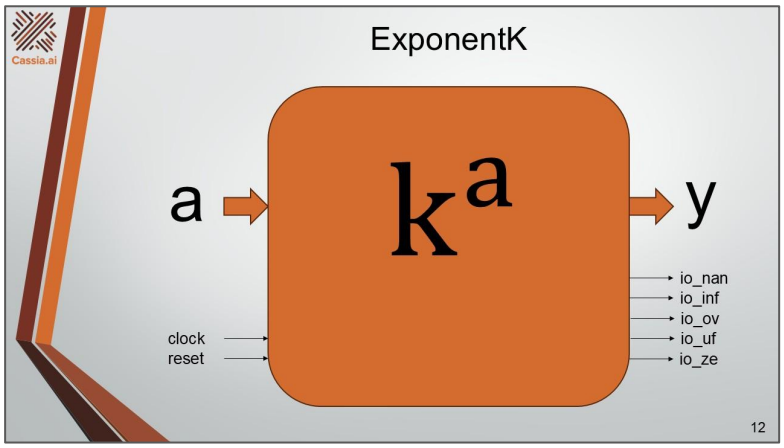
Cassia.ai

Exponent (Base K)

Exponentiation in Base-K

Overview

This highly accelerated and low-gate-count mathematical function is perfect for building or accelerating custom AI accelerators



Key Features

- y=K^x
- Fast and low power Base K Exponentiation, highly accelerated, for standalone use or integration into higher-order functions. This function is fast even in a single cycle, or may be pipelined for even higher speed.
- Uses advanced mathematics to improve performance, power and area
- Single-cycle or pipelined option

Benefits

- May be accelerated 10X-20X depending on precision and system requirements
- Significant gate count reduction compared to traditional math
- Uses standard input and output formats - user definable to be IEEE Floating Point, Brain Float, Tensor Float, or custom implementations

Applications

- Use in AI accelerators for improved performance, power and area
- All RTL implementation with no custom circuits - portable to any foundry or FPGA
- Cassia.ai functions are designed to replace your existing mathematical functions, allowing you to use pre-trained models without additional training, just instantiate and recompile verilog. They also work for training workloads!

Available Standard Precisions

BF16(8E7M), TF32(8E10M), FP19(8E10M), FP32(8E23M)

Precisions with extra lead time

FP16(5E10M), FP4(2E1M), FP6(3E2M), FP8(3E4M), FP8(4E3M), FP8(5E2M), FP16(6E9M), FP64(11E52M), Others by request

Maturity: Available on request

Deliverables

- Verilog RTL
- C++ or CUDA on request
- PyTorch and Docker containers
- Full documentation
- Licenses to Cassia.ai patents

Email Contact: info@cassia.ai