

n=4979

Percentage of Scenarios (%)

50

40

30

20

10

0

27.2%

41.7%

Detected by
Full-Stack

Missed by
Output-Only (H1)

