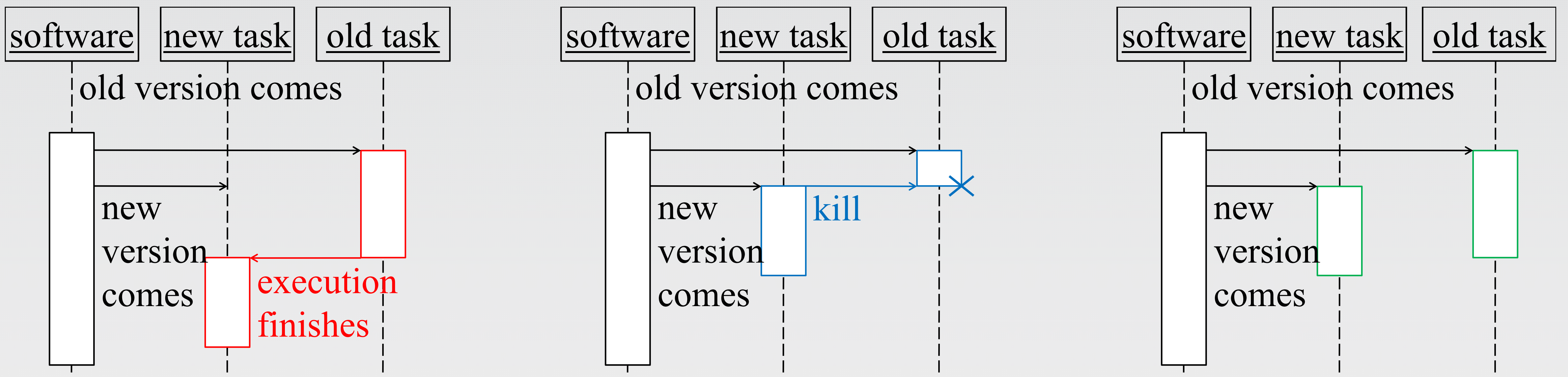
Testing in Parallel

Zhenyu Zhang[1], Zijian Tong[2], and Xiaopeng Gao[3]

[1]*Institute of Software, Chinese Academy of Sciences* [2]*Sohu.com* [3]*Beihang University*

Motivation

- We are one of the largest IT companies in China and maintain one of the largest search engines in China.
- In our searching component projects, it is common that the average integration period is merely two hours with a reduced test suite containing more than 2000 test cases, which execution needs three hours or even longer.
- A build version is compiled and is ready to subject for a new round of regression test before the previous round of regression test has been completed over the last build version.
- Previously, we have three strategies to address this problem.

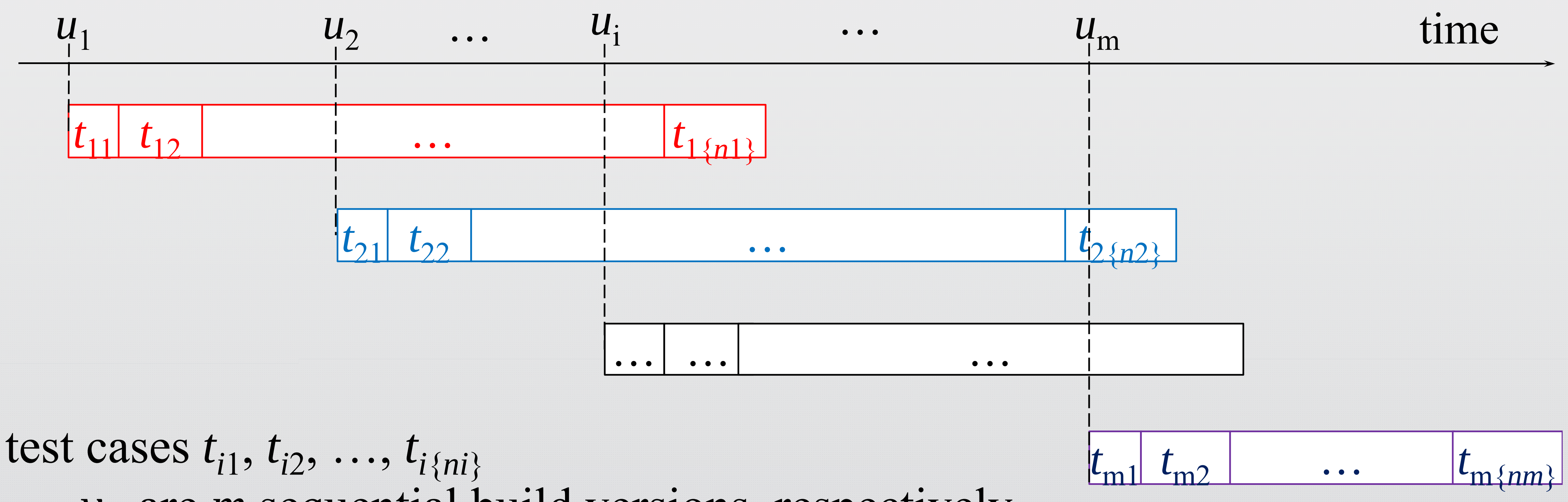


(a) Strategy 1: Wait & Create (b) Strategy 2: Kill & Create (c) Strategy 3: Keep & Create

Effectiveness *(in terms of number of run test cases)*	***High***	*Low*	***High***
Efficiency *(in terms of speed to run test cases)*	*Low*	*Low*	***High***
Limitation *(in terms of number of paralleled tasks)*	***Less***	***Less***	*More*

Our Proposed Preliminary Solution

- Paralleling the run of test suite is necessary since it may increase the probability of revealing failure and thus increases the effectiveness of regression testing.
- Scheduling the test case priorities among different test suites is important because the information obtained from test case that runs on old build versions may provide optimization opportunities on the regression test of later versions.



Problem Settings

- In each test suite S_i, a permutation of its test cases $t_{i1}, t_{i2}, \ldots, t_{i\{ni\}}$ is expected as an output. Suppose $v_1, v_2, \ldots, v_m$ are m sequential build versions, respectively released at time $u_1, u_2, \ldots, u_m$.
- The problem of testing in parallel can be deemed to be a parallelized test case prioritization. Suppose that test cases in test suite S_i are organized to run in some order of $t_{i1}, t_{i2}, \ldots, t_{i\{ni\}}$ with respect to version v_i; where the order $<i1, i2, \ldots, i\{ni\}>$ is represented by O_i, which is a permutation of $<1, 2, \ldots, ni>$. The set of permutations $O_1, O_2, \ldots, O_m$ is denoted by P.
- For convenience, we further suppose that every test case takes identical time t_u to finish execution. At a given time t, the set of executed test cases is $S(t, P) = \{t_{i,j} \mid u_i+t_u\times j\leq t\}$. If we further use a term $F(t, P)$ to denote the set of faults revealed by $S(t, P)$, our aim is to find the optimal set of permutations P_{opt} so that $F(t, P') \subseteq F(t, P_{opt})$ for any P' and any t.

Selected References

[1] G. Rothermel, R. H. Untch, C. Chu, and M. J. Harrold (2001). Prioritizing test cases for regression testing. *TSE*.

[2] B. Jiang, Z. Zhang, W. K. Chan and T. H. Tse (2009). Adaptive random test case prioritization. In *ASE 2009*.