BLAZAR BE3-BURST Accelerator Engine IC
Intelligent In-Memory Computing
1Gb Memory

PRODUCT BRIEF: MSR630

Acceleration Engines give Software and Hardware System Architects Acceleration Options not previously available

BANDWIDTH ENGINE (BE) INTRODUCTION

The **BLAZAR Family of Accelerator Engines** are high capacity, high-speed memories that support high bandwidth, fast random memory access rates and includes optional *embedded **I**n-**M**emory **F**unctions (**IMF**)* that solve critical memory access challenges for memory bottlenecked applications like network search, statistics, buffering, security, firewall, 8k video, anomaly detect, genomics, ML random forest of trees, graph walking, traffic monitoring, AI and IoT.

All Accelerator Engines have two ways to be uses in a system:
1. As a standard parallel like QDR, SyncSRAM or RLDram
2. As a high density, high bandwidth memory with Optional In-Memory Acceleration Functions.

Both modes are independent. You do not have to use the In-Memory functions if they are not useful in your application. In that case, you would use the Accelerator Engine like any other memory.

BASE FEATURES

The Bandwidth Engine 3 BURST (BE3-BURST)

- 1Gb of tRC of 2.7 ns memory
  - Replaces 8 QDR type memories
- In-Memory BURST functions
- RTL Memory Controller

APPLICATIONS FOCUS

- Slower speed applications needing high capacity
- SRAM with high capacity and high speed
- High bandwidth data access application where low latency and Movement of Data is a critical
- FPGA Acceleration for Xilinx and Intel

KEY FEATURES - Memory

- 1 Gb SRAM (16M x 72b)
  - User defined WORD width
  - Typical 8x, 16x, 32x, 36x, … 72x
- High Bandwidth, low pin count serial interface
  - Highly efficient reliable transport command and data protocol optimized for 90% efficiency
  - Eases board layout and signal integrity, minimal trace length matching required, operates over connectors
  - Reduction of I/O pins from 5x to 45x depending on equivalent memory density and type
- High access rate SRAM class memory
  - Up to 6.5 Billion transactions/sec
  - 2.7 ns tRC
- Highest Single Chip Bandwidth – up to 640 Gb/s throughput (320 full duplex)

KEY FEATURES – In-Memory Functions

The Acceleration function are optional and do not impact the device used as a memory only.

BURST In-Memory Function
- For sequential Read or Write functions for DATA MOVEMENT
- Burst length: 1, 2, 4 or 8 words
- Can double or triple QDR bandwidth

MoSys ACCELERATOR ENGINE Elements of BE3-BURST

MoSys Engines have a Unique Memory Architecture that can replace SyncRAM/RLDRAM memories and Embeds In Memory Functions (IMF) that execute many times faster. A single function replaces many traditional memory accesses.

Bandwidth Engine Device Key Features

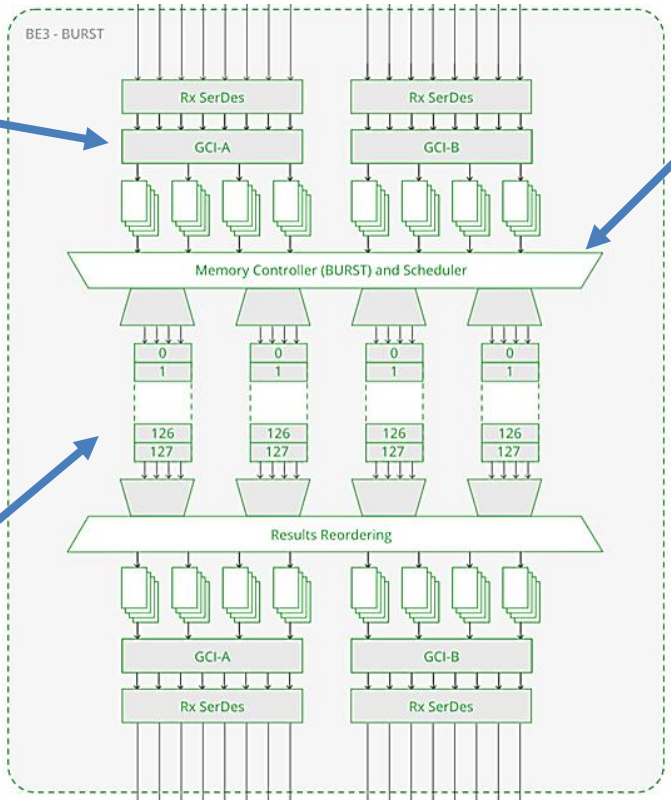
BE3-BURST

High speed serial I/O

- GCI serial I/O versions of 12.5, 15 and 25 Gbps for high bandwidth (up to 640 Gbps)
- Device can operate with a minimum of 4 lanes.
- Has two, full duplex 8 lane ports that operate independently
- Reduces number of signal pins over traditional memories, increases signal integrity allowing longer board traces to ease board signal routing
- Operates across connectors

Main Memory

- 1Gb (BE2 has 576Mb))
  - 4 partitions/128 banks
  - 16 READ & 16 WRITE ports
- 2.7 ns tRC
- Allows parallel partition & Bank execution



Memory/Function Controller

- Directs function execution to selected bank of memory
- Manages all random access read write
- Manages the sequence of In-Memory functions
- **BURST** – Sequential read or writes
  - Up to 8 read or writes
- Controls simultaneous memory access to partitions and banks

Common Key Features for BANDWIDTH Engines (BE2 & BE3)

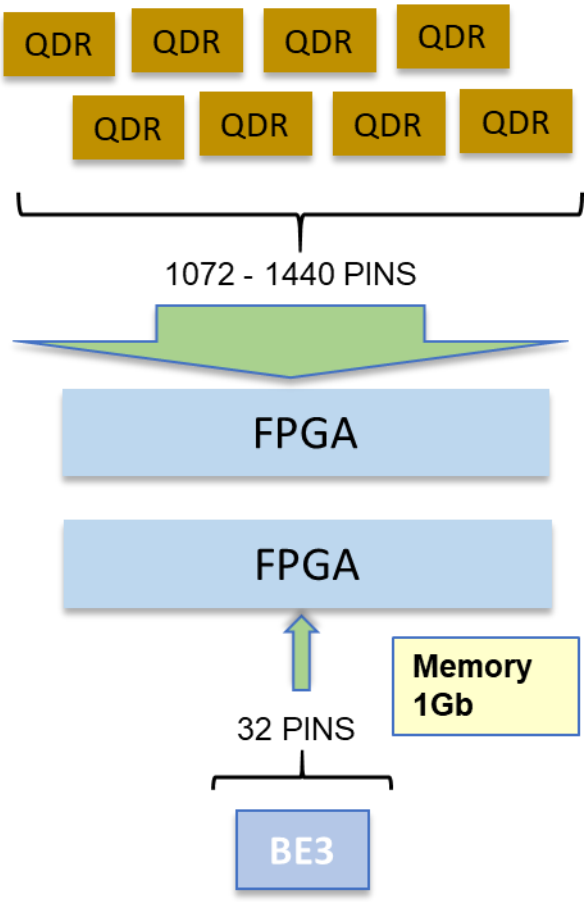
- **High capacity High-Speed Memory on a single device**
  - BE2 576Mb
  - BE3 1Gb
- **High speed tRC access**
  - BE2 with 3.2ns
  - BE3 with 2.7ns
- **Achieve the highest bandwidth possible**
  - Use of serial interface with high efficiency GCI protocol
  - BE2 with 3.3B access per sec
  - BE3 with 6.5B accesses per sec
- **Reduce number of interface pins compared to other memories**
  - Typical system uses 8 lanes or 32 pins
  - Highest bandwidth uses 16 lanes or 64 pins
  - Minimum use of 4 lanes on one port or 16 pins
- **Eliminate external components required for signal integrity**
  - On device Auto-Adaption to eliminate external signal conditioning board components
  - Signals will work over a backplane
- **Make the interface like a parallel QDR**
  - *MoSys supplies an RTL memory controller that handles the memory and serial interface*
  - *Serial device interface is transparent to user*
  - *Provides a parallel QDR like Read/Write RTL interface*
- **Make interface easily adapted to an AXI or Avalon bus**
  - Minimal RTL logic required

Benefits of BE3 vs QDR

SUMMARY OF BENEFITS

- ***Capacity … 1Gb memory … Replaces 8 QDR/SyncSRAM devices***
- ***Costs … One BE-3 is approx. the price of 3 QDR memories with 8x the memory***
- ***Pins … Typical application uses only 16 signals (32 pins) with signal Auto-Adaptation***

More High-Speed memory generally allows acceleration options for software and hardware architects/Designers

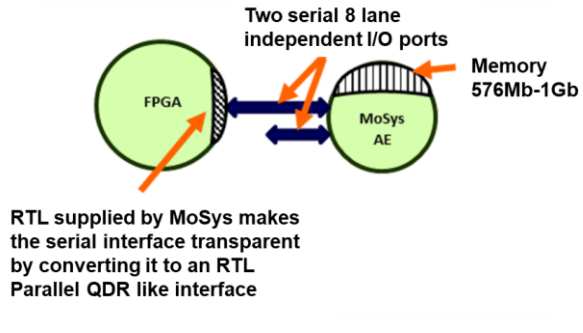


OVERVIEW COMPARISION BE vs. QDR

- Memory size
  - BE3 with 1Gb equivalent to 8 QDRs with 144Mb per device
- Device PCB board Space Saving
  - 1 BE3 device vs 8 QDR devices
- Signal Pins Reductions
  - 8 QDR…1Gb requires 1072-1440 pins
  - ***1 BE3 …1Gb…BE3 typical system uses 8 lanes or 32 pins.***
  - All BE devices have *Auto-Adaptation which handles on-board signal tuning*, eliminating the need for any external components to insure a clean, reliable signals
- ***Costs***
  - ***One BE-3 with 8x the memory capacity is approx. the price of 3 QDR memories***
- Applications Benefits
  - Larger Buffers, High Bandwidth
  - Allows Realtime operations and analysis at Line rate
  - Eliminates need for complex parallel operations using RLDRAM, HBM, or slow DRAM

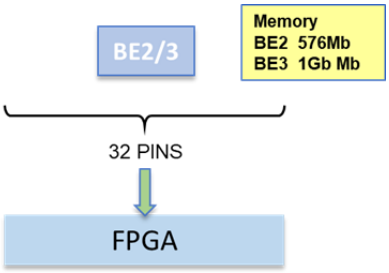
Simplifying the User Interface to BE2/3 with MoSys RTL Controller

INTERFACE PORTS



- Device has two, 8 lane independent ports
- Typical system uses one port (shown), 8 lanes or 32 pins.
- Can use as few as 4 lanes on one port or 16 pins
- High bandwidth systems use both ports, 16 lanes, 64 pins
- Independent ports can operate as a dual port with simultaneous access between two FPGAs.

SERIAL INTERFACE

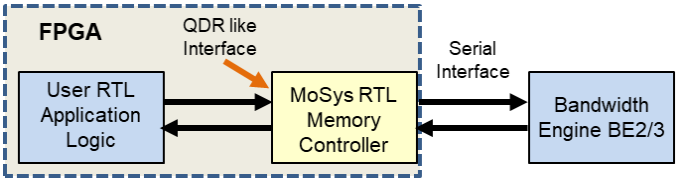


- Benefits of a serial interface allows high bandwidth over very few pins
- Key to bandwidth is the MoSys GCI interface that is transparent to the user with the MoSys supplied FPGA RTL Memory controller

MOSYS FPGA Parallel RTL INTERFACE

MoSys suppled RTL Controller simplifies the user interface with the BE.

MoSys supplied FPGA RTL Memory Controller interfaces with the MoSys Bandwidth Engine. This controller is between the User Application RTL logic and the BE device.

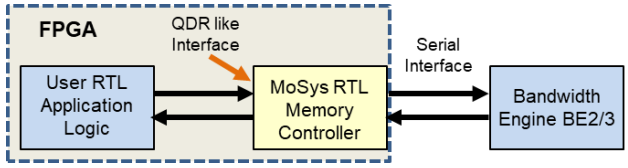


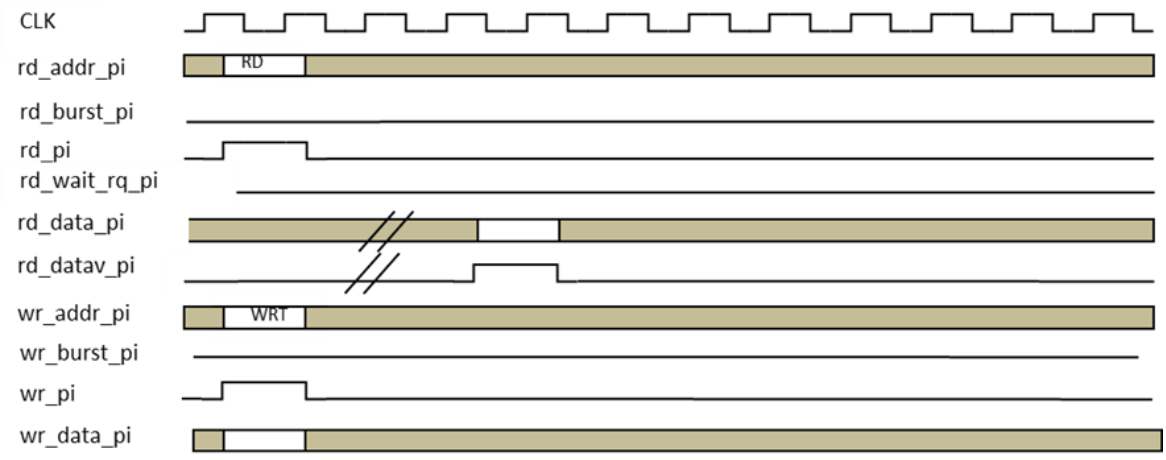
MoSys suppled RTL Controller simplifies the user interface with the BE.

- It handles all the logic for the Serial GigaChip Interface (GCI) in the FPGA
- Eliminates the user having to design a serial interface by making it transparent and providing a QDR parallel like interface.
- Memory WORD width is user definable in RTL
  - Typical word widths are 8, 16, 32, 36, 64 …
- While the memory on the BE2 is organized as 8Mx72b and the BE3 is 16Mx72b, the address conversion mapping from the selected WORD width to the BE memory is handled by the RTL.
  - Address translation to BE memory organization is transparent to the application
- All memory addressing and commands are presented to the QDR like parallel interface.
- If the optional In-Memory functions are use, the RTL controller will mange their execution

The signal interface at the User Application is a simple SRAM memory Address, Data, Control structure. This simple interface shields the users from the BE commands, serial interface and the scheduling logic and memory partition timing.

High Speed GCI Serial Interface





SIGNAL NAME	WIDTH	DIR	DESCRIPTION
Read Interface			
rd_p	1	In	Assertion of this signal indicates that this is a read transaction.
rd_addr_p	32	In	Read address. Please refer to the Address section of this specification to see the detail of this address field.
rd_partsel_p	1	In	Indicates the BE-2 partition that this read command will be operated upon: 0 = Partition 0 for GCI port A, Partition 1 for GCI port B 1 = Partition 2 for GCI port A, Partition 3 for GCI port B
rd_data_p0	*	Out	Returned data from BE-2 memory. This data is qualified by the “rd_datav_p0” signal
rd_data_p1	*	Out	Returned data from BE-2 memory. This data is qualified by the “rd_datav_p1” signal. Note that rd_data_p1 will only have valid data if rd_data_p0 is valid as well. rd
rd_datav_p0	1	Out	The Memory Controller asserts this signal to indicate the current data in the “rd_data_p0” bus is valid
rd_datav_p1	1	Out	The Memory Controller asserts this signal to indicate the current data in the “rd_data_p1” bus is valid. Note that rd_data_p1 will only have valid data if rd_data_p0 is valid as well
rd_wait_rq_p	1	Out	The Memory controller asserts “rd_wait_rq_p” to indicate that it cannot accept the current read request from user. The User Application should hold all the request signals (rd_p, rd_addr_p …) until the de-assertion of this signal.

SIGNAL NAME	WIDTH	DIR	DESCRIPTION
Write Interface			
wr_p	1	In	Assertion of this signal indicates that this is a write transaction.
wr_addr_p	32	In	Write address of the memory for this transaction. Please refer to the Address section of this specification to see the detail of this address field.
wr_partsel_p	1	In	Indicates the BE-2 partition that this write command will be operated upon: 0=Partition 0 for GCI port A, Partition 1 for GCI port B 1=Partition 2 for GCI port A, Partition 3 for GCI port B
wr_data_p	*	In	Write data from the User Application logic.
wr_wait_rq_p	1	Out	The Memory controller asserts “wr_wait_rq_p” to indicate that it cannot accept the current write request. The User Application should hold all the request signals (wr_p, wr_addr_p ...) until the de-assertion of this signal.

Accelerator Engine Family Overview

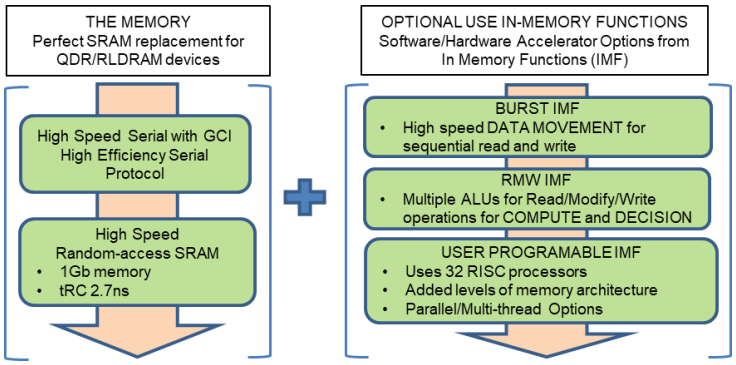
Software Defined - Hardware Accelerated

In-Memory	Part Number	Description	Package	Interface						Memory		Access Rate	In-Memory Functions		
			Pkg Size	Lanes	Rate per Lane Gb/s				BW MAX.	tRC	Size	Billion Transaction/s	BURST for Data Movement	RMW / ALU for Compute and Decision	Custom & User Functions with 32 RISC
			mm	Tx/Rx	10.3	12.5	15.6	25	Gb	ns	Gb				
BURST	MSR622	Bandwidth Engine 2 Burst Serial 0.5Gb High Access Memory	FCBGA 19x19	16	✓	✓			320	3.2	0.5	3.3	✓		
BURST	MSR630	Bandwidth Engine 3 Burst Serial 1Gb High Access Memory	FCBGA 27x27	16		✓	✓	✓	640	2.7	1	6.5	✓		
RMW	MSR820	Bandwidth Engine 2 RMW Serial 0.5Gb High Access Memory with ALU for RMW functions	FCBGA 19x19	16	✓	✓			320	3.2	0.5	3.3	✓	✓	
RMW	MSR830	Bandwidth Engine 3 RMW Serial 1Gb High Access Memory with ALU for RMW functions	FCBGA 27x27	16		✓	✓	✓	640	2.7	1	6.5	✓	✓	
Program	MSPS30	Programmable HyperSpeed Engine Serial Interface, 1Gb Memory, 32 RISC Processor cores for custom algorithms, compute, functions	FCBGA 27x27	16		✓	✓	✓	717	2.7	1	24 Internal	✓	✓	✓
RTL	RTL-AE	RTL Memory Controller for Bandwidth Engine and Programmable HyperSpeed Engine. Manages memory and the serial interface signals. Presents a QDR like parallel RTL interface to the user.	FPGA RTL Code		✓	✓	✓	✓			576Mb & 1Gb	6.5	✓	✓	

BW MAX. = At the highest serial interface speed

CONTACT MOSYS TO LEARN ABOUT THESE ADVANCED FEATURES

IN-MEMROY ACCELERATION FUNCTIONS



The Acceleration function are optional and do not impact the device used as a memory only.

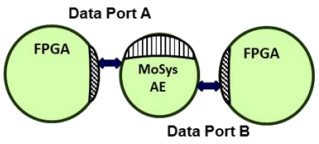
BURST In-Memory Function

- For sequential Read or Write functions for DATA MOVEMENT
- Burst length: 1, 2, 4 or 8 words
- Can double or triple QDR bandwidth
- Simultaneous execution of read and writes

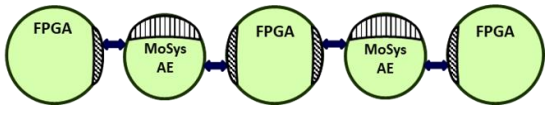
RMW In-Memory Function

- RMW are Read/Modify/Write functions
- Includes may functions for COMPUTE and DECISION
- Examples: ADD, SUB, Compare, INC plus 15 other functions
- Increases execution speed and bandwidth

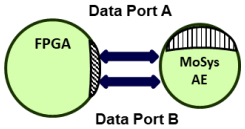
FLEXIBLE CONFIGURATIONS USE



DUAL PORT



PIPELINE



HIGH BANDWIDTH

www.mosys.com https://mosys.com/products/blazar-family-of-accelerator-engines/

2309 Bering Drive, San Jose, CA 95131
Tel: 408-418-7500 Fax: 408-418-7501
www.mosys.com