Reducing I/O Bottlenecks
with Numerical Compression

Al Wegener, CTO
Samplify Systems
awegener@samplify.com

Samplify Investors

MOOG	IDT	Schlumberger	mamiya
2009	**2010**	**2011**	**2011**
Computed Tomography	**Wireless**	**High Performance Computing**	**Imaging**
▪ Prism CT ▪ Exclusive licensee	▪ Prism IQ ▪ Strategic investor ▪ Acquired wireless infrastructure patent portfolio	▪ APAX ▪ Strategic investor ▪ Exclusive licensee for oil & gas exploration	▪ Prism 3 ▪ Strategic investor
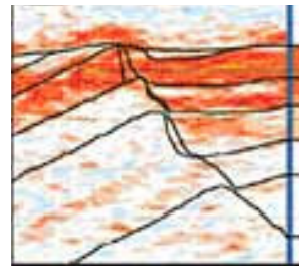
…simply the bits that matter®

Intel VPFPU – ISSCC Paper *(Feb 2012)*

ISSCC 2012 / SESSION 10 / HIGH-PERFORMANCE DIGITAL / 10.3

10.3   A 1.45GHz 52-to-162GFLOPS/W Variable-Precision Floating-Point Fused Multiply-Add Unit with Certainty Tracking in 32nm CMOS

Himanshu Kaul, Mark Anders, Sanu Mathew, Steven Hsu, Amit Agarwal, Farhana Sheikh, Ram Krishnamurthy, Shekhar Borkar

Intel, Hillsboro, OR

controlled by up to
overflow carries (c
ing signals at 24b,
Overall support fo
area by 12% for 4:

To realize performa
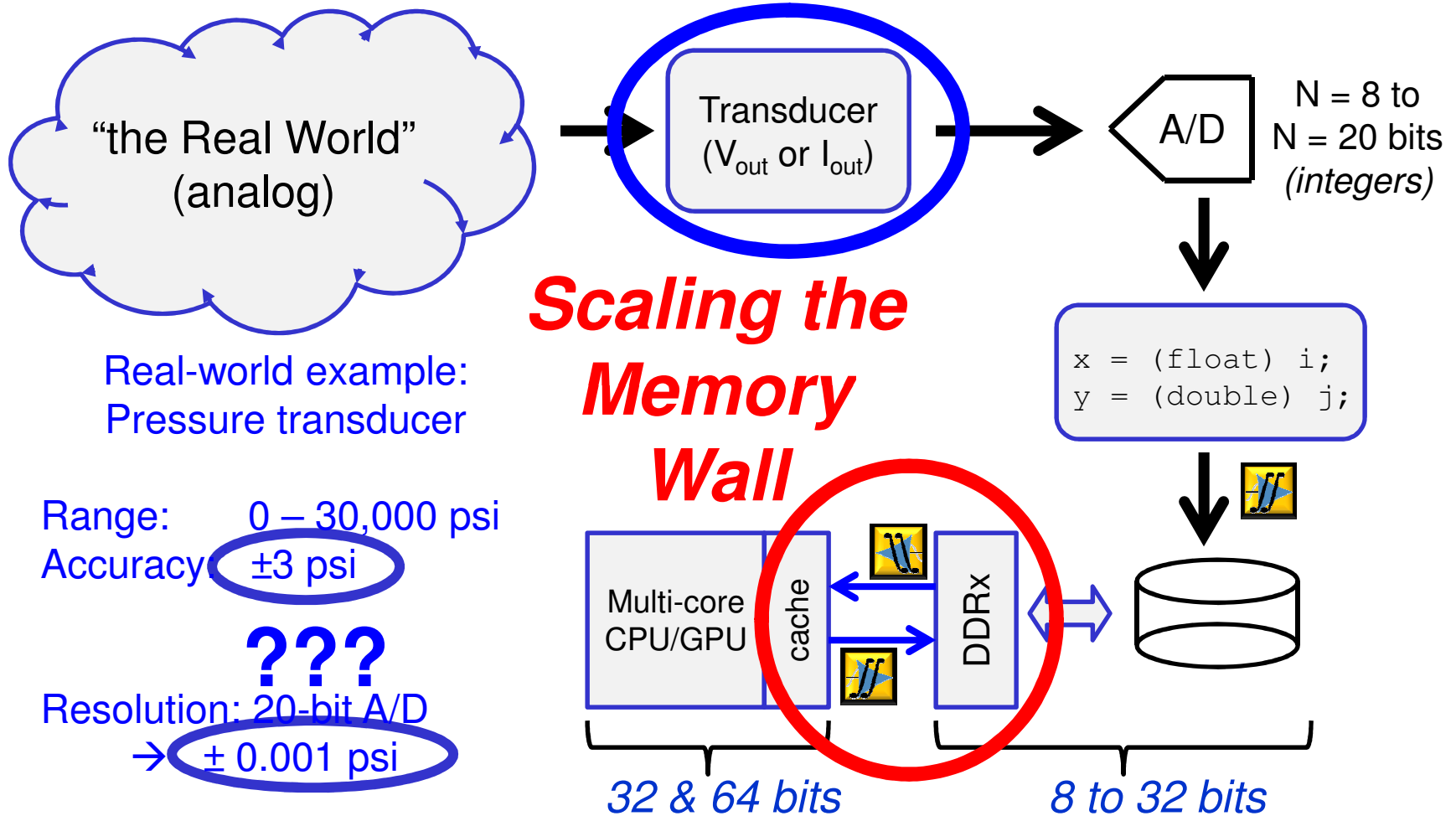sion accuracy, cer
cuits to calculate c

Variable Precision Floating Point Unit

(10.3) A 1.45GHz 52-to-162GFLOPS/W Variable-Precision Floating-Point Fused Multiply-Add Unit With Certainty Tracking in 32nm CMOS
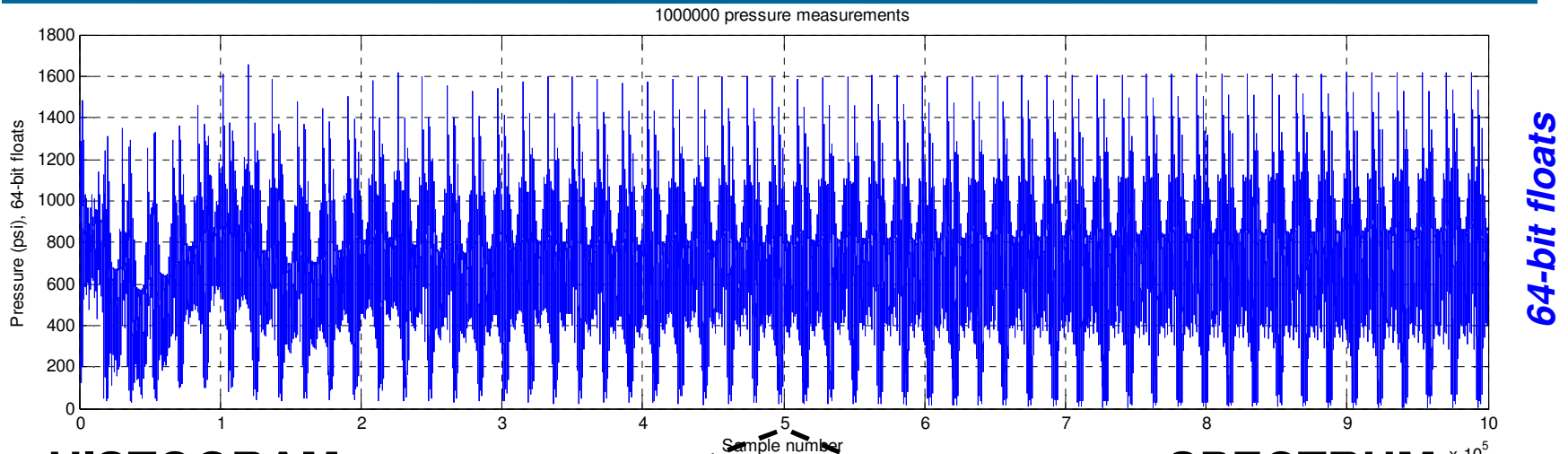
- 1st reported variable-precision floating point unit with accuracy tracking for multiply-add
- Today's floating-point math wastes energy, time, and storage by using worst-case precision everywhere
- Using variable precision (24-bit→12-bit→6-bit) as needed can cut energy by 50%
- Uses NTV circuits for up to 7x further efficiency gain
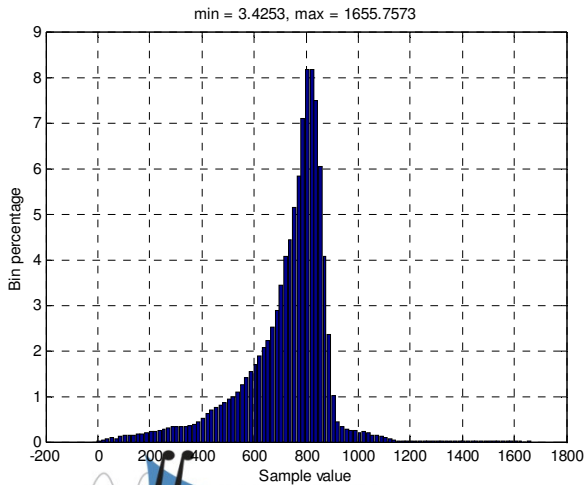
…perhaps IEEE-754 is sometimes overkill ?

…simply the bits that matter®

APAX's "Just Right" Numerical Resolution

"the Real World"
(analog)

Transducer
(V_{out} or I_{out})

A/D

N = 8 to
N = 20 bits
(integers)

Scaling the Memory Wall

```
x = (float) i;
y = (double) j;
```

Real-world example:
Pressure transducer

Range: 0 – 30,000 psi
Accuracy: ±3 psi

???

Resolution: 20-bit A/D
→ ± 0.001 psi

Multi-core
CPU/GPU

cache

DDRx

32 & 64 bits   *8 to 32 bits*

samplify
SYSTEMS®

…simply the bits that matter®

Example: How Much Compression?

1000000 pressure measurements

64-bit floats

HISTOGRAM

min = 3.4253, max = 1655.7573

SPECTRUM

1500 pressure measurements

max 79.7 dB

dyn range 51.0 dB

noise floor 28.7 dB

5

...simply the bits that matter®
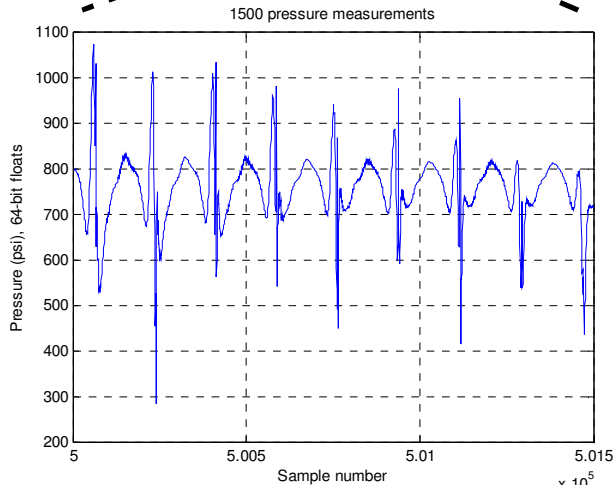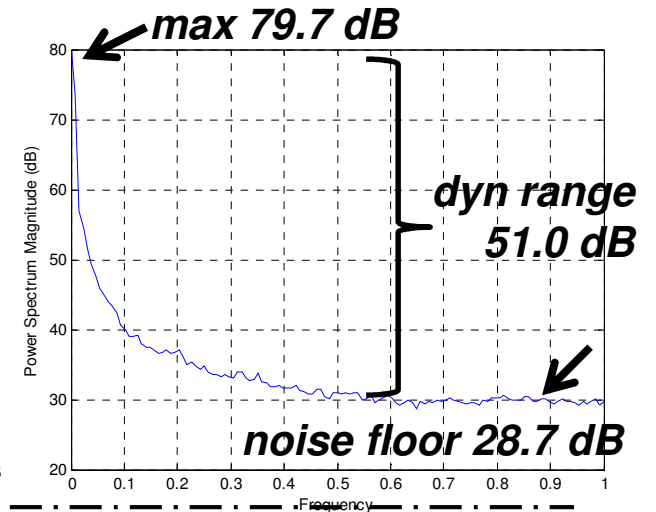
Lossy Error = Original - Decompressed

compress

decompress

decompressed signal

Input samples (ints or floats)

original signal

+

Error

psd = "power spectral density"

…an often-used Matlab function

FFT (psd)

Error spectrum

samplify SYSTEMS®

…simply the bits that matter®
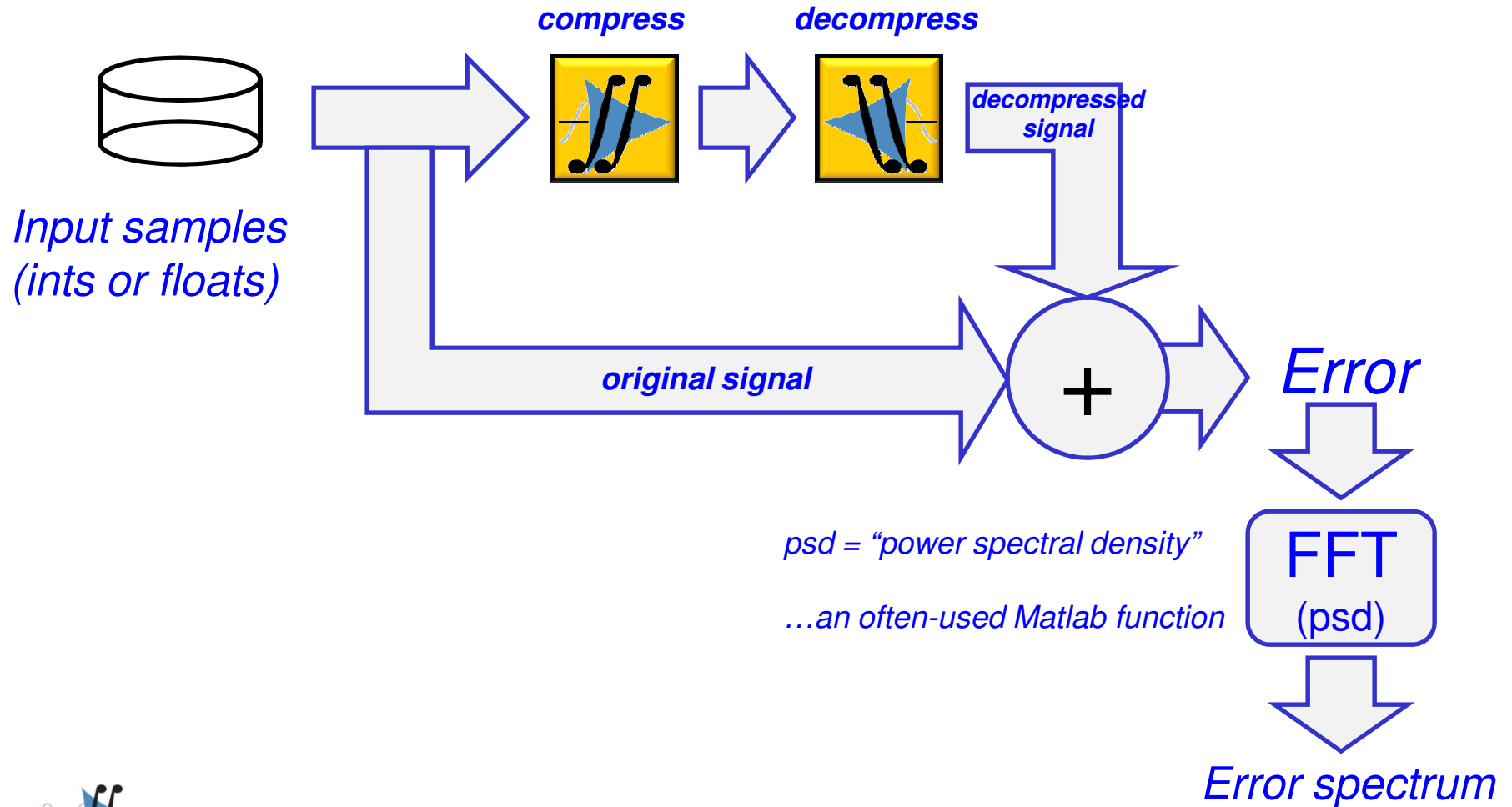
Example: Error at 2:1 and 3:1 Compression

Original, from 281.6599 to 1071.1343

Error at 2:1, , from -0.16361 to 0.18899

Error at 3:1, , from -1.0989 to 1.1132
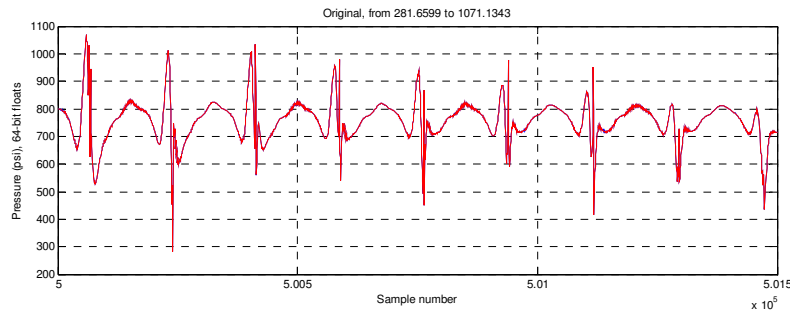
Range: {281.66, 1071.13 psi}
Blue: Original
Red: Decompressed
(do you see any blue ?)
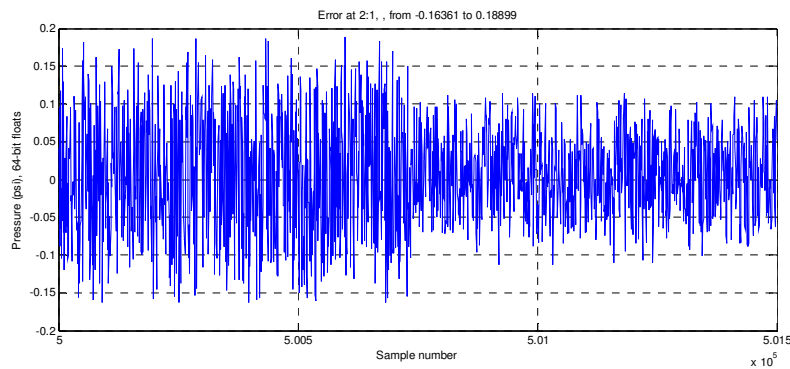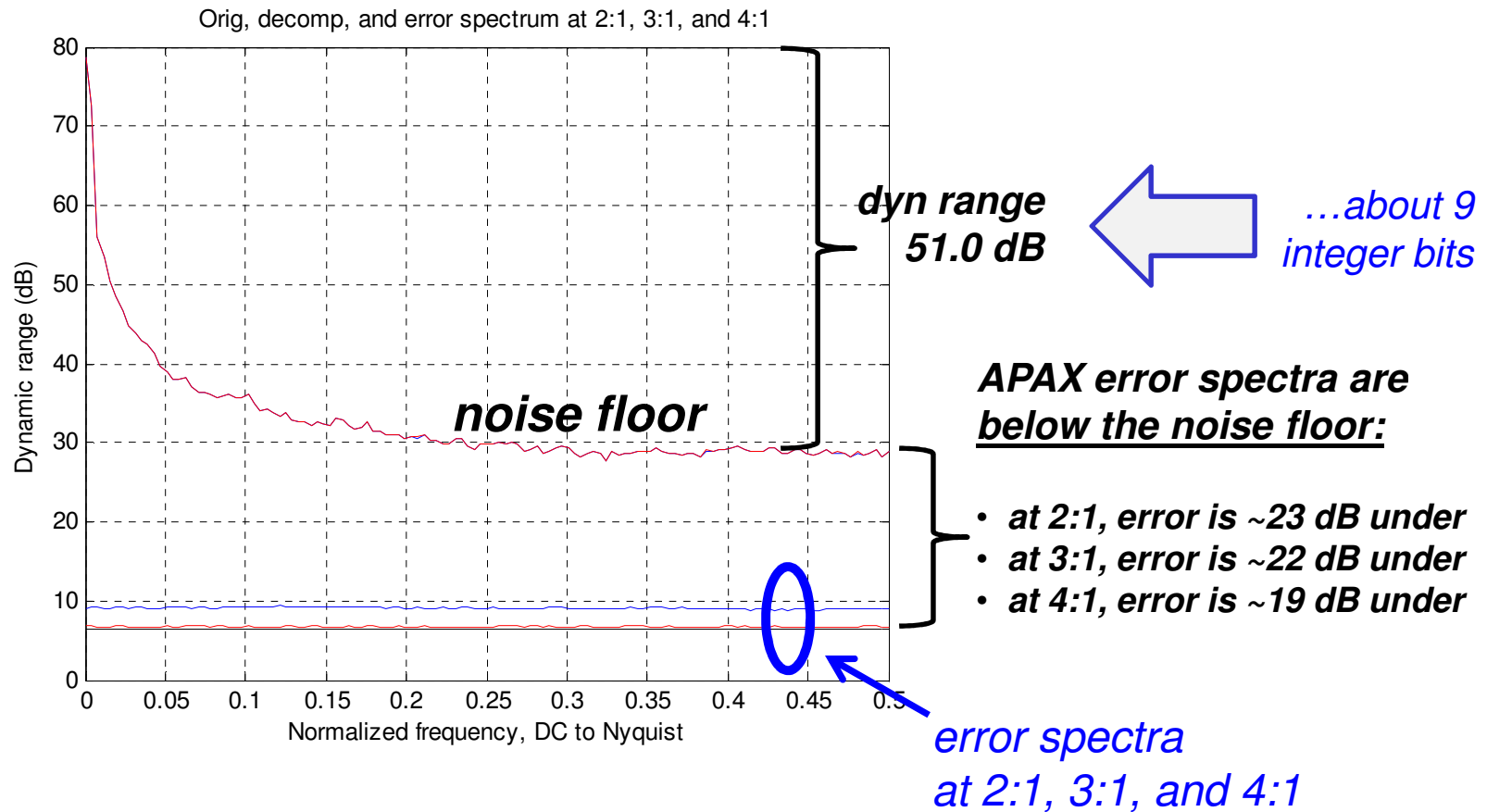
Error after 2:1 compression:
{-0.164, +0.189}, or
{-0.02%, +0.02%}
…vs. accuracy of ±3 psi ?
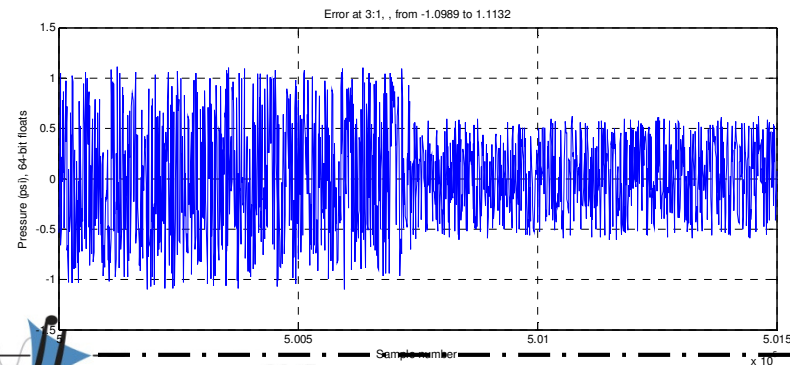
Error after 3:1 compression:
{-1.099, +1.113}, or
{-0.13%, +0.14%}
…vs. accuracy of ±3 psi ?

…simply the bits that matter®

samplify
SYSTEMS®

Example: Spectra at 2:1, 3:1, and 4:1

Orig, decomp, and error spectrum at 2:1, 3:1, and 4:1

(Plot: Dynamic range (dB) vs. Normalized frequency, DC to Nyquist)

- Y-axis: Dynamic range (dB), 0 to 80
- X-axis: Normalized frequency, DC to Nyquist, 0 to 0.5

noise floor

dyn range 51.0 dB

…about 9 integer bits

APAX error spectra are below the noise floor:

- **at 2:1, error is ~23 dB under**
- **at 3:1, error is ~22 dB under**
- **at 4:1, error is ~19 dB under**

error spectra at 2:1, 3:1, and 4:1

…simply the bits that matter®

Lossless: OK. But LOSSY ? Yes … *2:1 to 6:1*

INTEGERS

| Signal Type	Lossless C. R.	Fixed rate C. R. & quality metrics
Wireless baseband (3G, LTE)	1.2:1 – 1.5:1	1.6:1 – 2.3:1 EVM, PCDE, ACLR
Wireless RF (3G, LTE)	2:1 – 3:1	3:1 – 5:1 EVM, PCDE, ACLR
Computed tomography	1.6:1 – 2.7:1	3:1 – 4.5:1 Radiologists & SSIM
Ultrasound (ADC)	1.5:1 – 2:1	2:1 – 3:1 Sonographers & SSIM
Ultrasound (beamformer)	2:1 – 3:1	3:1 – 4:1 Sonographers & SSIM
Images & video	1.5:1 – 2:1	2:1 – 3:1 viewers, PSNR, SSIM
Oscilloscope (SerDes & LVDS)	1.3:1 – 2:1	2:1 – 4:1 BER, rise/fall time
Radar	2:1 – 3:1	3:1 – 5:1 p_d, p_{fa}
High-speed Imaging	1.4:1 – 2:1	2:1 – 4:1 Photographers, SSIM
Motion est. frames	1.6:1 – 2.5:1	2:1 – 6:1 Viewers, SSIM
Graphics textures	1.8:1 – 2.8:1	3:1 – 8:1 Gamers, SSIM

FLOATS
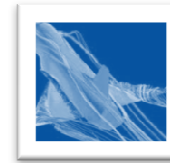
Aerospace/ Defense

Automotive
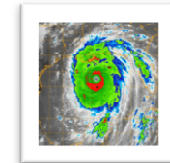
Weather Forecasting
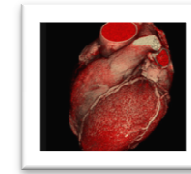
Medical Imaging

Financial Analysis

Genomics

GPU Application	Fixed Rate Comp Ratio
Hydrodynamics	3.3:1
Car crash simulation	1.5:1
Computational fluid dynamics	1.5:1
Financial models	1.45:1
Image recognition	4.2:1

…simply the bits that matter®

APAX: a Numerical Compressor for I/O

Ints, floats, & HD frames

(from faster memory)

Number aligner

Floating-pt Processor

Image Pre-Processor

Signal Monitor

Redundancy Remover

Bit Packer

Packet Header Generator

APAX Control

Compressed APAX packets

(to slower memory)

User chooses:

- **lossless**
- **fixed rate (0.05)**
- **fixed quality (0.5 dB)**

DMA descriptor

~100k gates

Xilinx Zynq FPGA latency ~30 clocks
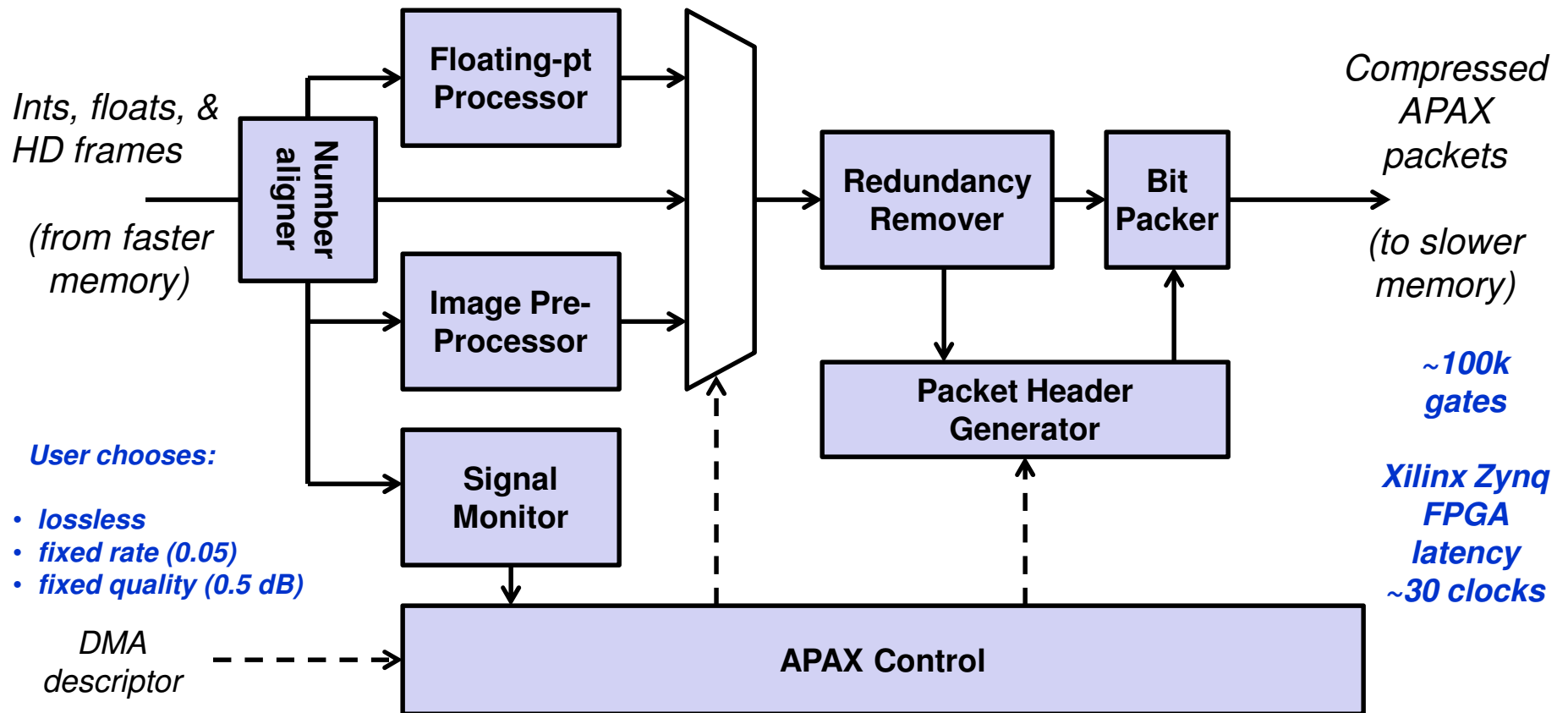
Memory Wall Take-away:
Compression is easier than designing faster DDR

…simply the bits that matter®

samplify SYSTEMS®

APAX Software Products

Software Product	Details
APAX library (.dll, .so)	APAX_profile( ) APAX_compress( ) APAX_decomp( )
APAX MPI extensions	MPI_WRITE_APAX_C( ) MPI_READ_APAX_D( )
APAX file system extensions	fwrite_apax_c( ) fread_apax_d( )
APAX for CUDA & OpenCL	cudaDMAcustom( )

©201

11
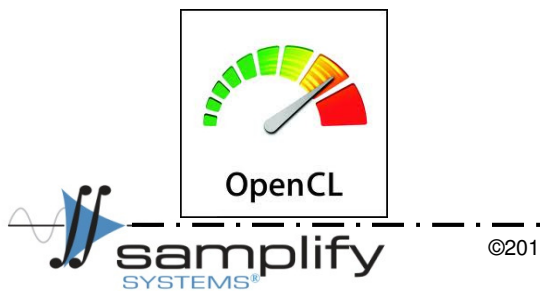
To try APAX on your datasets…

Samplify Systems, Inc: Intellectual Property Solutions for Data Acquisition and Data Processing Bottlenecks

Samplify Systems, Inc. is a provider of intellectual property solutions which solve data acquisition and data processing bottlenecks. Samplify's Prism Compression addresses data acquisition bottlenecks in industrial, scientific, medical markets, while Samplify's APAX technology accelerates software applications which are performance-limited by I/O, storage, or memory bandwidth bottlenecks in the computing, consumer, and mobile device markets.

Samplify is a private company based in Silicon Valley, with backing from leading venture capitalists, Charles River Ventures and Formative Ventures and strategic partners Schlumberger, IDT, and Cosmo.

Samplify's APAX Technology Featured on EngineeringTV

From DesignWest 2012, Bill Wong interviews Samplify's Founder and CTO, Al Wegener, on the company's APAX technology for High Performance Computing, Big Data, and consumer electronics applications.

Al Wegener

Samplify Systems
Santa Clara, CA

awegener@samplify.com
408-260-4000

www.samplify.com

…simply the bits that matter®

12