A Satellite-Based Architecture for High-Throughput Storage

Flash Memory Summit

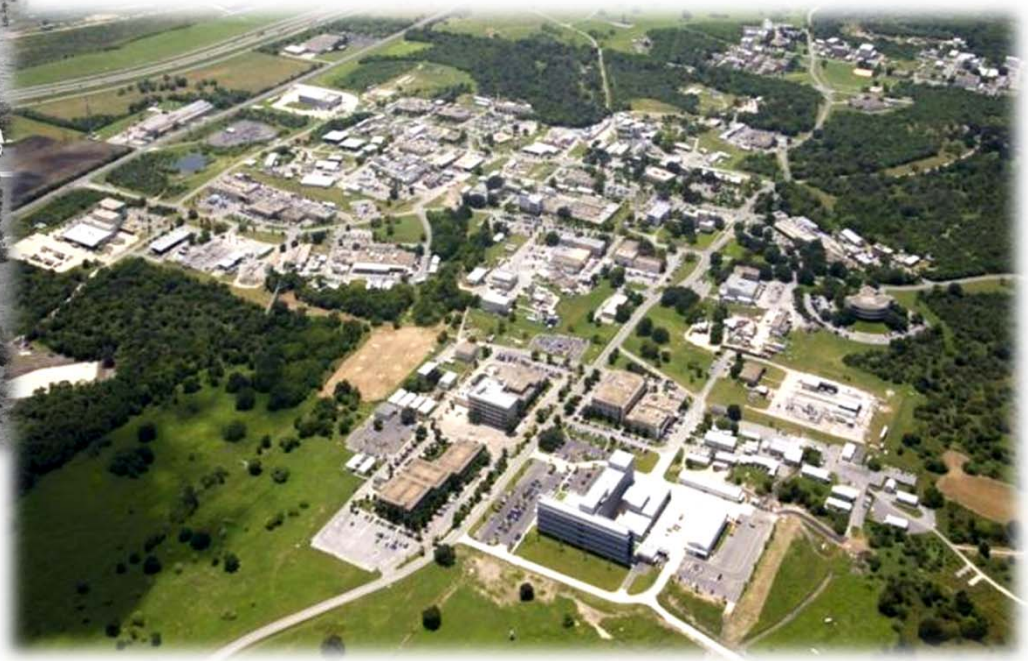
Southwest Research Institute®

August 8, 2018

Michael Koets

Southwest Research Institute

- Founded in 1947
- 20 Employees
- Budget of $100,000



- Over 2,800 employees
- Over 1,000 patents
- 37 *R&D 100* awards
- Over 1,200 acres / 4.86 km^2 facility
- 2.2 million ft^2 / 204,400 m^2 of labs & offices

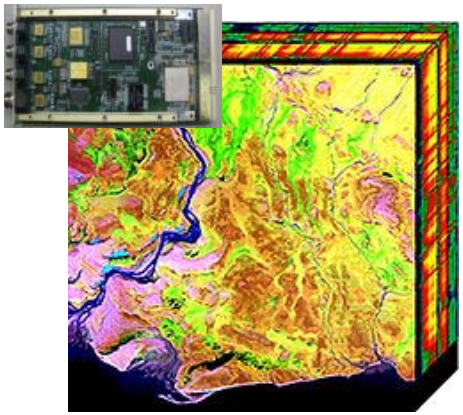
SwRI Operational Characteristics

- Nonprofit
- Unaffiliated with government lab or university
- Independent and unbiased
- Revenue from contracts
- Applied R&D services
- Physical sciences and engineering
- Broad technology base
- Extensive internal research program

Space Science and Engineering

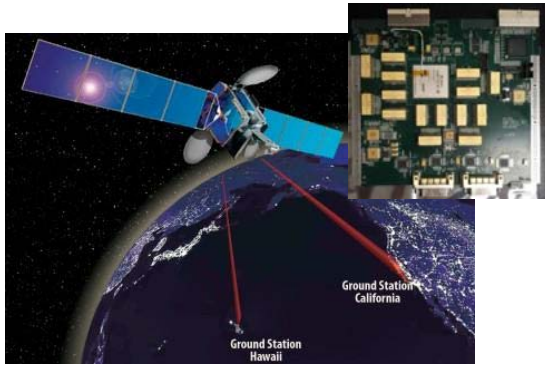
- Payloads and Instruments
  - Science and technology driven development
  - Technical expertise over a broad range of application areas
  - Expertise in payload integration and accommodation on spacecraft
- Spacecraft Avionics
  - Turnkey systems
  - Single board solutions
  - ASICs
  - Transceivers (K-Band, S-Band)
  - Signal of opportunity receivers
- Spacecraft Development and Integration
  - Advanced concepts for Small, Micro, Nano, and CubeSat class vehicles
  - Program Management
  - Systems Engineering
  - Parts and Reliability Engineering

Key Areas of SwRI Technical Innovation

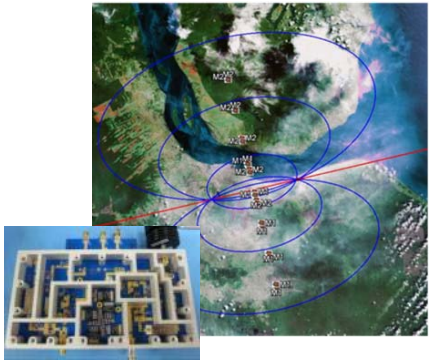


Hyperspectral Sensing

High-Reliability, Fault-Tolerant Spacecraft Avionics

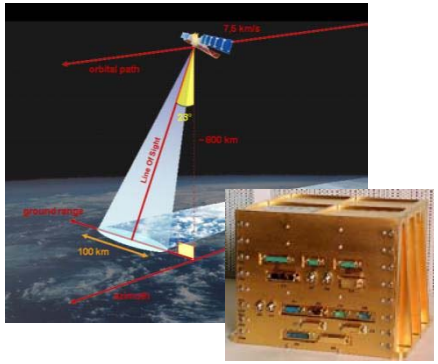


High Rate Communications (Optical and RF)



Advanced Signal and Image Processing

Low Cost, High Reliability, Responsive Space Avionics and Instrumentation



Synthetic Aperture Radar (SAR)

Demand for High Capacity, High Throughput Data Storage Systems for Space Applications

- Rapidly emerging need for high performance data recorders
  - Expanding capability of sensors and instruments
    - High resolution image sensors
    - Multispectral and hyperspectral sensors
    - Synthetic aperture radars
    - Scientific instruments
  - Continuing limitations on communications bandwidth
    - Limited opportunities for communication with low earth orbit spacecraft
    - Crowded spectral environment
- Typical Throughput Requirements:
  - 10's of Mbps to 10's of Gbps
- Typical Storage Capacity Requirements:
  - 100's of Gb to several Tb

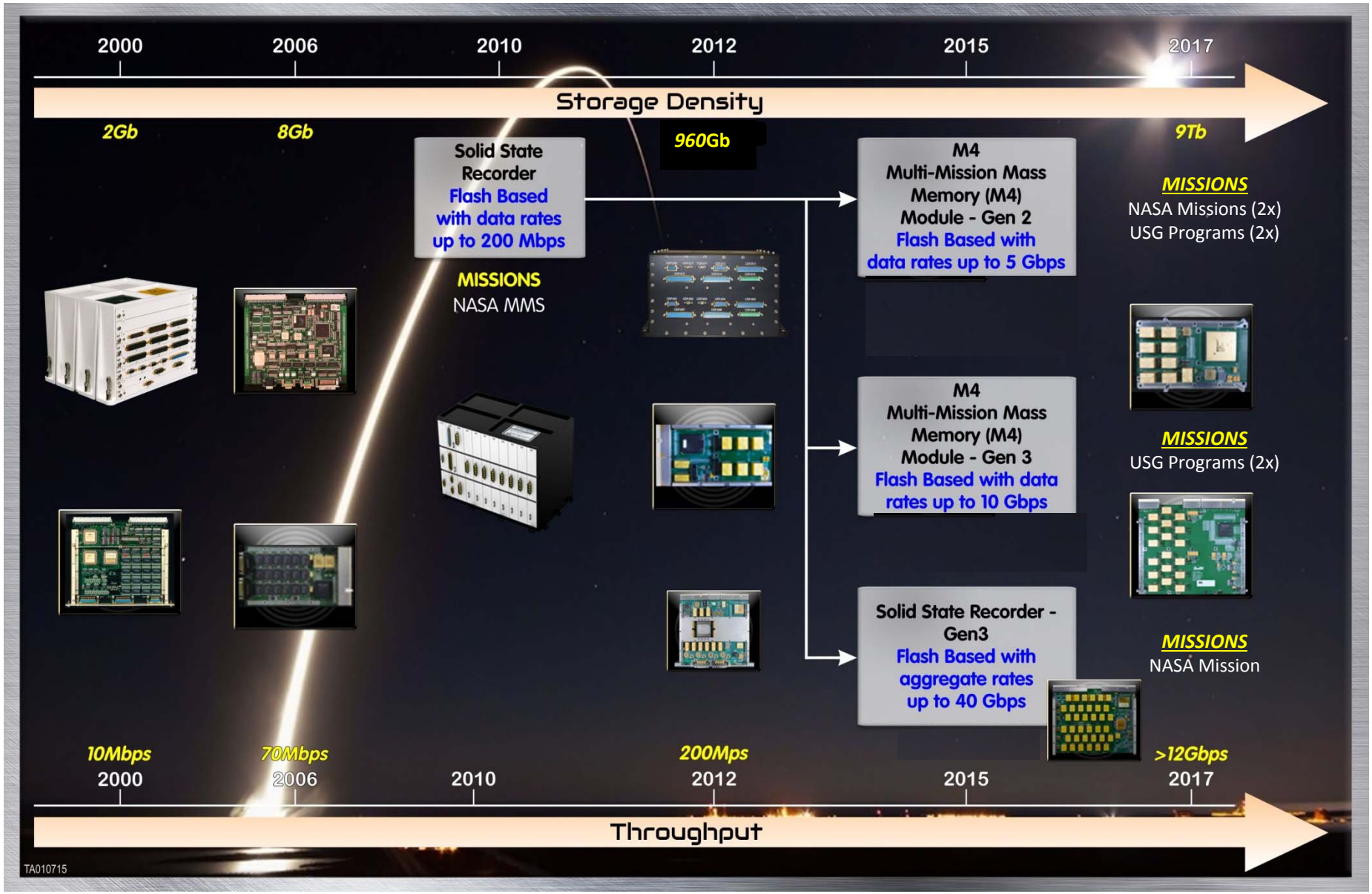
Driving Technical Characteristics

- Support for Multichannel Recording and Playback
  - Multiple active sensors
  - Storage of metadata and spacecraft management information
  - Multiple concurrent communications links
    - Different data sources for each link
- Data Rate Variation
  - Wide range of nominal data rates for different data flows
  - Continuously variable rates
- Simultaneous Recording and Playback
  - Concurrent read and write access to a single pool of storage
- Automatic Erasure of Stored Data
  - Just in time erasure of old data to make room for new
- Demand for small size and low power
  - At least for the aerospace industry

Example Missions and Applications

- NASA/ISRO Synthetic Aperture Radar (NISAR)
  - Multi-band synthetic aperture radar
  - Dedicated data channels for different polarizations
- NASA Joint Polar Satellite System 2
  - Imagery for weather forecasting
  - Multiple scientific instruments
- NASA Landsat-9
  - Earth imagery for resource management
  - Multiple scientific instruments
- Deep Space Missions
  - Long term storage of data between high rate data collection and low rate data transmission to Earth

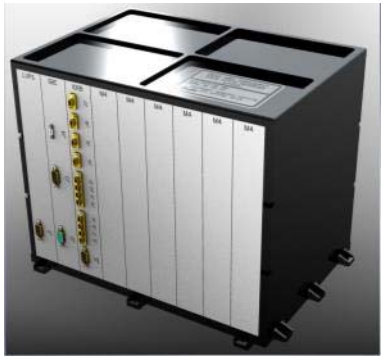
SwRI Solid State Recorders



SwRI SSR Technology

- SwRI SSRs provide a modular, scalable architecture for a wide range of input/output (I/O) rates and data storage capacities

SSR Generation	Total Density	Bandwidth
1st (2010-2013)	1Tb	Up to 200Mbps
2nd (2013-2015)	Up to 12Tb	Up to 20Gbps
3rd (2016 -)	Up to 28Tb	Up to 40Gbps



- Low size, mass, and power for ease of spacecraft integration
- High heritage design based directly on solid state recorder technology currently on orbit as part of the NASA MMS mission for reduced development risk
- Early operations planning through SwRI's Mass Memory ConOp Developer (MMCD)

Multi-Mission Mass Memory (M4) Module

- Gen 3 M4 module provides
  - Up to 3.5 Tb of user memory storage per board
    - Guaranteed density – spare memory ensures that bad blocks do not detract from total density
  - Ultra-fast storage and playback via Multi-Gbps transceivers using SwRI's patented approach:
    - Storage and playback at up to 10 Gbps
    - Simultaneous storage and playback at up to 5 Gbps
    - Excellent SEU performance
    - EDAC / ECC provided on top of 3.5Tb memory
  - Aggressive power management delivering low power dissipation for high throughput applications

Generation 1 M4 module

Generation 2 M4 module – 3U

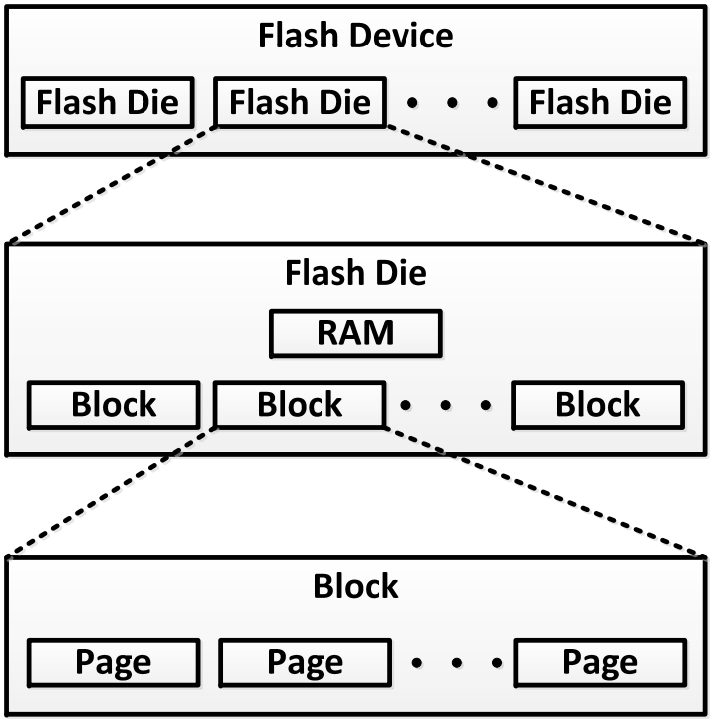
Generation 3 M4 module

Space-Grade Flash Memory Components

- 3DPlus Stacked Die
  - 4-8 independent flash components
  - 1-2 control/data busses
    - Data busses shared by 4-8 components
  - Opportunities for component level pipelining and parallelization
- Control via Open NAND Flash Interface (ONFI)
  - Granular operations
    - Page read/write
    - Block erase
  - Multi-plane operations
- Asynchronous Physical Interface
  - Limited data rates
  - Implications for FPGA interface logic

Flash Memory for Solid State Recorders

- Characteristics of Flash Memory
  - Nonvolatile
  - High density
  - Variability in access times
    - Read: 10s of microseconds
    - Write: 100s of microseconds
    - Erase: milliseconds
  - Must be erased before new data is recorded
  - Bad blocks develop over time
  - Limited endurance to program/erase cycles
- Flash Components for Space Applications
  - High density stacked packages
  - Shared data busses
  - Opportunity to pipeline operations over flash die

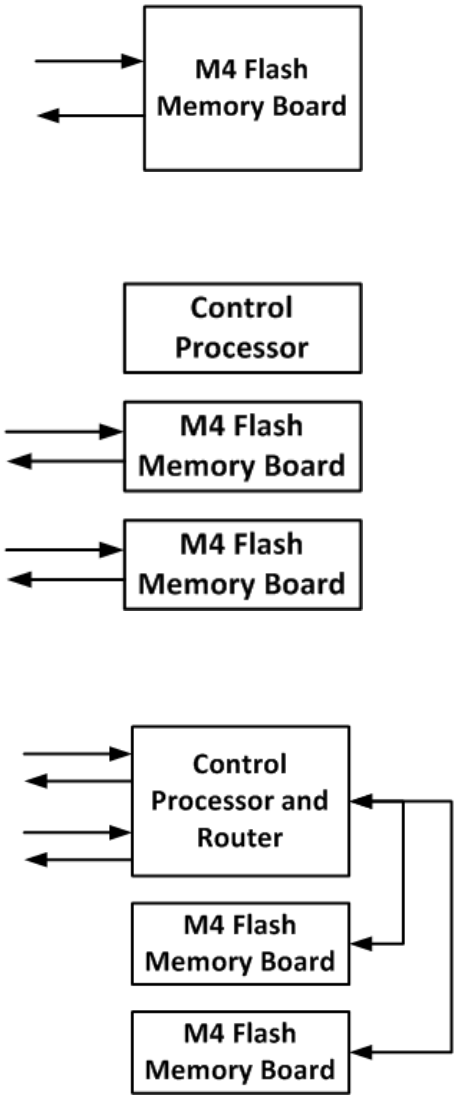


SSR Architecture Considerations

- High Level System Architecture
  - Number of independent flash memory boards
  - Presence and function of controller and data router
- Hardware Topology
  - Organization of flash memory components
  - Shared signals and busses
- Access Plans
  - Low level scheduling of reads and writes to flash
- Buffering
  - Data management to support concurrent data flows
- Application Level Data Management
  - Organizational abstractions: files, queues
- Environmental tolerance
  - Shock, vibration, thermal variation, radiation

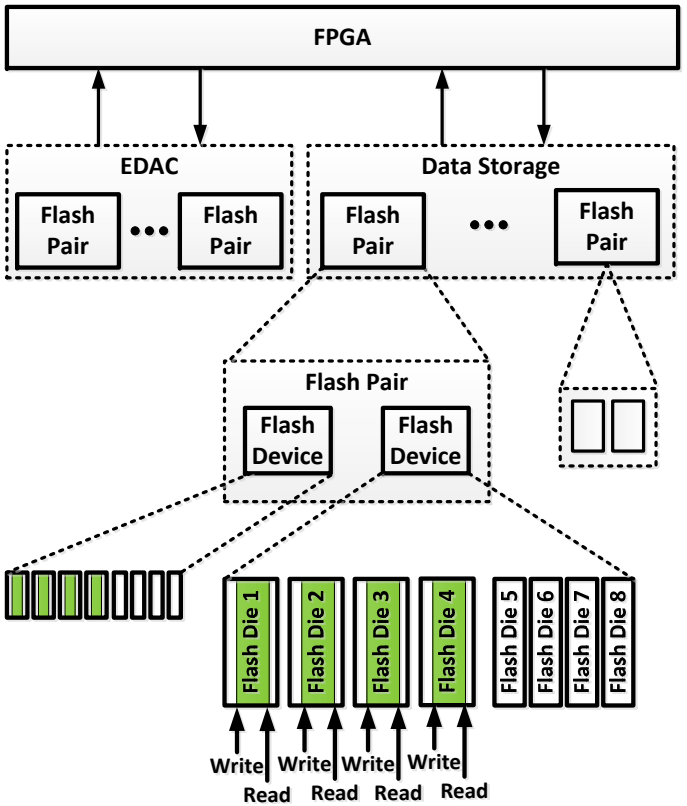
High Level System Architecture

- Various high-level SSR configurations
  - Number of M4 flash memory boards
  - Control processor
  - Data router
- Design Drivers
  - Overall SSR storage capacity
  - Total concurrent data rate
  - Data management strategy
  - Data flow management



Planning Flash Accesses in Time and Space

- Utilize ONFI for access to flash
  - Write Page
  - Fetch Page
  - Read Page
  - Erase Block
- Spatial Planning
  - Multiple flash devices
  - Multiple flash die per device
    - Shared data bus requires data transfer to only one die in a device at any time
- Temporal Planning
  - Operations utilizing device interface
    - Read/Write data transfers
  - Autonomous operations
    - Executed by flash die independently
    - Page load, page program, block erase
    - Allow pipelining of operations over multiple die within a device

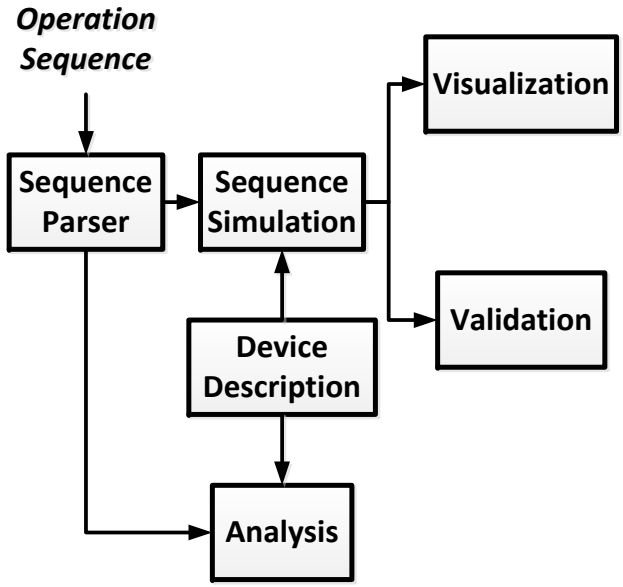


Motivation for a Design Framework

- Tools to Explore Design Space
  - Hardware Configurations
    - Number of flash devices
  - Interface Topologies:
    - Shared signals
    - Concurrent interactions with multiple components
  - Mapping and Scheduling of Operations
    - Read and write transfers
    - Erasure
    - Device busy times
- Analysis Objectives
  - Verify correctness of a design
  - Accurately predict performance characteristics with respect to application level requirements
  - Optimize Design
    - Minimize number of components to reduce size, cost, power consumption
    - Minimize clock rates to provide design margin and reduce power consumption

Development Tools

- Simple formal language for specifying time/space schedule of accesses
- Parser to interpret access plan
- Detailed description of device characteristics
  - Worst case access times
  - Clock rate
- Simulator executes access plan using detailed device timing
- Validation tools ensure functional correctness
  - No collisions on flash device interfaces
  - Operations complete under worst cased conditions
- Performance analysis
  - Accurate prediction of data rates
- Visualization tools provide graphical view of flash activity
- All tools implemented in MATLAB



Specifying Memory Access Plans

- Specify flash access using ONFI commands
  - Write page
  - Fetch page
  - Erase block
  - Check device status
- Temporal operators
  - Delays
    - Allows scheduling time for operations to complete
  - Pipeline advance
    - Specifies when operations are initiated on other die within a flash device
- Definition of device characteristics
  - Device structure
    - Number of die
    - Number and size of blocks and pages
  - Access execution times
    - Based on interface clock rates
  - Operation execution times
    - Page program, block erase, etc.
    - Worst case execution times from device specifications

Simulation and Visualization

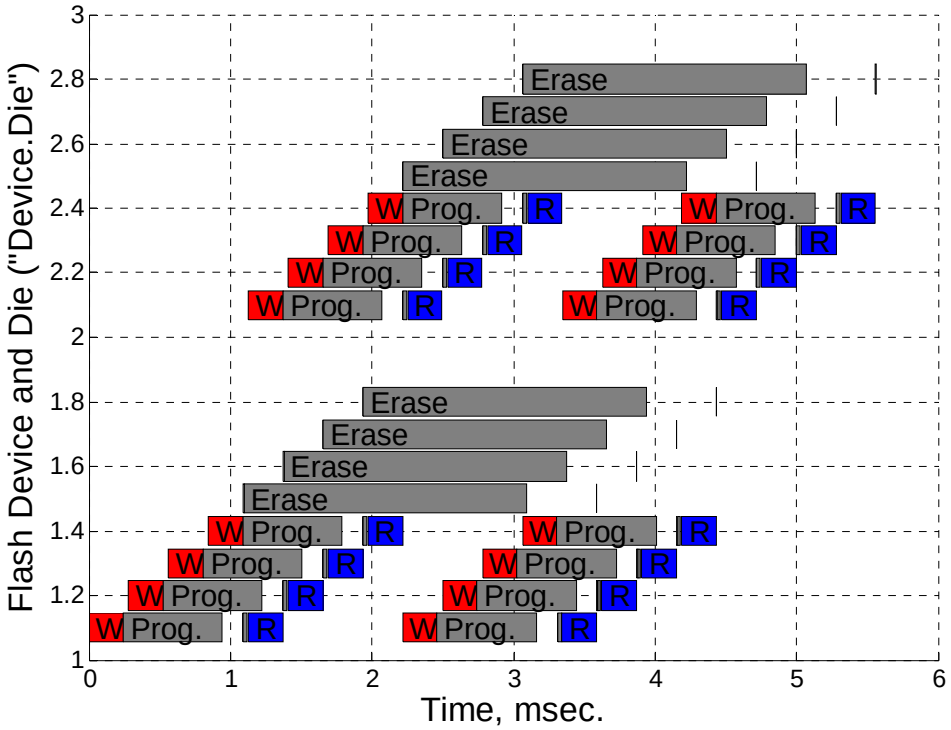
- Simulation
  - Execute time and space access plan
  - Elaborate operation sequence with execution time information for flash device
  - Determine activity on flash die over time
- Visualization
  - Provide precise visual depiction of flash activity
  - Shows activity across flash die and over time
  - Visualization can be explored to understand access conflicts and to measure timing

Achieving Design Objectives

- Data Rates
  - Automated analysis reports data rates achieved by design
- Flash Utilization
  - Maximum data throughput achieved by increasing duty cycles on interface busses
- Unequal Record and Playback Rates
  - Access schedules can repeat operations to control overall data transfer rates
- Handling Data Transfer Channels
  - Different operations may be associated with different logical data flow channels
- Automated Erasure
  - Erasures represented in data access schedule ensure memory capacity is freed before it is needed

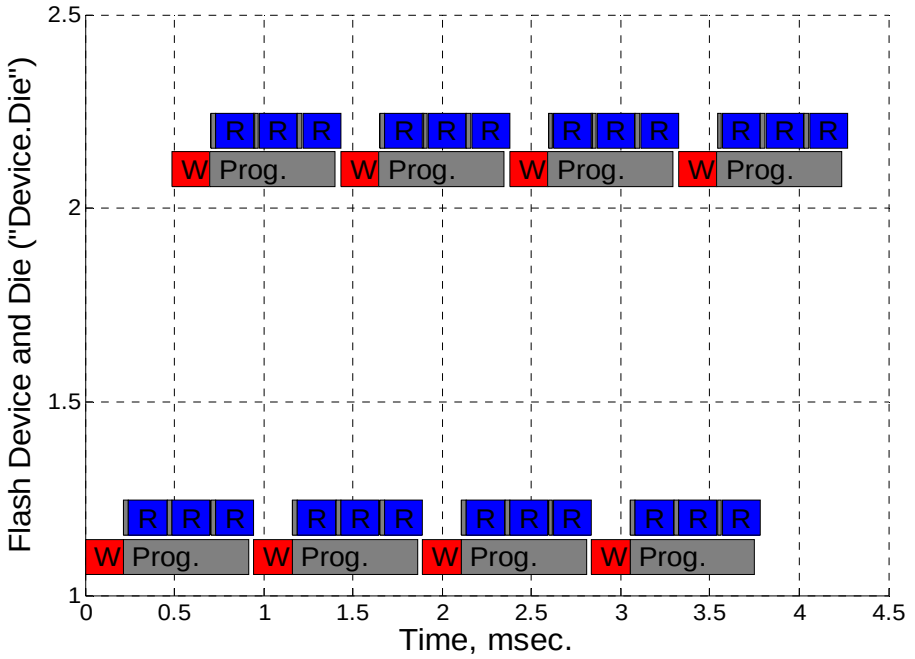
Case Study 1: Equal Record and Playback Rates, Automated Erasure

- Operations scheduled over eight die on each of two flash devices
  - Pairs of devices may be operated in parallel to further increase data rates
- Simultaneous storage and retrieval of data
- Near continuous recording and playback
  - Minimizes requirements for buffering
- Heavily utilized data busses
  - Maximum performance from hardware configuration
- Erasure of old data occurs concurrent with recording and playback



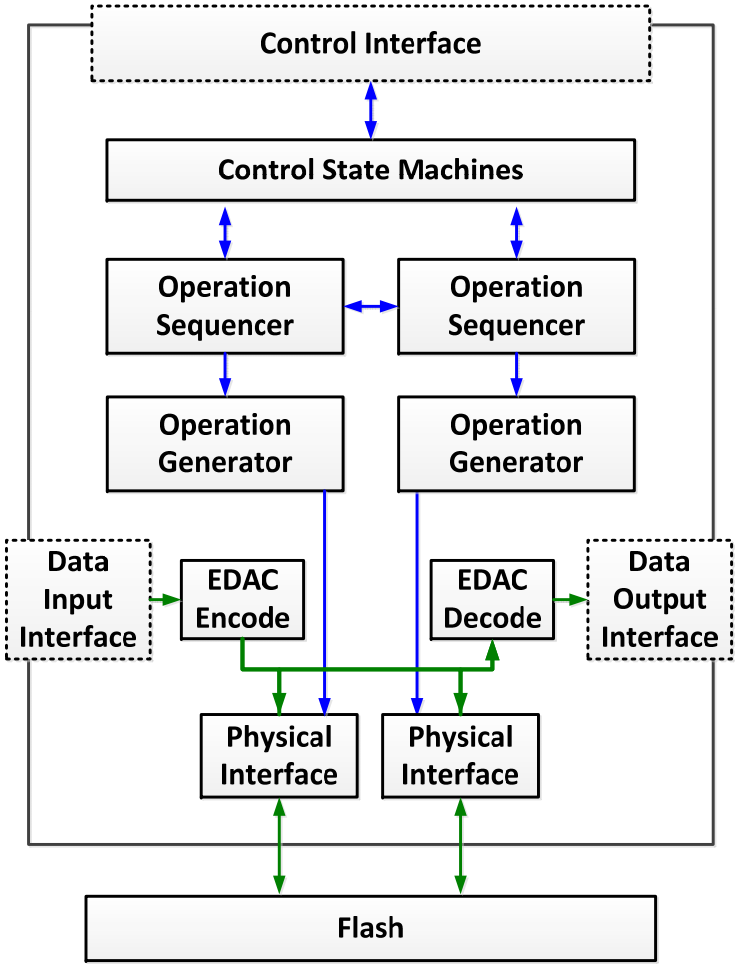
Case Study 2: High Rate Playback

- Operations scheduled on two die each of two devices
- Data playback at a much higher rate than recording
- Flash device bus activity maximized
- Operation pipelining over two die can be mapped to any two die within a device



Implementing a Design

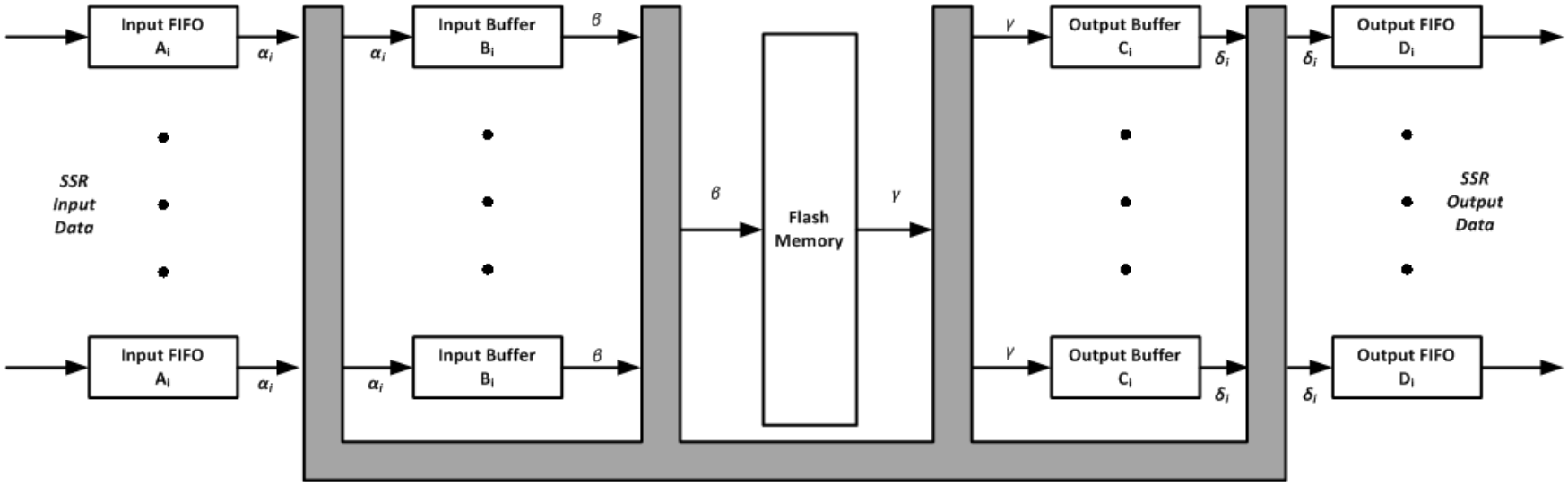
- Access plans implemented using custom FPGA logic
- Design Elements
  - High level control state machine manages modes
    - Record, playback, etc.
  - Operation sequencers implements the access plan
  - Operation generators implement low level operations when triggered by sequencer
  - Physical interface handles signaling to flash devices
  - Error control coding may be integrated into data flow



Data Buffering

- Data flow requires temporary storage of data
  - Flash memories are inaccessible for extended times
  - Incoming data continues to arrive
  - Need for continuous data playback
  - Addressing flexibility excludes deconfliction
- Memory hierarchy
  - Input and output FIFOs
    - Dual-port FPGA block RAM
    - Dedicated to interface
  - Input and output buffers
    - Fast external memory: SRAM, SDRAM
  - Flash memory
    - Long term, nonvolatile storage

Buffering Design



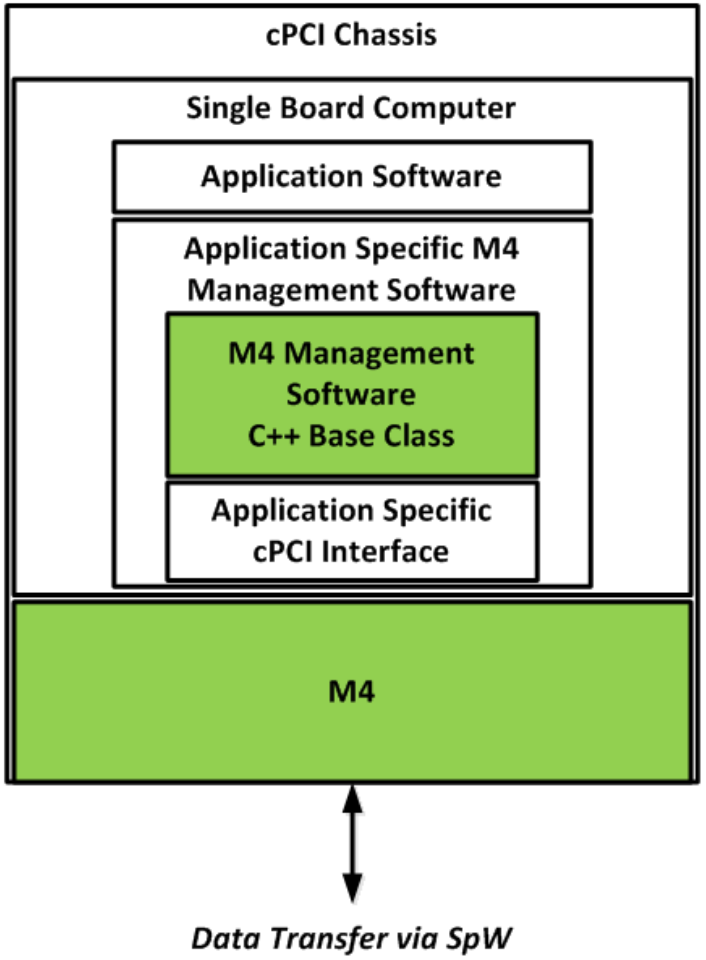
- FIFO memory very limited in capacity
- Buffer memory accesses
  - Potential for very high throughput
  - Scheduling of transfers over single data interface
    - Buffer to/from FIFOs
    - Buffer to/from flash memories

SSR Management Software

- Software algorithms and data structures provide mechanisms to organize and access data
  - Allocate memory for storage of data
  - Select data for playback
  - Multiple record and playback interfaces
  - Control of data erasure
  - Reporting on status of stored data
- Utilizes control link to M4 hardware
  - Command mode changes
  - Provide address information
  - Hardware management
- Integration options
  - Embedded single board computer integrated into SSR
  - Data management software hosted on system processor

SSR Management Software

- Provides application specific data management abstractions via software API
- Transparently manages key SSR operations
  - Address management
  - Bad block handling
  - Wear leveling
  - Automated erasure
  - Real time control of M4 hardware
- Generic C++ implementation provides flexible software integration
- Mass Memory Conop Developer
  - Software-only SSR simulators enable early application integration
- Record and playback data does *not* utilize the software library or processor resources
  - Tape recorder, not hard disk



SSR Management: Data Queue Abstraction

- SSR Storage Organized as Multiple Queues
  - Independent queues for each input data interface
  - Reconfigurable allocation of capacity to queues
  - Data stored in order received
- Flexible Data Playback
  - Arbitrary routing of queues to playback interfaces
    - Each queue supports playback on one interface at a time
  - Data playback in order received
  - Playback pointer may be repositioned within queue
    - Supports retransmission of data
- Data Erasure
  - Oldest data erased first
  - Just in time automated erasure or erase on command

SSR Management: File System Abstraction

- Light-weight File System Operation
  - Data stored to “files”
    - Memory allocated at time of file creation
  - Partial or complete playback of files on command
  - Data erased at file level
  - Not to be confused with full Linux-like file system
- More flexible, efficient use of capacity
- Greater data management flexibility
- Drawbacks
  - More complicated data management software
  - Greater need for nonvolatile storage of file system state information
  - Wear leveling

Handling Bad Blocks

- Bad Blocks in Flash Memory
  - Localized, permanent failure of a region of flash memory
  - Bad blocks exist at manufacture
  - Additional bad blocks may develop due to radiation or aging
  - M4 hardware detects bad blocks and reports to software
- Operation with Bad Blocks
  - Pool of reserve blocks on each flash device
  - Reserve block used in place of a bad block
  - Replacements handled at flash device level
- Logical to Physical Address Translation
  - Logical addressing used for interaction with application software
  - Data management software associates a physical block on each concurrently accessed flash device with a logical address
  - Constant usable SSR capacity
  - Fixed, uninterrupted logical address space

Address and Bad Block Management

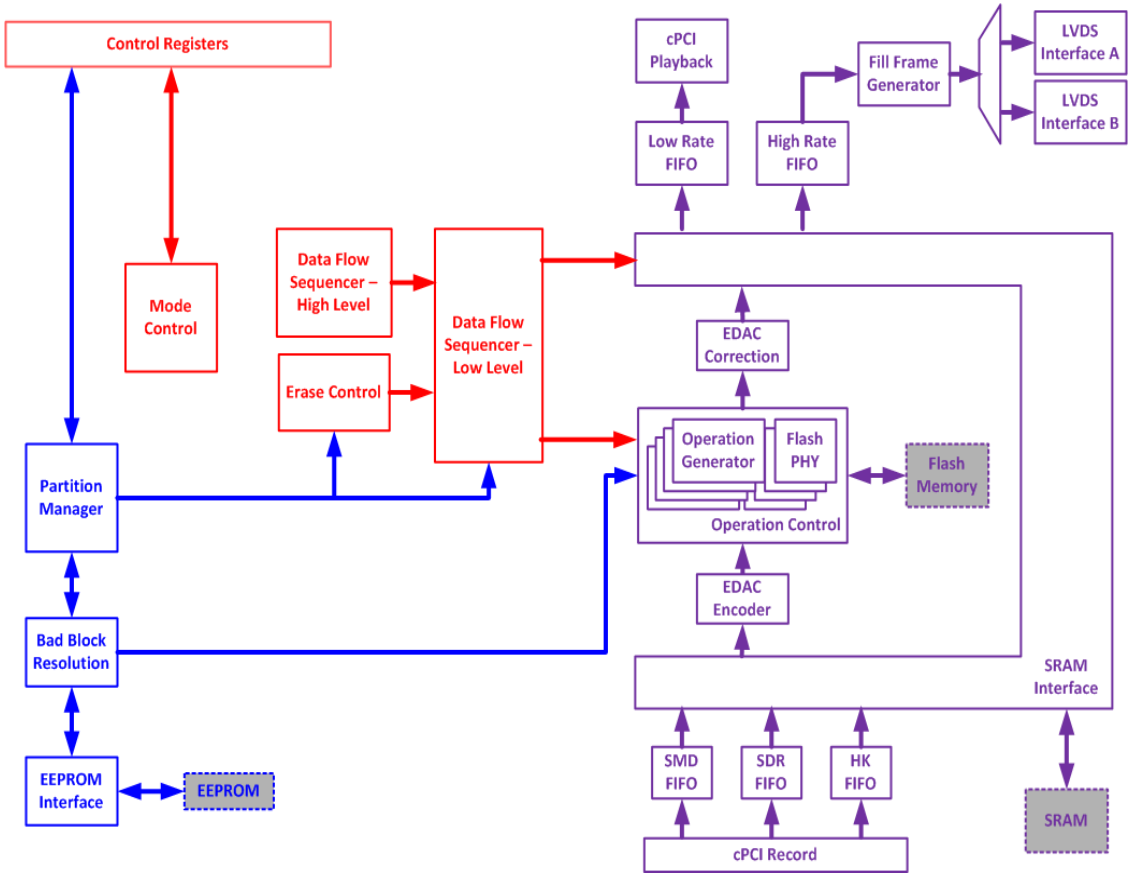
- M4 hardware periodically requests physical addresses from software for recording, playback, and erasure
  - Requests delivered via interrupt or message
  - Pipeline processing tolerates software response latency
- Software computes next logical address
- Bad block table used to provide address translation
  - Map single logical address to concurrently accessed physical addresses
  - Use reserve blocks to avoid known bad blocks
- Bad block table maintenance
  - Initialized with factory data
  - Updated as bad blocks are found
    - Errors during recording logged, but replacement not deployed
    - Upon erase of logical block, bad block is permanently removed from service
  - Bad block table exported to nonvolatile storage to preserve bad block table through a software restart

Environmental Tolerance

- Thermal
  - Wide temperature variation in space environment
  - No convective cooling
- Shock and Vibration
  - Rockets are never gentle
- Radiation
  - Devices selected to tolerate radiation without permanent failure
  - Numerous radiation-induced errors in stored data
  - Extensive, aggressive data protection coding
- Operational lifetime
  - Uninterrupted operation for years without maintenance

System Architecture: Stand-Alone M4

- Capacity: 100s of Gb
- Data Interfaces:
  - Recording
    - 2x at 10s of Mbps
    - 1x at 1s of Mbps
  - Playback
    - 1x at 100s of Mbps
    - 1x at 2s or Mbps
- Control Interface: cPCI
- Data Management
  - Realized in M4 FPGA
  - Multiple queue data organization
  - Hardware bad block management



System Architecture: Software Supported M4

- System Architecture
  - Multiple M4 boards
  - Central control software
  - SSR management software hosted on application processor
- Capacity: Multiple Tb
- Data Rates: Multiple Gbps record and playback
- File system based data management
- Multiple nested data encoding schemes
  - Error free operation under radiation-induced errors
  - Error free operation under single device failure

System Architecture: Integrated Recorder

- System Architecture
  - Multiple M4 boards
  - Dedicated processor executing data management software
  - Extensive data routing and buffering
- Capacity: Multiple Tb
- Data Interfaces:
  - Multiple recording channels, 1s to 100s of Mbps
  - Multiple playback channels, 1s to 100s of Mbps
- File system based data management
- Redundant hardware configurations support graceful degradation of capacity while supporting full data rates

Ongoing Research and Development

- Flash Memories
  - Need for higher density, higher data rates
  - Use of MLC flash in a radiation environment
  - Coding approaches to compensate for higher bit error rates
- Interface Technologies
  - High speed interfaces to space-grade components
- Buffer Memories
  - Advanced data buffering management techniques
  - Higher speed memories
- Integration of advanced, radiation-tolerant FPGAs
- Additional formal modeling and simulation tools

Conclusion

- SwRI solid state recorder development approach allows precise tailoring of design to application requirements
- Analysis tools allow highly efficient designs
  - Minimize flash components and FPGA clock rate
  - Reduces cost, size, weight, and power requirements
- Designs cay be easily revised to address changes in data management requirements
- Automated, early analysis ensures robust designs which function under worst case conditions