Chapter 08: Database CAP Theorem

CAP Theorem কী?

CAP Theorem (Consistency, Availability, Partition Tolerance) হলো **Distributed Database System**-এর একটি মৌলিক নীতি। এটি ২০০০ সালে কম্পিউটার বিজ্ঞানী **Eric Brewer** প্রস্তাব করেন এবং পরে ২০০২ সালে **Seth Gilbert** ও **Nancy Lynch** এটি গাণিতিকভাবে প্রমাণ করেন।

CAP Theorem অনুযায়ী,

> **একটি Distributed Database একই সময়ে Consistency (C), Availability (A) এবং Partition Tolerance (P)—এই তিনটি বৈশিষ্ট্য একসাথে শতভাগ নিশ্চিত করতে পারে না। Network Partition ঘটলে সিস্টেমকে Consistency অথবা Availability—এই দুটির মধ্যে একটি বেছে নিতে হয়।**

CAP বলতে কী বোঝায়?

১. Consistency (C)

Consistency নিশ্চিত করে যে, একটি ডেটা আপডেট হওয়ার পর সব Database Node একই এবং সর্বশেষ (Latest) ডেটা দেখাবে।

অর্থাৎ—

- সব ব্যবহারকারী একই ডেটা পাবে।
- কোনো Node-এ পুরোনো (Stale) ডেটা থাকবে না।

উদাহরণ

ধরুন, একটি ব্যাংক অ্যাকাউন্টে **১,০০,০০০ টাকা** আছে।

আপনি **১০,০০০ টাকা** উত্তোলন করলেন।

Update হওয়ার পর সব Server-এ Balance হতে হবে—

৯০,০০০ টাকা

যদি কোনো Server এখনো **১,০০,০০০ টাকা** দেখায়, তাহলে Consistency বজায় থাকেনি।

---

২. Availability (A)

Availability নিশ্চিত করে যে, Database-এর একটি বা একাধিক Node বন্ধ থাকলেও প্রতিটি Request-এর উত্তর (Response) পাওয়া যাবে।

তবে সেই Response সর্বশেষ ডেটা নাও হতে পারে।

উদাহরণ

ধরুন, Facebook-এ আপনি একটি ছবি আপলোড করেছেন।

আপনার বন্ধু ২–৩ সেকেন্ড পরে ছবিটি দেখতে পেল।

এটি গ্রহণযোগ্য।

কিন্তু যদি Facebook-ই সম্পূর্ণ বন্ধ হয়ে যায়, তাহলে সেটি Availability-এর সমস্যা।

৩. Partition Tolerance (P)

Partition Tolerance নিশ্চিত করে যে, Database Server-গুলোর মধ্যে Network Communication বিচ্ছিন্ন (Network Partition) হলেও সিস্টেম কাজ চালিয়ে যাবে।

উদাহরণ

ধরুন, ঢাকা Data Center এবং সিঙ্গাপুর Data Center-এর মধ্যে Network Link বিচ্ছিন্ন হয়ে গেছে।

তারপরও Database Service চালু থাকবে।

Network Partition কী?

Network Partition হলো এমন একটি পরিস্থিতি, যেখানে Distributed Database-এর এক বা একাধিক Node পরস্পরের সাথে যোগাযোগ করতে পারে না।

কারণ হতে পারে—

- Network Failure
- Router Problem
- Switch Failure
- Cable Cut
- Data Center Outage

এই অবস্থায় CAP Theorem কার্যকর হয়।

CAP Theorem-এর মূল ধারণা

Normal অবস্থায় অনেক Database-ই C, A এবং P তিনটিই দিতে পারে।

কিন্তু **Network Partition** ঘটলে Database-কে দুটি বিকল্পের মধ্যে একটি বেছে নিতে হয়—

- **Consistency (C)**
- **Availability (A)**

কারণ **Partition Tolerance (P)** বজায় রাখা প্রায় বাধ্যতামূলক।

CAP-এর তিনটি সমন্বয়

১. CP (Consistency + Partition Tolerance)

এখানে Database সব Node-এ একই ডেটা নিশ্চিত করে।

যদি Network Partition হয়, তাহলে কিছু Request সাময়িকভাবে Reject করা হতে পারে।

সুবিধা

- Data Accuracy বেশি।
- Data Consistency নিশ্চিত।

অসুবিধা

- Availability কমতে পারে।

উপযুক্ত ক্ষেত্র

- Banking System
- Payment Gateway
- Inventory Management
- Healthcare

উদাহরণ

- MongoDB (Majority Write Concern ব্যবহারে)
- HBase
- ZooKeeper

২. AP (Availability + Partition Tolerance)

এখানে Database সবসময় Request-এর উত্তর দেওয়ার চেষ্টা করে।

তবে কিছু সময়ের জন্য বিভিন্ন Node-এ ডেটা ভিন্ন হতে পারে (Eventually Consistent)।

সুবিধা

- Service সবসময় সচল থাকে।
- High Availability নিশ্চিত হয়।

অসুবিধা

- কিছু সময়ের জন্য পুরোনো (Stale) Data দেখা যেতে পারে।

উপযুক্ত ক্ষেত্র

- Social Media
- Messaging App
- News Feed
- Content Delivery Network (CDN)

উদাহরণ

- Apache Cassandra
- Apache CouchDB
- Amazon DynamoDB

৩. CA (Consistency + Availability)

CA System একই সঙ্গে Consistency এবং Availability নিশ্চিত করে।

কিন্তু এটি তখনই সম্ভব, যখন **Network Partition না থাকে**।

এ কারণেই বড় Distributed System-এ প্রকৃত অর্থে CA অর্জন করা সম্ভব নয়।

উদাহরণ

- Single Server Oracle Database
- Single Server PostgreSQL
- Single Server MySQL

> **দ্রষ্টব্য:** অনেক বইয়ে Oracle, PostgreSQL ও MySQL-কে CA Database বলা হয়। বাস্তবে এগুলো Distributed Mode-এ চালালে Partition ঘটলে CAP Trade-off প্রযোজ্য হয়।

CAP Theorem-এর বাস্তব উদাহরণ

উদাহরণ ১: Banking System

ধরুন—

একজন গ্রাহক এটিএম থেকে **১০,০০০ টাকা** তুললেন।

হঠাৎ Network Partition হয়ে গেল।

এখন Database-এর দুটি বিকল্প—

বিকল্প ১ (CP)

Transaction সাময়িকভাবে অপেক্ষা করবে বা ব্যর্থ হবে, কিন্তু ভুল Balance দেখাবে না।

ফলাফল: Data সঠিক থাকবে।

বিকল্প ২ (AP)

Transaction চলতে থাকবে, কিন্তু কিছু Server পুরোনো Balance দেখাতে পারে।

ফলাফল: একই টাকা দুইবার তোলার ঝুঁকি তৈরি হতে পারে।

তাই Banking System সাধারণত **CP** বেছে নেয়।

উদাহরণ ২: Social Media

আপনি Facebook-এ একটি ছবি পোস্ট করলেন।

Network Partition হওয়ার পরও সবাই Facebook ব্যবহার করতে পারছে।

কিছু ব্যবহারকারীর কাছে ছবিটি কয়েক সেকেন্ড পরে দেখা গেল।

এটি গ্রহণযোগ্য।

তাই Social Media সাধারণত **AP** নীতি অনুসরণ করে।

CAP Theorem-এর সারসংক্ষেপ

বৈশিষ্ট্য	অর্থ
Consistency (C)	সব Node একই ও সর্বশেষ ডেটা দেখাবে।
Availability (A)	প্রতিটি Request-এর উত্তর পাওয়া যাবে।

বৈশিষ্ট্য	অর্থ
Partition Tolerance (P)	Network বিভাজন হলেও সিস্টেম চলবে।

কখন কোনটি বেছে নেবেন?

Application	অগ্রাধিকার
Banking	CP
Payment System	CP
Hospital	CP
Inventory	CP
Facebook	AP
Instagram	AP
WhatsApp	AP
IoT Data Collection	AP

CAP Theorem কেন গুরুত্বপূর্ণ?

- Distributed Database নির্বাচন করতে সাহায্য করে।
- Database Architecture ডিজাইন সহজ করে।
- Consistency ও Availability-এর মধ্যে সঠিক ভারসাম্য নির্ধারণ করতে সাহায্য করে।
- High Availability System ডিজাইনে গুরুত্বপূর্ণ ভূমিকা রাখে।
- Big Data ও Cloud Database বুঝতে সহায়তা করে।

উপসংহার

CAP Theorem হলো Distributed Database Design-এর অন্যতম গুরুত্বপূর্ণ নীতি। এটি আমাদের শেখায় যে, Network Partition-এর সময় একটি Distributed Database একই সঙ্গে **Consistency**, **Availability** এবং **Partition Tolerance**—এই তিনটি বৈশিষ্ট্য সম্পূর্ণভাবে নিশ্চিত করতে পারে না। তাই প্রজেক্টের ব্যবসায়িক চাহিদা অনুযায়ী **CP** অথবা **AP** পদ্ধতি নির্বাচন করতে হয়। একটি সফল Database Architecture তৈরির জন্য CAP Theorem সম্পর্কে পরিষ্কার ধারণা থাকা অত্যন্ত গুরুত্বপূর্ণ।