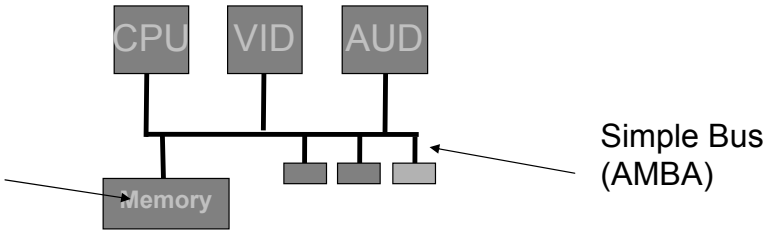
EDP WorkShop: Processor Centric Designs Are Dead April 17, 2008

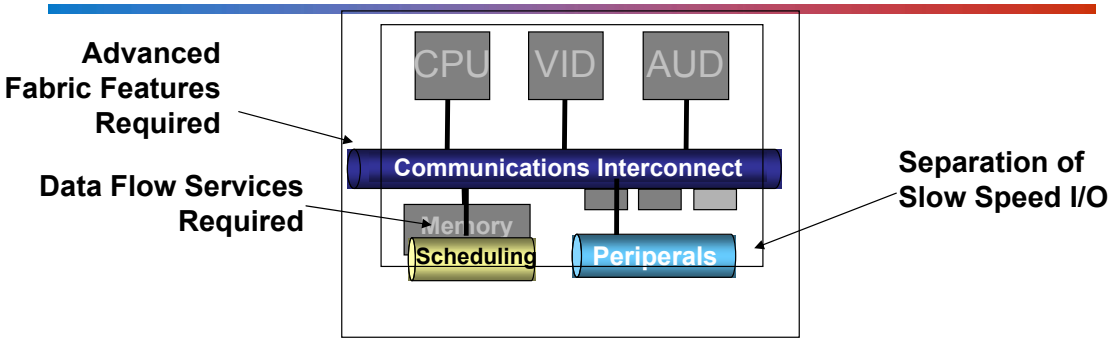
Ray Brinks, Sonics

Circa 2000 …



Processor Selection Was Biggest Design Challenge

Circa 2008 Designs got a lot tougher



Processor Selection Was Biggest Design Challenge

- Good news for my fellow panel members
  - Lots more processors on chip
  - Heaven forbid: ARC, Tensilica, MIPS AND ARM could all co-exist on a single die
  - Driven by tile based design re-use
  - Facilitated by advanced automation techniques and standards based integration flows
- Bad news for my fellow panel members
  - SoC world no longer revolves around the processor

Touch a Couple Examples

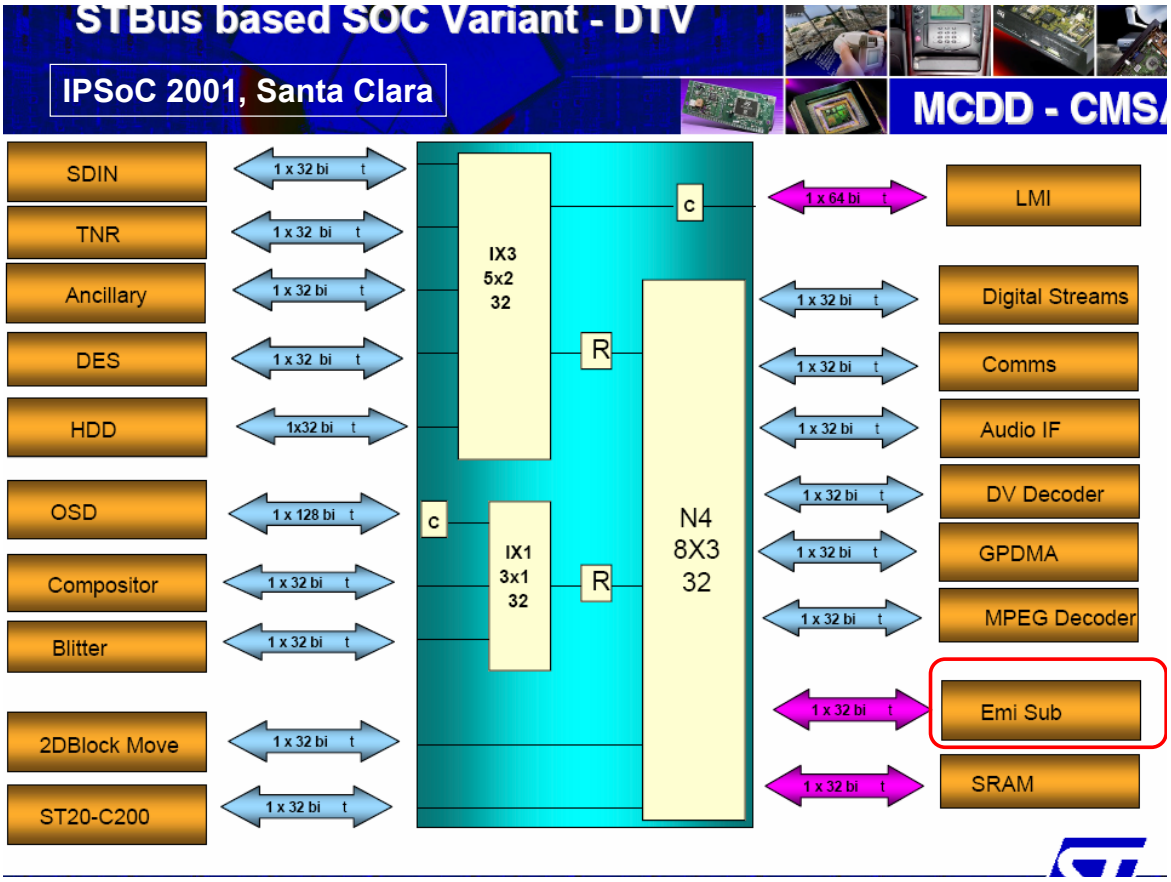
- DTV chips
  - What makes them hard to do?
- Application processor chips
  - What are the challenges there?
- Is there a difference?

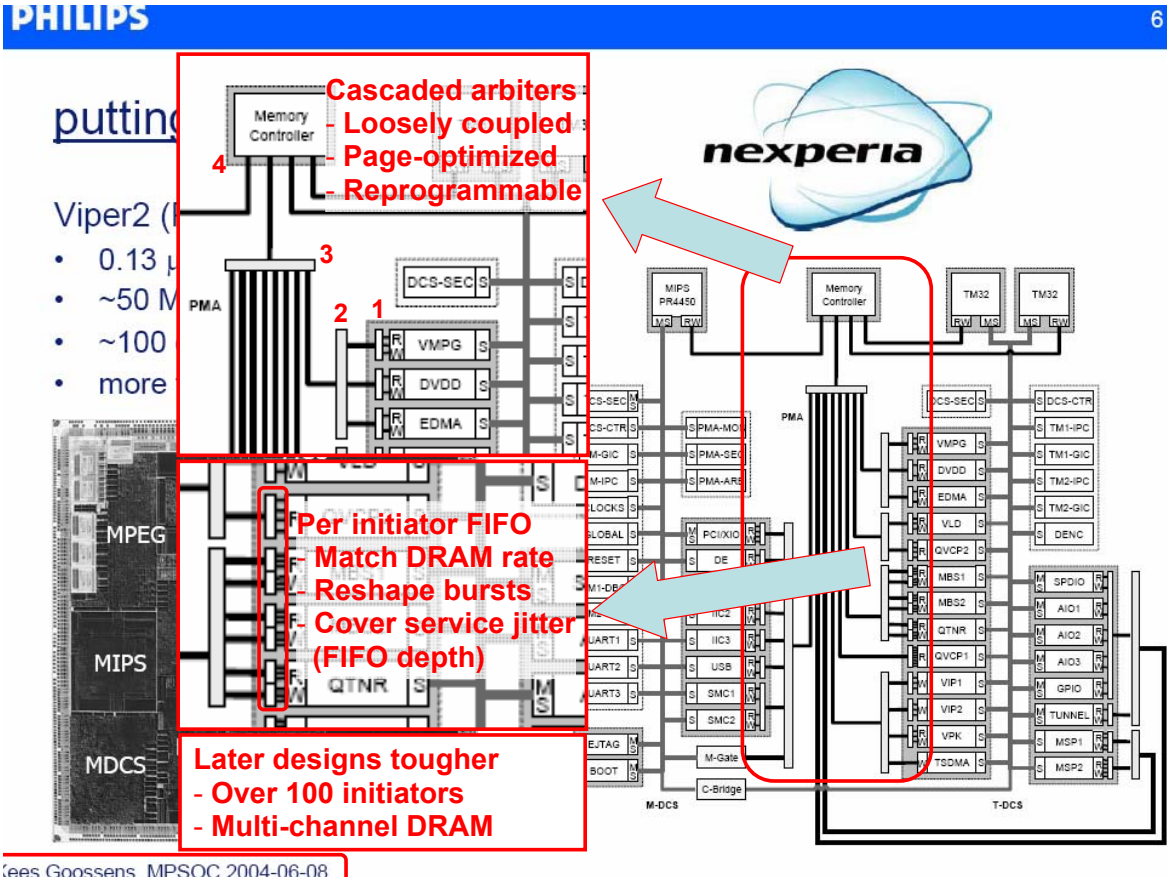
What Makes DTV SoC's Difficult?

- Multiple channels of multi-format video decoders
  - MPEG 2, H.264, …
- Lots of scaling/resampling requirements
  - PIP, screen sizes, …
- High-bandwidth image enhancement features
  - Reverse pull-down, 120 Hz, motion judder, …
- Increasing CPU & graphics acceleration
  - Richer user experience, 2D/3D menus, slide shows, …
- Wide variety of I/O
  - Video, audio, PC, USB, 1394, Ethernet, WiFi, HDD, …
- Huge SW challenge to make it all work together in real time: latency sensitivity, isochronous, quasi-isochronous
- All for under $10!

DTV SoC Architecture: DRAM-optimized

- Achieving required functionality and system BOM requires shared memory architecture
- Working set sizes of most subsystems too large for on-chip RAM
  - Shared memory is predominately external DRAM
- ***System performance defined by DRAM performance***
  - Optimizing DRAM transfer efficiency, while guaranteeing real-time behavior, is key requirement
  - ***Memory subsystem dominates DTV SoC architecture***
- Most DTV SoC's use in-house interconnects & DRAM controllers
  - Carefully optimized to maximize DRAM performance
  - *Tightly coupled to DRAM technology, frequency & configuration*


STBus based SOC Variant - DTV
IPSoC 2001, Santa Clara
MCDD - CMS
SDIN
TNR
Ancillary
DES
HDD
OSD
Compositor
Blitter
2DBlock Move
ST20-C200
1 x 32 bi t
1x32 bi t
1 x 128 bi t
IX3
5x2
32
IX1
3x1
32
N4
8X3
32
C
R
1 x 64 bi t
LMI
Digital Streams
Comms
Audio IF
DV Decoder
GPDMA
MPEG Decoder
Emi Sub
SRAM


PHILIPS
6
putting
Viper2 (
• 0.13 µ
• ~50 M
• ~100
• more
Memory Controller
Cascaded arbiters
- Loosely coupled
- Page-optimized
- Reprogrammable
4
3
2
1
PMA
DCS-SEC
VMPG
DVDD
EDMA
Per initiator FIFO
- Match DRAM rate
- Reshape bursts
- Cover service jitter
(FIFO depth)
QTNR
MPEG
MIPS
MDCS
Later designs tougher
- Over 100 initiators
- Multi-channel DRAM
nexperia
MIPS PR4450
TM32
M-DCS
T-DCS
Kees Goossens, MPSOC 2004-06-08

What Makes Apps Processors Difficult?

- Multiple channels of multi-format graphics
  - Graphics/Video accelerators, 2D/3D, image capture, …
  - DRAM utilization challenge
- Wide variety of I/O
  - Video, audio, USB, timers, GPIO, I2C …
- Concurrent management of various tasks while maintaining a 3G cell connection:
  - *Watching live TV while receiving an incoming phone call*
    - *DVB-H radio tuner -- H.264/WMV video decode -- 64-voice polyphonic MIDI ring tone*
  - *Video conferencing while recording*
    - *MPEG4 or H.264 video encode and decode -- AAC+ audio encode & decode -- MMC record*
  - *Over the air synchronization while listening to MP3 songs*
    - *Synchronization protocol -- MP3/WMA audio playback*
- Incredibly low power
- All for under $20!

TI

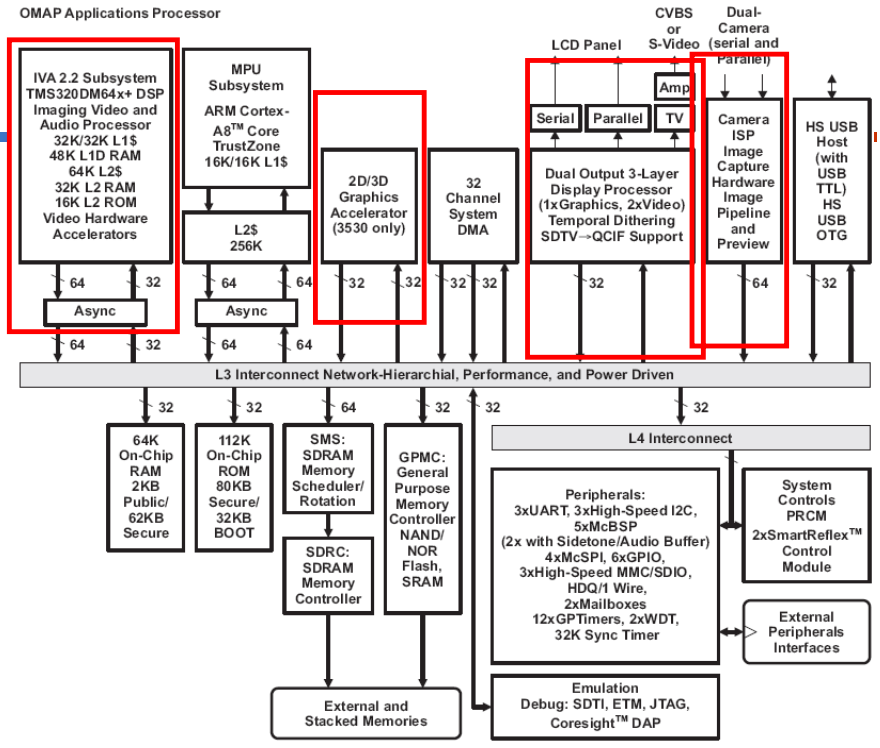


Figure 1-1. OMAP3530/25 Functional Block Diagram

SoC Architecture Summary

- Critical factors:
  - Need highest DRAM transfer efficiency…
  - While ensuring real time requirements are met
  - … to get to low power FAST!
- Multi-functional accelerators explode the SW complexity
  - Real-time constraints that vary from core to core: latency (asynchronous), synchronous (isochronous and quasi-isochronous) flows
  - Traffic contention at the memory interface that is difficult to predict. Caused by varying data chunk sizes and access methods as well as demanding bandwidth requirements.
  - Error recovery and diagnostic capabilities competing with gate count requirements and in-band timing requirements.
  - Late- or fast-changing market requirements driven by new standards and rapid consumer obsolescence.
- Preservation of IP core re-use
  - Datapath widths, operating frequencies, interface protocols, FIFO sizes, burst lengths, arbitration algorithms

What is Needed?

An optimized architecture that allows direct comparison of architectural choices and produces the complete solution:

- Accommodate full variety of partitioning approaches and services: active decoupling
- Automation to enable rapid model configuration & optimization
- Re-use of legacy initiators/data flows as a starting point
- Performance tooling to qualify results
- Metrics to determine success:
  - DRAM efficiency while ensuring real time
  - Address pattern dependencies
  - Concurrency requirements
  - etc

Multicore SoC Interconnects Require

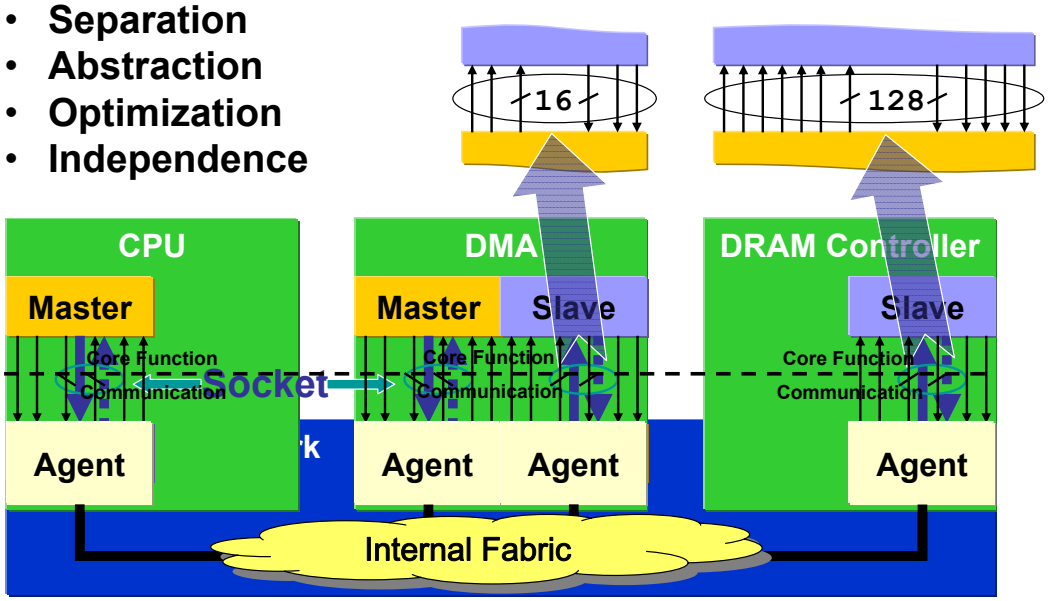
- *System communications* infrastructure as core fabric to achieve the flexibility needed to manage multiple data traffic flows simultaneously
- *Advanced Fabric Features* that facilitate heterogeneous multiprocessing using distributed architectures
- *Data Flow Services* that minimize new system management challenges that add exponential design load and risk to SoC development

Advanced Fabric Features

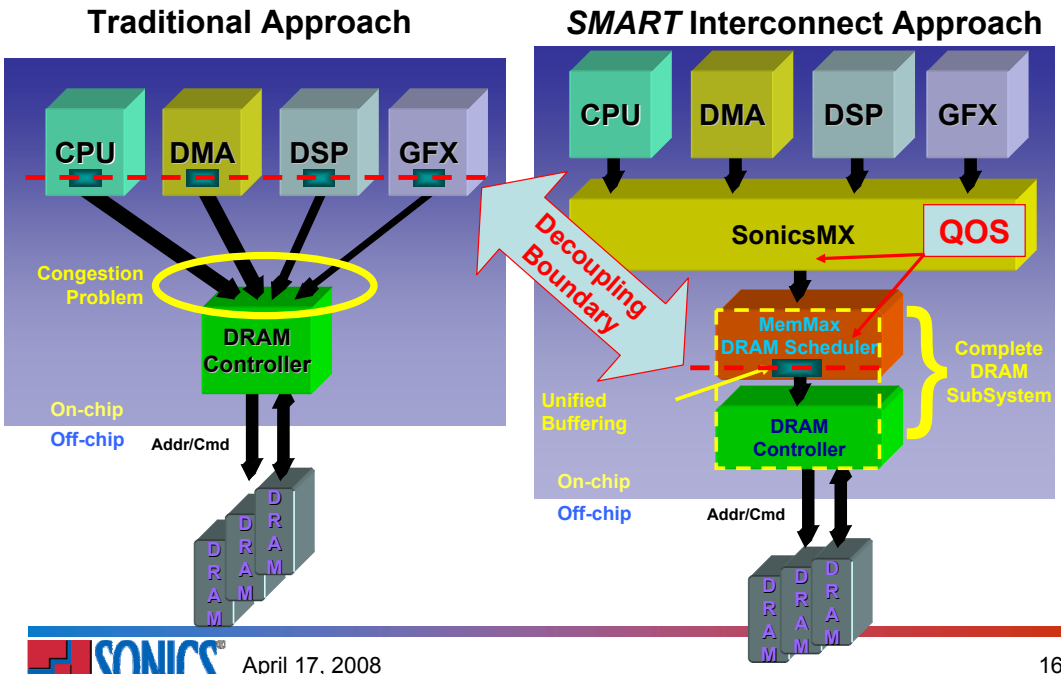
- Increase Performance
  - Non-blocking architecture with networking enables ultra low latency data flows for Multi-core applications
- Maximum IP core Reuse
  - Decoupling IP cores from Interconnect minimizes impact of incremental changes to platform architecture
- Maximum IP Library Flexibility
  - Universal Connectivity (OCP, AHB, AXI, APB)
- Preserves Previous Investments
  - Architecture consistency across Interconnect solutions
- Low Project Risk
  - Ability to model Interconnect during architecture phase of SoC designPerformance tooling to qualify results
  - Metrics to determine success:
    - DRAM efficiency while ensuring real time
    - Address pattern dependencies
    - Concurrency requirements
    - etc

Network-based SoC: Active Decoupling

- Separation
- Abstraction
- Optimization
- Independence

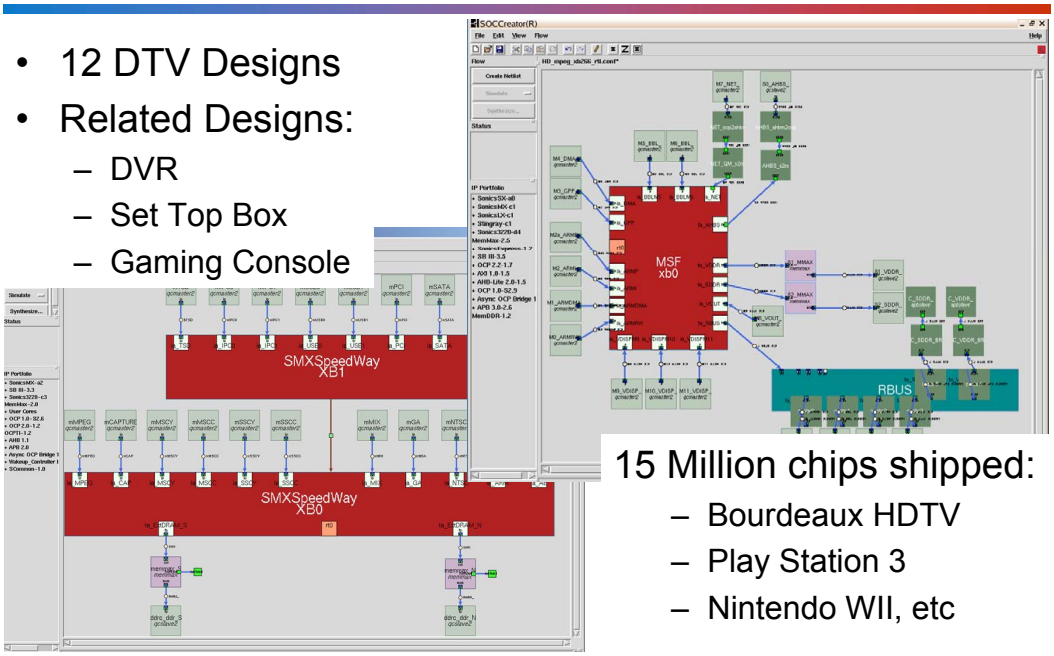


Decoupled Memory Subsystems



Automated Configuration Management

- 12 DTV Designs
- Related Designs:
  - DVR
  - Set Top Box
  - Gaming Console



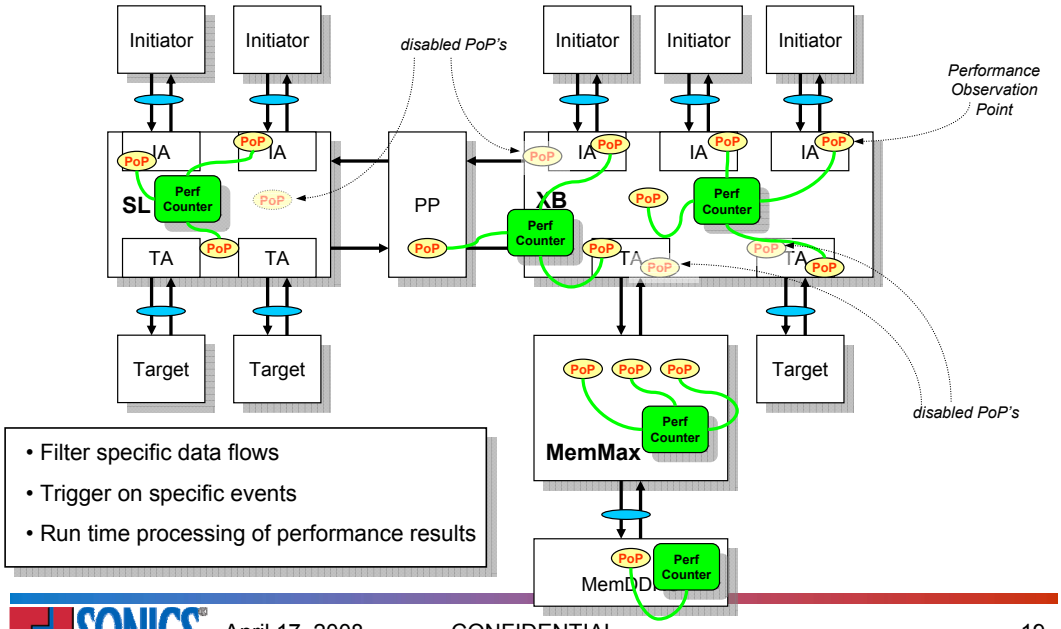
15 Million chips shipped:

- Bourdeaux HDTV
- Play Station 3
- Nintendo WII, etc

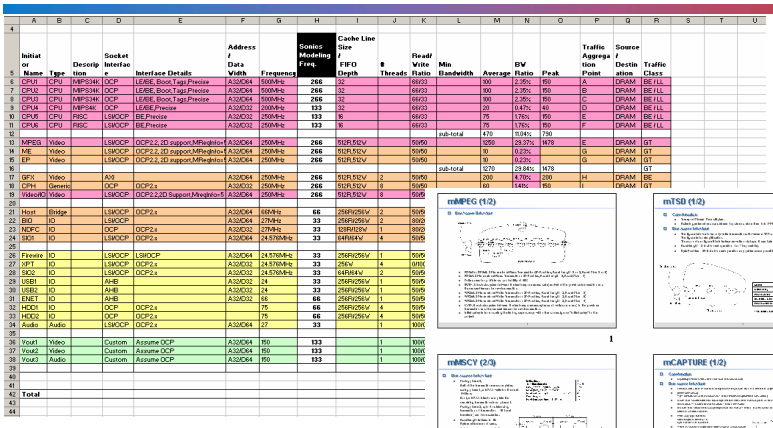
Application Specific Stimulus

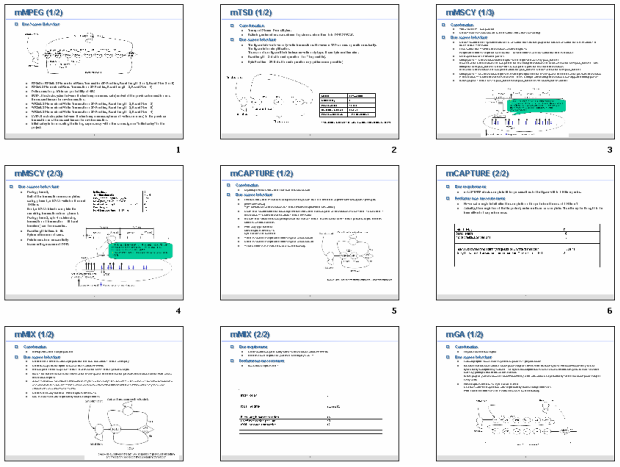
- Stimulus based on real world video traffic
- Estimated data flows created with spreadsheet
  - Model complex interdependent data flows
  - Simple worst case validation
  - Begin architectural tradeoffs early
- Convert emulator traces to master-side transactions
- Generate traffic and re-play on abstracted models

Performance Instrumentation



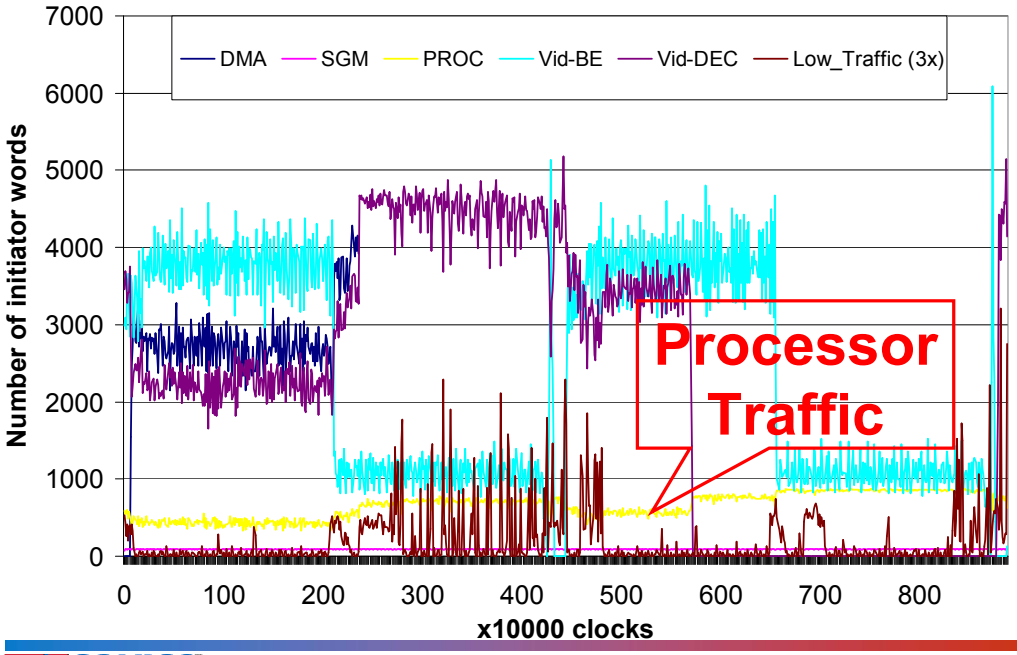
Extensive, Varied Data Sets



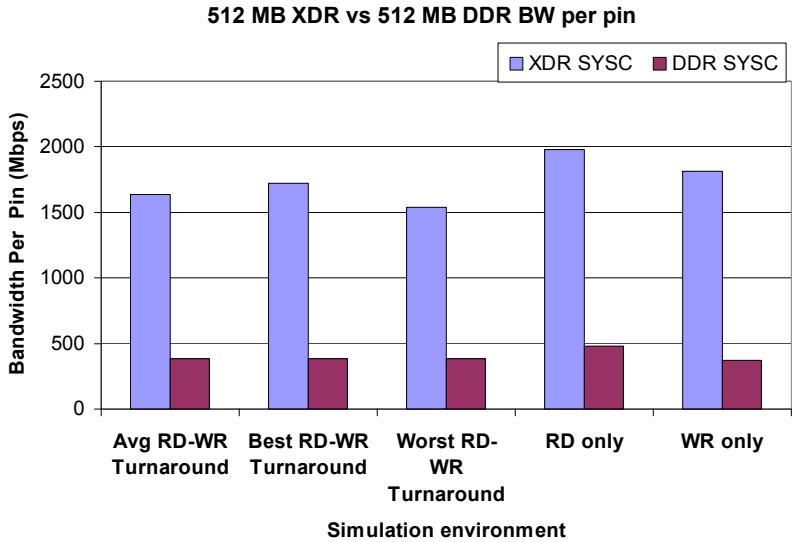


- Trace based
- Customer algorithm
- Worst case validation

DTV Traffic Pattern (2 Frames Data)

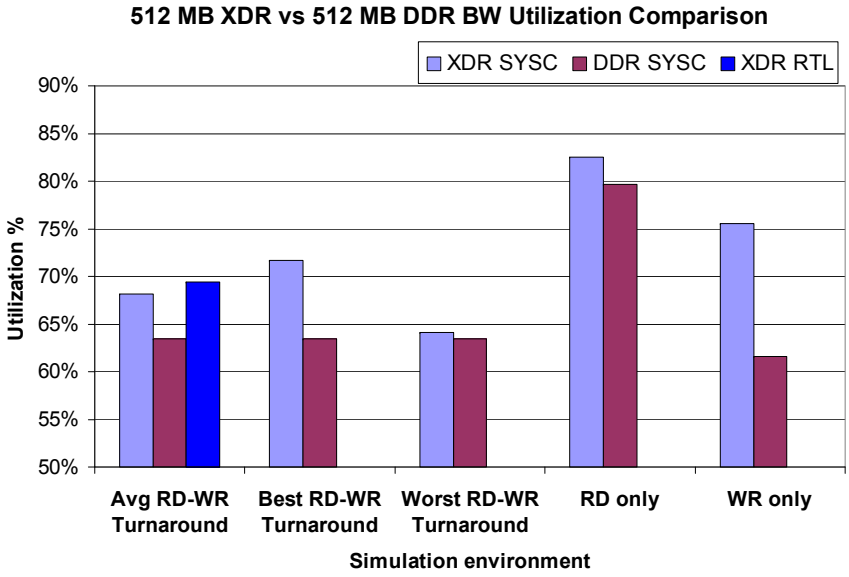


Analysis: XDR vs. DDR-2 for 512MB

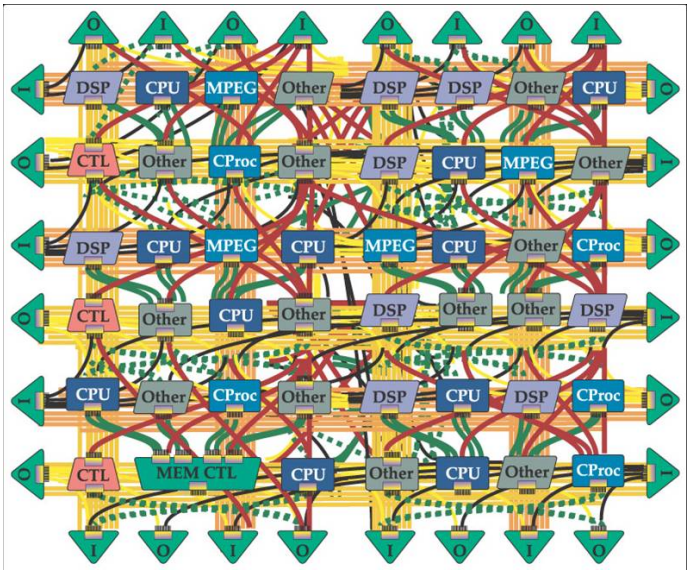


Stimulus used was derived from " real customer" traffic patterns

Analysis Validates Designs



The Old Way



The Sonics Way

From Computing
- Low Latency
- High Bandwidth
- Low Chip Area

From Networking
- Efficient Decoupling
- Real-time Guarantees
- High Scalability

Just a cool picture!

Just a cool picture!