

Cortex II

Data Store for Large and Complex Data.

Cortex II is a scalable parallel processor specifically designed for storage and access to large volumes of complex, unstructured data.

Introduction

Many of today's datasets are large, complex, unformatted and contain errors. Cortex II and the AURA technology is the only technology that has hardware designed *specifically* to handle such data. The result is optimum performance, scalability and compact implementation at a cost effective price.

Cortex II is a flexible parallel computer designed specifically to support Cybula's AURA technology for the storage and retrieval of large complex data. AURA solves the problem of accessing large complex data sets using a suite of powerful data access tools. Cortex II has been specifically designed to support AURA-based applications, hosted on industry standard enterprise servers.

Cortex II systems may be clustered via optional pattern match control software allowing large collections of distributed data to be searched. Based on Grid and cluster software, Cybula's clustering technologies allow complete scalability.

Providing unrivalled access to large complex datasets, Cortex II is the new standard in high performance computing. The system is primarily aimed at applications using large amounts of data that can be noisy and potentially incomplete. It provides a very compact platform at a competitive cost when compared to other systems.

Platform

Cortex II consists of a set of pattern match nodes, AURA software and a host machine. Shown here is a typical implementation on a Sun platform. However, the Cortex II system may be hosted on any processor with a PCI bus.

AURA Application areas include:

- Financial data analysis
- Diagnostic systems
- Medical signal analysis
- Address management and lookup
- Image databases
- Document management
- Access to pharmaceutical data



Data Types

Cortex II complements existing database servers by providing fast and effective access to both complex data (i.e. graphs, numbers & text) as well as data that is incomplete or noisy. Typically this type of data is difficult to index with conventional databases.

Applications

Cybula's application adapters build on the base data types to allow searching of specific data. Existing systems include trade mark searching (AURAtm), molecular searching (AURAmol), Signal data searching (Signal Data Explorer), text searching (AURAtse), document searching (MinerTaur), Facial recognition (FaceEnforce). Other adaptors may be developed by the user or by Cybula.

Scalability

Utilising Cybula's own high performance pattern match engine, AURA, this highly flexible parallel computer applies a unique design with methods that allow access to large, complex and uncertain data. A system can be configured with any number

Cortex II

Data Store for Large and Complex Data.

of these unique nodes depending on the performance required by the user, offering scalability that can meet current and future demands for any organisation.

The AURA technology is aimed at customers wishing to integrate flexible access to complex data within large data-rich applications. Developers are presented with an easy-to-use API that can be used to integrate with existing applications, and Cortex II is the cost-effective parallel system designed to support these applications.

Typical performance

One node in a Cortex II system is capable of searching in the order of 30 million text records for a matching text sample (i.e. unformatted address) in under 500ms.

Applications and Training

Cybula provide comprehensive training and documentation as well as integration consultancy to allow users to quickly apply the system to solve their data problems.



Example configuration for very high performance computing:

- Sun enterprise server host
- 4TB RAID
- 20 PRESENCE 2 Pattern Match Cards
- AURA data store
- AURA graph matcher
- AURA text matcher
- AURA signal matcher
- Cybula PMC Grid-based control for pattern access over multiple nodes
- Cortex II System monitor

FEATURES

- Cortex II can be used effectively for large and small datasets.
- Cortex II AURA software supports three main types of data:
 - Numbers: i.e. signals, measurements.
 - Text: i.e. documents, lists, web pages.
 - Graphs: i.e. images and knowledge structures.
- Cortex II can be used stand alone, on a network or within a distributed computing environment (Grid's).
- Cortex II is hosted on industry-standard platforms.
- A Cortex II system uses Cybula's unique high-performance pattern match PCI hardware, PRESENCE 2. A system may be configured for any number of PRESENCE 2 cards and AURA Pattern Match Engines.
- Each PRESENCE 2 node contains up to 4GB of on board memory for data structures.

For more information, please contact Cybula at the address below.

powered by AURA

Cybula Ltd
IT Centre, York Science Park,
Heslington, York YO10 5DG

t: +44 (0) 1904 567686
f: +44 (0) 1904 567685
e: enquiries@cybula.com

www.cybula.com

Company Registered Number 3972962
ISSUE 1.1 2004

CYBULA
> high performance pattern recognition systems