uniserver

D3.1 Definition of the UniServer Board

Contract number	688540
Project website	http://www.uniserver2020.eu
Contractual deadline	Project Month 3 (M3): 30th April 2016
Actual Delivery Date	1st June 2016
Dissemination level	Public
Report Version	1.0
Main Authors	Sundeep Gupta (APM), Yoon Chae (APM),
Contributors	Kris Purcell (APM), Kaustubh Gondkar (APM), Rihcard Zenkert (APM)
Reviewers	Georgios Karakonstantis (QUB), Dimitris Gizopoulos (UoA), Pedro Trancoso (UCY), Mustafa Rafique (IBM)
Keywords	Micro-servers, Voltage Scaling, Frequency Scaling, Refresh-Rate, X-Gene board

Notice: The research leading to these results has received funding from the European Community's Horizon 2020 Programme for Reseach and Technical development under grant agreement no. 688540.

© 2016. UniServer Consortium Partners. All rights reserved

Acknowledgements
The work presented in this document has been conducted in the context of the EU Horizon 2020. UniServer is a 36-month project that started on February 1st, 2016 and is funded by the European Commission. The partners in the project are:

The Queen's University of Belfast (QUB)
The University of Cyprus (UCY)
The University of Athens (UoA)
Applied Micro Circuits Corporation Deutschland Gmbh (APM)
ARM Holdings UK (ARM)
IBM Ireland Limited (IBM)
University of Thessaly (UTH)
Worldsensing (WSE)
Metitorious Audit Limited (MER)
Sparsity (SPA)

More information
Public UniServer reports and other information pertaining to the project are available through the UniServer public Web site under http://www.uniserver2020.eu.

Change Log

Version	Description of change
0.1	Initial draft
0.2	Added X-Gene features
0.3	Added SHADOWCAT features
0.4	Added TIGERSHARK features
0.5	Added a detailed introduction section
0.6	Feedback from partners added
0.7	Described the specific UniServer features
0.8	Compiled the confidential information
0.9	Feedback from partners added
1.0	Final version for release to EC

Table of Contents

Index of Figures

Executive Summary

This document describes the specifications of the UniServer board as expected in the task T3.1 within the work package 3 (WP3). In particular, the general features of the TIGERSHARK X-Gene2 board that is the chassis of the UniServer platform, as well as the characteristics unique to this in-house board that will be exploited for supporting the objectives and functionalities of the UniServer hardware and software ecosystem, are analyzed. The specification of the board characteristics that are desired and planned to be included in the follow-up version of the board, the X-Gene3 are also discussed.

The intention of the document is to specify the unique characteristics of the UniServer platform, lay out the supported functions and suggest desirable characteristics that may be included in a future version of the UniServer X-Gene based board. The document serves as a reference for all partners on the specifications of the UniServer platform.

The document is organized as follows. Section 1 discusses the characteristics of the Micro-servers and introduces the unique features of the UniServer platform. Section 2 describes in detail the generic functionality supported by the X-Gene board, while Section 3 discusses some unique features build within a limited version of the commercial X-Gene board on which UniServer technologies are being integrated on. Finally, Section 4 discusses the planned extension of the UniServer platform with the exploitation of a brand new version of the X-Gene board anticipated on 2017.

1. Introduction

In this section, the characteristics of the Micro-servers are discussed and the unique features of the UniServer platform are being introduced.

1.1. Micro-servers

Micro-servers are servers which are based on a single System-on-a-Chip (SoC). These SoCs integrate most of the peripherals usually required as part of the server rack onto a single silicon die (except for the DRAM, the boot Flash and power circuits) along with the multi-processor cores, their associated caches and on-chip coherent bus. This renders itself to a highly efficient board design in terms of area, power and cost. A micro-data-centre may integrate multiple of these Micro-servers. Micro-servers satisfy the 3P requirements namely, Price, Power and Performance ideally required for a medium performance setup.

As the internet traffic is growing, there are multiple applications which do not need the high performance that is typically offered by the high-end servers. However, these applications do demand a huge number of small tasks to be performed. One of the biggest factors for this kind of growth in the Internet traffic will be Internet of Things (IoT) devices. IoT applications typically will require a lot of non-CPU intensive data handling. Micro-servers are ideally suited for these types of applications.

Micro-servers are typically used in applications where the workloads are predictable and are scale-out in nature. 'Scale-out' workloads imply that the workloads are inherently parallel and can be split across multiple computing units. This also implies growing workloads can be countered by simply adding more computing units. Overall, any data-centre is optimized around the three-Ps, i.e. Price, Power, and Performance and Price,

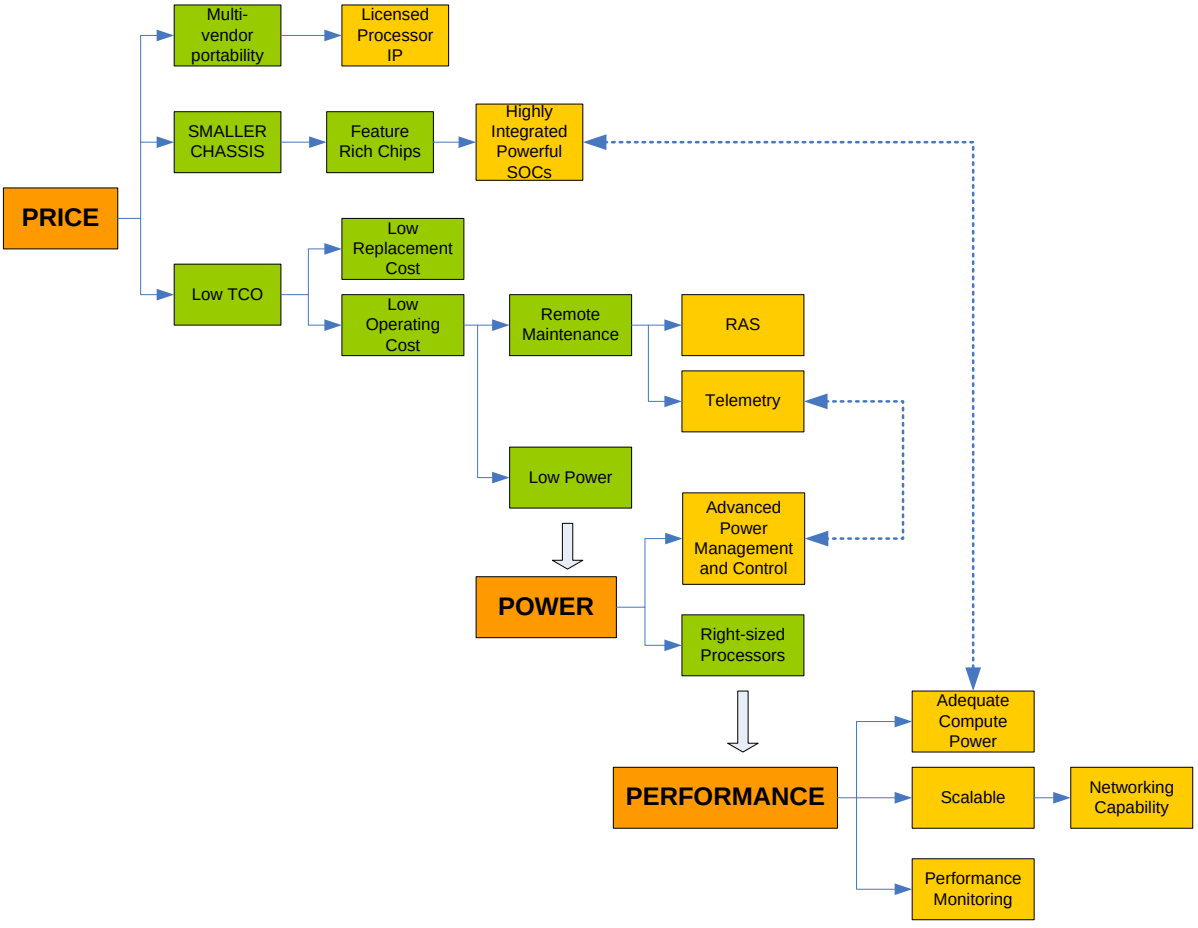

Figure 1: The three Ps for a Micro-server architecture

which need to be balanced as per the required application. Figure 1 below shows the relation between the various factors and how they translate to a required feature set at a broad level.

1.1.1. Price

The Price of the server is directly proportional to:

- *Multi-vendor portability:* Any design should not be solely dependent to a particular vendor. The overlaying software stack should be portable across multiple vendors to allow for healthy competition leading to lower costs. From this perspective the ARM Instruction Set Architecture (ISA) is the most widely available and most popular CPU ISA which is keeping up with the latest processing requirements as demanded by the technological innovations.

- *Board real estate:* The price of a Micro-server board is directly proportional to the size of the board. Moreover, since data-centres have thousands of such boards housed in a single location, the size of the board directly translates to the size of the data-centre facility. Hence, it becomes imperative to reduce the size of the board as much as possible. This effectively translates to highly integrated SOCs with powerful multi-core CPU design thus reducing not only the number of peripheral chips required on the board but also, the overall Bill of Material.

- Low Total Cost of Ownership (TCO): This requires that:
    - The replacement cost should be low in case of fatal system failures.
    - Low operating cost which in turn translates to:
        - Enable remote maintenance thus reducing manual labour costs. For remote maintenance the system design should lend itself to a robust RAS (Reliability, Availability and Serviceability) architecture. Also, it should support Telemetry to enable Remote measurement of various vital statistics such as Power consumption, Performance and Temperature.
        - Low power which in turn implies lower cooling costs as well.

1.1.2. Power

For reducing the overall Power consumption of any system, the SOCs should have advanced power measurement capabilities and control. This enables the system designer to tailor the voltage, frequency and DDR refresh rate as per the transient workload requirements. In addition, the processing power should be just about adequate for the required workloads.

1.1.3. Performance

The third and the most important parameter is the performance of the entire Micro-server system. The performance of the system is dictated by:

- Adequate compute power for the required workload.

- Scalability to ensure that as workloads increase additional systems can be easily added to meet the increased throughput demand. To enable this, the SOCs should feature fast and configurable networking interfaces in order to be able to create a fully interconnected data-centre.

- Performance monitoring: Based on the system workloads it is necessary to continuously monitor the system performance and analyse whether any system parameters need to be changed to improve the performance of the system. Conversely, if the system workload is low at a given point of time, the performance of the system can be tweaked down for power/energy savings. This requires that adequate system performance metrics are available to the software to make these decisions. This ensures that the system parameters, such as Voltage, Frequency and DDR refresh rates can be tweaked as per the transient workload and thus reducing the average power consumption.

Table 1 below summarizes the broad feature-set requirement for any Micro-server System-On-Chip (SoC).

Table 1: Micro-server broad feature set

Feature	Comment
Multi-vendor CPU ISA	Reduces monopoly and thus overall cost
Powerful CPUs	To ensure adequate compute power
Highly Integrated SOCs	Reduce the chassis size, and thus the overall cost
RAS	To ensure robustness and ease of maintenance
Power Management and Control	Reduces overall Power consumption and indirectly the overall TCO
Telemetry	Ensures ease of monitoring capabilities and ease of maintenance
Scalability	Integrated network interfaces for adding more servers as workload increases
Performance Monitoring	To ensure that the Power/Performance balance is maintained at all times

1.2. UniServer platform

These objectives are put forth considering certain challenges in mind. The challenges that are faced by today's Big Data world are as follows:

- Stagnant power scaling: As the transistor density is increasing it is not translating to proportional power scaling. This leads to extremely dense chips with high Thermal Design Power (TDP). [HEsm.ISCA11, ITRS13, JKoo.AHC11]

- Variations and Pessimistic Margins: Static and dynamic variations [KBow.JSSC11] result in adding voltage safety margins to ensure correct processor operation. Thus, pessimistic nominal power and voltage numbers are quoted to ensure operation of the chip across all process, temperature and voltage corners. This leads to increased power for even the parts which could safely be operated at lower power supplies without any malfunctioning.

- Scalability: The IoT revolution is going to skyrocket the demands on the Cloud Architecture. [FORRESTER14]. This makes it imperative to define architectures which would satisfy these requirements.

- Programmability: Need to ensure that the software stack is portable and does not rely on specialized hardware accelerators which will require re-programming when the underlying hardware is changed [ABar.TPDS13].

- Availability and Dependability: The system needs to ensure the availability of the resources even when some of them are down due to hardware failures. Moreover, data integrity needs to be guaranteed under such circumstances.

- Privacy and security: This requires local cloud at known locations rather than a public cloud at an unknown location.

UniServer aims to vertically integrate the overall work in a proof-of concept prototype. The project consortium members are planning to have an end-to-end development tool chain running real world applications on top of the ARM-based development platform (X-Gene2 which may eventually be replaced with X-Gene3). This prototype will serve as a common point of reference, and will foster cooperation among partners. It will also be used to demonstrate the envisioned functionality as part of the dissemination and exploitation effort. The X-Gene2 Validation board (codename TIGERSHARK) enhanced with automated power, error reporting and handling facilities, as well as the whole data-centre software stack will be able to support a large repertoire of emerging scale-out workloads. Its expected low cost, low power and cooling demands makes it suitable for environments with strict power constraints such as embedded, home users and small enterprises allowing to drive exciting new applications at the edge of the cloud. The developed platform will be the building block of

any system providing cloud services at different scales. Different constraints will allow for different configurations of the same building block, while the powerful software stack allows it to be adopted out of the box from current data-centre operators and help setup new data-centres closer to the data sources. UniServer has three SME partners which develop relevant applications for the IoT and Big Data paradigms: smart traffic control (IoT), financial services (Big Data) and social networking (IoT, Big Data). These real life applications will be ported to the UniServer prototype so as to evaluate the benefits of the real applications on real hardware which features the UniServer concept for improved energy and performance efficiency.

UniServer envisages an off-chip and an on-chip approach to achieving this as detailed in the WP description. These approaches translate to the following additional feature requirements from the Server SoC from what was described for a typical Micro-server earlier.

- Support for supply-voltage sense pins that connect directly to the supply voltage rails on-chip and are routed to accessible test-points on the PCB.
- Integrate a supply-voltage measurement macro to sense on-die voltage rails.
- Ability to have a fine-grain control (in the order of tens of mV) on the various VRMs supplying power to the different power rails of the SoC.
- Ability to have a fine-grain control on the Operating Frequency of the processors on a per-pair of processor basis at approximately 15% granularity.
- Ability to have a fine-grain control on the DRAM refresh rates at a bit granularity level which translates to the order of hundreds of micro-seconds.
- Ability to measure the on-die temperatures.
- Ability to measure voltage droop across the die.
- Detailed performance monitoring unit.
- Comprehensive and Robust Error logging, reporting, and handling capabilities for the most important hardware blocks.

The main objectives of the UniServer platform are

- To define a high performance design that is able to satisfy quality of service (QoS) requirements of any Internet Service.
- To be easily configurable and deployable to both traditional cloud data-centres as well as to data-centres in a box deployed at the edge of the cloud.
- Be highly energy efficient to reduce the capital and operating costs of traditional cloud data-centres but also to facilitate the utilization near the sources of data using a home power distribution network without the need for expensive cooling infrastructure.
- To support an easily programmable system software stack consisting of a virtualization environment that can seamlessly combine fog resources with the resources in a backbone cloud.
- To process the data reliably and securely without compromising their privacy and potential confidential nature.

As mentioned earlier, these objectives will be hit by revealing and harnessing the inherent pessimistic design margins of the hardware components of the platform.

1.3. X-Gene – the chassis of UniServer

The UniServer board will be based on the X-Gene™ Server-on-a-Chip™ that is developed by APM, core partner of UniServer. Applied Micro Circuits Corporation (APM) is a global leader in computing and connectivity solutions for next-generation cloud infrastructure and data-centres. The X-Gene™ Server-on-a-Chip™ product is a flagship platform that leads the entire product portfolio in hyperscale data-centers, increasing by 2016 the overall addressable market to as much as 10x that of 2008 levels.

It is anticipated that the new addressable market will consist of server solutions for cloud data-centers and enterprise applications, data-center connectivity solutions and a variety of related products including routers,

network-attached storage appliances, top of rack switches, and wireless access points. No other small-cap semiconductor company can match the explosive addressable market growth that APM has engineered.
X-Gene builds on APM's history as a global leader in providing energy-efficient, sustainable solutions to process and transport information for current and next generation networks. Being a leader in processor design, embedded processor System-on-Chip (SoC) design, mixed-signal design, high speed signal processing, Internet protocol, OTN and Ethernet packet processing gives AppliedMicro the foundation to compete and excel in the next several generations of solutions. High speed, low power connectivity and computing solutions provide the ideal balance for systems and products that conserve energy and lower costs while maintaining high performance levels for the target workloads.

As the need for delivery capability, density and lower total cost of ownership emerge as key differentiators of their products and services, data-centers will require high-speed connections to transfer information between systems more efficiently to enable virtualization and cloud computing applications. X-Gene, the world's first and currently only commercial ARM® 64-bit Server-on-a-Chip solution, features APM's enterprise class high performance Cloud Processor® cores coupled with switch fabric and high-speed networking capabilities. Combined, they deliver substantially lower power consumption and substantial total cost of ownership (TCO) savings. X-Gene represents a new, grounds-up Cloud Server® solution tailored for the rapid growth of structured and unstructured compute requirements in next generation data centers. APM has 3 generations of X-Gene products namely X-Gene1, X-Gene2 and X-Gene3. X-Gene1 and X-Gene2 have already been in the market and have been proven to be solid products.

1.4. APMs role

APM, on the basis of its proven leadership in the ARM-64 bit based Micro-server segment, has been selected as the hardware platform provider for the UniServer project. The UniServer project aims to define a universal Micro-Server ecosystem by exceeding the energy and performance scaling boundaries. The UniServer project is broken into multiple work packages and sub-tasks. This document specially addresses the work detailed in WP3 under task T3.1.

1.5. APM solution suitability for UniServer

The UniServer requirements as described above call for an above-par Server SoC, which has all these features. APMs X-Gene2 is ideally suited for this purpose. X-Gene2 features a ***custom*** 8-core ARMv8 Server SoC with integrated IOs required for Micro-server applications. APM was the first-to-market with the ARMv8 based Server SoC solution. It leads the competition in terms of server feature implementation and this has given it the unique advantage of 4 years of customer engagements and design wins enabling them to smoke out all issues and enhance their design.

Due to APM's proprietary custom ARMv8 processors, it has the unique advantage of tuning the processor design to achieve highest possible frequency of operation. As of today, X-Gene2 processors can achieve a highest frequency of 2.4GHz which is unmatched by any other ARM v8 processors available in the market today.

This alone, gives the advantage to squeeze out the maximum possible performance within the best possible power envelope. The following sections provide details of the X-Gene2 feature set which are relevant for the UniServer project and the future roadmap for APMs X-Gene products.

2. X-Gene2 Feature Set: Salient points

This section describes the salient features of the X-Gene2 processor SoC from the ***UniServer perspective***.

2.1. X-Gene2 Processor Complex (PCP)

The X-Gene 2 Processor Complex consists of four Processor Modules (PMDs). Each PMD contains two high-performance X-Gene 2 cores, each of which has its own 32 KB L1 I-cache, 32 KB L1 D-cache, and Floating-Point Unit (FPU). The pair of X-Gene 2 cores in a PMD shares a 256 KB L2 cache unit which interfaces to Central Switch (CSW) interconnect. All four PMDs share an L3 cache (8 MB), which is also attached to the CSW. The X-Gene 2 Processor Complex features include:

- Eight X-Gene 2 cores operating at up to 2.4 GHz; each core contains
  - an FPU / SIMD Unit,
  - 32 KB L1 Data Cache 8-way set-associative, write-through to L2 cache, parity-protected,
  - 32 KB L1 Instruction Cache 8-way set-associative, parity-protected.
- 256 KB L2 cache per pair of processors inclusive of L1 write-through data caches, 32-way set associative, ECC-protected. Since L3 is essentially a victim cache
- Shared 8 MB L3 cache (attached to CSW) protected with SECDED ECC
- Hardware cache coherency
- Cache snooping and invalidation of I/O accesses

2.2. ARM GICv2m features

The GIC (ARM standard GICv2m) interrupt controllers are composed of three sets of parts:

- Interrupt Distributor
- Per-processor Physical CPU Interfaces
- Per-processor Virtual CPU Interfaces
- 16 Software-Generated Interrupts (SGI)

GICv2m is ARM specification that augments GICv2 functionality to enable it to work with MSI/MSI-X based interrupts. This layer includes control registers that allow specifying mapping of MSI/MSI-X to SPIs of GICv2.

2.3. SMMU

The key features include translation support for Stage 1 v7 VMSA/v8 AArch32, AArch64 with 4 k and 64 k granules and Stage 2 v8 AArch32/AArch64 4 k/64 k granules.

2.4. Four DDR3 SDRAM Memory Controllers

The features of each of the DDR3 SDRAM memory controllers are as follows:

- Supports up to DDR3-1866 (933 MHz DDR interface @ 1.5 V, 1.35 V)
- 64-bit DRAM interface (72 bits with ECC); 8-bit ECC calculated across 64 bits of data
- Single channel support per memory controller
- Support for eight Ranks / Chip Selects and four output clocks per memory controller, using x4, x8 memory modules
- Support for maximum of 64 independent Banks of memory (8 Banks per Rank, 8 Ranks Total) • Supports up to 128 GB of DRAM memory per controller
- Support for Self-refresh / Auto-refresh / Auto-power-down on a per-rank basis
- Utilizes 2T signal timing (one command every two DDR cycles; address and command held for two cycles)

2.5. SLIMpro

SLIMpro™ is AppliedMicro's Scalable Light-weight Intelligent Management processor. Features include:

- 32-bit low-power, 3-stage pipeline ARM Cortex-M3 processor
- Operates at up to 250 MHz • Implemented in Standby power plane (separate from CPU complex)

It is capable of:

- Wake on Interrupt
- Secure software upgrades

2.6. PMPro

The APM883408X2 has a dedicated 32-bit low-power processor called PMpro. The PMpro processor provides the following advanced power management capabilities:

- Multiple power planes and clock gating
- Thermal Protection Circuit
- ACPI standard power management states
- Active power management
- External power throttling support

Table 2 below summarizes the feature requirements for UniServer (as indicated in previous sections) and indicates all those supported by X-Gene2.

Table 2: Comparison of UniServer requirements and corresponding X-Gene capabilities

UniServer Feature Requirement	X-Gene2 Support
Multi-vendor CPU ISA	Supports ARM v8 ISA which is openly licensed to multiple vendors by ARM
Powerful CPUs	2.4GHz operation. Highest in class
Highly Integrated SOCs	All required IOs integrated making it a truly Server On Chip™ product
RAS	End to end support for RAS. This includes robust Error handling and reporting
Power Management and Control	Supports ARM specified ACPI states.
Telemetry	Yes
Scalable	Yes. Has PCIe, 10G Ethernet and SATA interfaces to enable scalability.
Performance monitoring	Detailed PMU unit
Support for supply-voltage sense pins	All power rails support Vsense pins.
Fine-grain control of the various VRMs supplying power to the different power rails	Yes. This detailed access information will be provided in the Hardware Exposure Interface (HEI) Specification as part of the D4.1 deliverable.
Fine-grain control of the Operating Frequency of the processors	Yes
Fine-grain control of the DRAM refresh rates	Yes.
On-die temperature measurement	Yes
Robust Error Reporting and handling capabilities	Yes. This is the same feature as RAS stated above.

2.7. Additional information

Given the confidential nature of many of the above features the details are being discussed in internal documents available to the UniServer consortium. The Deliverable D4.1 is going to include many more details on the Hardware Exposure Interface (HEI) and on the Error Reporting and Handling. D4.1 will also describe the software application interface to access the hardware states and various error values. In addition, the details on the Linux-firmware framework application programming interface which is defined to get alert and error notification from X-Gene processor to monitoring application will be described in D4.1.

X-Gene processor provides a Register Map with set of registers for the application to query various parameters in the system, override the default parameter values and to command the system to perform particular function. It can query different thermal sensors, VRD telemetry data such as VRD output power, VRD temperature, or sensors/system event configuration as supported by the X-Gene software.

X-Gene processor and software framework will notify of critical or catastrophic errors system by triggering an alert. In brief, the list of errors/events that would trigger an alert is given below:

- System Events such as platform booting/reset/etc.
- Thermal and Power events
- PMD/CPU Errors
- Memory Errors
- PCIe Errors
- SATA Errors
- Other I/O Errors
- ACPI state change

3. TIGERSHARK X-Gene2 – The Chassis of UniServer

This section, describes in more detail the TIGERSHARK design which consists the basic chassis of the UniServer platform. TIGERSHARK is a Shadowcat X-Gene2-based design which was designed for internal validation and evaluation and is not commercially available, thus providing to UniServer consortium a unique board equipped with many advanced features. Shadowcat is a member of the Applied Micro X-Gene family of high performance server class SoC SoC inheriting the basic features described above. Shadowcat provides a high-performance processor with multiple interfaces to a wide range of peripherals for key server markets.

3.1. SHADOWCAT X-Gene2 overview

Figure 2 below shows the X-Gene2 block diagram, where all the individual essential components and the supported functionalities can be distinguished.

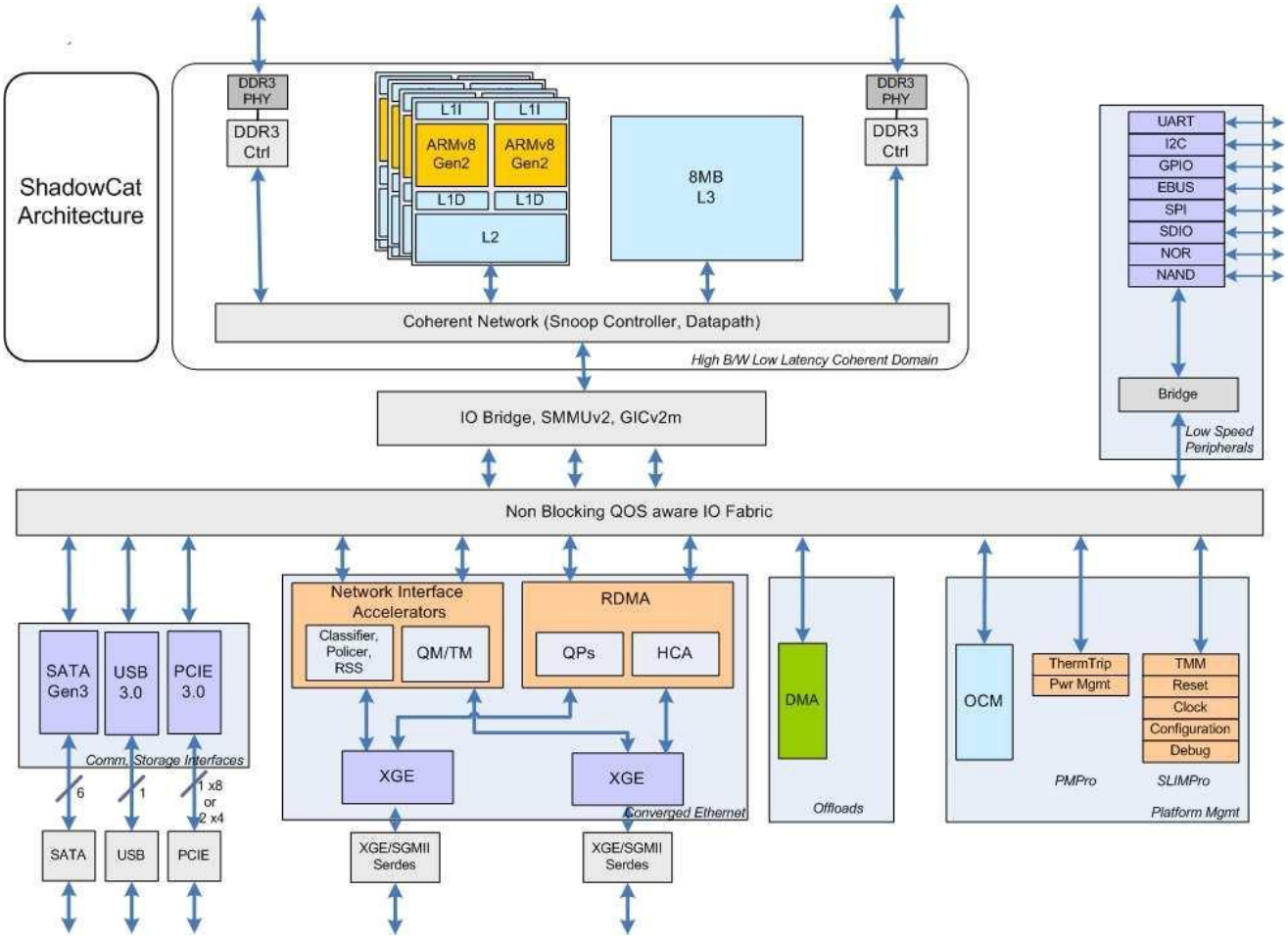

Figure 2: SHADOWCAT X-Gene2 Block Diagram

3.2. SHADOWCAT X-Gene2 feature brief description

The SHADOWCAT X-Gene2 board consists of the following:

1. Eight X-Gene™ processor cores operating at up to 2.4GHz ARM™ v8 compliant 64-bit cores

2. Floating Point (FPU)/Single Instruction Multiple Data (SIMD) Unit per core
    a. 32KB L1 Data cache per core
    b. 32KB L1 Instruction cache per core
    c. L1 Caches protected by parity (and write through to L2 cache)

3. Shared 256KB L2 cache per each pair of processors
    a. ECC protection on L2 caches

4. Shared 8MB L3 cache protected by parity

5. Hardware Cache Coherency. The coherence protocol is based on a standard MOESI protocol with an additional Forward state for clean copies

6. Four DDR3 memory controllers with SEC/DED ECC (72-bit)

7. Coherence interconnect across processor core, memory subsystems and bridge to I/O Fabric (IOF)

8. QoS capable High-Performance I/O Fabric (IOF)

9. Queue Manager / Traffic Manager
    a. Message passing architecture
    b. Manages Work, Free, and Virtual Queues

10. System Memory Management Unit (SMMU)
    a. Key components are Translation Buffer Unit (TBU), Translation Control Unit (TCU), and TBU-TCU Interconnect
    b. Key features include Soft error detection and recovery for the RAMs used in TBUs and TCU and Distributed translation support with up to 32 TBU's

11. High-Speed Networking Interfaces
    a. Two 10-Gbps Ethernet XGE MACs (XFI/SGMII) with in-line classification capability along with RSS and virtualization. Manages Work, Free, and Virtual Queues
    b. SFP+ Module (Limiting Mode) supported as defined in the SFF8431 specification. (SFP+Cu not supported)

12. PCI Express® Gen 3 ports with internal DMA
    a. 1x 8-lane or 2x 4-lane
    b. Hot Plug Support

13. High Speed Interfaces
    a. Up to two PCI Express® Gen 3 ports w/internal DMA, Hot Plug Support
        - 2x 4-lane
        - 1x 8-lane
    b. One USB 2.0 Hosts with integrated PHYs
    c. Six SATA Gen 3 ports

14. Other Interfaces
    a. I2C
    b. UARTs
    c. GPIOs
    d. SPI
    e. SDIO 3.0
    f. Debug support through JTAG and ETMv4 high speed serial interface (HSSTP)

g. NAND Flash Controller with ECC

15. Off-load Features
    a. True Random Number Generator (TRNG)
    b. Packet DMA Engine with RAID 5 offload

16. Power Management
    a. The PMPro provides support for multiple Advanced Configuration and Power Interface (ACPI) states. It features ACPI system registers which provide an interface to the power management software running on the PMDs and to ACPI software running on the PMPro.
    b. The PMPro also provides the Temperature Sensor Interface which provides access to on-die temperature sensors to accurately measure the on-die junction temperature.
    c. In addition, the PMPro, also provides 4 PWM outputs which can also be used for controlling the Fan speeds.
    d. The X-Gene2 also has multiple power planes and clock gating on a per block basis.

17. RAS: The X-Gene 2 processor supports Reliability, Availability, and Serviceability (RAS) with the following features:
    a. Reliability: Single-error correction and double-error detection (SEC/DED) in DDR memory, data-error poisoning, and error-reporting at the point of detection (separate from any immediate or later attempted uses of the data).
    b. Availability: End-to-end error propagation (including data poisoning), error exceptions only by consumers of corrupt data, and Gigabit and 10-Gigabit Ethernet ports.
    c. Serviceability: On-chip programmable platform controllers that can provide independent remote-management capability.

3.3. TIGERSHARK board features

The following sections describe TIGERSHARK rev. C board features, which further enhance the features of the SHADOWCAT designs mentioned above, allowing essentially more detailed validation experiments.

3.3.1. Physical

TIGERSHARK board is designed to fit into standard ATX chassis and made compatible to ATX form-factor: PCI Express expansion connector and mounting holes match ATX compatible chassis. Also, connectors are located in dedicated ATX I/O connector area.

3.3.2. TIGERSHARK supported features

- **Processor**
  - ARM 64bit X-Gene2 Processor
  - Speed: up to 2.4GHz.

- **PCIE Gen 3**
  - One PCIE x8 connector on board

- **Serial ATA-III**
  - Six ports of SATAIII on board

- **USB 2.0**
  - One USB2.0 host port at rear I/O connector area

- **DDR-III Memory**
  - 4 DDR3 Channels, with two DIMMs per channel (total 8 DIMMs on-board)
  - Speed: Support up to 1866MT for single DIMM per channel/ 1600MT for two DIMMs

- **Embedded Local Bus (Legacy Parallel Bus)**
  - 512 Mbit Parallel NOR Flash (can support legacy boot mode)

- **UART**
  - Four UART ports with transceivers
- **SDIO**
  - Two SDIO cards
- **GPIO and Pinstrap**
  - GPIO mux for pinstrap signals (after detecting pinstrap configurations GPIOs are released for other system function)
  - Pinstrap headers to set boot features like PCP clock configuration, Boot ROM type, device slave address
  - GPIO Header
- **Boot Options**
  - Bootstrap EEPROM
- **SPI and I2C**
  - EEPROM on SPI0 bus
  - SPI NOR Flash on SPI2 bus is main boot device
  - Five I2C buses used for boot
- **Ethernet**
  - Custom adapter slot to support SGMII or XFI mode on High-Speed Ethernet Interfaces. Adapter cards provide required reference clocks and miscellaneous signal support
- **Clocks**
  - CPU System Reference Clocks and all other (PCIe/SATA/USB reference clocks) from on-board clock generators
- **Power**
  - Board powered from standard ATX power supply connectors
- **JTAG and Trace Support**
  - Dedicated JTAG headers for X-Gene cores and SoC

3.4. TIGERSHARK block diagram

Figure 3 shows the TIGERSHARK Board Block diagram. In the following sections the various components on board are briefly described.

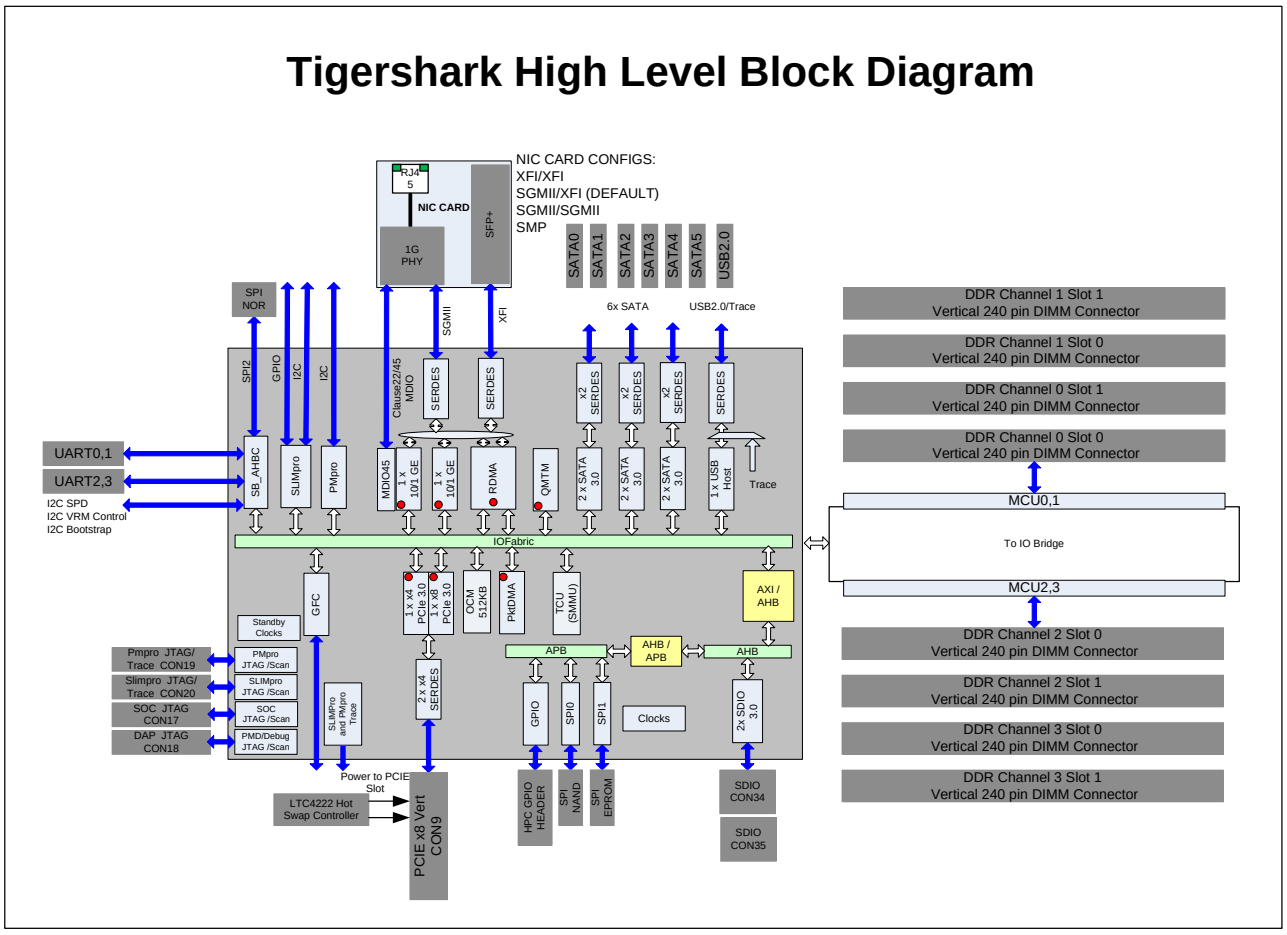

Figure 3: TIGERSHARK Block Diagram, UniServer's Chassis

3.5. Major building blocks

In this section, all X-Gene2's major functional blocks are described in detail.

3.5.1. Power management

The TIGERSHARK support three power domains: Standby, SoC and PCP.

3.5.2. SATA & PCIe

TIGERSHARK has six SATA-III Port (6Gbps/port) and backward compatible to SATA1 and SATA2. SATA interfaces are directly routed to the SATA connectors on the mainboard.

The PCIE is routed to a x8 PCIe Gen3 connector. This support PCIE Gen3 and backward compatible to Gen2 and Gen1.

3.5.3. DDR-III

TIGERSHARK supports four channel of DDR3 memory. As depicted on Figure 4, on X-Gene2 board, there are 2 DIMM per channel configuration with total of 8 DIMMs supported for 64-bit + 8 Bit ECC with data transfer rate up to 1866MT.

Each channel of DDR design will have two DIMM connector. The Slot0 is at the far end and the Slot1 is the near the end connector.

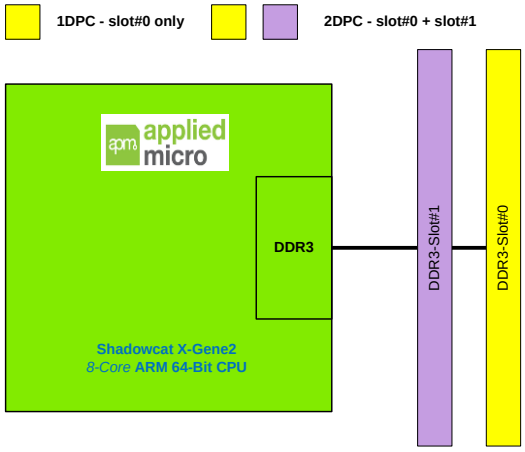

Figure 4: TIGERSHARK DDR3 Topology

3.5.4. X-Gene2 clock settings and strapping values

TIGERSHARK supports both pinstrap and bootstrap configuration, the details of which are only available to the consortium.

3.5.5. Ethernet(SGMII/XFI) interfaces

Two 10GE/1GE controllers and serdes are included in the X-Gene2 SoC. In order to facilitate the breakout to enable evaluating both XFI and SGMII interfaces on the board, a breakout board daughter card solution is chosen. There are three types of daughter card available to enable the network connection, namely DC with two SFP+ connectors, DC with one SFP+ and one SGMII and DC with two SGMII Ethernet.

3.5.6. USB

The USB interface on X-Gene2 SHADOWCAT is depicted on Figure 5. In particular, there is one USB2.0 port 0 on the board and the USB interface needs one 12MHz XTAL for USB2.0 IPs.

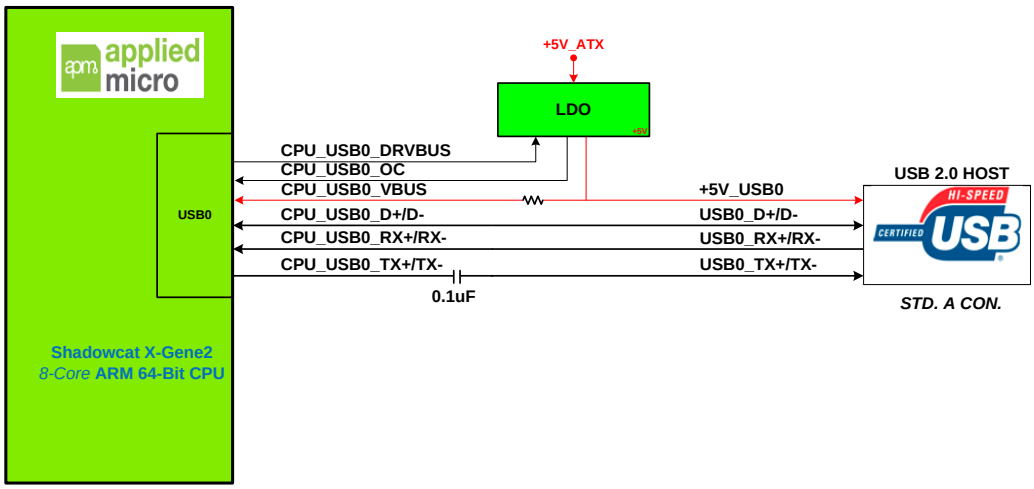

Figure 5: TIGERSHARK USB Host Interface

3.5.7. I2C bus

X-Gene2 needs to access various peripheral devices via I2C bus. Some of these devices are essential for proper system boot and must be allocated at specific I2C busses and addresses.

3.5.8. SPI interface

There are three SPI interfaces available for the X-Gene2 SOC. On TIGERSHARK, SPI2 is the boot interface. The flash provides 32MB memory space for firmware and boot code. The other two SPI interfaces, SPI0 and SPI1 are connected with SPI EEPROM and SPI NAND Flash respectively on the TIGERSHARK board.

Table 3: TIGERSHARK SPI devices

BUS NAME	FUNCTION
SPI0	SPI EEPROM
SPI1	SPI NAND Flash
SPI2	SPI NOR Flash (Boot Device)

3.5.9. UART

There are four UART interfaces on TIGERSHARK board. See UART connector identification in IO Connector Area section.

3.5.10. SDIO interface

The SDIO on X-Gene2 is connecting to two SD card connectors.

3.5.11. Buttons function

On TIGERSHARK board there are two Buttons:

Power button

- Single, push button power switch located on the board.
- When AC voltage provided to ATX power supply already, if this button is pressed, the system will power up.

Reset button

- Single, push button reset switch located on the board.
- Pressing reset will result in a system reset.

3.5.12. LEDs

TIGERSHARK provides up to twenty-eight indicator LEDs located on the board and nine Ethernet LEDs at two RJ45 connectors. These LEDs are used to indicate whether the various power rails have been turned ON or not. They are also used to indicate Boot Failure and SD Card status.

3.5.13. GPIO operation

X-Gene2 has 3 set of GPIOs:

- GPIOs belonging to Standby domain – GPIO_DS[21:0]
- GPIOs belonging to SoC GPIO_FL[47:0]
- GPIOs shared with other functions – GPIO[31:0]

3.5.14. Clocking scheme

X-Gene2 requires that each serdes be sourced with its own individual reference clock input. The clock inputs, namely PCIE, SATA, SGMII, USB2.0, SYSTEM, XFI reference clocks are CML differential voltage level.

3.5.15. Power requirements

TIGERSHARK mainboard can be powered with any 250W ATX compatible power supply. The system gets power from AC-DC ATX PC power supply, or Mini ITX power supply with 250W output capacity. Then generating other powers as: +1.5V, +1.8V, +2.5V etc. by using Buck converters, LDOs, and VRDs. The TIGERSHARK board uses two-regulator power supply configurations, one for SoC rail and one for PMD rail.

4. X-Gene Future Roadmap

As APM has already established its market leadership in the ARM v8 server space, it endeavours to continue its leadership by enhancing the X-Gene2 capabilities and ultimately the UniServer platform. APM is already planning the next two generations of the X-Gene products. X-Gene3 is slated to be available within the current year i.e. 2016 when it will be also customized and provided to UniServer consortium for integrating the developed technologies. The X-Gene3 improves the processor performance by many folds and also will have a significantly higher frequency of operation. X-Gene3 is slated to be a true server class product and will rival many of existing non-ARM v8 based products. X-Gene3 will provide a true alternative solution to the existing server class processors and break the monopoly. This will eventually lead to a more competitive environment and thus cheaper servers.

X-Gene3 builds on two generations of enterprise class server on chip products that are currently shipping to leading edge enterprise and data center end customers.

The newest member of the X-Gene family, fabricated in TSMC's 16 nanometer FinFET process, will feature 32 brawny cores operating at speeds up to 3.0 GHz, eight DDR4-2667 memory channels and 42 PCIe Gen 3 lanes. Performance is expected to be four-to-six times that of the currently shipping X-Gene family of products. This will ultimately provide an enhanced chassis to UniServer on which the developed technologies will be integrated and demonstrated.

Given that the X-Gene3 will be fabricated in 16nm process technology, the overall leakage power is significantly less. This implies that the clock-gating of logic not in use is sufficient to achieve significant power savings unlike previous technologies where the leakage power was a big factor and power-gating was essential to achieve savings in power.

At a base speed of 3.0GHz, X-Gene3 (code-named Skylark) will score about 550 SPECint_rate2006, according to the APM's estimates. X-Gene3's total performance is well within the range of Xeon E5 products. On this metric, it's similar to the fastest parts typically used in cloud servers and trails models such as the Xeon E5-2699v4. In addition to rip-roaring per-socket performance, X-Gene3 will have respectable per-thread performance. While most other ARM server-processor vendors are using smartphone-class CPUs, APM designed its own high-performance quad-issue microarchitecture backed with robust cache and memory subsystems. X-Gene3's biggest advantage will be on memory-intensive applications such as web search, big data processing, machine learning, and certain high-performance-computing (HPC) applications. These workloads require high per-thread and per-socket performance coupled with a large memory subsystem and are representative of applications that will benefit from enabling computing at the edge of the cloud as UniServer technologies aim at.

5. References

[HEsm.ISCA11] H. Esmaeilzadeh et al. "Dark silicon and the end of multicore scaling." In Proc. 38th Annual International Symposium on Computer Architecture (ISCA), 2011, IEEE.

[ITRS13] ITRS 2013. Online. http://www.itrs.net//2013ITRS/ 2013TableSummaries

[JKoo.AHC11] J. G. Koomey et al. "Implications of historical trends in the electrical efficiency of computing." Annals of the History of Computing, IEEE 33.3, 2011.

[KBow.JSSC11] K. Bowman et al., “A 45 nm resilient microprocessor core for dynamic variation tolerance,” IEEE J. Solid-State Circuits, vol. 46, no. 1, Jan 2011.

[FORRESTER14] “The Public Cloud Market Is Now In Hypergrowth: Sizing the Public Cloud Market, 2014 To 2020”, Forrester Research, April 24, 2014.

[ABar.TPDS13] A. Bartolini et al. “Thermal and energy management of high-performance multicores: Distributed and self-calibrating model-predictive controller,” IEEE Trans. Parallel Distrib. Syst., vol. 24, no. 1, Jan 2013.

[END OF DOCUMENT]