

Data Dwarf – Shrinking a Petacube

Definition:

The **Dwarf structure** is a **highly compressed multidimensional data structure** that unifies **prefix and suffix redundancies** to significantly reduce storage space. It is suitable for **handling large-scale OLAP data cubes**, making it feasible to store **petabyte-sized cubes in a few gigabytes**

Features of Data Dwarf (Referenced from PDF)

Feature	Description
Compression	Dwarf eliminates prefix and suffix redundancies in OLAP cubes
Efficiency	A petabyte-sized cube (25 dimensions) can be reduced to 2.3 GB
Sparse & Dense Handling	Uses prefixes for dense areas and suffixes for sparse areas
Storage Optimization	Reduces storage size and computational overhead
Distributed Processing	Suitable for mobile networks and distributed systems