

Opteron und I/O

Toni Schmidbauer

11. Mai 2005

Zusammenfassung

Eine kurze Beschreibung der AMD Opteron Architektur

Inhaltsverzeichnis

1	Allgemeines	2
2	Was ist ein Interconnect?	2
3	Traditionelles PC Chipset	2
4	Xeon mit proprietärer NUMA Architektur	3
5	Opteron: NUMA und HyperTransport	4
5.1	HyperTransport	4
5.2	Opteron	5
6	Konklusio	6

1 Allgemeines

AMD's neue Opteron Architektur hat auch bei uns zu einigen Missverständnissen geführt. Es war / ist zum Teil unklar was diese CPU jetzt wirklich von standard Intel Systemen unterscheidet. Ich versuche daher einen kurzen Einblick in die I/O Architektur der CPU zu geben. Im Punkt I/O unterscheidet sich der Opteron wesentlich von bisherigen Intel CPU's und dieser Unterschied kann auch als seine großer Vorteil betrachtet werden.

2 Was ist ein Interconnect?

Als Interconnect bezeichnet man die Verbindung einzelner Komponenten in einem Server/PC. Im wesentlichen werden folgende Teile miteinander verbunden:

- CPU, Memory, Bridges
- I/O Komponenten (Audio, Netzwerk, Storage, Video)

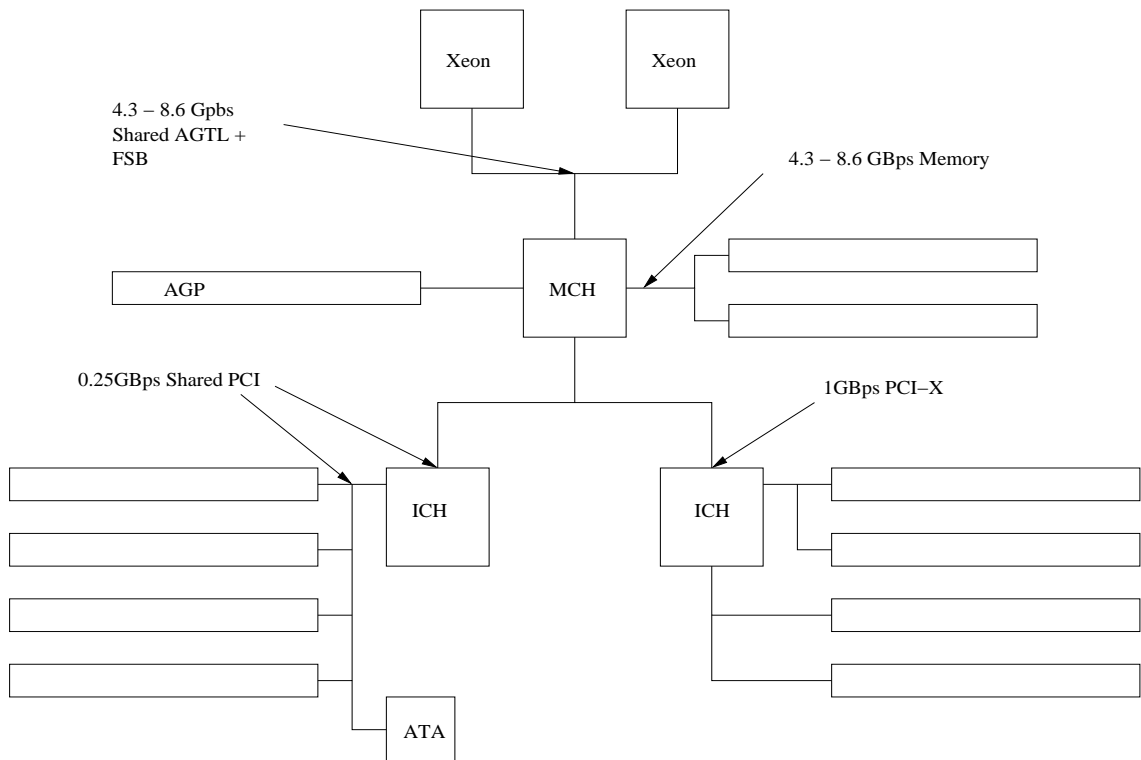
Bis jetzt gab es eigentlich nur sehr wenige Standards wie diese Teile miteinander verbunden werden. Es gab keine universelle Lösung, jede Plattform hatte ihre eigenes Süppchen gekocht.

3 Traditionelles PC Chipset

Das traditionelle PC Chipset besteht aus:

- Northbridge: System Interconnect (CPU, Memory)
- Southbridge: I/O Interconnect

Diese beiden wurden wieder über eine I/O Bridge miteinander verbunden. Hier ein Grafik die das ganze verdeutlicht:



MCH (Memory Controller Hub) ist in diesem Fall die Northbridge, ICH (I/O Controller Hub) bezeichnet die Southbridge. Beide sind typischer Weise via PCI oder PCI-X verbunden.

Das Hauptproblem dieser Architektur ist der sogenannte Front Side Bus oder FSB. FSB bedeutet, daß es nur **eine** Verbindung zwischen der CPU und dem restlichen System gibt.

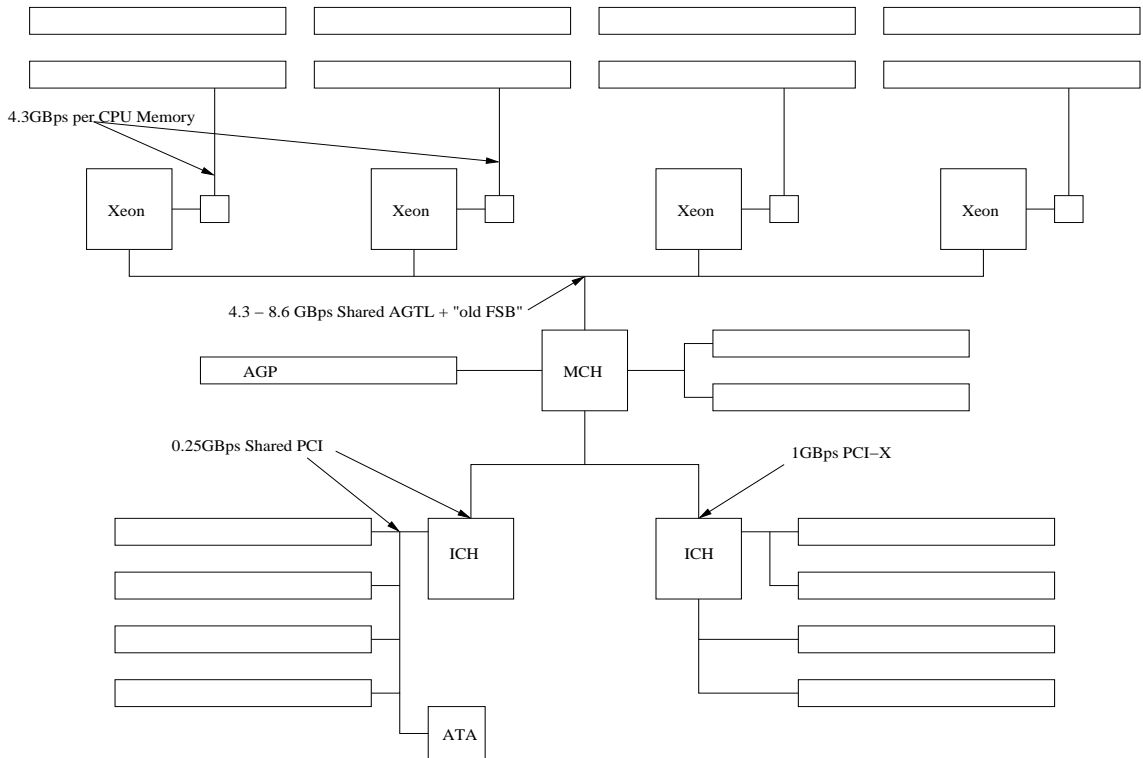
Wie anhand der Grafik deutlich erkennbar ist teilen sich CPU, Memory und I/O Board den gleichen Bus. Bei modernen System mit Gigabit Ethernet Karten, diversen RAID Controller usw. kann es daher leicht zu einer Auslastung dieses Buses kommen.

4 Xeon mit proprietärer NUMA Architektur

NUMA bedeutet Non-Uniform Memory Access. Das heißt jede CPU hat eine direkte Verbindung zu ihrem eigenen Speicher. Es gibt keinen gemeinsamen Bus zwischen den CPU's zum Arbeitsspeicher. Einziger Nachteil dieser Lösung ist, wenn eine CPU auf den Speicher einer anderen CPU zugreifen muss, ist dies mit einem Overhead verbunden. Eine elegante Lösung zu diesem Problem ist ein OS, welches mit dieser Architektur umgehen kann.

Speicher der in einer Applikation allokiert wird, sollte immer lokaler Speicher der CPU sein, auf dem die Applikation gerade läuft (Processor Affinity).

Hier eine Grafik zur Veranschaulichung:



Trotz aller Vorteile, gibt es auch ein paar gravierende Nachteile:

- NUMA mit Xeon CPU ist ein "Hack". Es gibt keine Standard Chipset's die dies beherrschen.
- CPU, System Interconnect und Speicher, befinden sich meist auf einer Tochterkarte, welches die Kosten in die Höhe trieb
- Die CPU selbst hat nur einen FSB, d.h. es ist nur ein Transfer nach dem anderen möglich (CPU <-> Memory oder CPU <-> I/O).
- CPU <-> CPU bzw. CPU <-> I/O Transfers laufen über einen gemeinsamen Bus.

5 Opteron: NUMA und HyperTransport

5.1 HyperTransport

Zwei Firmen waren ausschlaggebend an dem neuen Interconnect beteiligt:

- API Networks (gehört jetzt zu AMD)
- AMD

API Networks ist aus Digital's Halbleiter Abteilung heraus entstanden. Die meisten Designer der Alpha Plattform gingen nach dem Exodus von Digital entweder zu API Networks oder zu AMD.

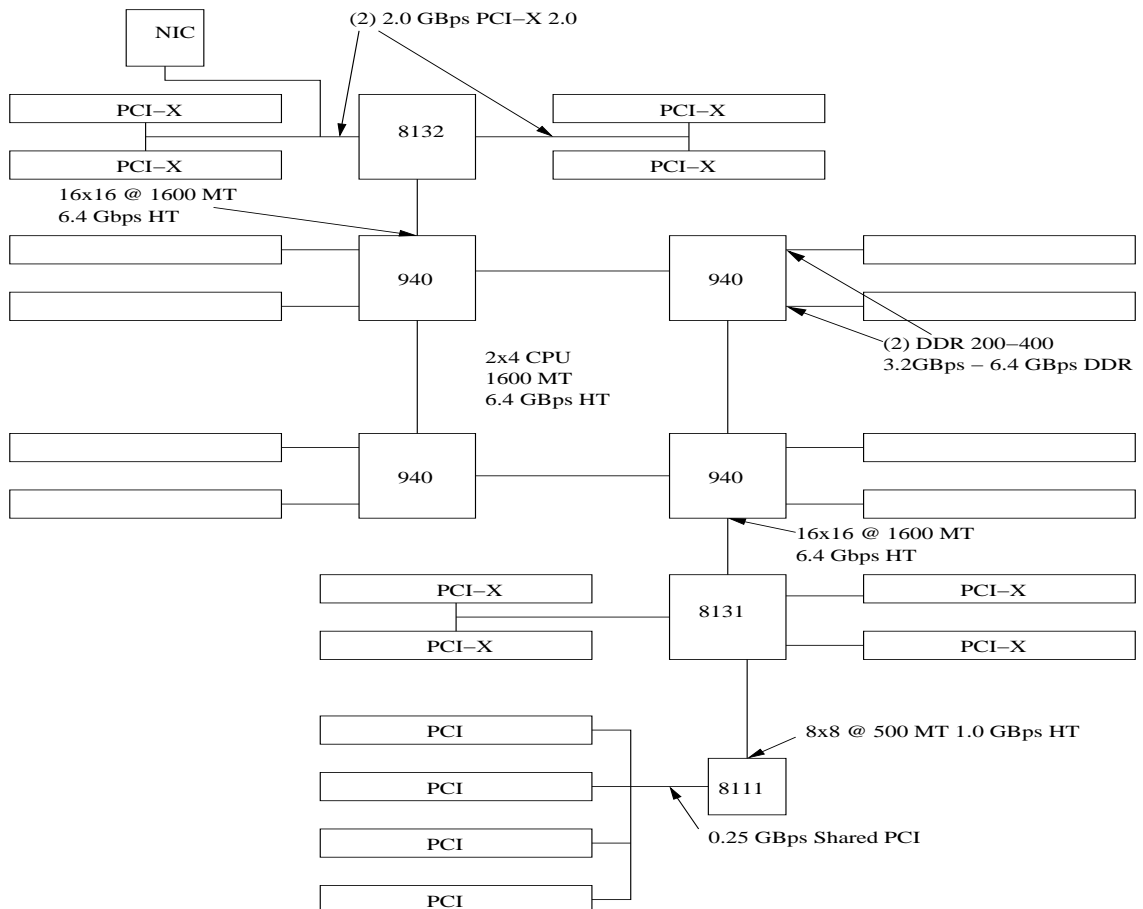
Ziel der HyperTransport Architektur:

- Universeller, nahtloser Interconnect zwischen CPU's
- Ein skalierbarer Interconnect zwischen Bridges (North, South)

Derzeit haben IBM, Sun und Apple Lizenzen von AMD zur Verwendung des HyperTransports in den jeweiligen nicht x86 Plattformen.

5.2 Opteron

Zu Beginn wieder eine Grafik:



Wie man auf den ersten Blick erkennt, gibt es keinen gemeinsamen FSB. Jede CPU hat bis zu 5 unabhängige Anschlüsse für einen Interconnect. Die Northbridge wurde ebenfalls in der CPU implementiert. Was bedeutet, daß jede CPU eine direkte Verbindung zum Memory hat.

Untereinander werden die CPU's via HyperTransport verbunden, auch der Zugriff auf I/O erfolgt über den HyperTransport. Jede CPU hat eine Verbindung zu anderen CPU's, zu ihrem Speicher und zum I/O.

Ein weiterer Pluspunkt ist das der Takt mit dem auf den Speicher zugegriffen wird, unabhängig vom Takt der I/O Zugriffe ist. Somit ist es egal welcher Speicher verwendet wird. Zugriff auf andere CPU's oder den HyperTransport geht immer über einen eigenen Takt (Clock).

Da die MMU (Memory Management Unit) bereits in die CPU integriert ist, bewirkt das jeder Transfer von einem I/O Gerät in den Speicher lokal auf eine CPU beschränkt ist. "Processor Affinity" ist somit nicht nur für Applikationen und Speicher gegeben, sondern auch für jeden I/O Transfer.

Bei Intel EM64T System ist die MMU nach wie vor in der Northbridge implementiert. Auch ist nach wie vor ein gemeinsamer FSB vorhanden.

6 Konklusio

Ich hoffe ich konnte die neuen Opteron System etwas näher bringen und verdeutlichen, warum sie derzeit in der x86 Welt das non plus ultra darstellen. Besonders für Anwendungen die sehr stark I/O Lastig sind, ist die AMD64 Plattform beim heutigen Zeitpunkt die beste Wahl. Interessant wäre ein Vergleich der nicht nur die reine CPU Leistung misst (openssl speed), sondern ein Test der auch I/O einbezieht.