Data Centre Design Report

1. Beginning The A5.xlsx spreadsheet gives the specifications for the design of a data centre infrastructure, which this paper describes. To satisfy operational needs for the next five years, the design takes into account server configurations, computation resources, storage needs, and rack structure. The system can run 100 applications on 25 clusters. Its main goals are to be scalable, reliable, and make the best use of resources. We will now show you the main requirements, how they are met, and pertinent data tables with comments.

2. Things you need Summary There are five sheets in the spreadsheet that give detailed specifications: Summary, Compute, Storage, Application, and End of Row. The main prerequisites are:
    - Support for 100 applications across 25 clusters over 5 years.
    - Each cluster contains 5 servers, with 4 applications per cluster.
    - Each server occupies 2 Height Units (HU), with 10 HU per cluster and 40 HU per rack, accommodating 4 clusters per rack.
    - Management clusters: 2 clusters.
    - Topology: End of Row.
    - DRAM per server: 512 GB (with some discrepancies noted).
    - Storage configurations with RAID for reliability.
    - Rack layout with power, networking, and equipment placement.

3. System Design The design is structured to meet the specified requirements through a modular and scalable architecture. Below, we address each component of the system:
    3.1. Server and Cluster Configuration
    Each cluster comprises 5 servers, with each server occupying 2 HU, resulting in 10 HU per cluster. With 4 clusters per rack, the total HU per rack is 40, which aligns with the rack capacity specified. Over 5 years, 25 clusters support 100 applications, with each cluster hosting 4 applications. Two additional management clusters are included for administrative tasks, ensuring operational oversight.
    3.2. Compute Resources
    The compute resources are designed to support various server types, including load-balanced web servers, program code servers, clustered database servers, AAA servers, and logging servers. Each server type has specific CPU and memory allocations, with a total of 160 CPUs and 1472 GB of memory per cluster. A noted discrepancy exists in the DRAM per server: the Summary sheet specifies 512 GB, while the Compute sheet lists 64 GB or 295 GB. For this design, we assume 512 GB as the target to ensure sufficient memory capacity, with the lower values possibly reflecting specific server configurations or minimum requirements.
    3.3. Storage Configuration
    The storage system is designed with RAID configurations to ensure high availability and data protection. Clustered database servers, for instance, use RAID10 for OS and log files, RAID1 for applications, RAID6 for databases, and RAID5 for backups. This multi-tiered approach balances performance and redundancy. Logging servers and AAA servers also utilize RAID10 and RAID5 for log files and backups, respectively.

3.4. Rack Topology

The rack topology follows an End of Row configuration, with identical layouts in the Application and End of Row sheets. Each rack includes servers, switches (Top of Rack and Out-of-Band Management), SAN units, KVM, PDUs, and UPS units. Networking includes 1GB and 10GB connections to support high-speed data transfer. Power consumption is managed with dual 7kW UPS units per rack, ensuring uninterrupted operation.

4. Data Tables and Commentary Below, we present key tables from each sheet with commentary to illustrate how the design meets the requirements.

4.1 Summary Sheet

Table 1: Summary of Key Parameters

Table 1: Summary of Key Parameters

Parameter	Value
Servers per Cluster	5
Applications per Cluster	4
Total Apps over 5 Years	100
Clusters over 5 Years	25
Server HU	2
HU per Cluster	10
Clusters per Rack	4
HU per Rack	40
Management Cluster	2
Topology	End of Row
DRAM per Server	512 GB

Commentary: The Summary sheet sets the basic rules for the data centre. The design can handle 100 applications spread out over 25 clusters, with each cluster hosting 4 apps and 5 servers. The HU calculations (2 HU per server, 10 HU per cluster, and 40 HU per rack) show that 4 clusters may fit in a single rack, which is in line with the rack's capacity. The 512 GB of DRAM per server makes sure there is enough memory for applications that use a lot of processing power, but this number is looked at again in the Compute sheet analysis.

4.2 Compute Sheet

Table 2: Compute Resource Allocation

Table 2: Compute Resource Allocation

Server Type	Quantity	CPUs per Server	Memory per Server (GB)	Total Memory (GB
Load Balanced Web Servers	2	4	8	16
Program Code Servers	2	4	8	16
Clustered Database Servers	2	4	64	128
AAA Servers	2	4	4	8
Logging Servers	2	4	8	16
Total per Cluster		160		1472

Commentary: The Compute sheet shows how resources are divided across different types of servers. The 160 CPUs and 1472 GB of RAM in each cluster are enough to meet the needs of high-performance applications. There is a difference in the amount of DRAM per server (512 GB in Summary vs. 64 GB or 295 GB in Compute). This could mean that the 512 GB value is for a maximum or upgraded setup, while the Compute sheet shows the baseline or certain types of servers. To satisfy the highest specified need, the architecture uses 512 GB per server for consistency.

4.3 Storage Sheet

Table 3: Storage Configuration

Table 3: Storage Configuration

Server Type	OS		Log Files		Application	Database			Backup			
	Qty	Cap (GB)	Qty	Cap (GB)	Avail	Qty	Cap (GB)	Avail	Qty	Cap (GB)	Avail	
Load Balanced Web Servers	2	250										
Clustered Database Servers	2	250	6	500	RAID10	2	1000	RAID1	10	1000	RAID6	10
500	RAID5											
AAA Servers									3	500	RAID5	
Logging Servers			10	500	RAID10				3	500	RAID5	

Commentary: The Storage sheet shows a strong storage architecture with RAID setups to make sure that data is always available and safe.
The most complete storage system is on clustered database servers. They use RAID10 for the OS and log files, RAID1 for applications, RAID6 for databases, and RAID5 for backups. This approach makes sure that important database operations are always available and run well. Logging and AAA servers use RAID10 and RAID5, respectively, to balance performance and redundancy. They focus on log data and backups.

4.4 Application/End of Row Sheet

Table 4: Rack Layout (Application/End of Row)

Table 4: Rack Layout (Application/End of Row)

HU	Power (W)	FRONT				BACK		
		PDU-A	ToR1	ToR2	PDU-B	BES1		BES2
47	100		ToR Switch					
45	100		ToR Switch					
43	100		OOBM Switch					
41	700		1	Server		1		1
39	700		1	Server		1		1
37	700		1	Server		1		1
35	700		1	Server		1		1
33	700		1	Server		1		1
31						Back End Switch		
29	700			SAN1		4	Back End Switch	4
27	700			SAN2		4		4
20	100	1		KVM	1			KVM
6	7000			UPS1				UPS1
3	7000			UPS2				UPS2

The Application and End of Row sheets show the same rack architecture, which supports the End of Row topology. There are 5 servers (HU 33–41), 2 SAN units (HU 27–29), and networking equipment

(ToR and OOBM switches at HU 43–47, Back End Switches at HU 29–31) on each rack. Adding two 7kW UPS units and PDUs makes sure that there is always electricity. The 1GB and 10GB networking connections let data move quickly, which is important for the clustered applications.

5. Final Thoughts The suggested design for the data centre fits all of the requirements, such as being able to run 100 applications across 25 clusters, with 5 servers per cluster and 4 clusters per rack. The computational resources have plenty of CPU and memory, and each server has 512 GB of DRAM to fix any problems. The storage system uses RAID setups to make sure everything works, and the rack architecture organises equipment in a way that makes the most of power and networking. This design makes sure that the data centre will be able to grow, be reliable, and work well for the next five years.