

AI demo on an AI Accelerator-Intel Gaudi April 2025

Dr. Yehuda Singer Dr. Joshua Gur

yehuda.singer@llp.co.il Joshua.gur@llp.co.il



Parallel Solutions Technology main Applications

- Parallel Solution has developed a unique application - *Auto Parallel*™ - with the following features:
 1. Automatic **acceleration** of execution of any application on several core computer. (Speedups are of the order of the number of cores)
 2. Automatically **saving** about 90% of the **power** for a multi-core system not supporting idle management techniques.
 3. Novel **distributed computing systems**
 1. **Multi-core computer:** Every computer runs a portion of the program; utilizing its cores, a tremendous power saving & execution acceleration is achieved
 2. Applicable to various HW platforms:
 1. Servers,
 2. Embedded System
 3. Cloud.
 4. Portability: the parallelizing application will work without any change on Windows/ Linux platform & Generic: using the standard compiler, linker and debugger.
 5. Automatic maintainability.
 6. Maximization of Cache utilization FOR Reverse Engineering immunity and Enhance performance.



Why current technology cannot compete

*Present technology faces **A Major Problem:***

In current State Of The Art Software containing **tightly coupled** threads, the required decomposition to independent threads involves inter-thread Communication and synchronization code.

This is a very difficult and **time-consuming effort**

Using **our technology**, this is performed automatically!



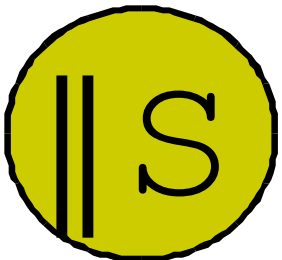
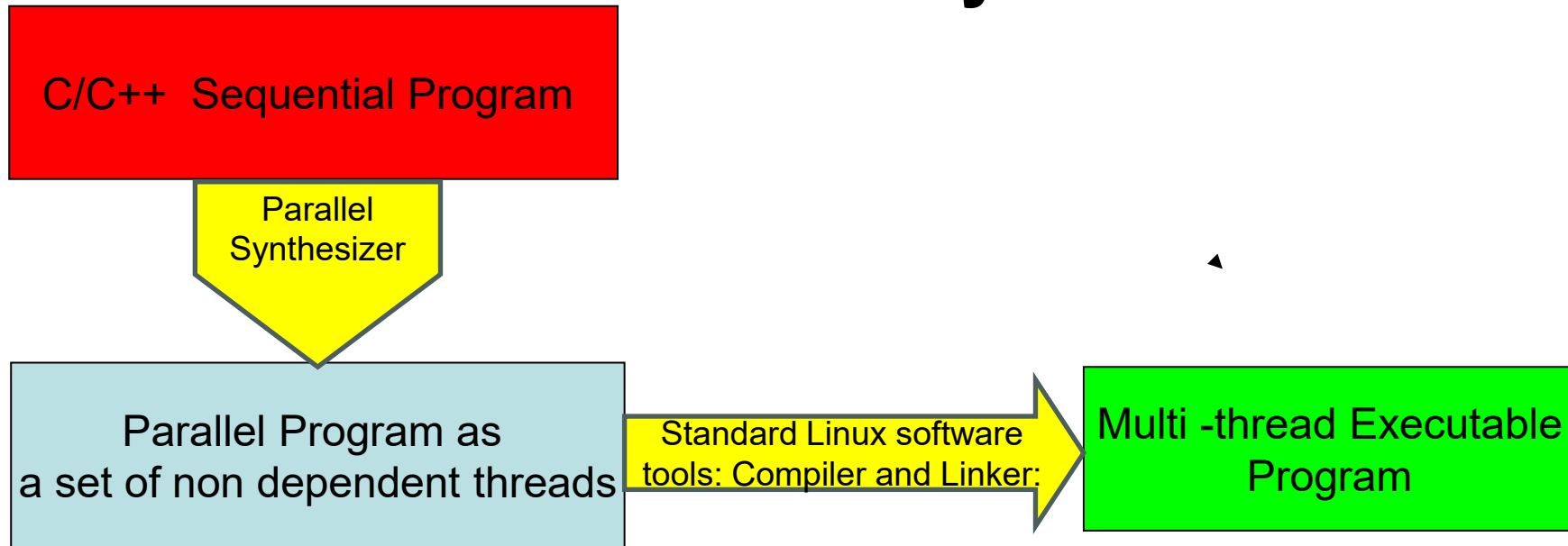
Our Solution is the *Auto Parallel*™

Our Auto-parallel - Breaking Dependencies Paradigm can:

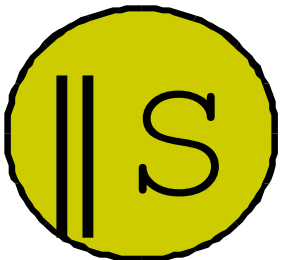
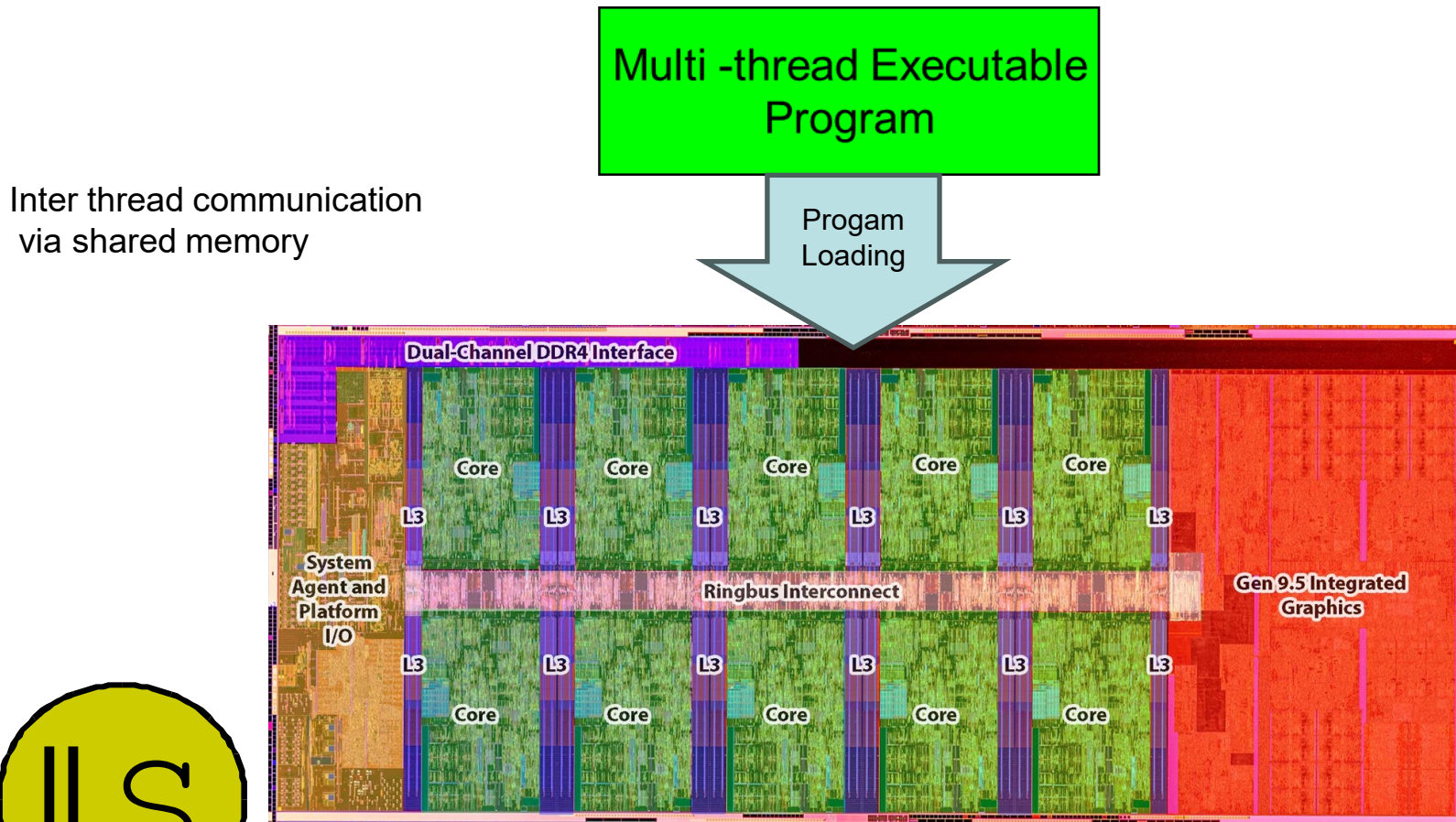
- Automatically parallelizes any **SERIAL** high-level program.
 - Automatically accelerates the execution of any application on a multi-core computer.
 - Automatically accelerates execution of programs on a set of multi-core computers connected via a communication network on a distributed computer system.
- Remark: Currently C/C++ are supported.



Transforming a sequential C/C++ program to an executable parallel file running on a multi-core system.

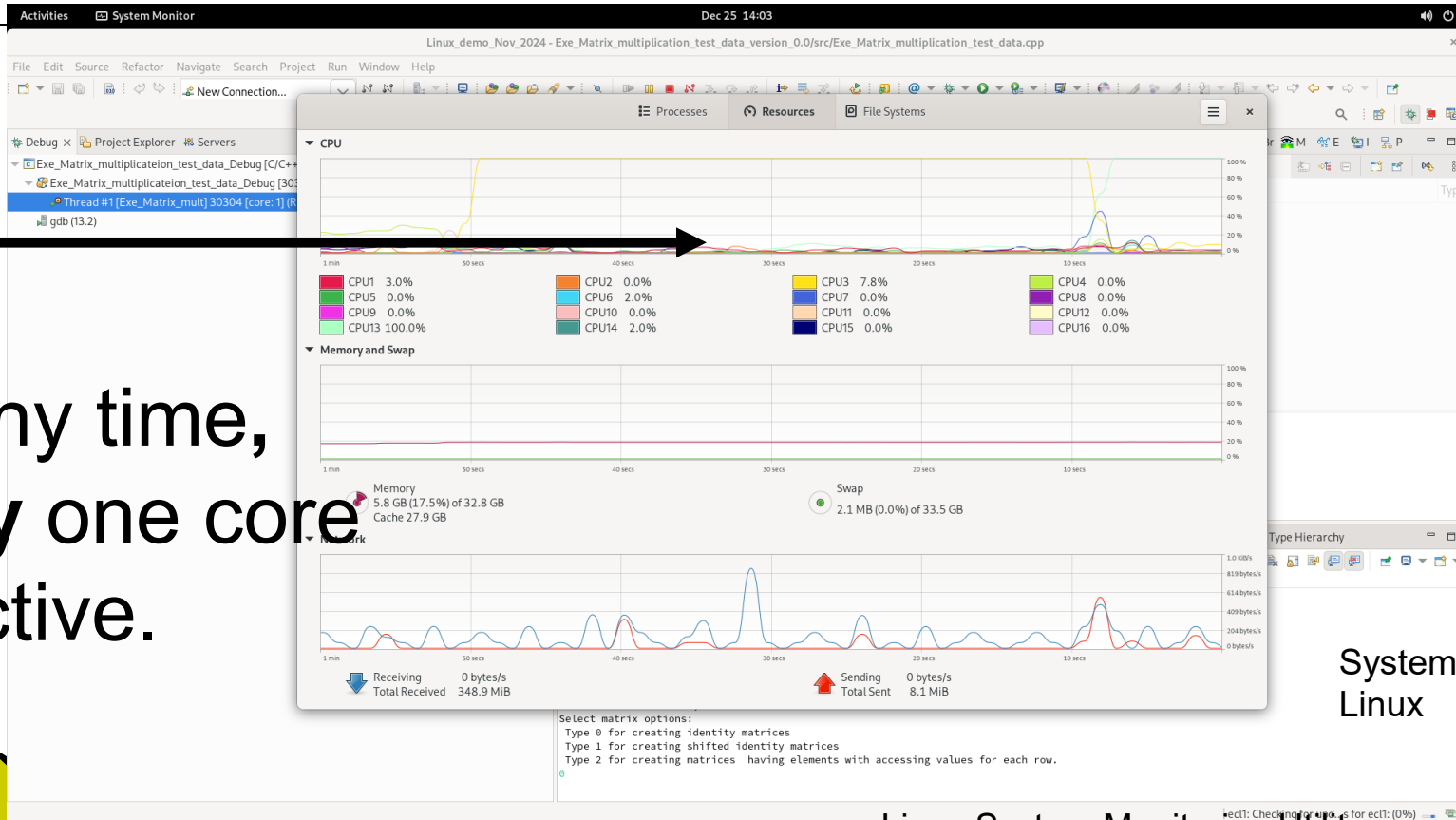


Running the program on an i9 Processor using Shared Memory



Sequential: 1 thread Running on an i9-16 cores Processor

At any time,
Only one core
is active.



System Monitor-
Linux

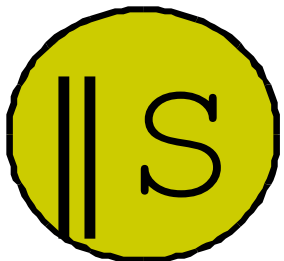
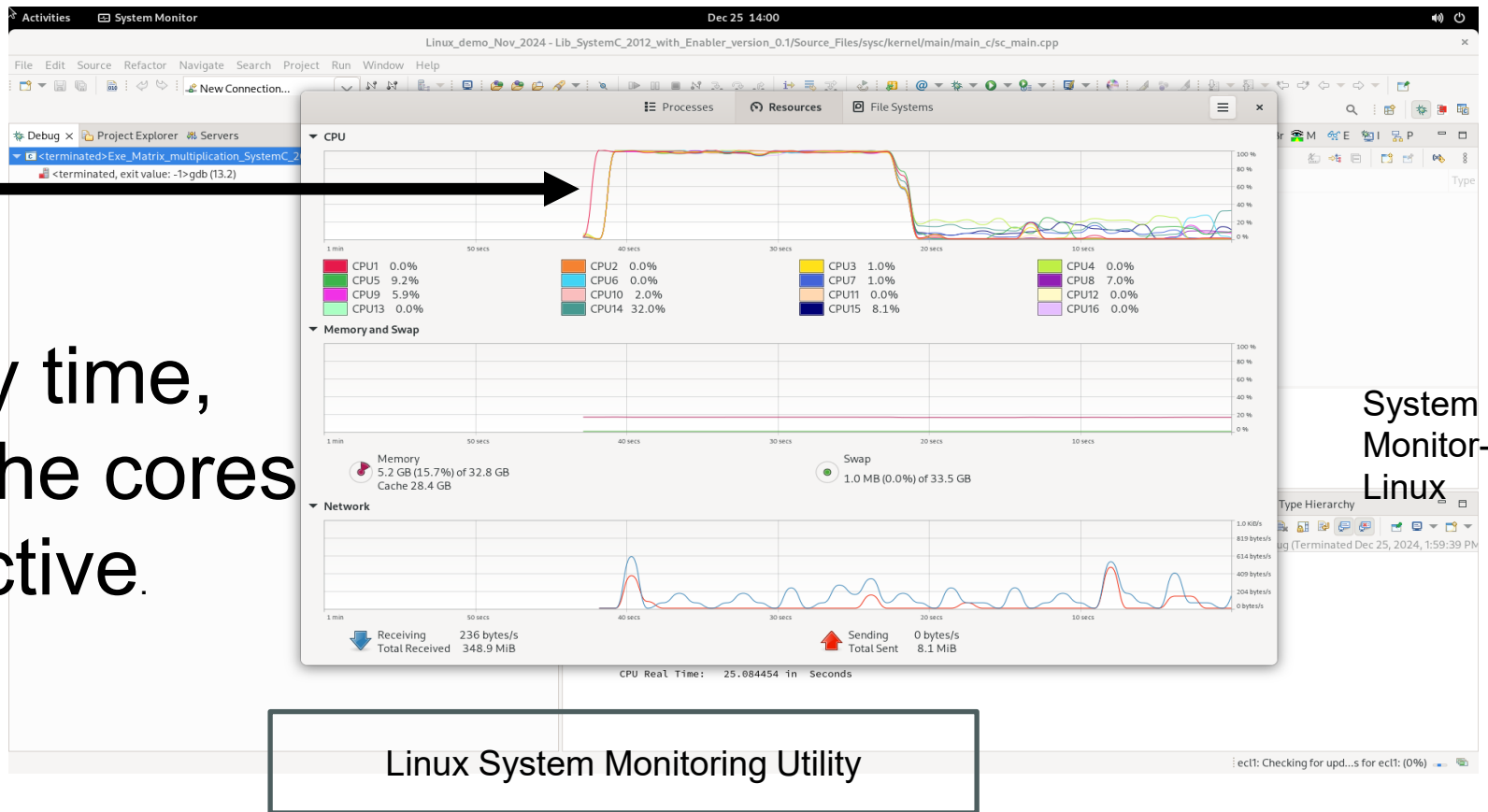
Linux System Monitoring Utility

22 April 2025

Proprietary Information Parallel Solutions Ltd under
NDA with Intel Inc. www.llp.co.il

Parallel: 16 threads Running on an i9-16 cores processor

At any time,
ALL the cores
are active.



Proprietary Information Parallel Solutions Ltd under
NDA with Intel Inc. www.llp.co.il

Performance of Sequential & Parallel programs

Measured by the internal clock of the Linux system.
Accuracy 1ns.

Matrix size	Serial Time	Optimized Parallel time	Acceleration Ratio
1024X1024	2.03451	0.044720	45.4944
2048X2048	65.056	3.011256	21.159
4096X4096	589.051	25.0819	23.485



Energy Saving

1. Energy saving because of reducing execution time by at least number of cores.
2. Options for further reducing the energy:
 - 2.1. Enhancing cache memory utilization automatically to reduce additionally execution time.
 - 2.2. Option: Reducing the clock frequency at the order 10-20% and save energy by 20-40%.



Demo on AI Accelerator- Intel Gaudi

- Goals:

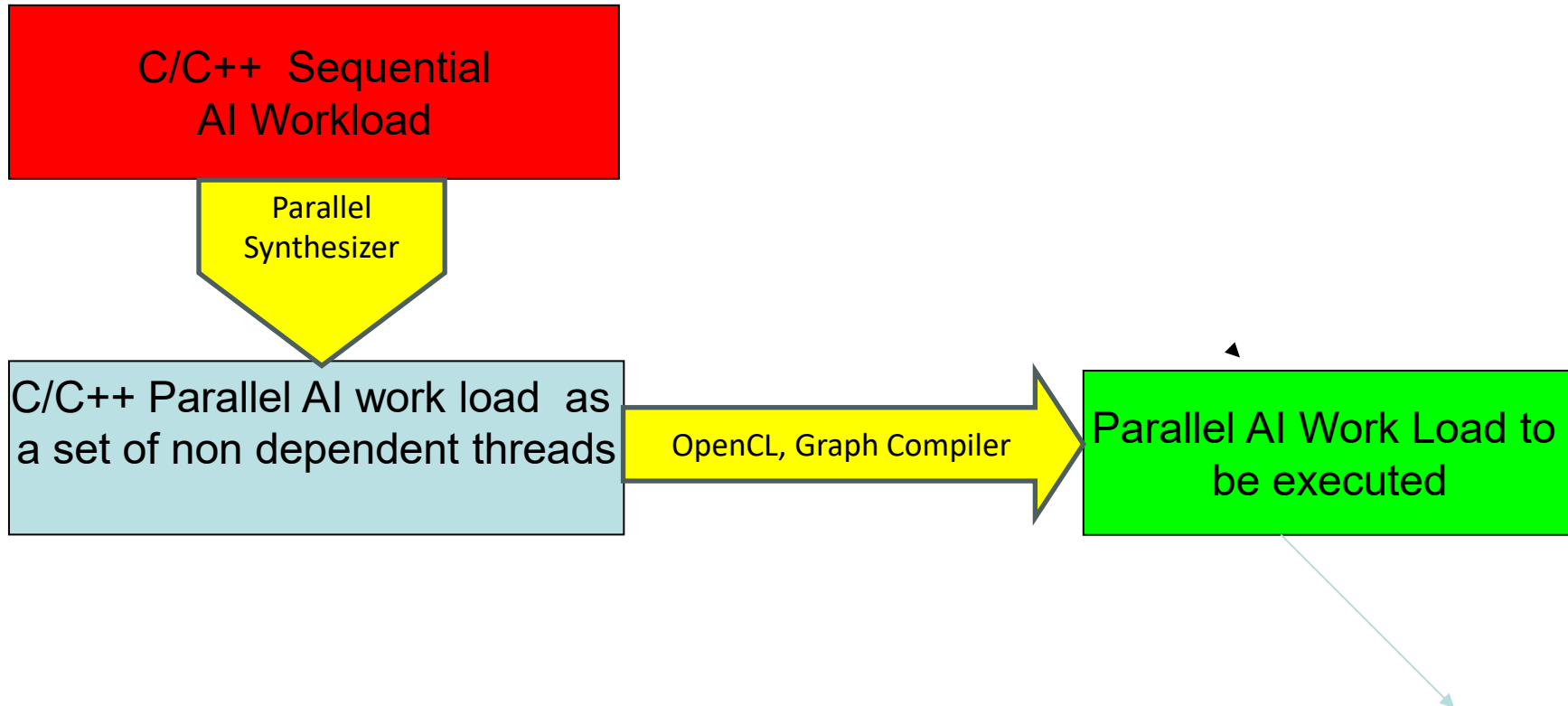
1. Reducing the execution time on a **Gaudi** based system by the factor of number of processors.
2. Reducing energy consumption by the same factor.

- Method

1. Running the original sequential AI Work Load.
2. Running the revised multi-threaded Work Load using PS technology.
3. Measuring execution time using OS clock.



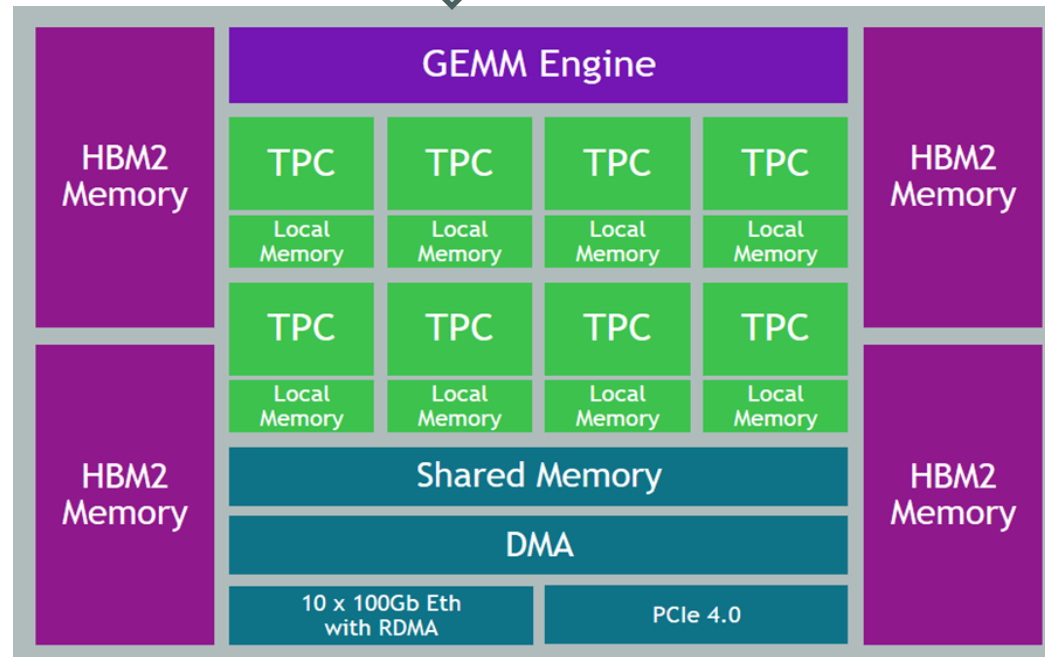
Producing an executable file for a Guadi Server



Running the program on A Gaudi Server

Parallel AI Work Load to be executed

PCI Express



Original Intel figure
AI accelerator

Inter thread communication
via shared memory



Proprietary Information Parallel Solutions www.llp.co.il 01-Jun-20

Parallel Solutions: Leading Team

- **Dr. Yehuda Singer**

Ph.D. in computer science, specialized in computer architecture, **Real-Time** and high performance. Over 46 years of experience in the development of embedded systems and FPGAs in various multi-disciplinary applications.

Dr. Singer was the leader of the Computer Studies of the extension of Derby university in Israel. Dr. Singer holds an M.Sc. and Ph.D. degrees in computer science from Weizmann Institute and Bar-Ilan university respectively.

- **Dr. Joshua Gur**

Joshua Gur – Ph.D. in Physics & Optics, specialized in multidisciplinary computerized system design as a chief Display & Video engineer at a security plant in Israel - the Israeli Aircraft Industries. and other security plants. Over 48 years of experience including 9 patents. Dr. Gur holds a B.Sc. in Physics and Mathematics, M. Sc in Electro- Optics from the Hebrew University of Jerusalem, and a Ph.D. in Physics & optics from the institute of optics at Rochester University in Rochester NY at USA.

