On-Premise FPGA Demo

https://fires.im

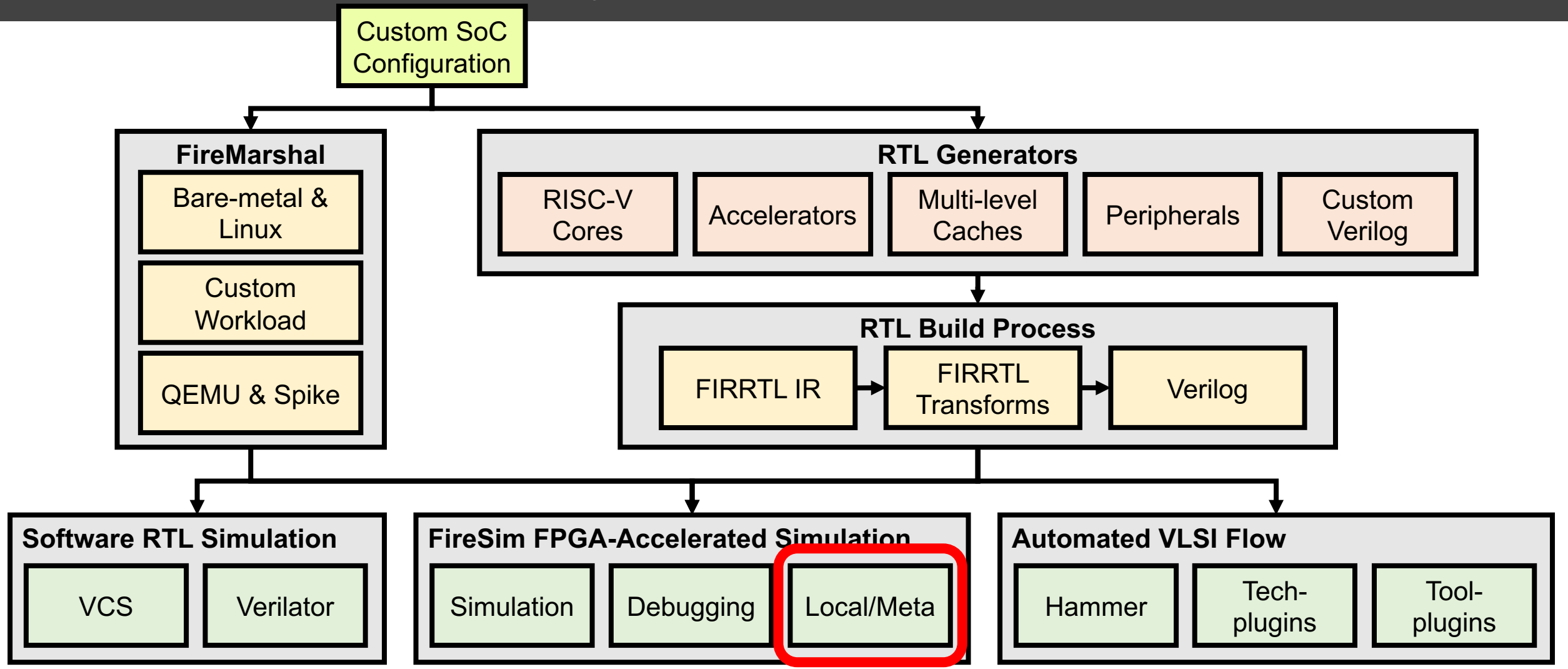
@firesimproject

HPCA Tutorial 2023

Abraham Gonzalez

Berkeley Architecture Research

Tutorial Roadmap



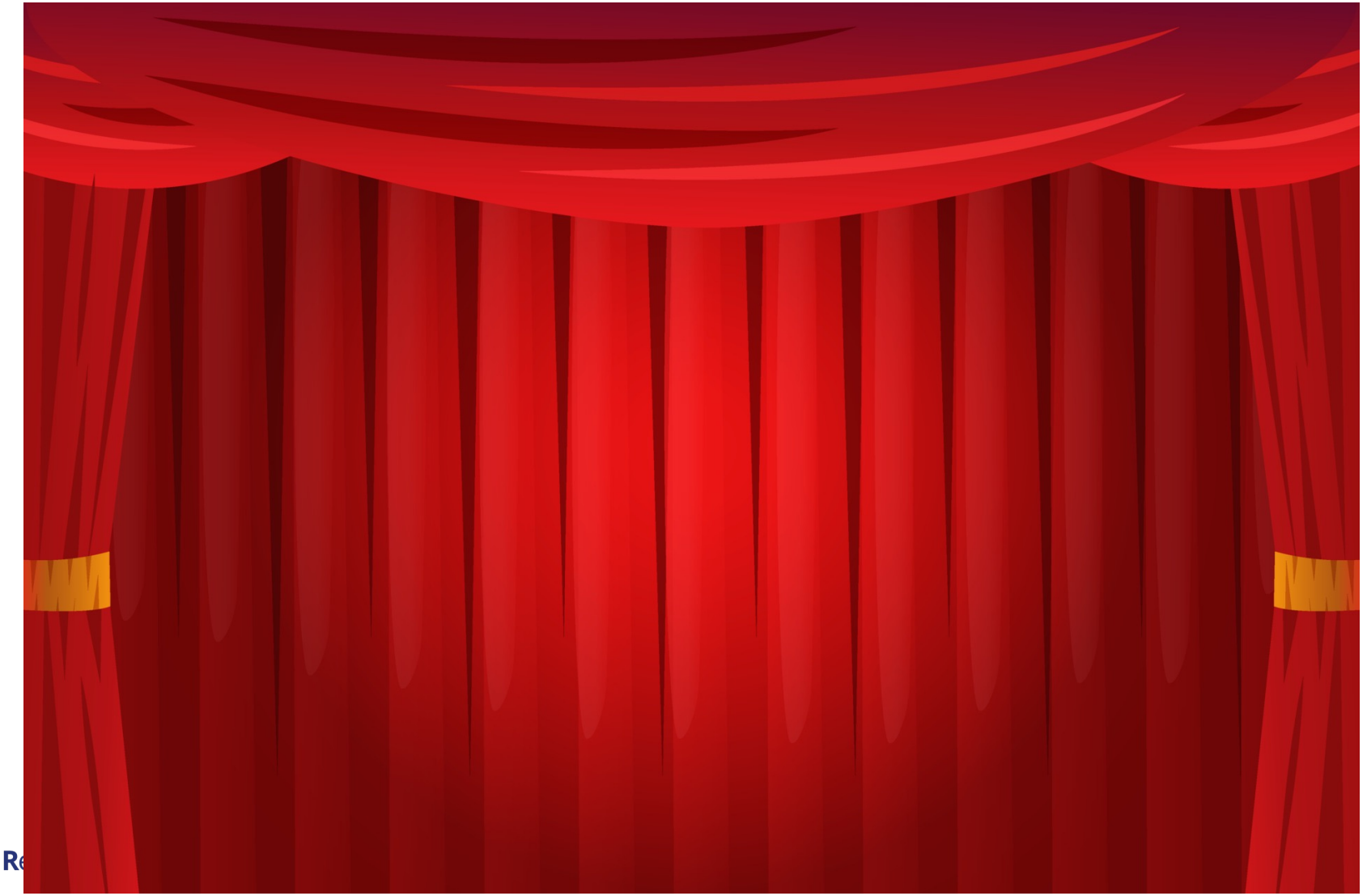
Two new FireSim features!

- Local FPGA support!
  - Re-architecting of FireSim Manager, Drivers, and Shells to support adding new FPGA and host platforms
    - Now supports Xilinx Alveo XRT-enabled FPGAs (e.g. U250)
- Distributed Meta-simulations
  - Early-stage accelerator development requires running many parallel verilator/vcs sims, but these are traditionally accessed via Make-system in Chipyard/FireSim
  - FireSim manager now supports distributing metasims using the same machinery as distributing FPGA simulations, on both EC2 and local machines
    - Same user-interface for workload/job specification/mapping, constructing heterogeneous systems, running sims, and collecting outputs (now including waveforms)

Two new FireSim features!

- **Local FPGA support!**
  - **Re-architecting of FireSim Manager, Drivers, and Shells to support adding new FPGA and host platforms**
    - **Now supports Xilinx Alveo XRT-enabled FPGAs (e.g. U250)**
- Distributed Meta-simulations
  - Early-stage accelerator development requires running many parallel verilator/vcs sims, but these are traditionally accessed via Make-system in Chipyard/FireSim
  - FireSim manager now supports distributing metasims using the same machinery as distributing FPGA simulations, on both EC2 and local machines
    - Same user-interface for workload/job specification/mapping, constructing heterogeneous systems, running sims, and collecting outputs (now including waveforms)

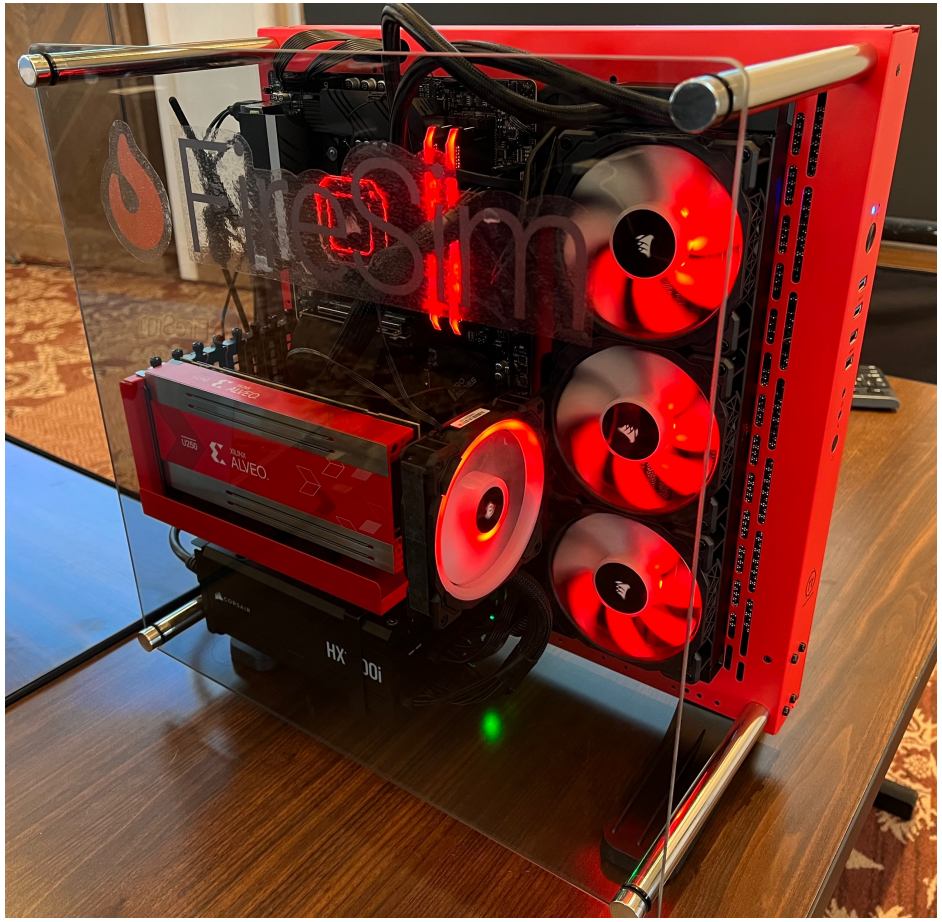
Drumroll…



FireTower v1 Machine Specs

- Intel Core i7 13700K
  - Liquid cooler (w/RGB)
- 32 GB DDR4 (w/RGB)
- **Xilinx Alveo U250 (active)**
- Motherboard spec'd for:
  - 2 U250 + GPU
  - OR
  - 3 U250
- 1500W PSU to support multi-FPGA/GPU
- Thermaltake Core P3 Red Case
- Ubuntu 18.04

$1500 without FPGAs or GPUs



What are we running?

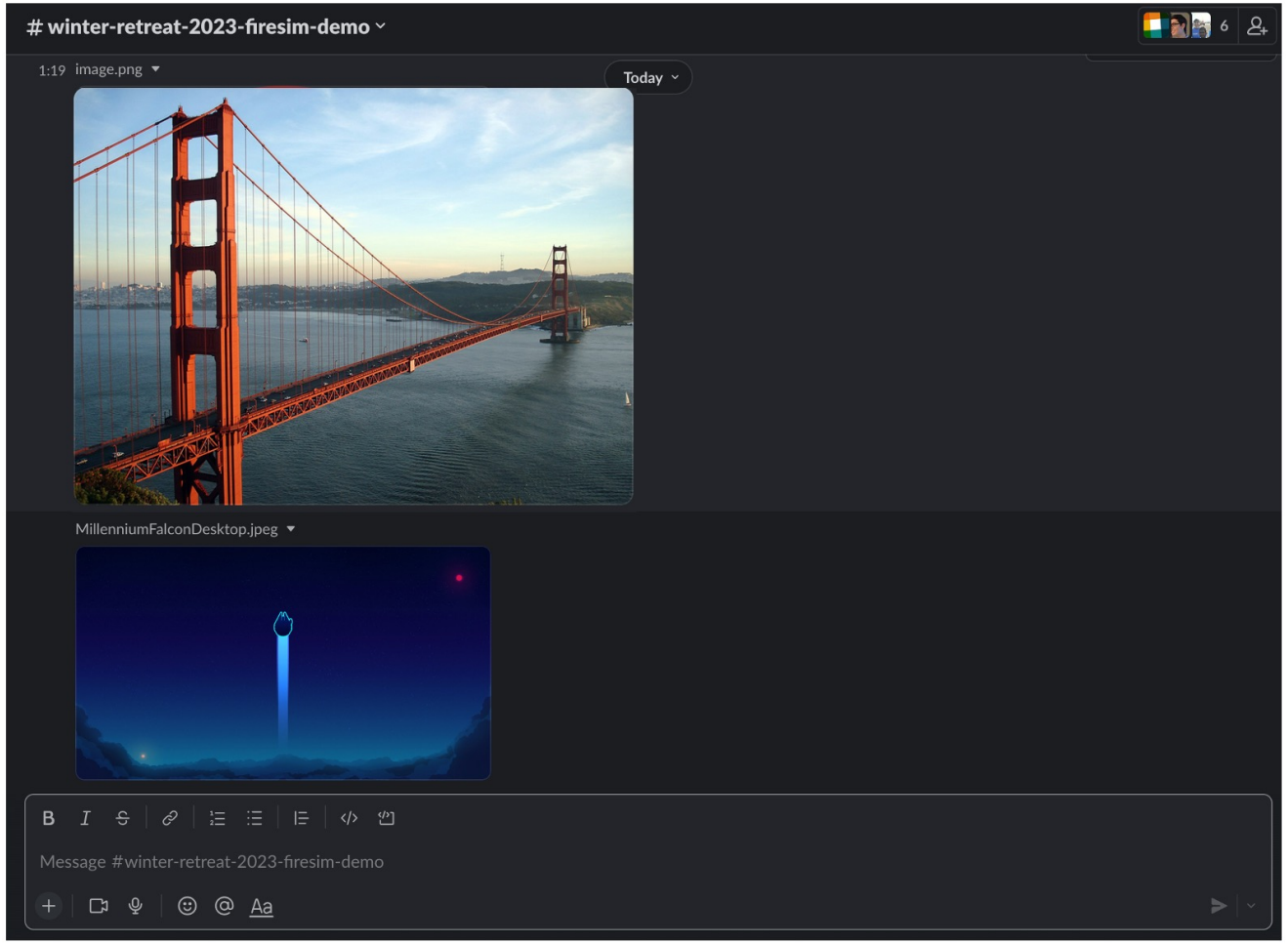
Running **MobileNet** image recognition

using the **Gemmini DNN accelerator**

on a **Chipyard Rocket-based SoC**

simulated with **a FireSim U250 FPGA-enabled desktop**

Demo Lifetime



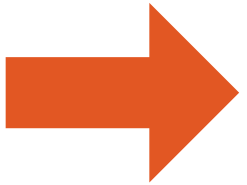
Demo Lifetime

Demo Lifetime

Demo Lifetime

Demo Lifetime

Demo Lifetime



Convert to .png

Demo Lifetime

Convert to .png

PyTorch

Image Preprocessing

Demo Lifetime

Convert to .png

PyTorch

Image Preprocessing

Compile target Gemmini C binary with image

Demo Lifetime

FireSim

Run FireSim
Rocket + Gemmini
simulation

Compile target
Gemmini C binary
with image

Demo Lifetime

FireSim

Run FireSim Rocket + Gemmini simulation

Compile target Gemmini C binary with image

Reset + Flash FPGA

Copying collateral

Running simulation

Demo Lifetime

FireSim

Run FireSim
Rocket + Gemmini
simulation

"It's a **bridge**"

Compile target
Gemmini C binary
with image

Demo Lifetime

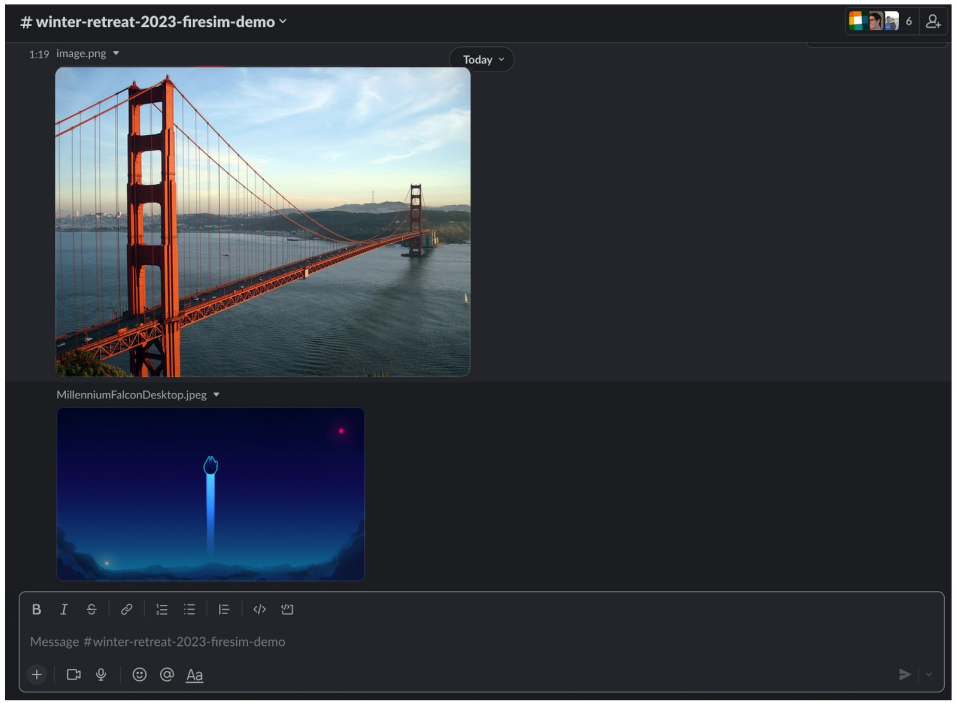
FireSim

Run FireSim
Rocket + Gemmini
simulation

"It's a **bridge**"

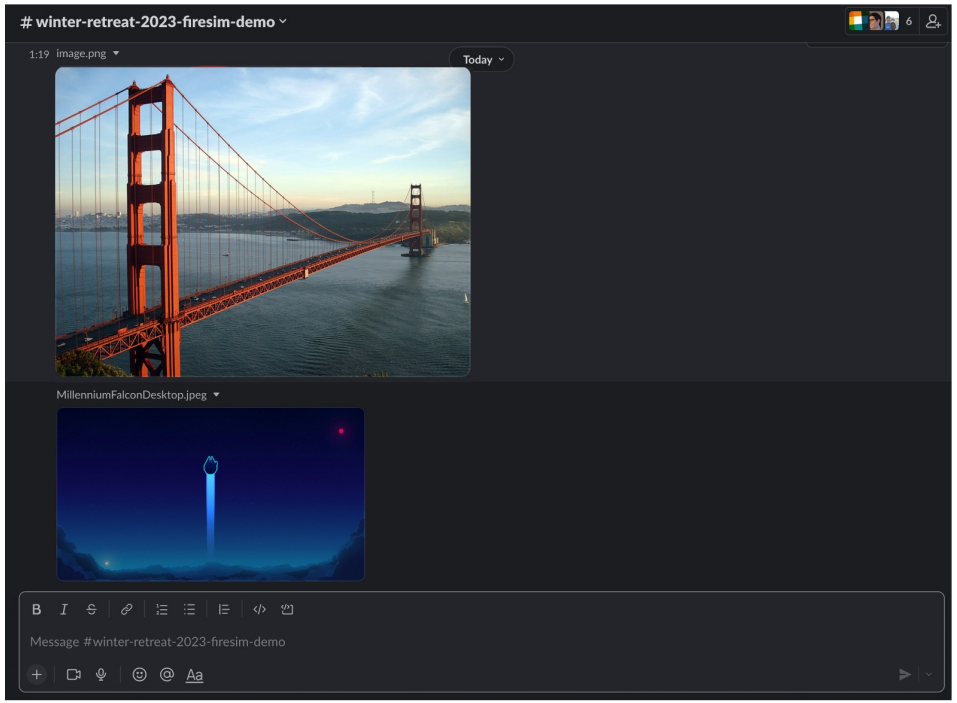
Compile target
Gemmini C binary
with image

Demo Lifetime



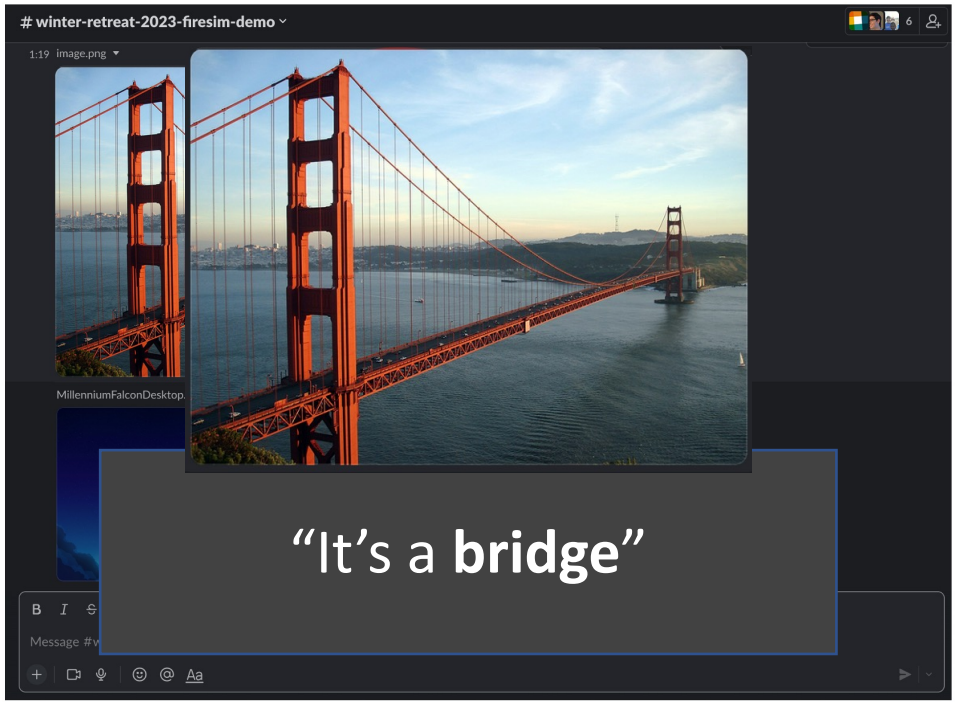
"It's a **bridge**"

Demo Lifetime

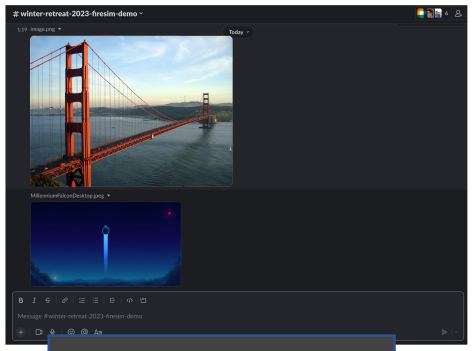


"It's a **bridge**"

Demo Lifetime



Let’s Run the Demo!



Slack UI

PyTorch

Image Preprocessing

Compile target Gemmini C binary with image

FireSim

Run FireSim Rocket + Gemmini simulation

“It’s a **bridge**”

Try it out yourself!

- You were added to a Slack called `firesim-chipyard-tutorial`
  - Done if you filled out the intro tutorial Google form
  - Link: fsim-cy-tutorial.slack.com
- Add your own photos in `demo-input` channel
- See the output in `demo-output` channel

!!! Reminder the photos uploaded are public !!!

On-premise FPGA support now available!

- High-level of automation/reproducibility enabled by FireSim on AWS F1 cloud now extended to local/on-prem FPGAs:
  - Went from new machine with no FPGA attached to working FPGA-accelerated simulation in 1 hour and 40 mins
- Use existing FireSim features at-scale and locally!
  - Cycle-accurate simulation
  - Debugging
    - Integrated logic analyzers, trace dumps, synth. assert/prints, co-simulation
  - Software support
    - FireMarshal workload management
  - … and more!