# Data Information Systems Management

The Architecture of Data Management - Database Systems (Distributed Databases), Grid Computing, Virtualisation & SAN architecture

LECTURE 2 – THE ARCHITECTURE OF DATA MANAGEMENT

CHAPTER OUTLINE

1. What is Distributed Database Architecture (Oracle view)

2. What is Grid Computing

3. Compare Grid computing to the Cloud

4. What is virtualisation

5. What is a SAN

6. The future of SAN technology
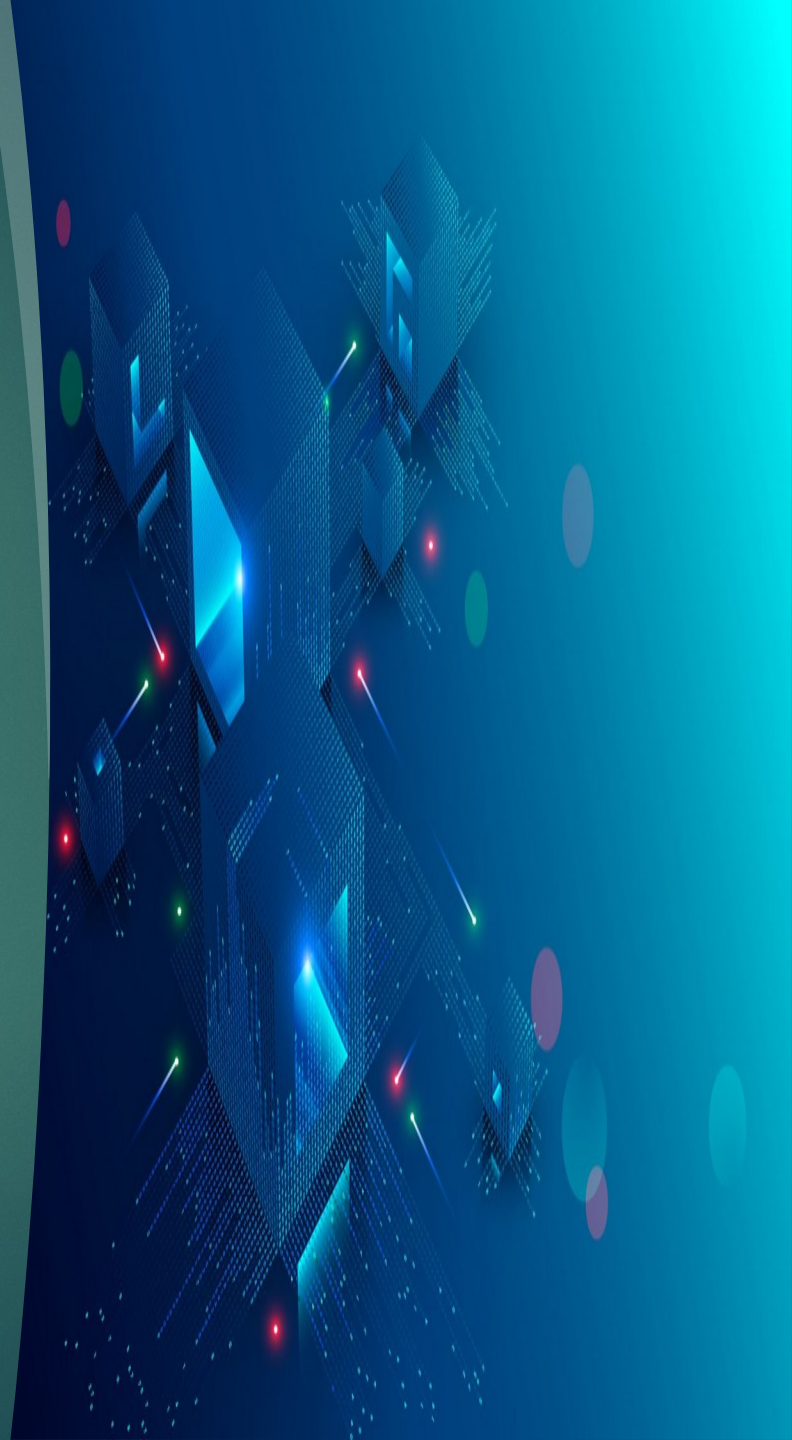
Distributed Database Architecture

▶ A **distributed database system** allows applications to access data from local and remote databases.

▶ In a **homogenous distributed database system**, each database is the same i.e. All Oracle.

▶ In a **heterogeneous distributed database system**, at least one of the databases are different. Distributed databases use a **client/server** architecture to process information requests.
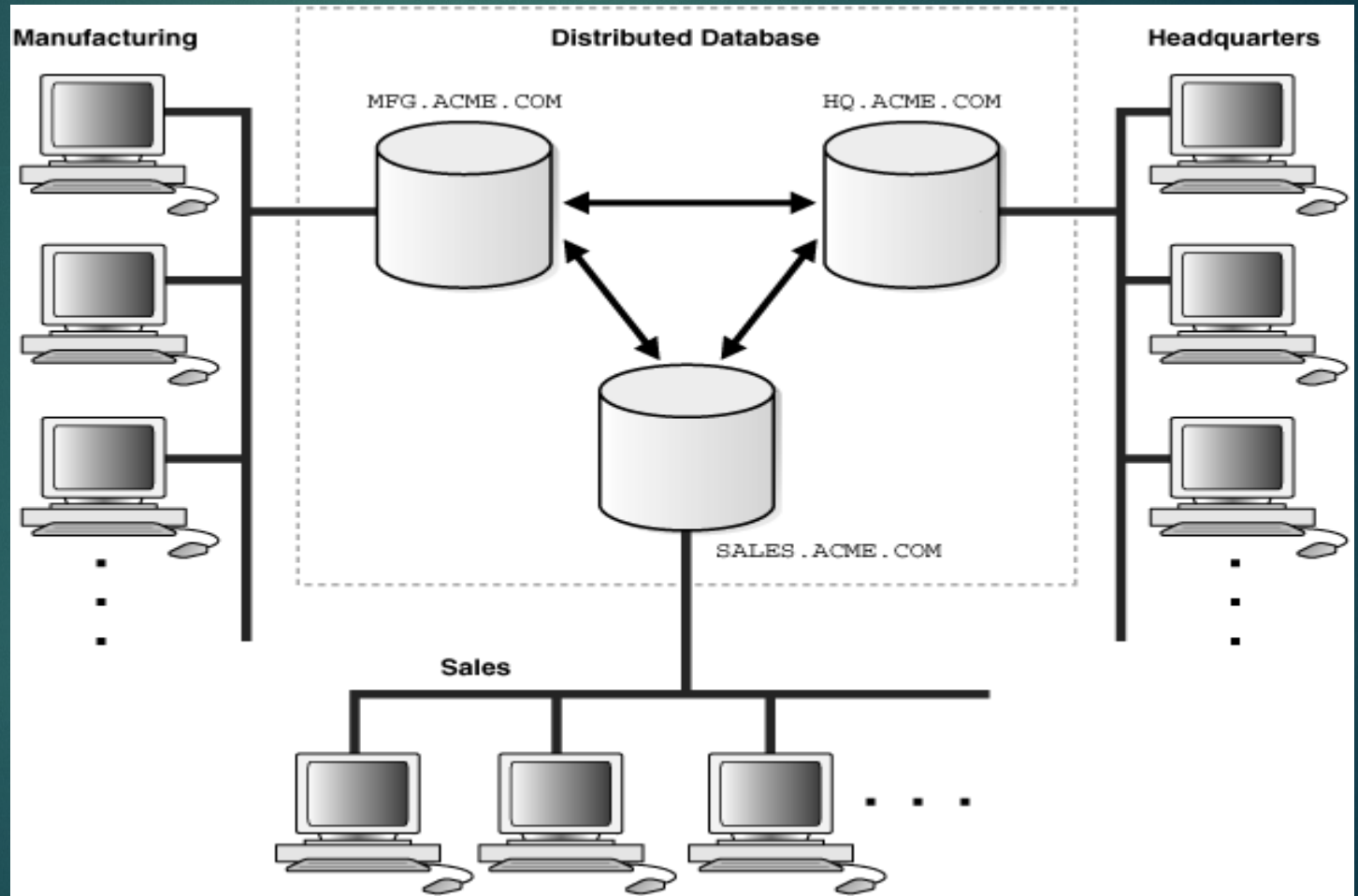
# Distributed Database Architecture

Homogeneous Distributed Database

Homogeneous Database: In a homogeneous database, all different sites store database identically. The operating system, database management system, and the data structures used – all are the same at all sites. Hence, they're easy to

Distributed Database Architecture

Heterogeneous Distributed Database
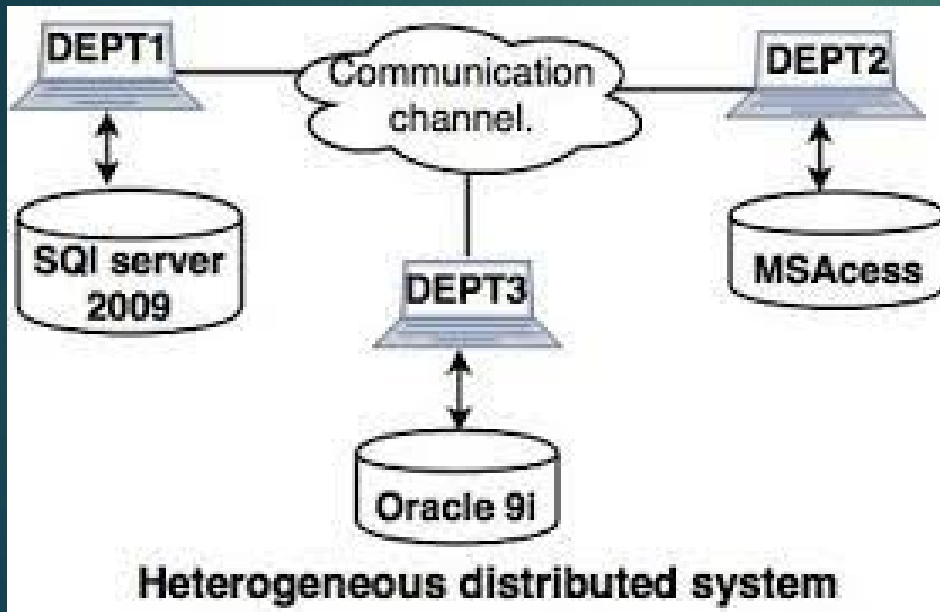
2. Heterogeneous Database: In a heterogeneous distributed database, different sites can use different schema and software that can lead to problems in query processing and transactions. Also, a particular site might be completely unaware of the other sites. Different computers may use a different operating system, different database application. They may even use different data models for the database. Hence, translations are required for different sites to communicate.

Geeksforgeeks.or (2022)

Grid Computing

What is "Grid Computing" and what can it do for you?

The idea of grid computing is to use many inexpensive devices to provide the total   resources required, thereby providing intrinsic redundancy (if one device fails, replace it)and scalability (add more devices as the load increases) – Could be databases, Infrastructure, Software

Grid Computing

The two tenets of Grid computing

Two core tenets uniquely distinguish grid computing from other styles of computing, such as mainframe, client-server, or multi-tier: virtualization and provisioning.

▶ • With *virtualization*, individual resources (e.g., computers, disks, application components and information sources) are pooled together by type then made available to consumers (e.g., people or software programs) through an abstraction. Virtualization means breaking hard-coded connections between providers and consumers of resources, and preparing a resource to serve a particular need without the consumer caring how that is accomplished.

▶ • With *provisioning,* when consumers request resources through a virtualization layer, behind the scenes a specific resource is identified to fulfil the request and then it is allocated to the consumer. Provisioning as part of grid computing means that the system determines how to meet the specific need of the consumer, while optimizing operation of the system.

The difference between Grid and the Cloud

The different between Grid computing and the Cloud

Clouds and grids compared

What is Virtualisation

What is virtualization, how it works and types … a where does cloud fit in..

What is Virtualisation

 What is VMware

What is Hyper V

What is a SAN?

▶ Storage Area Network (SAN) - An array of disk drives in a self-contained unit. In large enterprises, SANs serve as pools of storage for the servers in the network. Compared to managing disks attached to each server, SANs improve system administration. Treating all storage as a single resource makes disk maintenance and routine backups easier to schedule and control. To support disaster recovery, redundant SANs are deployed in separate locations, each a copy of the other.

The SAN transfers data between servers and disks at the same fast peripheral channel speeds as if directly attached, and Fibre Channel has been the traditional interface. Some SANs perform backup procedures without any processing overhead at the host computers

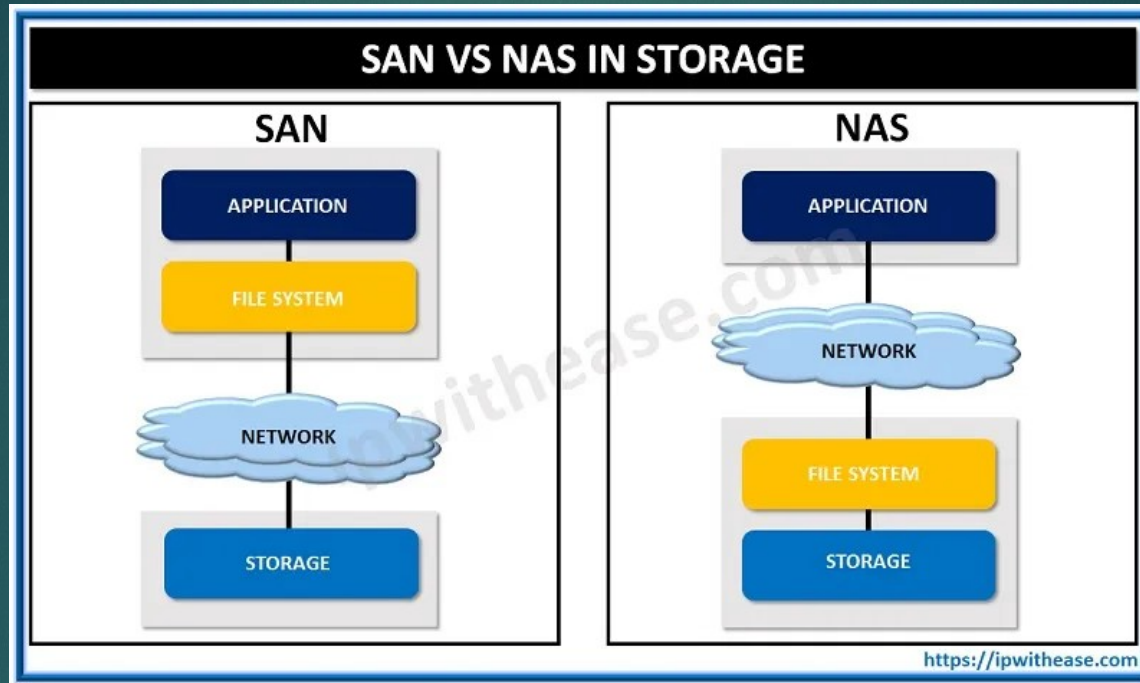(http://www.pcmag.com/encyclopedia/term/50794/san - 22/09/2015)
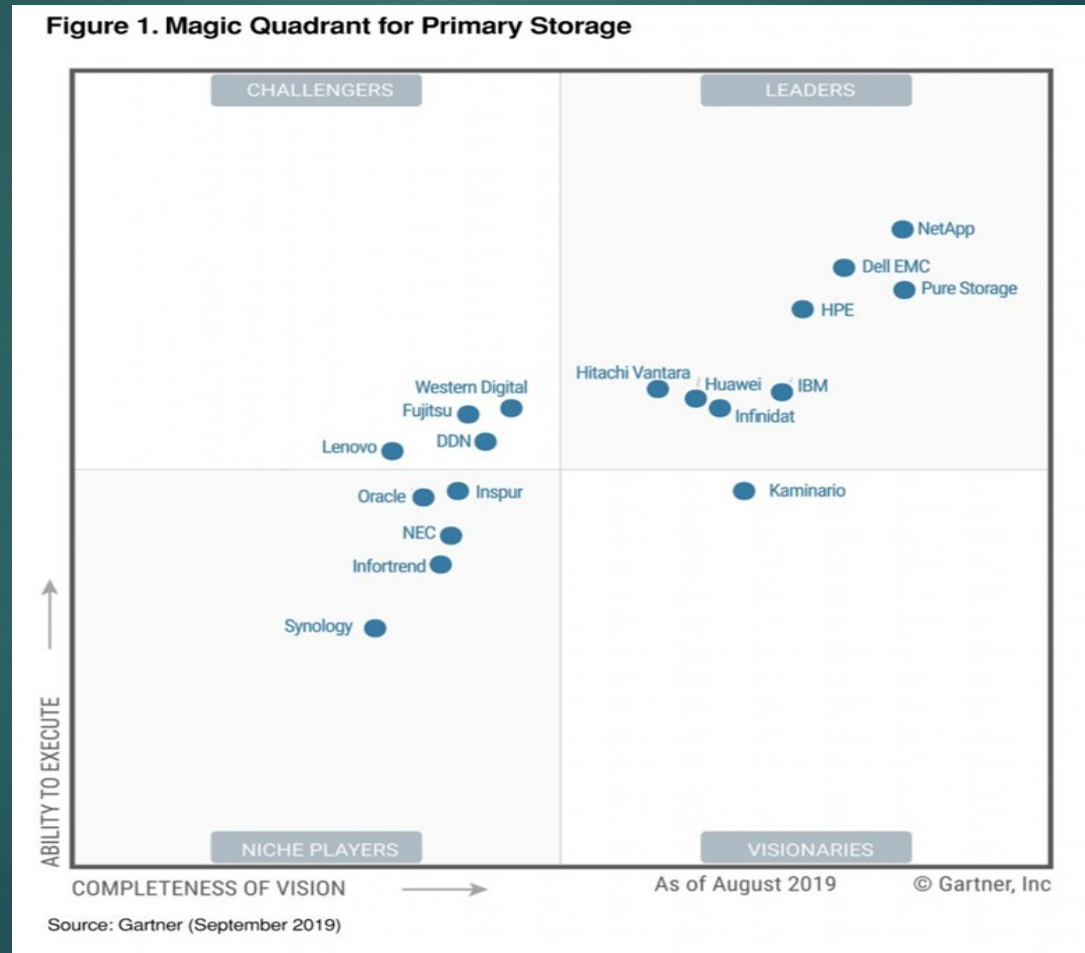
What is a SAN ?
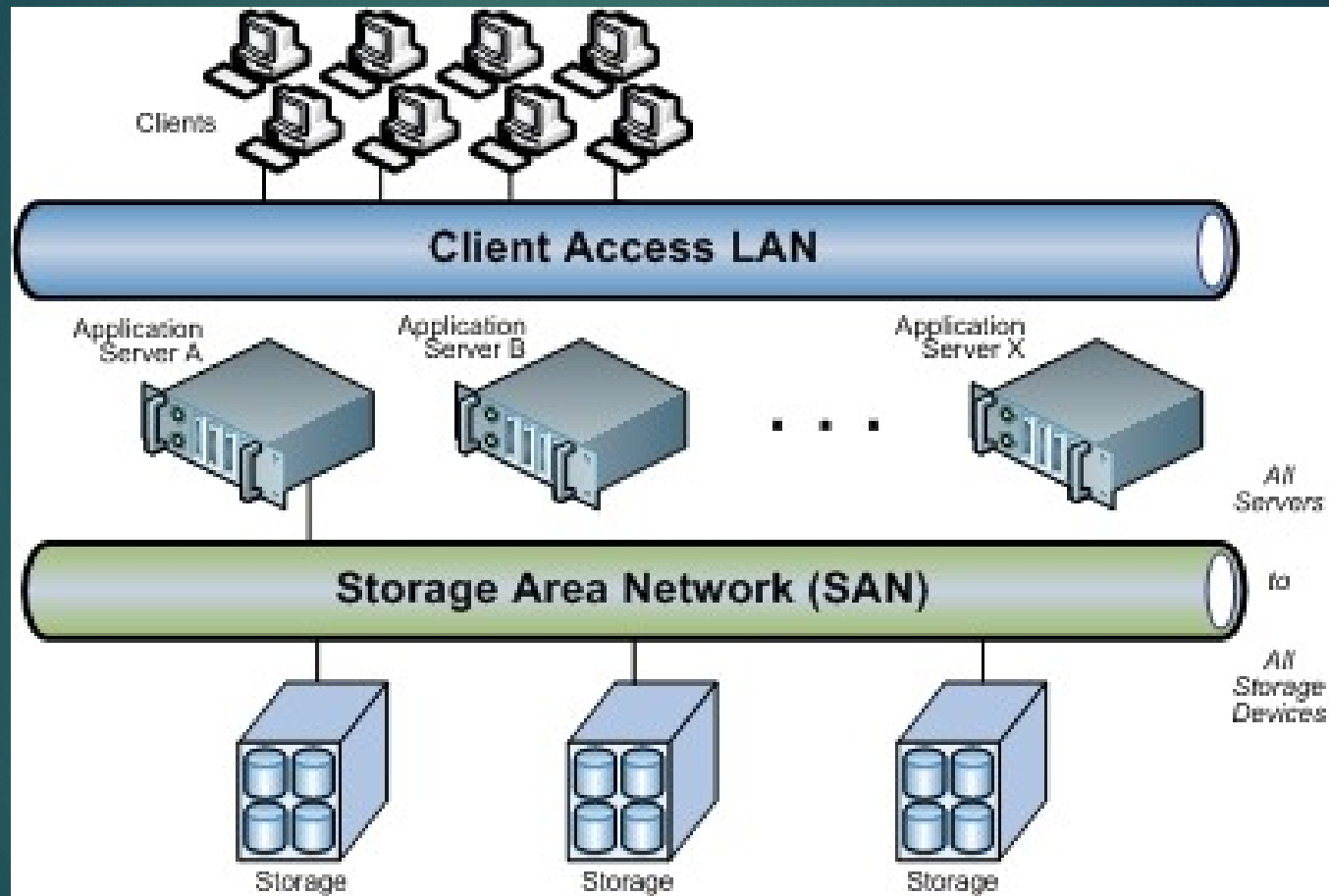
NAS versus SAN

SAN stands for **Storage Area Network** which is a centrally managed storage pool. With SAN, the storage traffic from the LAN can be separated which leads to an improved and efficient performance. Through SAN, the distribution and management of resources of the organizations is simplified. SAN is a dedicated host of interconnected storage devices and switches. The block level storage can be accessed by this high-speed network.

NAS stands for Network Attached Storage which is a computer device connected with a network. NAS provides file storage services over that network. NAS is the most effortless to set up. The standard protocols used by network attached devices are NFS (Network File System), Server Message Block (SMB) protocol or Common Internet File System (CIFS) protocol. The files that are stored in such devices are typically arranged in a logical order or storage containers. The NAS device is a node of the network itself, having its own IP address and is able to strike a communication with other devices on the same network.

Figure 29-2 An Oracle Database Distributed Database System

Who are the SAN market leaders – 2019

Figure 1. Magic Quadrant for Primary Storage

Figure 29-2 An Oracle Database Distributed Database System

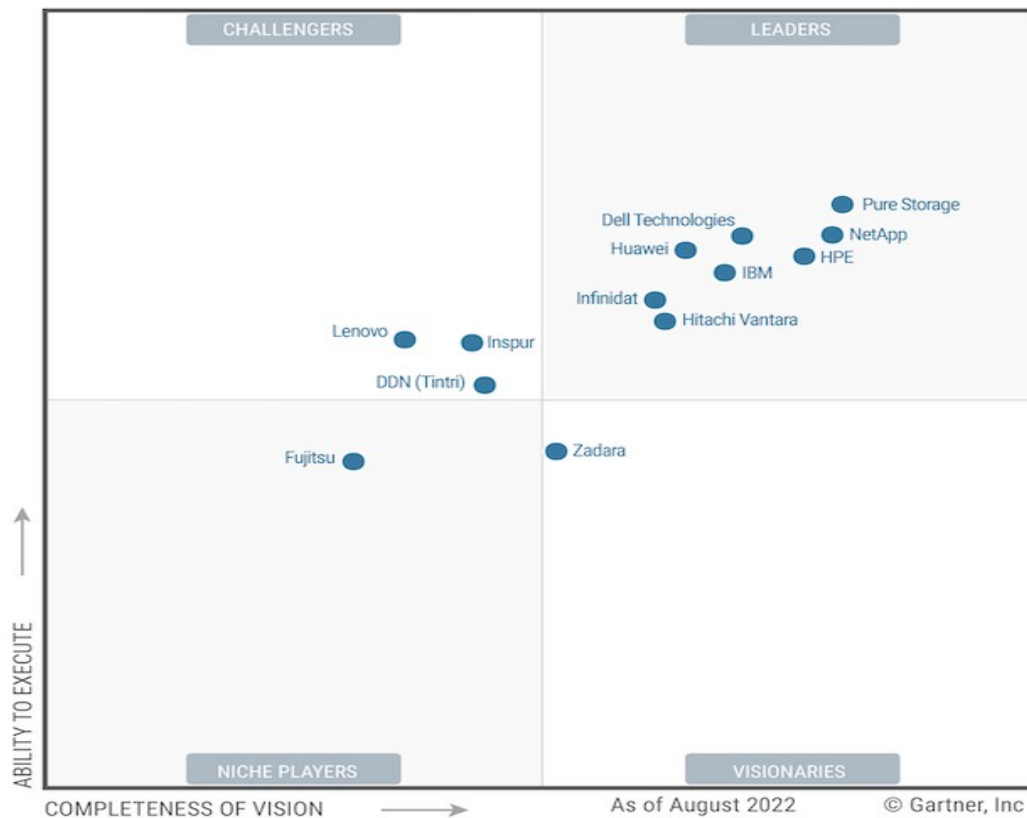# Who are the SAN market leaders – 2021

Figure 1: Magic Quadrant for Primary Storage

Source: Gartner (October 2021)
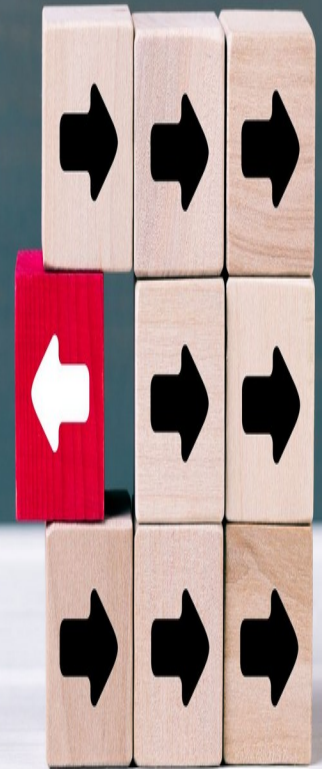
Who are the SAN market leaders – 2022

Figure 1: Magic Quadrant for Primary Storage

Object vs. File vs. Block Storage – What's the Difference

▶ What is the Difference

Why is Hyperconverged Infrastructure

▶ <u>What is Hyperconverged</u>

Future of SAN technology

- [The future of SAN technology](#)