Available online at www.prace-ri.eu

**Partnership for Advanced Computing in Europe**

# Redundancy and Reliability for an HPC Data Centre

Erhan Yılmaz*[1]

Ladina Gilly**[2]

** National High Performance Computing Center of Turkey, Uydu, Yolu, Maslak, 34469, İstanbul, Turkey*
*** CSCS- Swiss National Supercomputing Centre, Lugano, Switzerland*

**Abstract**

Defining a level of redundancy is a strategic question when planning a new data centre, as it will directly impact the entire design of the building as well as the construction and operational costs. It will also affect how to integrate future extension plans into the design. Redundancy is also a key strategic issue when upgrading or retrofitting an existing facility.

Redundancy is a central strategic question to any business that relies on data centres for its operation. In the traditional data centre reliant industries such as Internet Service Providers (ISP's), banks, insurances, or credit card services redundancy is of paramount importance, as a loss of availability has an immediate and sometimes drastic impact on revenue or legal due diligence for example. For this reason, the industry has formed a number of clear standards and guidelines that address the topic of redundancy and reliability.

Both these topics are of course just as important for HPC centres too, but not always in the same way given that some of the trade-off mechanisms may differ substantially and thus make it difficult for an HPC centre to rely fully on the existing standards used by the traditional data centre industry.

This white paper aims to discuss the key factors to be taken into account when selecting a level of redundancy and reliability for an HPC centre, providing managers with a set of topics that need to be considered when designing a new HPC centre or upgrading an existing one. These factors all have an impact on the design and cost of construction as well as on future operational costs for your centre.

[1] erhan.yilmaz@uybhm.itu.edu.tr
[2] lgilly@cscs.ch

## 1. Redundancy and reliability in the traditional data centre world.

Within the data centre industry, the most pervasive standard that allows us to measure redundancy and reliability is the Tier Classification[3] as defined by the Uptime Institute[4]. This classification looks at how a data centre must be set up in order to achieve a defined level of availability throughout a year, which is directly related to the level of redundancy and reliability. Given that the classification was developed with traditional data centres in mind, i.e. datacentres for banks, ISP's and other high-availability based industries, the top tier classification (Tier IV) offers the highest level of availability. Interestingly, these same datacentres usually run a variety of equipment at very low density – i.e. under 5KW per rack. Let us look at the two tables (Table 1and Table 2) that the Uptime Institute provides to describe the different Tier levels.

| | Tier I | Tier II | Tier III | Tier IV |
|---|---|---|---|---|
| **Active Capacity Components to support IT load** | N | N+1 | N+1 | N after any failure |
| **Distribution paths** | 1 | 1 | 1 active and 1 alternate | 2 simultaneously active |
| **Concurrently maintainable** | No | No | Yes | Yes |
| **Fault tolerance (single event)** | No | No | No | Yes |
| **Compartmentalization** | No | No | No | Yes |
| **Continuous Cooling** | Load density dependent | Load density dependent | Load density dependent | Yes (Class A) |

Table 1: Tier Requirements Summary

The Uptime Institute goes on to provide the following table (Table 2) of common attributes in data centres that are unrelated to Tier requirements.

[3] White Paper, Tier Classifications Define Site Infrastructure Performance by W. Pitt Turner IV, PE, John H. Seader, PE, Vince Renaud, PE, and Kenneth G. Brill (http://www.greenserverroom.org/Tier%20Classifications%20Define%20Site%20Infrastructure.pdf)

[4] www.uptimeinstitute.com

| | Tier I | Tier II | Tier III | Tier IV |
|---|---|---|---|---|
| **Building type** | Tenant | Tenant | Stand-alone | Stand-alone |
| **Staffing shifts**<br>**Staff/shift** | None<br>None | 1 shift<br>1 /shift | 1 + Shifts<br>1-2/ shift | "24 by Forever"<br>2+/shift |
| **Useable for Critical Load** | 100% N | 100% N | 90% N | 90% N |
| **Initial Build-out kW per Cabinet (typical)** | < 1kW | 1-2kW | 1-2kW | 1-3kW |
| **Ultimate Build-out per Cabined (typical)** | <1kW | 1-2kW | >3kW | >4kW |
| **Support Space to Raised-Floor ration** | 20% | 30% | 80-90% | 100%+ |
| **Raised-floor Height (typical)** | 30cm | 45cm | 75-90cm | 75-105cm |
| **Floor loading kg/m2** | 415 | 488 | 732 | 732+ |
| **Utility Voltage (typical)** | 208, 480 | 208, 480 | 12 – 15 kV | 12 – 15 kV |
| **Single Points-of-Failure** | Many + Human Error | Many + Human Error | Some + Human Error | Fire, EPO, Some Human Error |
| **Representative planned maintenance shut downs** | 2 annual events at 12 hrs each | 3 events in 2 yrs at 12 hrs each | None required | None required |
| **Representative site failures** | 6 failures over 5 years | 1 failure every year | 1 failure every 2.5 years | 1 failure every 5 years |
| **Annual site-caused, End-user downtime (based on field data)** | 28.8 hrs | 22 hrs | 1.6 hrs | 0.8 hrs |
| **Resulting End-User Availability based on site-caused downtime** | 99.67% | 99.75% | 99.98% | 99.99% |
| **Typical months to plan and construct** | 3 | 3 – 6 | 15 – 20 | 15 – 30 |
| **First deployed** | 1965 | 1970 | 1985 | 1995 |

Table 2: Common Attributes Found in Data Centres that are Unrelated to Tier Requirements

A traditional high-availability data centre is likely to include all or most of the following infrastructure items (Figure 1):

- Redundant HVAC Controlled Environment
- UPS Backup Generators
- Gas-based Fire Suppression System
- Biometric Access and Exit Sensors
- Continuous Video Surveillance
- Electronic Motion Sensors
- Security Breach Alarms
- On-premises Security Officers
- Server Operations Monitoring
- Seismically braced Server Racks

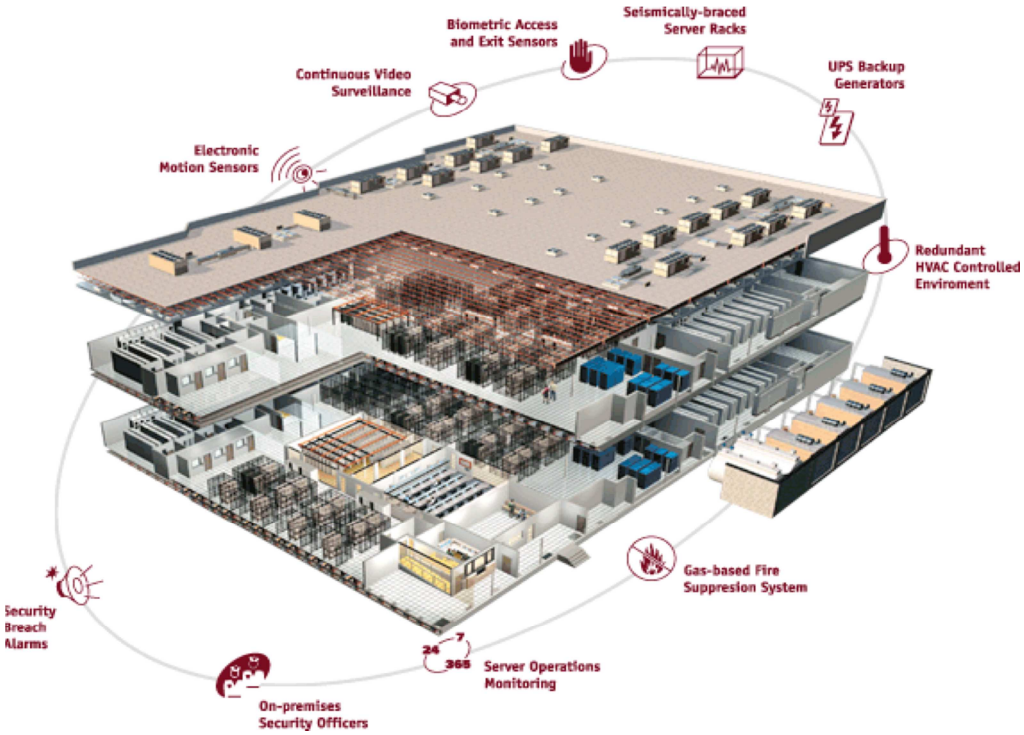


Figure 1: Shows an example of services considered critical for a general-purpose datacentre. Adapted from " The Exodus 1999 Annual Report", http://archive.icann.org/en/tlds/i1/REGOP/Exodus%20proposal.htm

The Uptime Institute is just one of several organisations that provide advice and standards for the construction of data centres but it is possibly the most widely recognised. Others include: 7x24exchange.org, datacentredynamics.com, datacenterknowledge.com as well as many of the infrastructure service providers and hardware vendors.

The Tier Classification provides an excellent guideline for someone building or extending a traditional data centre. It gives clear indications of what infrastructure is required in order to achieve a given level of availability. It is certainly useful to understand the Tier Classification if you are building or extending an HPC data centre so as to be able to take informed decisions as to where it may be important for you to apply the classification and where you may deviate from it.

## 2. Where an HPC centre may differ

The main mission of the data centres operated by PRACE partners is to provide compute services to simulation-based sciences within the public research institutions of their country in order to compete internationally. By virtue of this, the success of such centres can only be measured in the amount of valuable science that can be produced thanks to its services. In order to achieve best results centres are likely to have to take a number of decisions that may strongly differ from the decisions of a standard data centre operator.

In order to provide users with the tools to compete worldwide, HPC centres often invest in cutting edge technology. This can mean that the computer system will take longer to reach steady production once it has been installed but it will remain competitive for a longer time. This decision in itself initially introduces reduced reliability.

The leading HPC systems are also very large in order to facilitate the leading edge science and very large simulations. With size and number of parts increasing the risk of HW failures increases proportionally once again inherently reducing reliability.

Contrary to traditional data centres where rack density will vary between less than 1kW to around 6kW, HPC centres are dealing with much higher densities. It is not unusual to see machines currently on the market consuming over 30kW per rack and these numbers are expected to rise with the next generation of systems. This increases the importance of an efficient infrastructure and makes cooling much more complex.

Most of the large HPC centres across the world are government funded. For this reason, the costs for infrastructure and operation will always have to compete with the investment costs in new HPC hardware, as this is where the results come from. For this reason an HPC centre must minimize its infrastructure and operational costs in order to be able to maximise its investments in its core business. Electricity will be one of the largest cost contributors to the HPC data centre's budget. In order to control this expense it is critical to maximise the energy efficiency of the data centre. Reducing the redundancy and reliability requirements on power and cooling can help improve energy efficiency of the overall operation thus reducing operational costs as well as infrastructure investments. Therefore providing adequate availability with an acceptable number of failures and interrupts and maximum compute power is the key. For more information on cooling and electrical supply in an HPC centre please refer to the related white papers on these subjects.

A vast majority of the PRACE HPC centres focuses essentially on providing compute time to researchers from the public domain – universities and research centres. For this reason, their security requirements may well not demand the same level of investment, as would a banking data centre. This allows a number of cost-reductions to be made in areas such as physical access control, video surveillance and more.

Some services within an HPC centre do however require high-availability as we know it from the traditional data centre industry. These systems will usually comprise data storage (local and global file systems) but also safeguard against data corruption, the network backbone as well as computer management networks and the internet connection towards the outside world, corporate services (e-mail, work station backup, etc.) building automation systems and all infrastructure components that support these critical systems. These systems are fortunately of much lower density than the large HPC computers. Supporting these systems by UPS is therefore affordable but also critical to ensure service to the centres users and internal staff. Failures in this area will be problematic and could be damaging. Investing in higher redundancy and reliability for these areas of the service is important.

Some HPC centres will also provide service to 3rd party clients or host their machines for which the SLA's may require higher levels of availability than for the HPC systems. In this case it is important that the customer understand the costs – both in terms of investment but also operational costs – that are generated by their redundancy and reliability requirements. Given that such 3rd parties usually have to cover the costs for the operation of their system, they may well revise their requirements once they understand the cost implications. It is critical that an HPC data centre understand its costs for different levels of reliability and redundancy so this can be part of the contractual discussion with 3rd party customers.

Given that the systems within an HPC centre that require high-availability make up a much smaller load than the HPC systems, providing different levels of availability, redundancy and reliability is important. This may be achieved by forming physically distinct rooms or geographically distinct sectors within one room. In any case, an HPC centre will need to adjust its level of availability to different levels of requirements within its operation if it wants to optimise its infrastructure and operational costs.

## 3. Redundancy and Reliability requirements of the PRACE consortium sites

According to the survey responses received from PRACE members, the following have priority in terms of reliability.

- Preserving data integrity. As the computations scale up, they generally require or produce larger data sets, and preserving data consistency and integrity is important for HPC centre users.
- Users should be affected as little as possible by energy, network, hardware and software failures that may occur within HPC centres. HPC centres generally provide computation services to a large number of researchers, and minimizing the impact of such outages to researchers is important as they can otherwise rapidly cause a large loss in terms of productive working hours for their users.
- Generally HPC centres provide their users with IT services comprising compute time and storage services. It is important for centres to provide the services promised to the users.
- As the HPC centres get larger in computation power and storage space, the probability of errors

increases proportionally to the number of parts. Early detection and errors prevention is therefore becoming ever more important in order to keep the HPC centre operational. This requires an HPC centre to deploy good monitoring tools both on systems and building infrastructure.

In decreasing order of common answers the survey results show that the following common services are deemed by all respondents to require redundancy in order to provide these services reliably:

- Data storage services
- Network connections
- Core IT services (such as email, web, file/database servers)
- Backup services
- Building management and security services (such as security cameras, building access systems)
- Computing services/servers

It is important to note that even though the computation is assumed to be critical in an HPC centre, preserving data and accessing the data through the network has higher importance in terms of reliability. In fact computing services are deemed the least important in terms of reliability in HPC centres.

When analysing the survey responses, a pattern emerges that differentiates older and newer data centres as well as smaller and larger data centres.

Data centres that were built more than 15 years ago tend to have a very high level of redundancy and reliability across the centre, newer centres have worked to distinguish the diverse availability requirements within their centre and supply different levels of reliability and redundancy. This is driven by a number of factors. The cost for energy along with the density of HPC computers has risen steeply in the last 15 years. Energy has suddenly become one of the biggest cost items in an HPC centre operational budget. This has made it important to control and where possible reduce this cost item to avoid it cannibalising the budget for HPC investments. At the same time there has been a shift in attitude regarding the environmental impact of power hungry industries it is to be expected that energy costs will continue to rise in the near and medium-term future.

The same energy costs are probably also the main driving factor behind the fact that the larger HPC centres within the PRACE consortium have started to introduce the use of power directly off the grid (also know as “dirty power”) for the HPC systems. The cost and space requirements to provide UPS support have outweighed the cost of potential downtimes to the centre. For smaller centres on the other hand it can still be more cost-effective to provide a single level of availability throughout the centre given that the full power load is often an order of magnitude smaller that that of the large centres. The trade-off balance has not yet tipped in favour of lowering the level of redundancy and reliability.

## 4. Key recommendations

In order to determine the required level or redundancy and reliability when planning a new data centre or extending an existing one it is of utmost importance to first analyse which services require which degree of availability. Although it may be desirable for reasons of operational ease to provide high-availability to all services, the cost of this level of availability must be compared to the potential cost of failures of these services.

Providing a higher level of redundancy and reliability than a particular service requires will increase the inefficiency of your infrastructure and your operational costs will rise accordingly. HPC is a power-hungry industry and seemingly small inefficiencies therefore can have very large cost impacts.

The level of redundancy and reliability required for your services is directly inherent to the nature of these services and the requirements of the customers they serve. In this sense each HPC centre will have a slightly unique setup in terms of redundancy and reliability if it has done the analysis and optimised its infrastructure.

Make sure that you adapt all infrastructure parts supporting a service to the same level of availability. There is no point having redundant water supply if the electrical supply is not redundant. It is important to remember that the availability of a server will depend on the lowest level of availability within the entire infrastructure supply chain leading up to it. Define what the lowest acceptable threshold is and apply this throughout the infrastructure supply chain.

If you are planning to reduce the level of redundancy and reliability for some of your services make sure you manage the expectations of your customers and support them in adapting their workflow to this change in environment. If this is done well, the users will not perceive the reduction of availability levels as a problem as long as the number of failures remains acceptable.

Never loose sight of the main mission of the centre and consider the impact of all your planning decisions relative to this mission.

## 5. Conclusion

Contrary to the traditional data centre world, where your business is likely to define the Tier level your infrastructure should provide and your livelihood depends on the level of availability being met at all times, HPC data centres require a more differentiated approach in order to maximise scientific output and minimise infrastructure investment and operational costs. This will require a detailed analysis of the services you run and the customers you serve. Different services or customer groups may require different levels of redundancy and reliability and a trade-off needs to be made between the cost of providing higher availability against the cost of potential failures.

Given that no two HPC centres serve the exact same portfolio of customers, the mix of availability levels within each HPC centre will be unique. It is therefore usually not possible to simply define a Tier level and build to this although it is important to understand the Tier Classification in order to be able to make informed trade-off decisions during the planning phase.

Although there is a strong attraction to simplify planning by implementing just one Tier level across the centre, this is usually not the most cost effective way forward and will ultimately adversely impact your ability to invest in HPC computer HW. It is important to always bear in mind that the mission of an HPC centre is enable and support the maximum amount of research on their infrastructure and not to run an infrastructure that boasts a very high level of availability.

## References


- White Paper: *Tier Classifications Define Site Infrastructure Performance* - By W. Pitt Turner IV, PE, John H. Seader, PE, Vince Renaud, PE, and Kenneth G. Brill, 2008 Uptime Institute Inc.
- Alger D., *Build the Best Data Centre Facility for Your Business,* Cisco Press 2005.
- Presentation materials from the 3 HPC Infrastructure Workshops run in Lugano (2009), Dourdan (2010) and Garching (2011), with support from the PRACE 1IP project, are available on demand


## Acknowledgements

This work was financially supported by the PRACE project funded in part by the EUs 7th Framework Programme (FP7/2007-2013) under grant agreement no. RI-211528 and FP7-261557.