



Fault testing a synthesizable embedded processor at gate level using Virtex Ultrascale/Ultrascale+ FPGA emulation

Tom J. Mannos[†]

Rad-Hard/Trusted ASIC Products
Sandia National Labs
Albuquerque, NM
tjmanno@sandia.gov

Brian Dziki

Information Assurance Research
Department of Defense
Fort G. G. Meade, MD
bjdziki@tycho.ncsc.mil

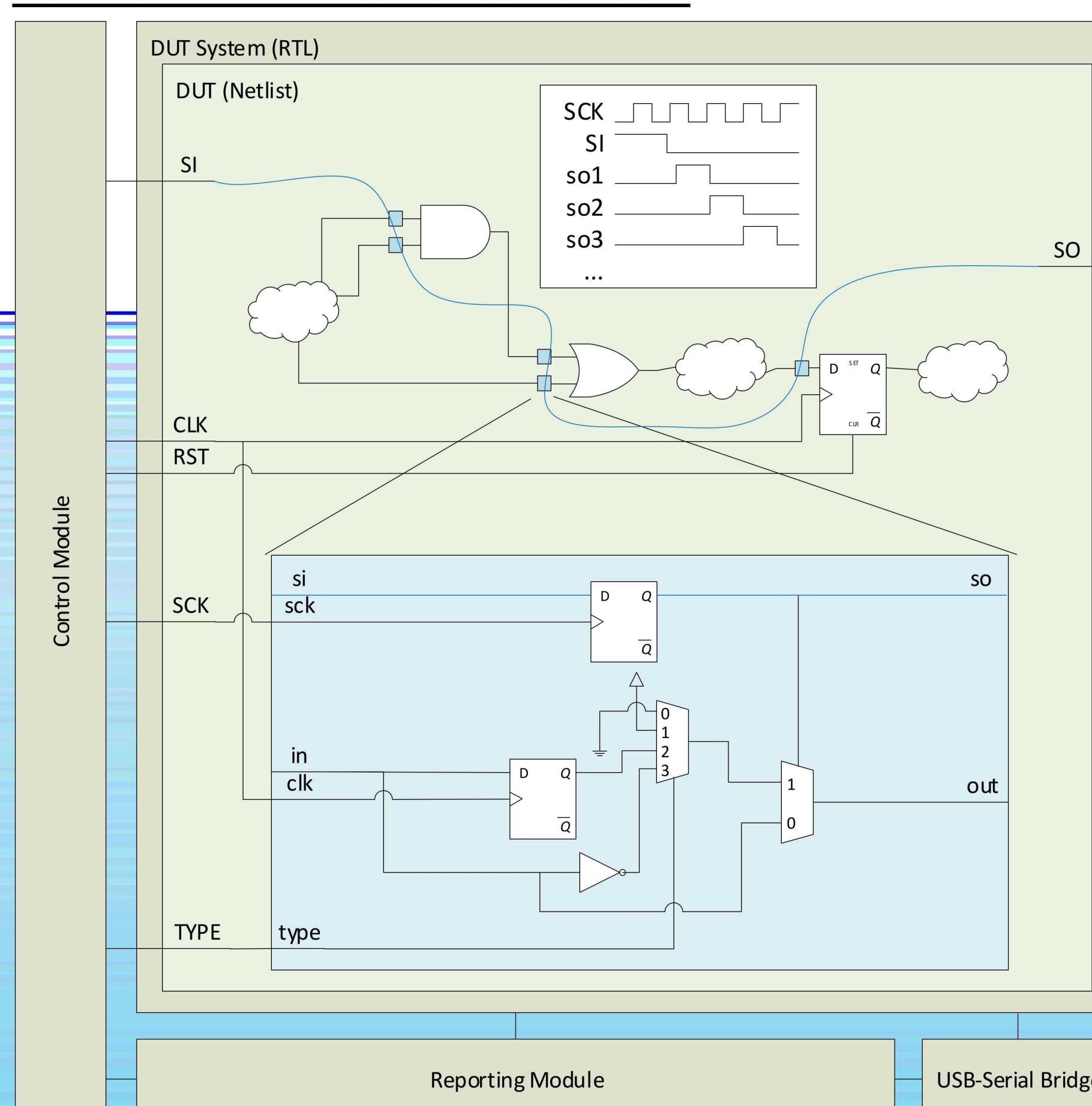
Moslema Sharif

Rad-Hard/Trusted ASIC Products
Sandia National Labs
Albuquerque, NM
msharif@sandia.gov

Abstract

While several applications exist to fault-test software at the abstract level, gate-level fault testing has traditionally been limited to injecting faults into small, dedicated circuits, due to the computational time required for gate-level simulations and the logic resources (LUTs and registers) required for FPGA emulation. In our work, we leverage the extensive resources of the Ultrascale and Ultrascale+ platforms to systematically test fault on all gate-level inputs of an ASIC or FPGA implementation of an embedded processor running a user application. Using a configurable, synthesizable saboteur circuit connected using a scan chain and controlled with a state machine, we tested four different types of faults on each logic gate of the LEON3 CPU running an AES 256 application. The entire process took just under two hours, a 7200X speedup from gate-level simulation. Though neither the hardware nor software employed fault mitigation techniques, we were surprised to discover multiple stuck-at-0, stuck-at-1, and delay faults that resulted in total or partial key and plaintext leakage through the communications serial port. Overall, we identified 22 unique faulty behaviors, four of which involve some form of crypto or memory leakage. These occurred in roughly the same proportions in ASIC and FPGA netlists of the LEON3, suggesting robustness of the analysis to different implementations. Of the four static fault types tested, delay faults were the most effective at uncovering behaviors of concern.

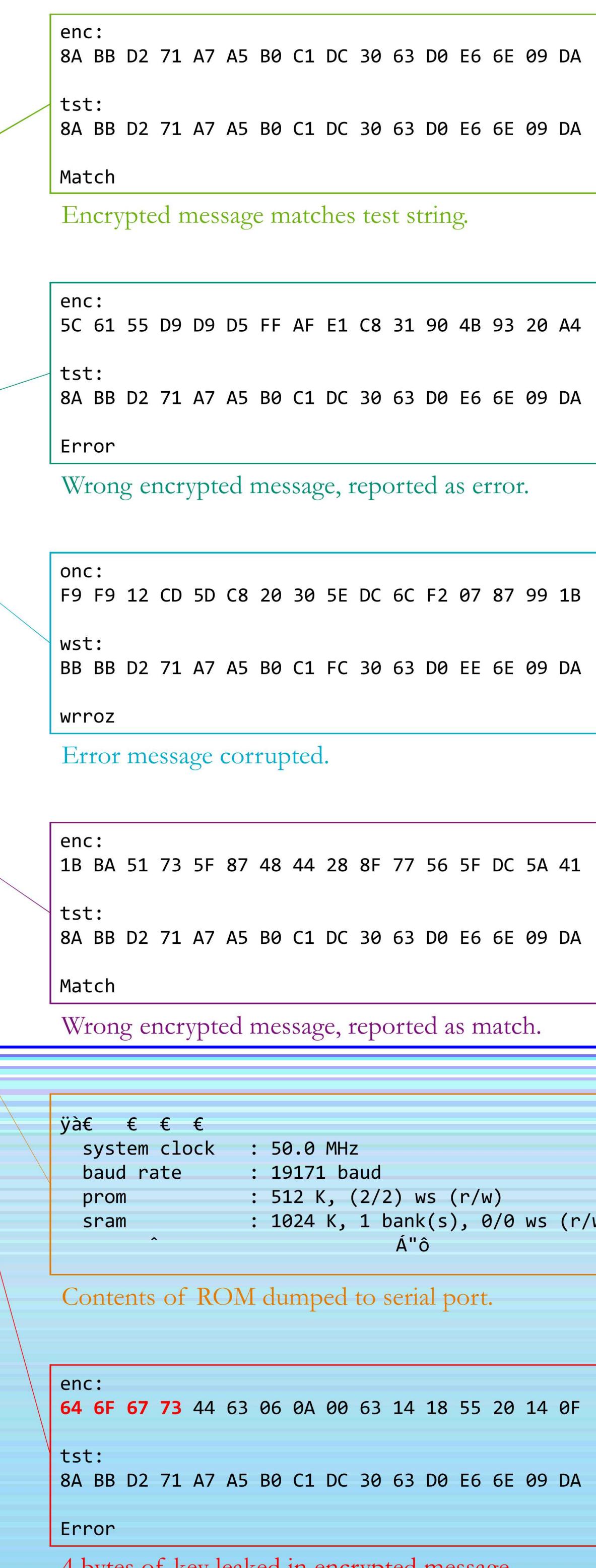
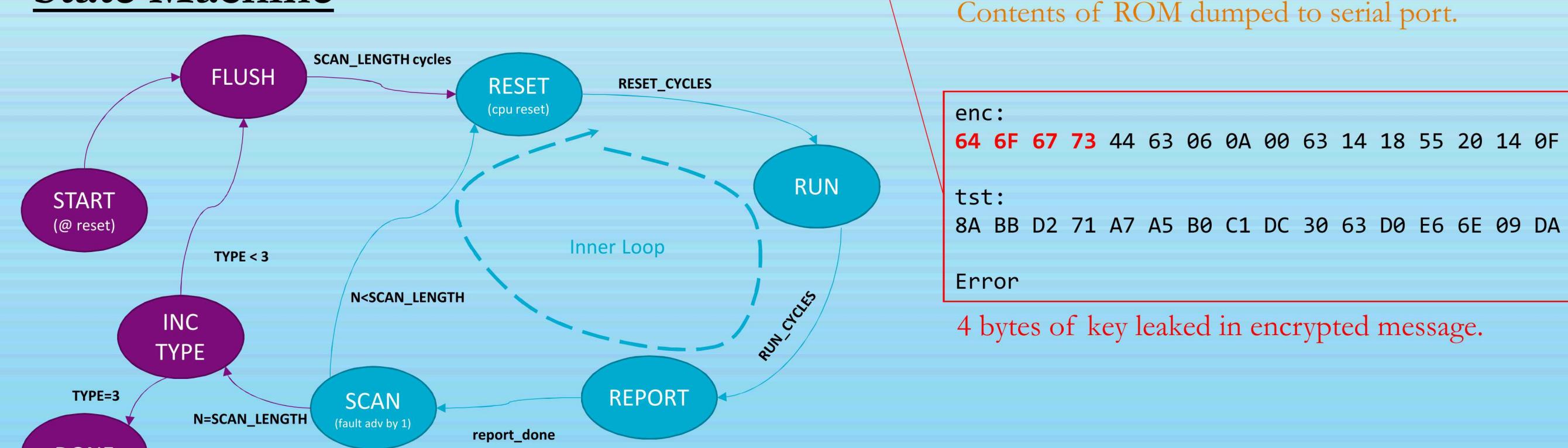
Scan chain and saboteur circuit



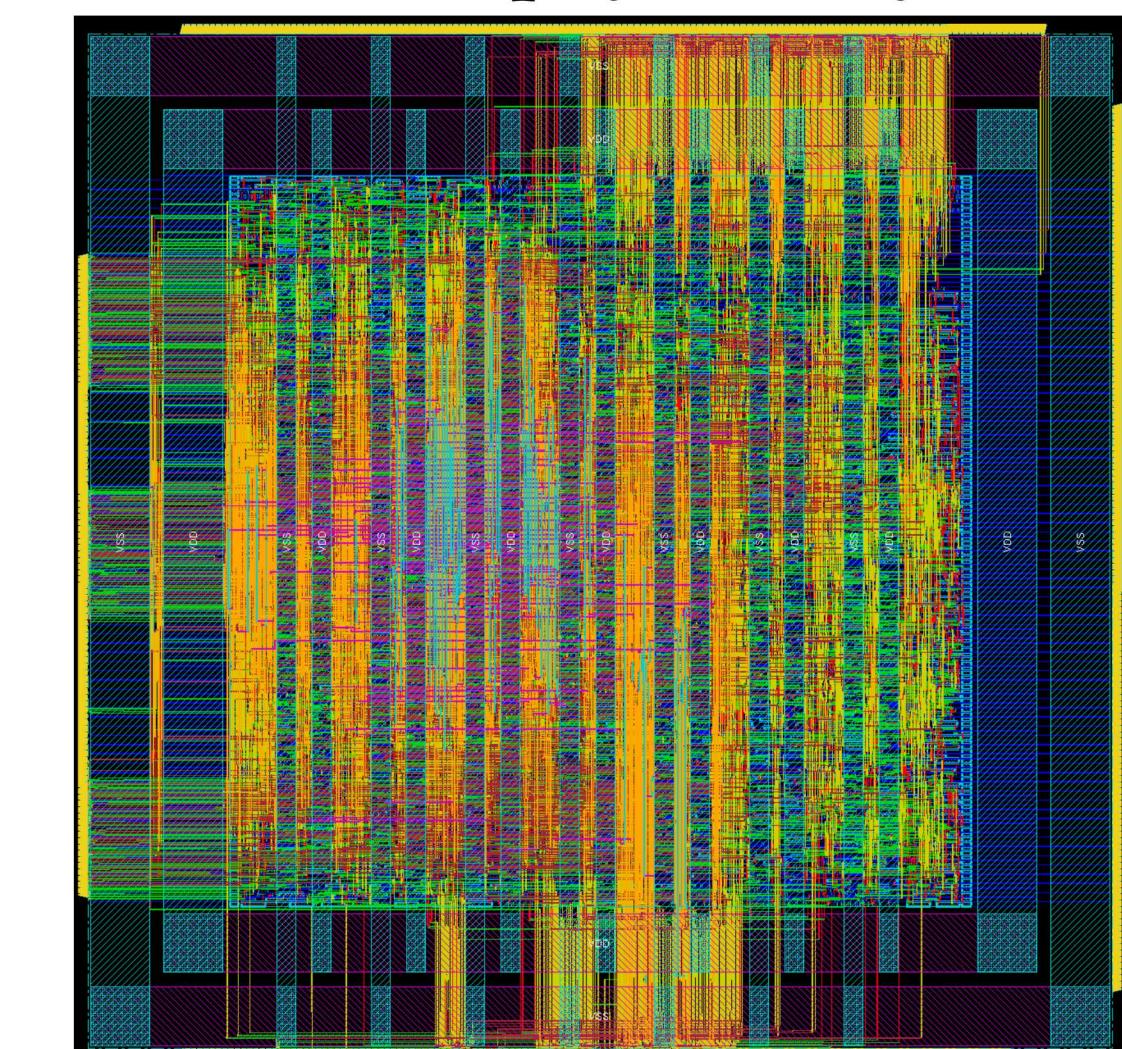
Fault behaviors observed

Severity/Behavior	Stuck 0	Stuck 1	Delayed	Inverted
0: output as expected; correct ciphertext received and validated				
perfect/match	13,732	13,357	15,249	9,681
1: partially corrupt; but correct ciphertext received and validated				
correct/match	416	377	514	315
repeat/match	10	6	12	8
2: wrong ciphertext received, but correctly reported as invalid				
agree/error	--	4	--	2
corrupt/error	18	10	16	4
disagree/error	615	400	776	275
3: no output, truncated or corrupt output				
agree/none	--	--	17	--
correct/error	77	79	73	115
correct/none	--	--	21	--
corrupt	212	232	182	198
disagree/none	8	8	12	2
empty	12,094	12,662	9,599	15,956
repeat/pattern	300	297	307	294
short	3,106	3,109	3,710	3,754
4: wrong ciphertext received, incorrectly reported as valid				
agree/match	60	56	48	34
corrupt/match	--	--	3	2
disagree/match	2	10	2	4
match/error	--	10	2	2
5: more output than expected, contents indiscernible				
long	2	8	24	10
6: output contains strings from memory				
leak/other	51	79	104	51
7: output contains strings from key and/or plaintext				
leak/crypto	13	11	43	9
leak/key	--	1	2	--

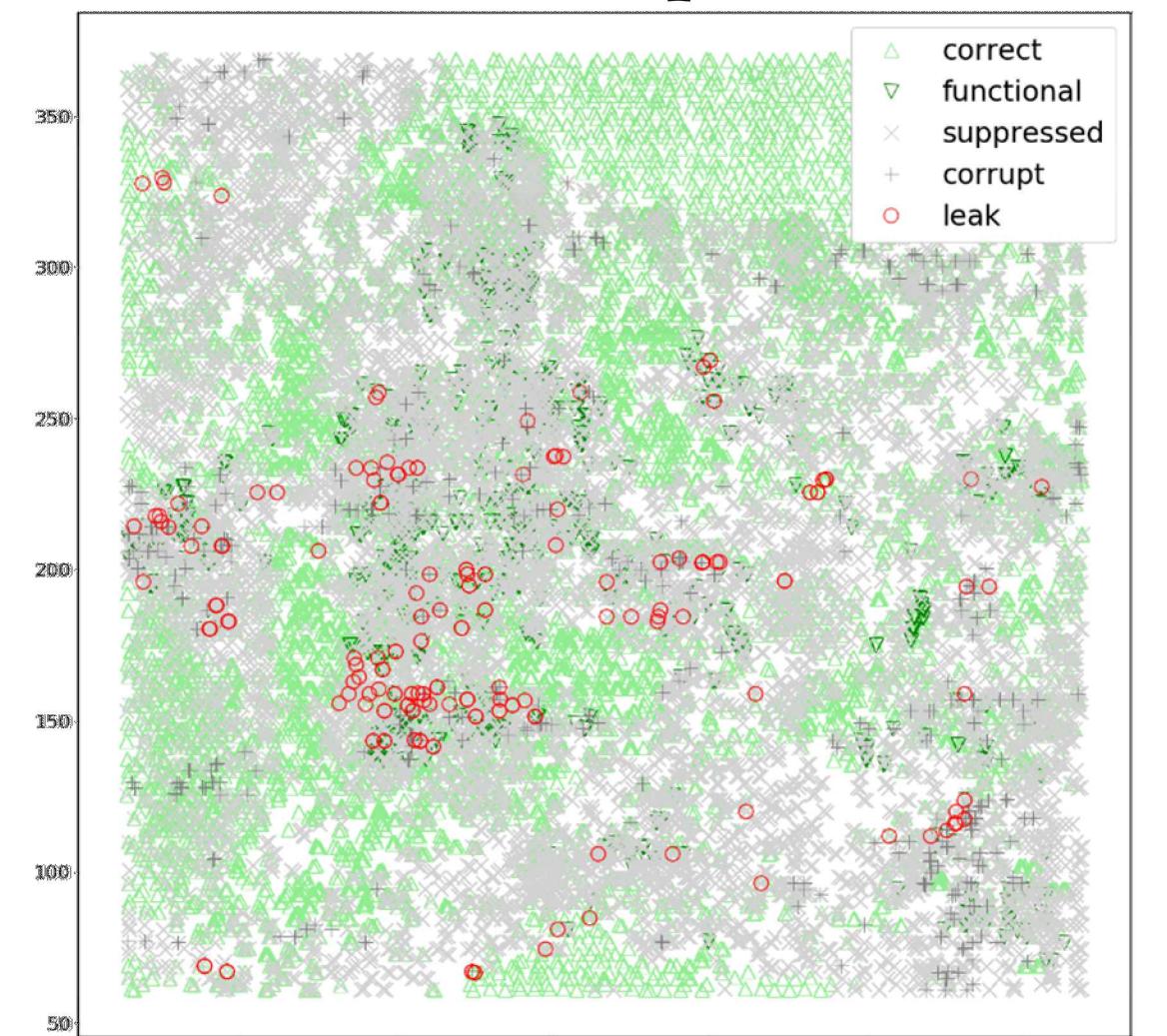
State Machine



ASIC macro physical layout



Location of fault points in ASIC



Scaling to larger processors

Design	Fault points	Simulation time	Emulation time	Register overhead
LEON3 IU	15,384	74 days	2 hrs	32,875
RISC-V Potato	35,264	171 days	4 hrs	149,687
RISC-V Rocket 32-bit	67,385	328 days	6 hrs	278,171
RISC-V Rocket 64-bit	335,834	4.5 years	30 hrs	1,360,598

Conclusions

Of the four types of faults we tested, delay faults proved the most useful in uncovering the widest range of behaviors, at the expense of two registers per fault insertion point rather than one. The inverted logic fault proved to be the least useful, not covering many of the high-severity behavioral categories.

The FPGA emulation platform allowed for fault-testing the LEON3 in 1/7200th the time required for gate-level simulation, even accounting for FPGA implementation and bitstream programming time. The emulated behaviors matched their simulated counterparts, except in 701 cases (2% of faults tested), for which the differences could not be explained. Nevertheless, we saw the same behaviors between simulation and emulation, in nearly equal proportions.

The extreme acceleration will allow testing and characterization of more complex processors and user code, which would be prohibitive to do in gate-level simulation.

Acknowledgments

We would like to thank Russell D. Miller for performing the ASIC synthesis and place-and-route used in this test and Sheng-Hao Huang for his early experiments on the RISC-V cores. We also thank Gerald Zuelsdorf for recommending security behaviors that should be investigated.

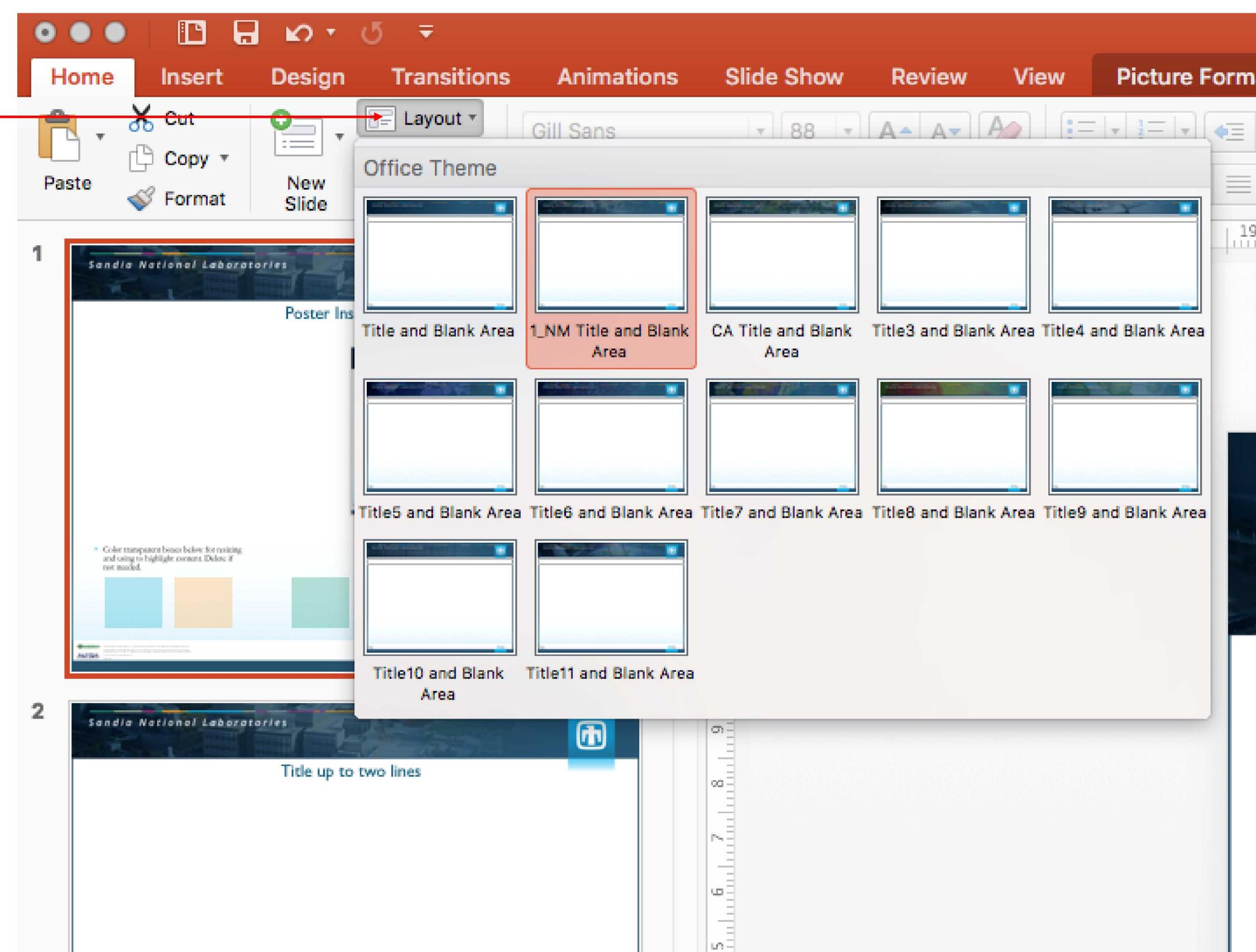
This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.



Title font: Gill Sans MT

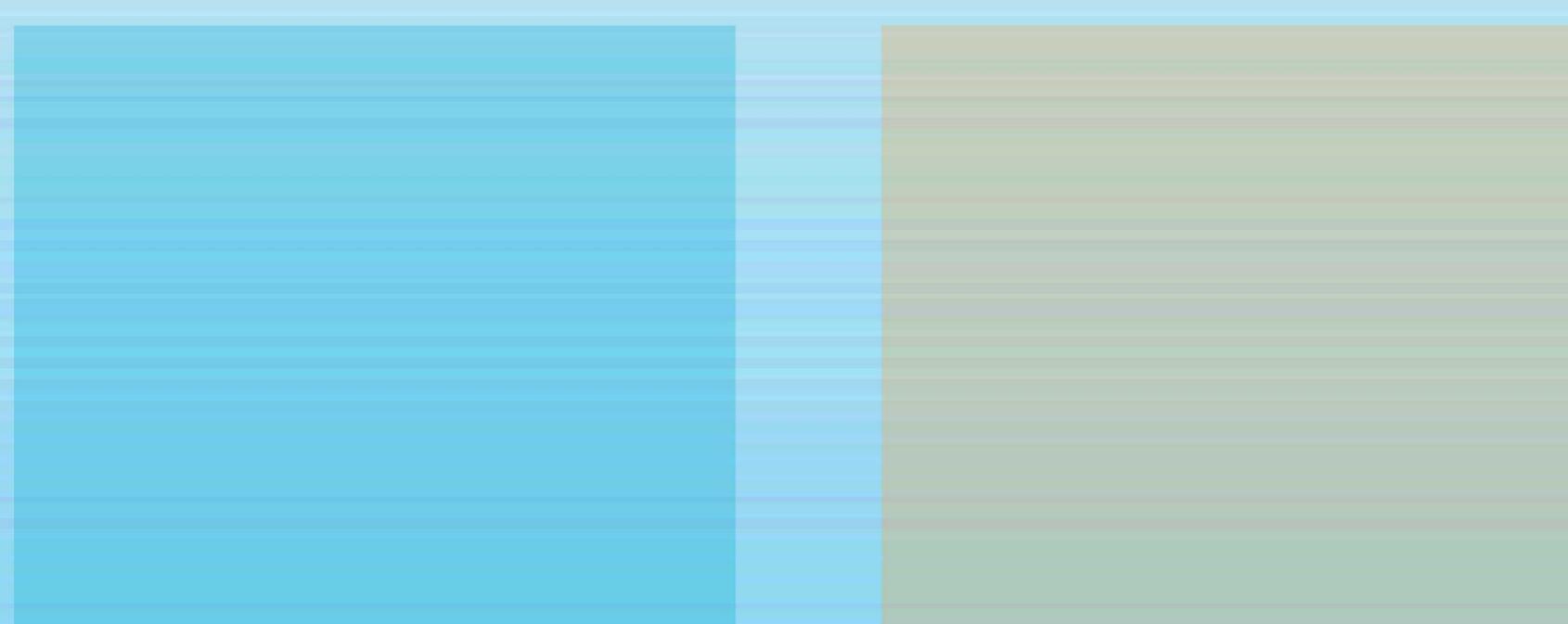
Poster Instructions

Choose from different headers by selecting the “Layout” option in the “Home” tab

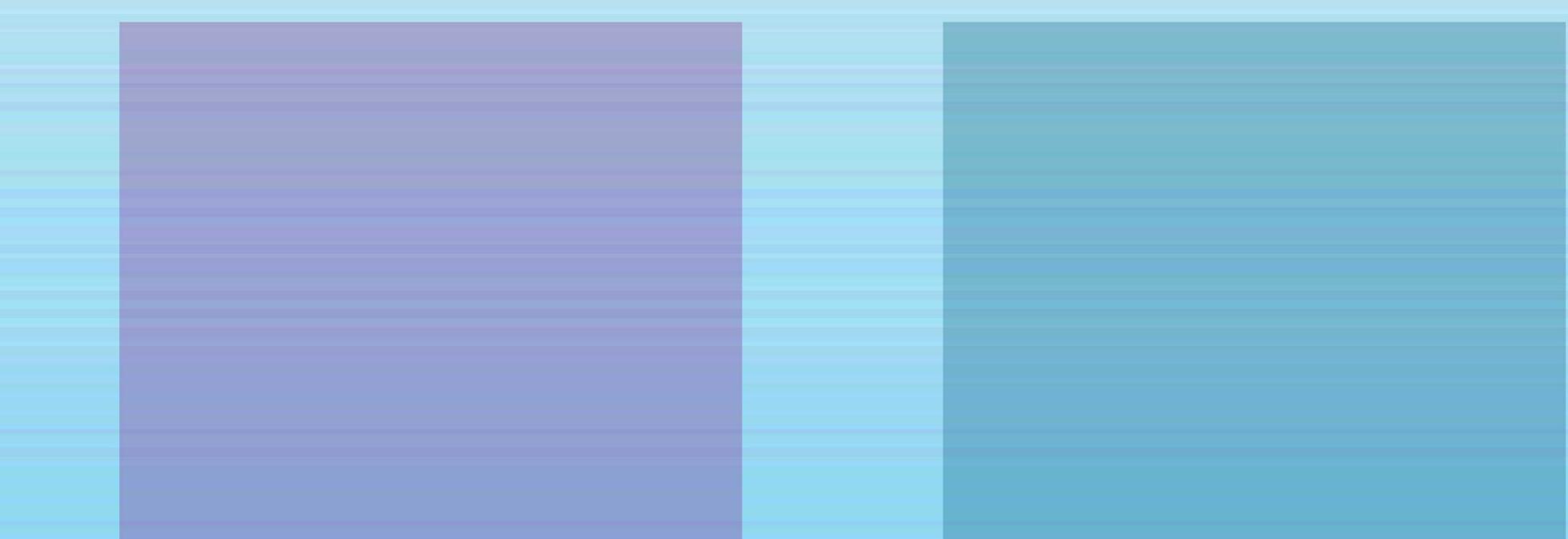
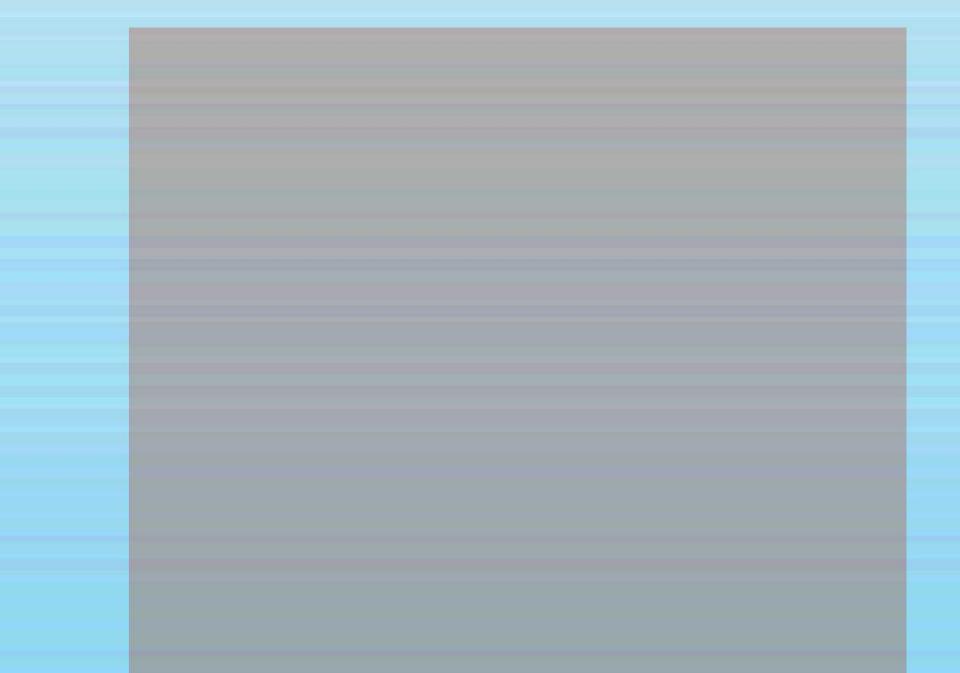


Body text/ support font:
Garamond MT

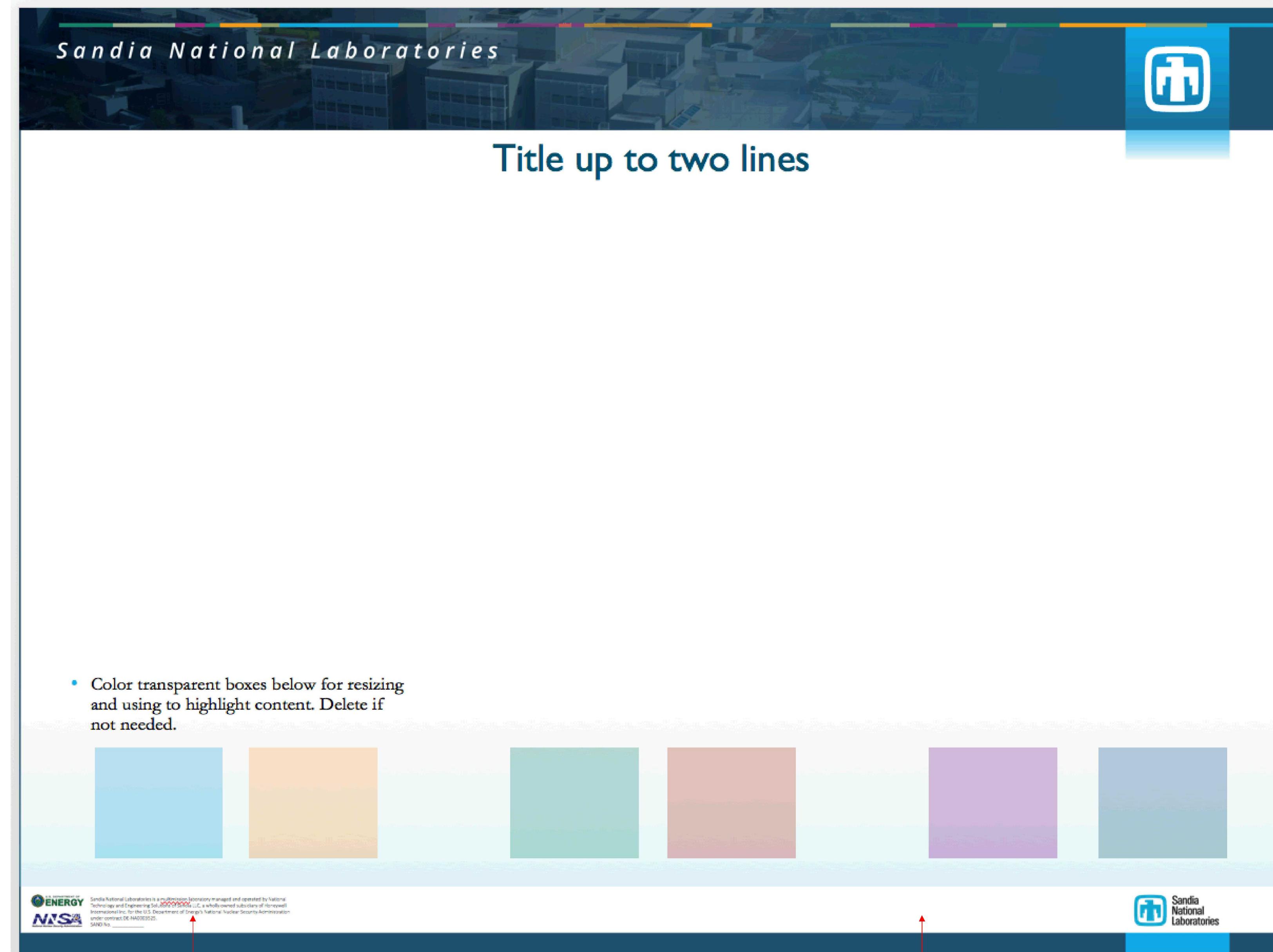
- Color transparent boxes below are for resizing and using to highlight content. Delete if not needed.



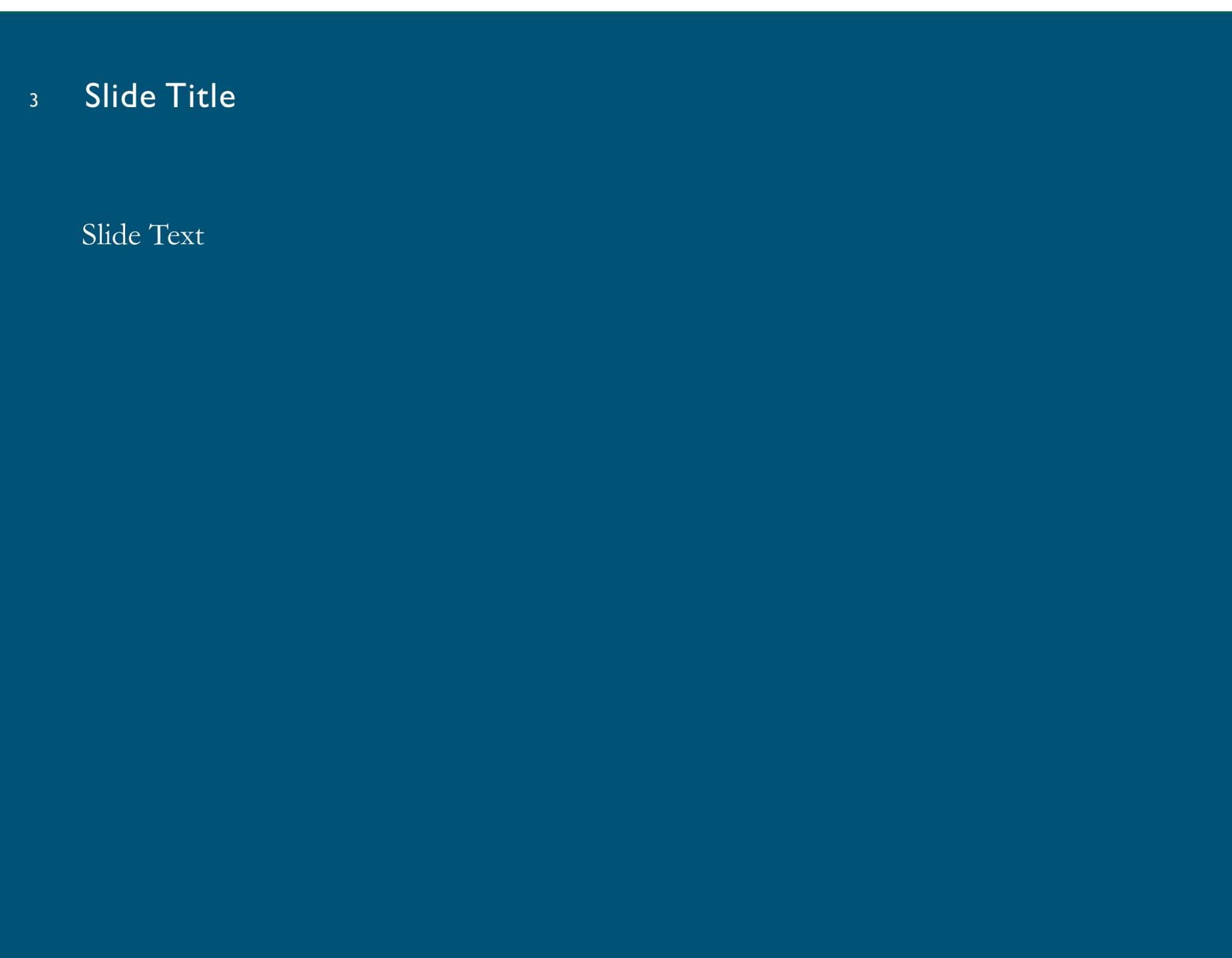
Add Sand Number to the funding statement within the Master Title slide



Additional program/partner logos can be added here

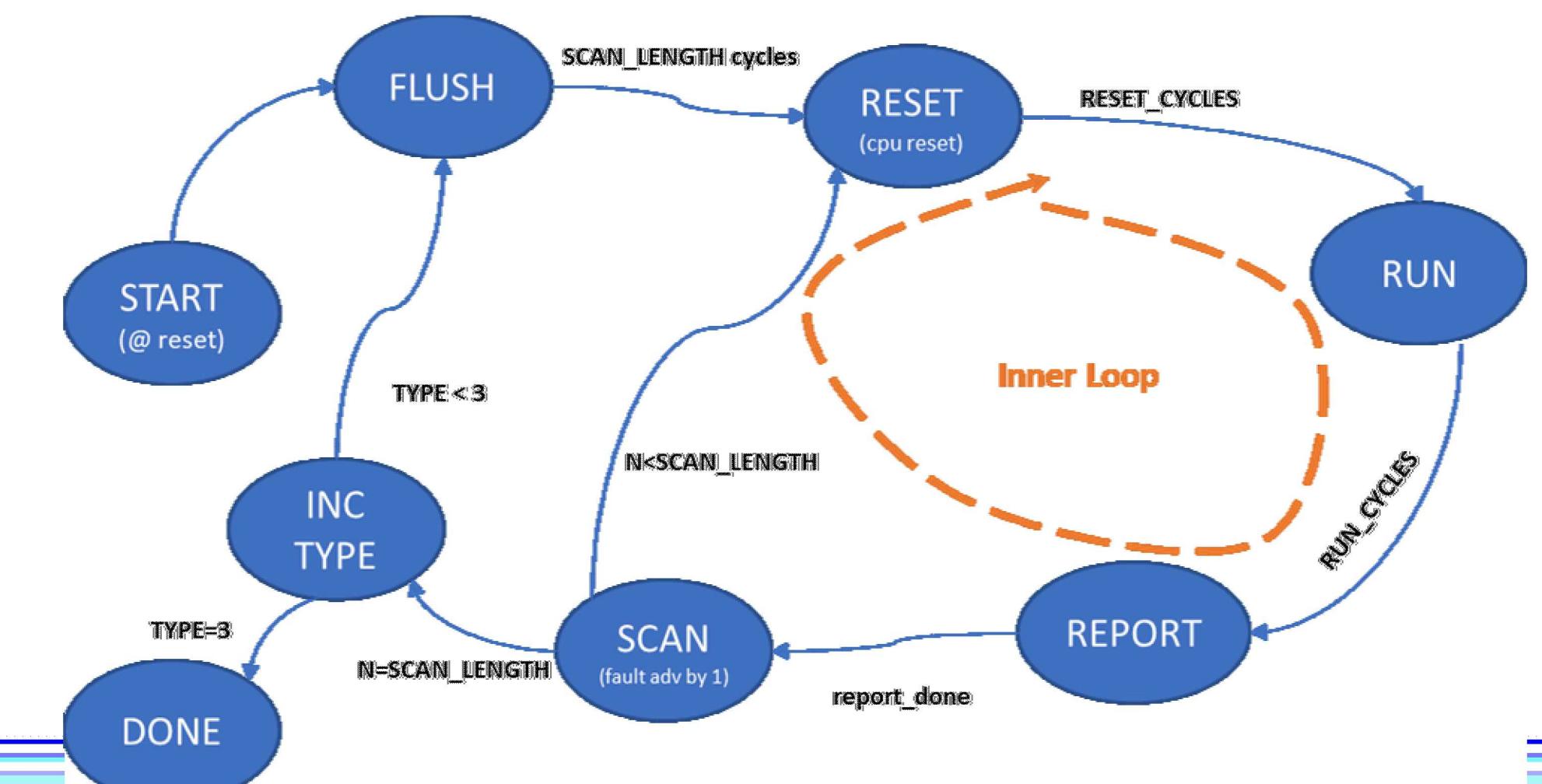


Scratchpad



3 Slide Title

Slide Text



Test	Average test time	Number of faults tested	Total test time
FPGA simulation	10 min	39,930	91 days
ASIC simulation	7 min	30,768	37 days
ASIC emulation	77 ms	61,536	79 min