Grid and Utility Computing

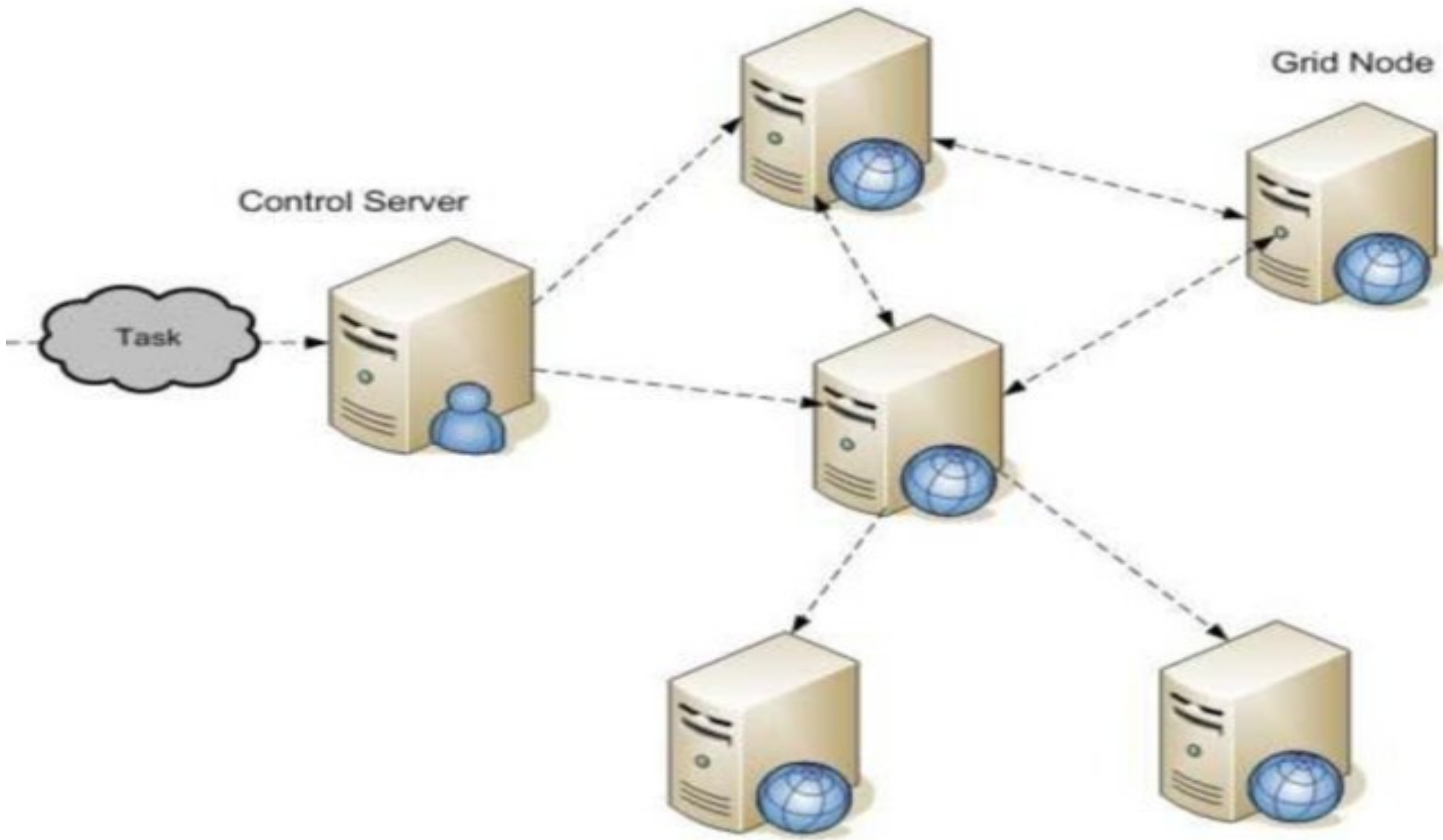
Review

- What is a grid?
- Features of the grid?
- Why do we need a grid?

Grid Computing

- **Grid computing** is the practice of leveraging multiple computers, often geographically distributed but connected by networks, to work together to accomplish joint tasks.
- It is typically run on a "**data grid**," a set of computers that directly interact with each other to coordinate jobs.

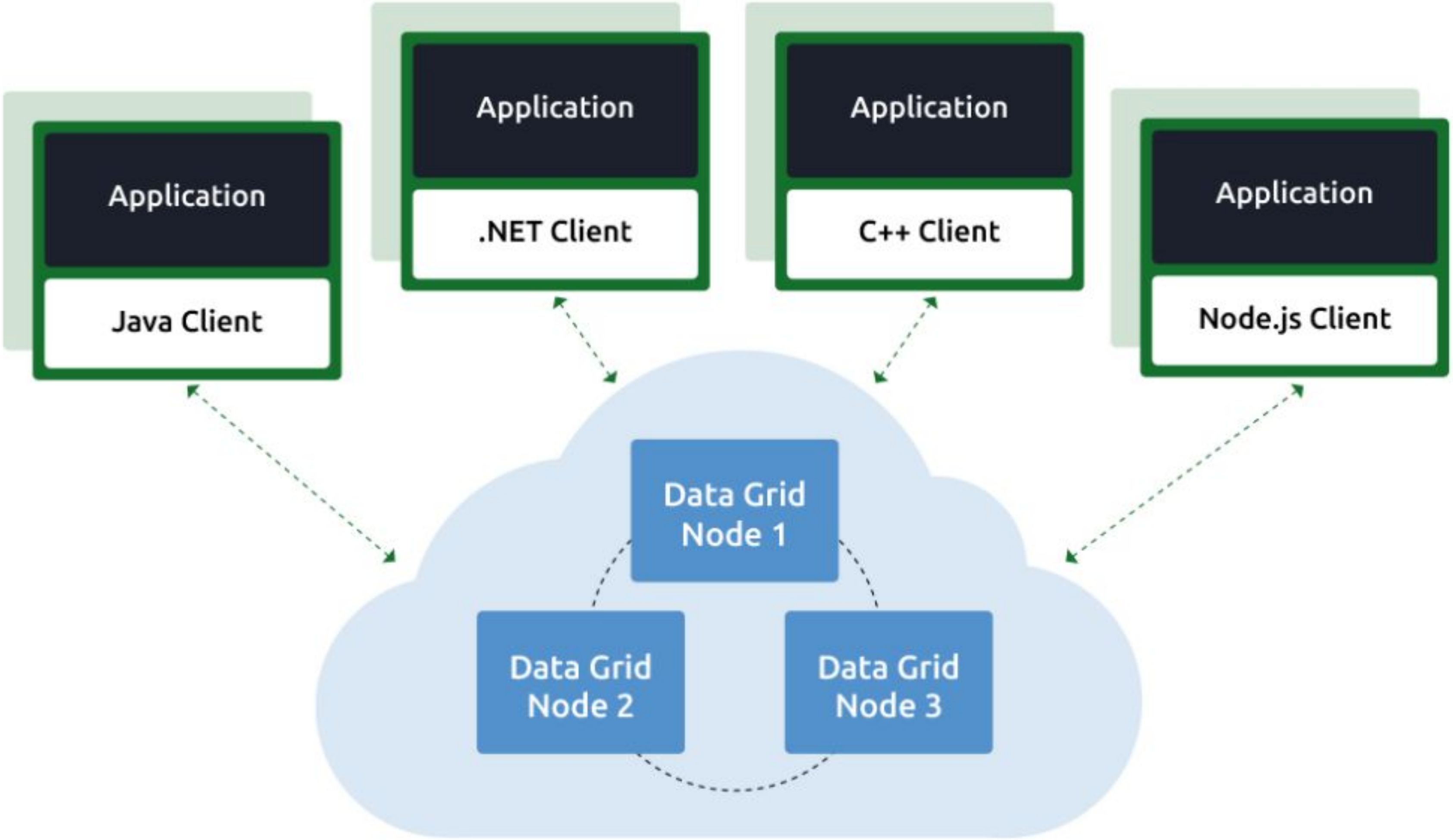
Grid Computing Architecture



How does Grid Computing Work?

- Grid computing works by running specialized software on every computer that participates in the data grid.
- The software acts as the manager of the entire system and coordinates various tasks across the grid.
- Specifically, the software assigns subtasks to each computer so they can work simultaneously on their respective subtasks.
- After the completion of subtasks, the outputs are gathered and aggregated to complete a larger-scale task.
- The software lets each computer communicate over the network with the other computers so they can share information on what portion of the subtasks each computer is running, and how to consolidate and deliver outputs.

Contd…


Application
.NET Client
Application
C++ Client
Application
Java Client
Application
Node.js Client
Data Grid
Node 1
Data Grid
Node 2
Data Grid
Node 3

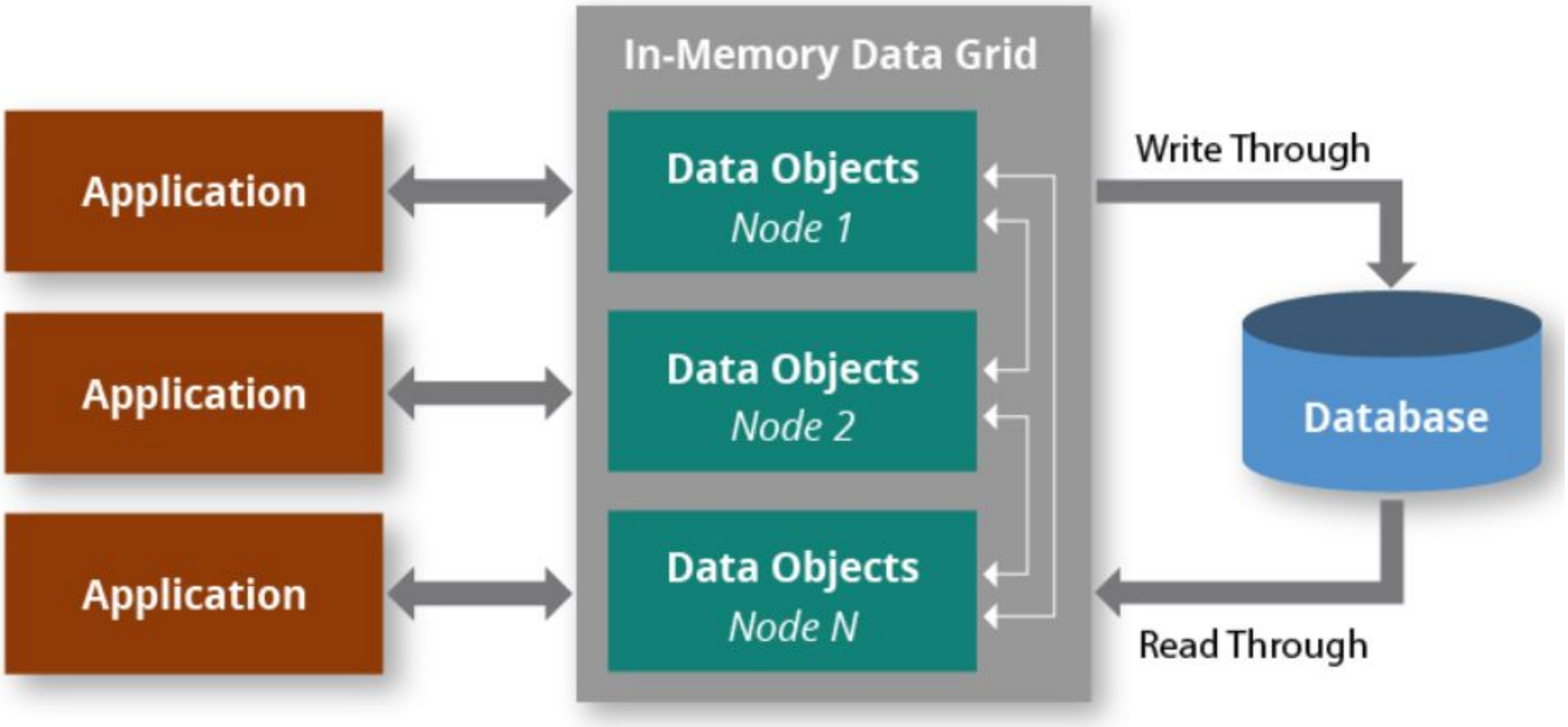
How is Grid Computing Used?

- Grid computing is especially useful when different subject matter experts need to collaborate on a project but do not necessarily have the means to immediately share data and computing resources in a single site.
- By joining forces despite the geographical distance, the distributed teams are able to leverage their own resources that contribute to a bigger effort.
- This means that all computing resources do not have to work on the same specific task, but can work on sub-tasks that collectively make up the end goal.
- *For example, a research team might analyze weather patterns in the North Atlantic region, while another team analyzes the south Atlantic region, and both results can be combined to deliver a complete picture of Atlantic weather patterns*.

In-memory Data Grid

- While often seen as a large-scale **distributed computing** endeavor, grid computing can also be leveraged at a local level.
- For example, a corporation that allocates a set of computer nodes running in a cluster to jointly perform a given task is a simple example of grid computing in action.
- A specific type of local data grid is an **in-memory data grid** (IMDG) in which computers are tightly connected via coordination software and a network connection to collectively process data in memory.
- The advantage is that the data is stored in memory across all computers in the data grid, so all data accesses are very fast.
- IMDGs are especially useful when the grid computing tasks require extremely high throughput and extremely low latency.

Contd…


Application
Application
Application
In-Memory Data Grid
Data Objects
Node 1
Data Objects
Node 2
Data Objects
Node N
Write Through
Database
Read Through

Types of Grids

- **Computational Grids**
  - Network of high performance servers
- **Scavenging Grids**
  - Network of desktop computers
  - Use idle CPU cycles ("scavenging") to execute tasks
- **Data Grids**
  - Access, modify or transfer huge amounts of data

Can you identify some potential design issues in Grid Computing?

Design Issues in Grids

- **Heterogeneity** – Connected systems may run different OS, have different architecture
- **Security** – How do we trust the solution returned by a node?
- **Dynamic Nature** – When nodes come and go dynamically, how do we obtain the result?
- **Network Connectivity** – What happens if nodes cannot periodically connect?

A Look Back at the “Grid” Analogy

- An electric grid possesses the following features:
  - Interconnected nodes acting like a single entity
  - Users do not know which node serves their requests
- Grid computing is a collaborative effort, and users usually do not pay for the services they receive.

Utility Computing

- Computing is considered as a utility, similar to how electricity is provided to homes.
- Every connection is metered.
- Users can utilize computing power according to their needs.
- Pay-as-you-use model.

Grid Computing	Utility Computing
It is a process architecture that combines different computing resources from multiple locations to achieve desired and common goal.	It is process architecture that provide on-demand computing resources and infrastructure on basis of pay per use method.
It distributes workload across multiple systems and allow computers to contribute their individual resources to common goal.	It allows organization to allocate and segregate computing resources and infrastructure to various users on basis of their requirements.
It makes better use of existing resources, address rapid fluctuations in customer demands, improve computational capabilities, provide flexibility, etc.	It simply reduces IT costs, easier to manage, provide greater flexibility, compatibility, provide more convenience, etc.
It mainly focuses on sharing computing resources.	It mainly focuses on acquiring computing resources.
It is of three types i.e., computational grid, data grid, and collaborative grid.	It is of two type i.e., Internal and external utility.
It is used in ATMs, back-end infrastructures, marketing research, etc.	It is used in large organizations such as Amazon, Google, etc., where they establish their own utility services for computing storage and applications.
Its main purpose is to integrate usage of computer resources from cooperating partners in form of VO (Virtual Organizations).	Its main purpose is to make computing resources and infrastructure management available to customer as per their need, and charge them for specific usage rather than flat rate.
Its characteristics include resource coordination, transparent access, dependable access, etc.	Its characteristics include scalability, demand pricing, standardized utility computing services, automation, etc.