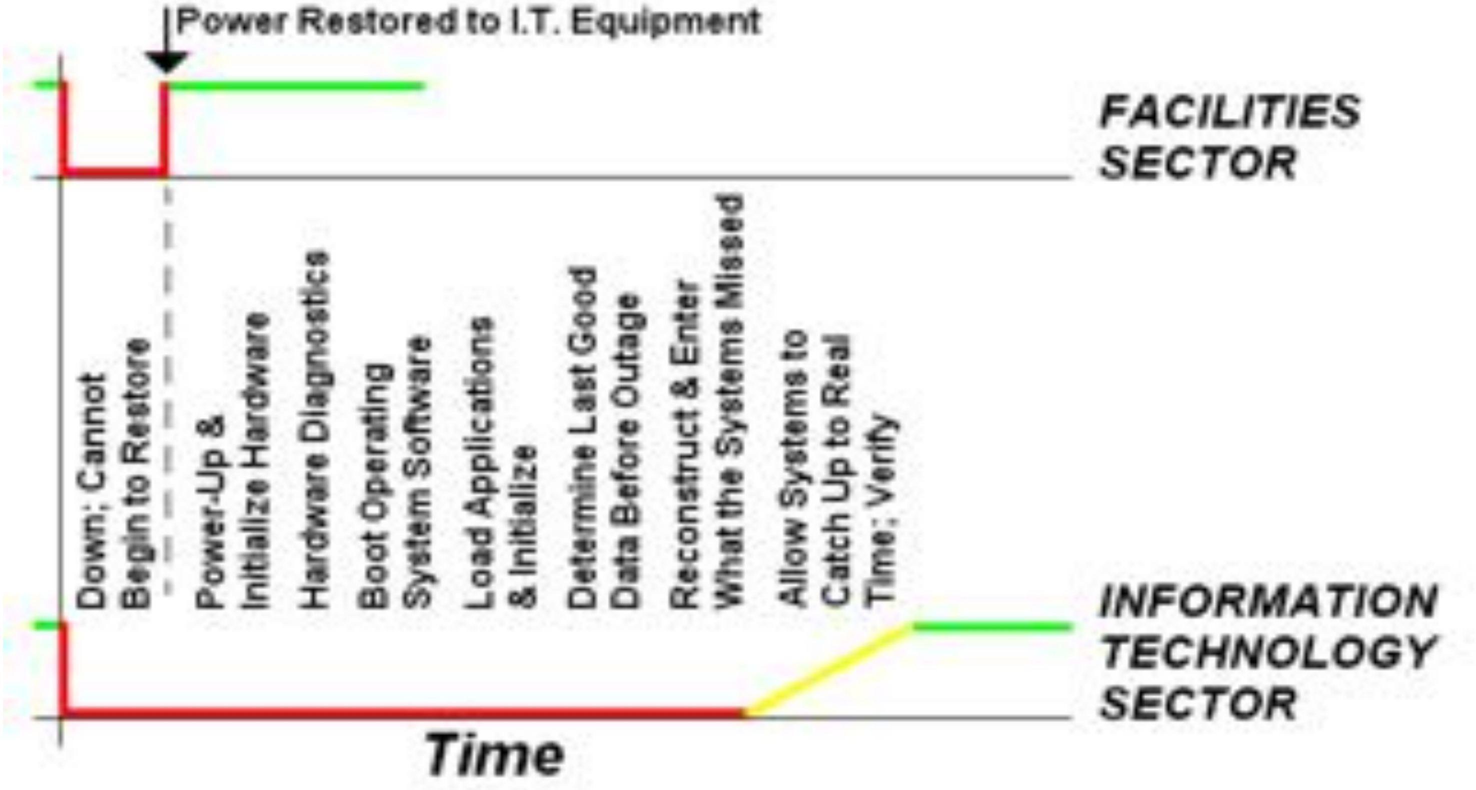
The '9s' become meaningless when a single 25ms break can potentially halt ALL IT services for hours

Additionally what is considered a service outage?

- 20 minute battery autonomy,
- Temperature rise,
- Where is this measured and how?



Availability – Facilities and IT mismatch


Power Restored to I.T. Equipment
FACILITIES SECTOR
Down; Cannot Begin to Restore
Power-Up & Initialize Hardware
Hardware Diagnostics
Boot Operating System Software
Load Applications & Initialize
Determine Last Good Data Before Outage
Reconstruct & Enter What the Systems Missed
Allow Systems to Catch Up to Real Time; Verify
INFORMATION TECHNOLOGY SECTOR
Time

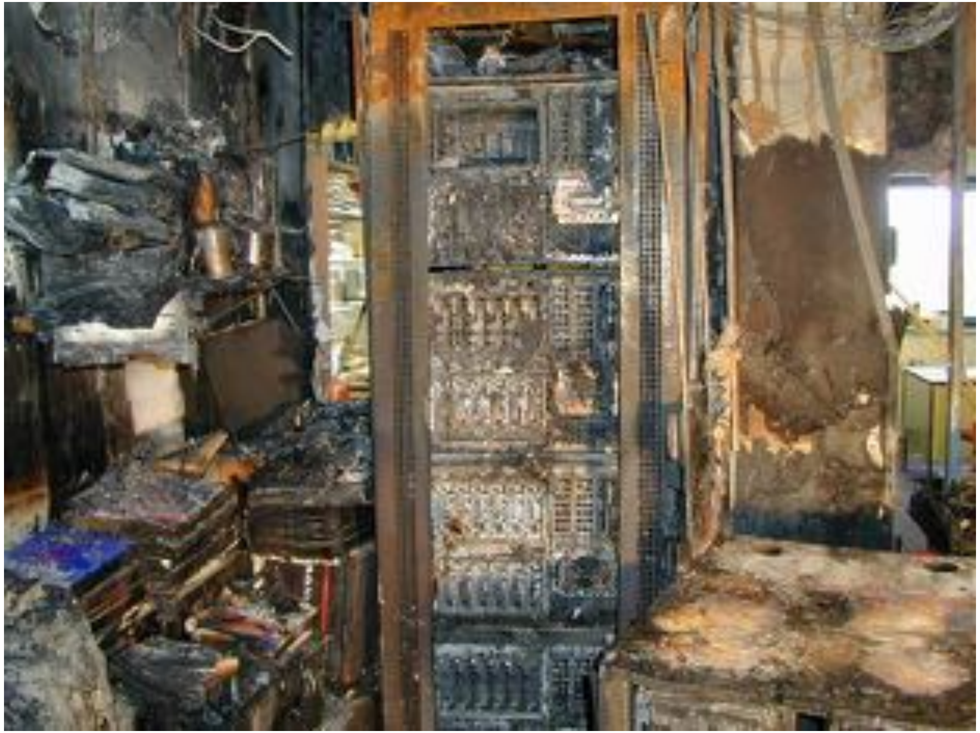
Perceived Protection based on Contract or SLA

- SLAs referencing “outages” and some environmental parameters frequently depend on ill-defined and poorly understood terms with inappropriate measurement
- E.G. a local UK government tender requiring EU CoC compliance and Endorser status on behalf of supplier, yet which goes on to specify an SLA based on a 18°C supply temperature....
- Are Business Expectations and Supporting Infrastructure truly aligned in this case?

Agenda

- Starting questions
- Structure of the EURECA courses
- Learning outcomes
- **Course contents**
  - What is the nature of the risks in data centres?
  - Why identifying these risks is important?
  - What risks exist and how do we identify them?
  - What is your uptime expectancy/SLA and/or resilience level?
  - **How to avoid and mitigate risk in the data centre?**
- Starting questions to be answered
- References and further reading

People and Attention to Detail are Key to Success



Could this happen in your site?

Managing Data Centre Risks

Data Centre site management should be different from non critical site management and requires the following:

- Extreme attention to detail as failure is a high impact to the business – either financially or by reputation damage

High Reliability Organisation

- Continual learning from Failures / Events and Incidents
  - Root cause analysis
  - Lessons learned
  - Corrective actions
- Continual evaluation of new technology and methods
- Not diluted by non critical requirements
- Clearly defined Roles and Responsibilities across all parties
- Regular practice of routine and emergency procedures

It is not enough just to fix problems - Critical Management should be ahead of problems by highlighting issues, taking preventative actions and both maintaining and improving supporting infrastructure to mitigate all risks

Develop a Risk Management Strategy

- **Identify** possible risks; recognise what can go wrong
- **Analyse** each risk to estimate the **probability** that it will occur and the **impact** (i.e., damage) that it will do if it does occur
- **Rank** the risks by probability and impact

  - Impact may be negligible, marginal, critical, and catastrophic
- **Develop** a contingency plan to manage those risks having high **probability** and **high impact**
- An effective strategy for dealing with risk must consider three issues (Note: these are not mutually exclusive)
  - Risk mitigation (i.e., avoidance)
  - Risk monitoring
  - Risk management and contingency planning
- **Risk mitigation** (avoidance) is the primary strategy and is achieved through a plan based on a an accurate risk register
- **Risk Assessment** is a continuous process!

How to Avoid data centre Risks ?

Do Not:

- Merely rely upon SLAs for service resilience
- Manage business critical spaces the same way as office spaces
- Rely solely on tools to make your site more efficient or reliable – the right people with the right training and experience are vital
- Forget business objectives and the implications for the data centre
- Have PUE / DCiE “Improvements” linked to your bonus
- Focus on “Green” to the detriment of reliability
- Perpetuate the IT Facilities divide
- Hide from the difficultly of predicting IT requirements over time

Improve the Chances of Success

Define and follow strict process and procedures

but

Sometimes individuals need to stand up, challenge rules and procedures, and suggest improvements or ways to simplify them etc. as a part of a team approach

Data Centre operations are never static and should be considered a continually evolving landscape

Agenda

- Starting questions
- Structure of the EURECA courses
- Learning outcomes
- Course contents
  - What is the nature of the risks in data centres?
  - Why identifying these risks is important?
  - What risks exist and how do we identify them?
  - What is your uptime expectancy/SLA and/or resilience level?
  - How to avoid and mitigate risk in the data centre?
- **Starting questions to be answered**
- References and further reading

Starting Questions -answered

1. What are the risks to my DC?

a) Assumption of contract protection. Contracts are frequently misaligned.
b) Inappropriate SLAs Violations (Based on incorrect assumptions of the resilience of the data centre and stringency of the SLA).
c) Data centre infrastructure or operational rigour not aligned with business needs.
d) Potential Service Outages even for Planned Maintenance (Likely in Tier I and Tier II data centres).
e) Assumption of Fault Tolerance (Only available in Tier IV type data centres).
f) Poor staff raining or misunderstanding of data centre operations or managing critical environments.
g) Lack of awareness of the importance of Human factors in the continuity of service availability from data centres.

2. How resilient is my DC and does it match my business need?

a) Tier I Type (Will require service outages for planned maintenance)
b) Tier II Type (Likely to require service outages for planned maintenance)
c) Tier III Type (Will not require service outages for planned maintenance but still vulnerable to unplanned events)
d) Tier IV Type (Will not require services for both planned and unplanned events)

3. What additional measures can I take to reduce the risks associated with my DC?

a) Ringfence critical resources (Prevent resource conflict and re-allocation of critical resources)
b) Build business cases for remedial action (Remedial action may be necessary but a case needs to be clearly illustrated)
c) Improve infrastructure or applications to reduce risk (Important but can be expensive and unnecessary to implement)

Agenda

- Starting questions
- Structure of the EURECA courses
- Learning outcomes
- Course contents
  - What is the nature of the risks in data centres?
  - Why identifying these risks is important?
  - What risks exist and how do we identify them?
  - What is your uptime expectancy/SLA and/or resilience level?
  - How to avoid and mitigate risk in the data centre?
- Starting questions to be answered
- **References and further reading**

References & Further Reading

- Uptime Insititute Tiers: https://uptimeinstitute.com/TierCertification/

- Greenberg, A., Hamilton, J., Maltz, D.A. and Patel, P., 2008. The cost of a cloud: research problems in data center networks. *ACM SIGCOMM computer communication review*, *39*(1), pp.68-73.

- Wustenhoff, E. and BluePrints, S., 2002. Service level agreement in the data center. *Sun Microsystems Professional Series*, *2*.