SIO2Jail

Wojciech Dubiel
Tadeusz Dudkiewicz
Przemysław Jakub Kozłowski
Maciej Wachulec

zamawiający:
Szymon Acedański
Olimpiada Informatyczna

Wyzwania konkursów algorytmicznych

Trzeba zapewnić:

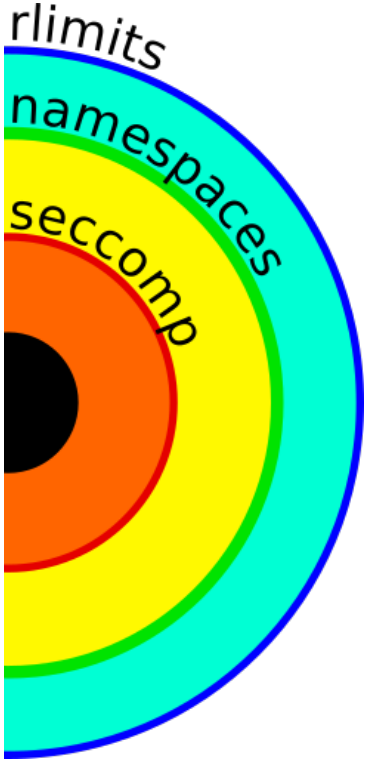
- powtarzalny pomiar wydajności
- powtarzalne środowisko wykonania
- ograniczenia zasobów
  - pamięć operacyjna
  - przestrzeń dyskowa
  - liczba procesów/wątków
- bezpieczeństwo
  - naruszenie zasad konkursu
  - atak na system sprawdzający

Obecne rozwiązanie

OITimeTool:

- używa biblioteki PIN
  - kilkukrotny narzut
  - własnościowa licencja
  - współdzielenie przestrzeni adresowej
  - znane problemy z bezpieczeństwiem
- bezpieczeństwo oparte na blokowaniu syscalli
  - mało elastyczne
- używany ze statycznymi binarkami (c/c++/pascal)

Użyte technologie



- pomiar czasu - perf
- ograniczenie zasobów - rlimits, seccomp
- bezpieczeństwo - namespaces

Perf

Nowoczesne procesory udostępniają liczniki m.in.:

- wykonanych instrukcji
- odwołań do pamięci
- trafienia i nietrafienia w cache

Linux umożliwia ich czytanie

- z podziałem na procesy / grupy procesów
- z podziałem na przestrzeń jądra i użytkownika

Do pomiaru wydajności programu używamy licznika instrukcji.

seccomp-bpf

Filtry wywołań systemowych

- zapisane w bajtkodzie BPF
- wykonywane po stronie jądra
  - mała powierzchnia ataku
  - duża wydajność

Używamy go do egzekwowania zasad konkursu.

rlimit

POSIXowy mechanizm określania maksymalnego rozmiaru zasobów zużywanych przez proces. W Linuxie pozwala ograniczyć m.in.:

- rozmiar przestrzeni adresowej i stosu
- czas działania programu
- liczbę otwartych plików
- maksymalny rozmiar utworzonego pliku

Ograniczamy nim dostępną pamięć i rozmiar pliku wyjściowego.

namespaces

Kontrola istniejących z punktu widzenia procesu zasobów:

- wirtualnego systemu plików
- uruchomionych procesów
- interfejsów sieciowych
- obiektów IPC
- nazwy hosta
- identyfikatorów użytkowników

Używamy ich do zapewnienia bezpieczeństwa.

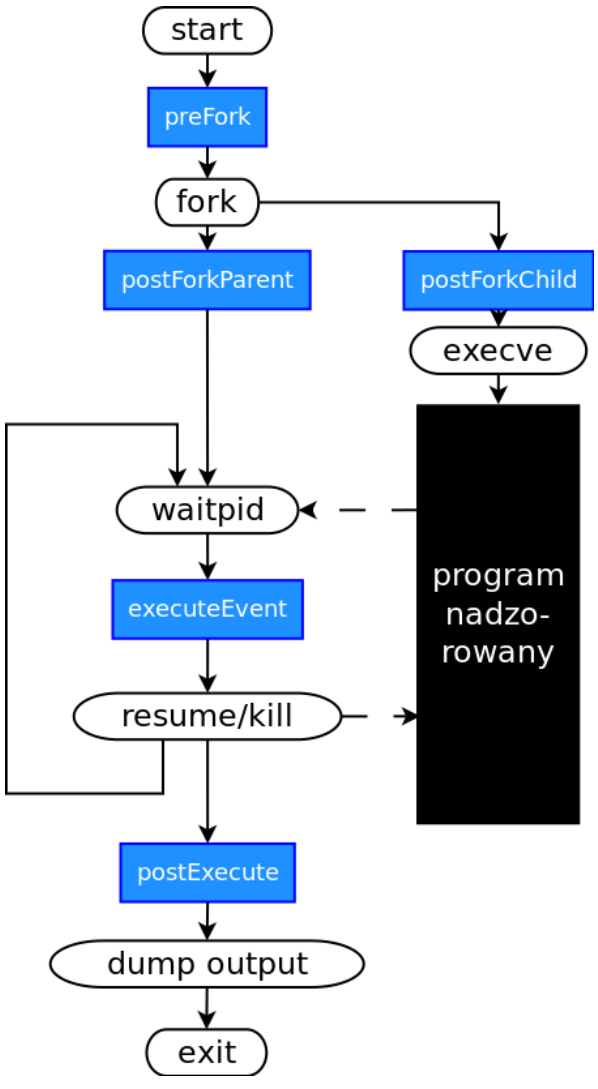
Cykl Życia

Cykl dzieli się na

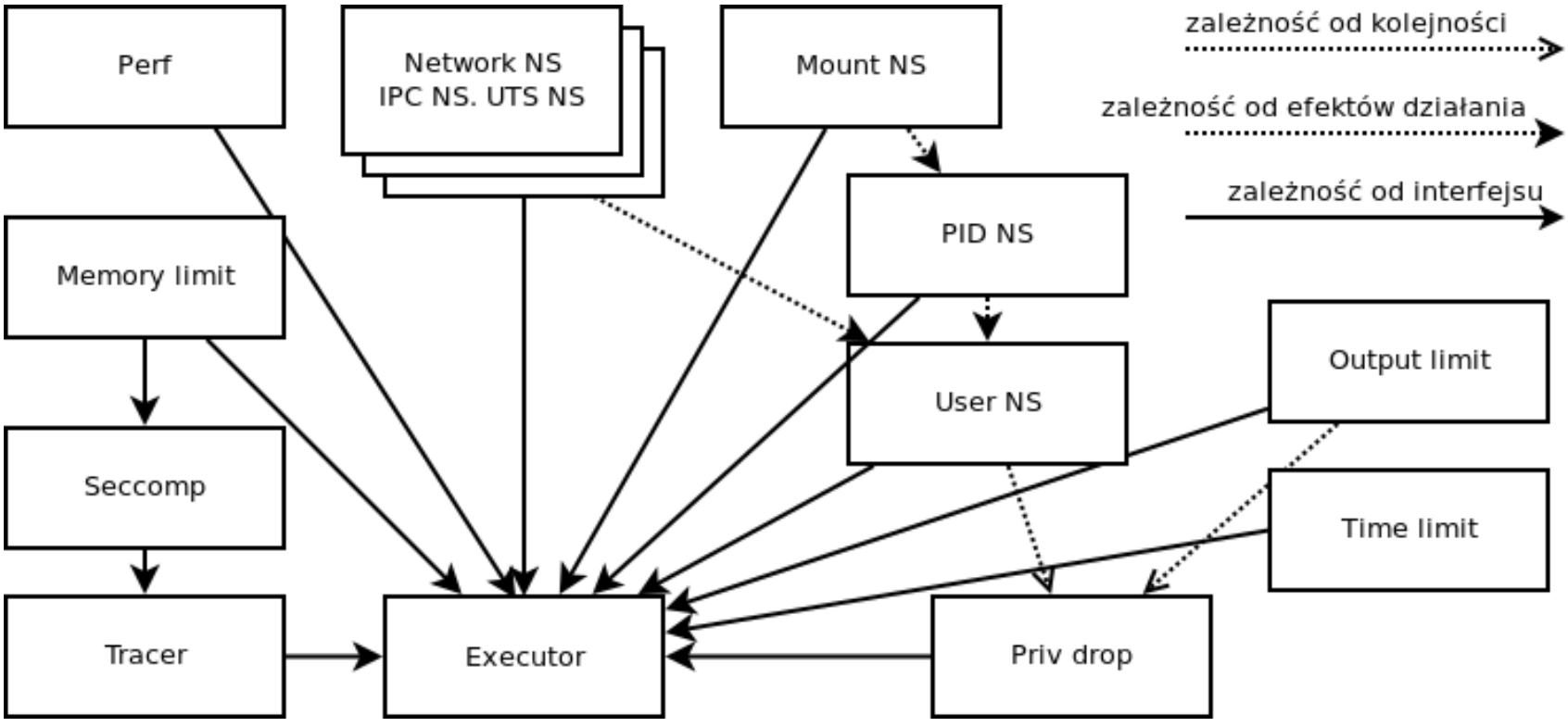
- etapy wykonujące się w pętli głównej
- etapy wywołujące callbacki modułów

Etapy:

- preFork, postForkParent, postForkChild: przygotowanie środowiska
- waitpid, executeEvent, resume/kill: reakcja na zdarzenia z procesu-dziecka
- postExecute: zebranie wyników, przygotowanie danych wyjściowych



Moduły


Perf
Network NS
IPC NS. UTS NS
Mount NS
zależność od kolejności
zależność od efektów działania
zależność od interfejsu
PID NS
Memory limit
User NS
Output limit
Seccomp
Time limit
Tracer
Executor
Priv drop

DEMO

Testy

Porównanie wyników SIO2jaila z OITimeToolem

- 500 000 losowo wybranych par (zgłoszenie, test) z II i III etapu 24. i 25. Olimpiady Informatycznej
- kompilacja gcc-4.9.2
- uruchomienie za pomocą obu narzędzi
- porównanie:
  - statusu sprawdzenia (OK, błąd wykonania, przekroczenie limitu)
  - zmierzonej liczby instrukcji
  - zmierzonej pamięci
  - czasu trwania sprawdzania

Liczba instrukcji

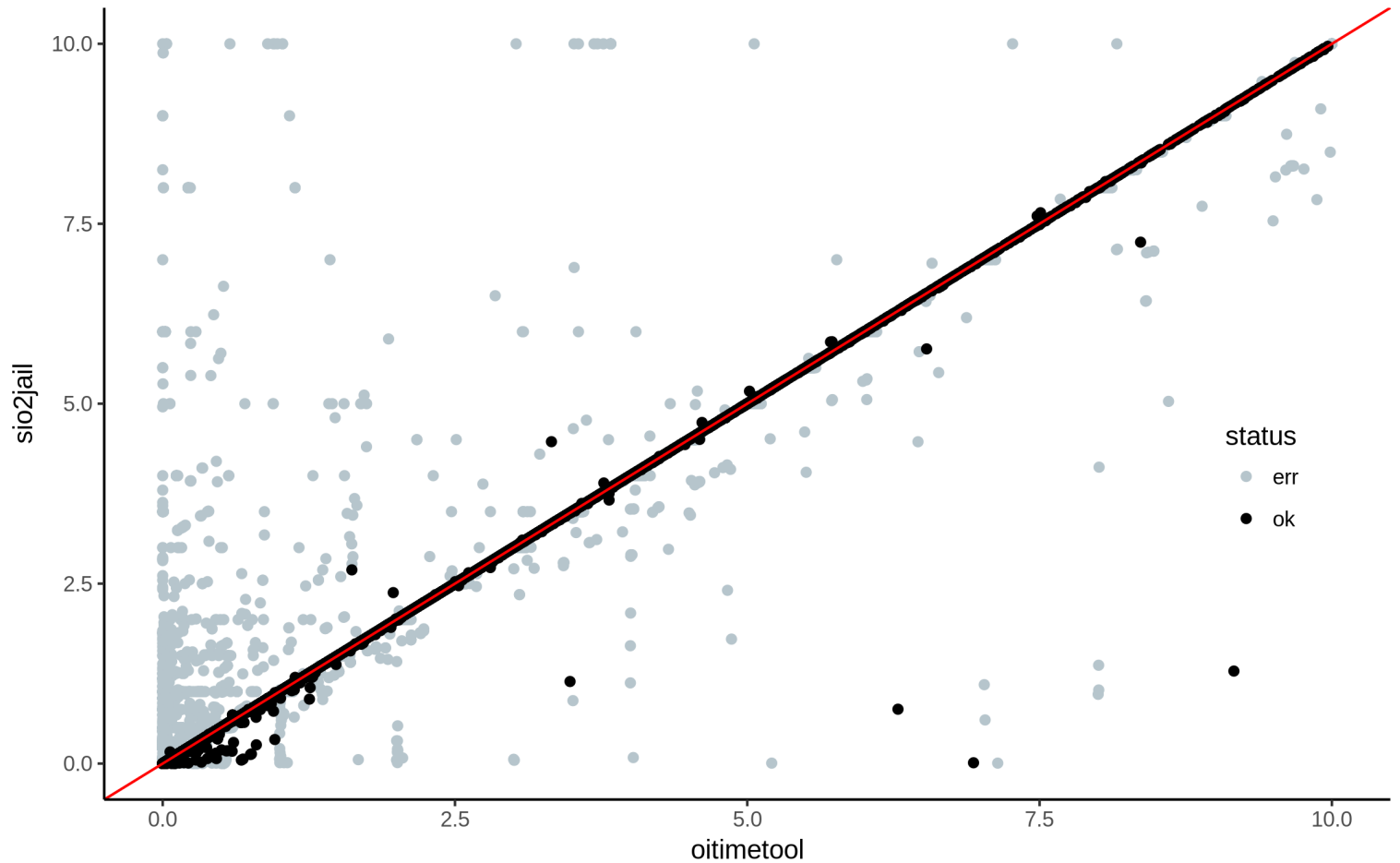


Figure 1: Liczba instrukcji zmierzona SIO2jail w zależności od liczby instrukcji zmierzonej oitimetool

Liczba instrukcji

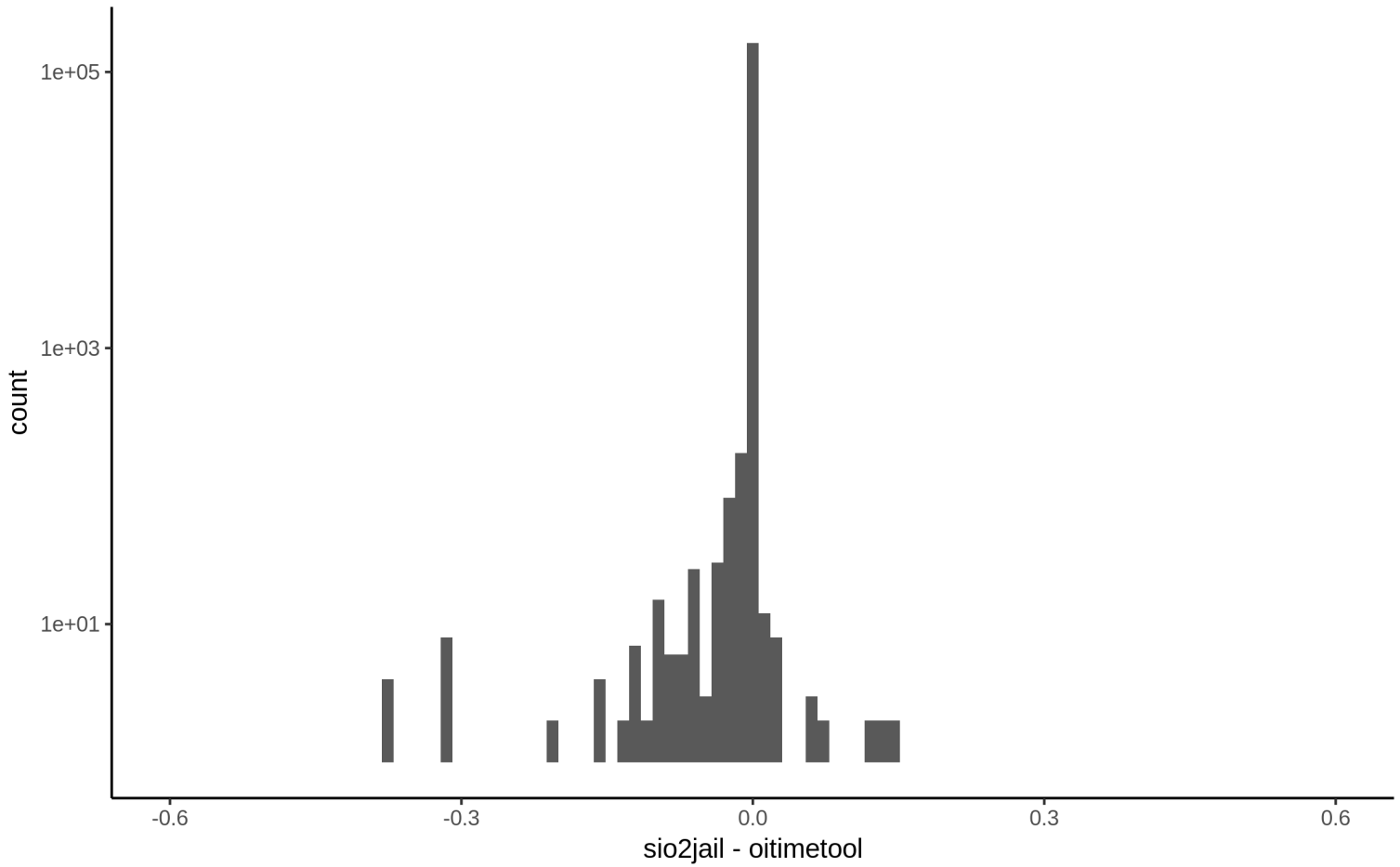


Figure 2: Histogram różnicy zmierzonych liczb instrukcji ($T_{sio2jail} - T_{oitimetool}$)

Zużycie pamięci

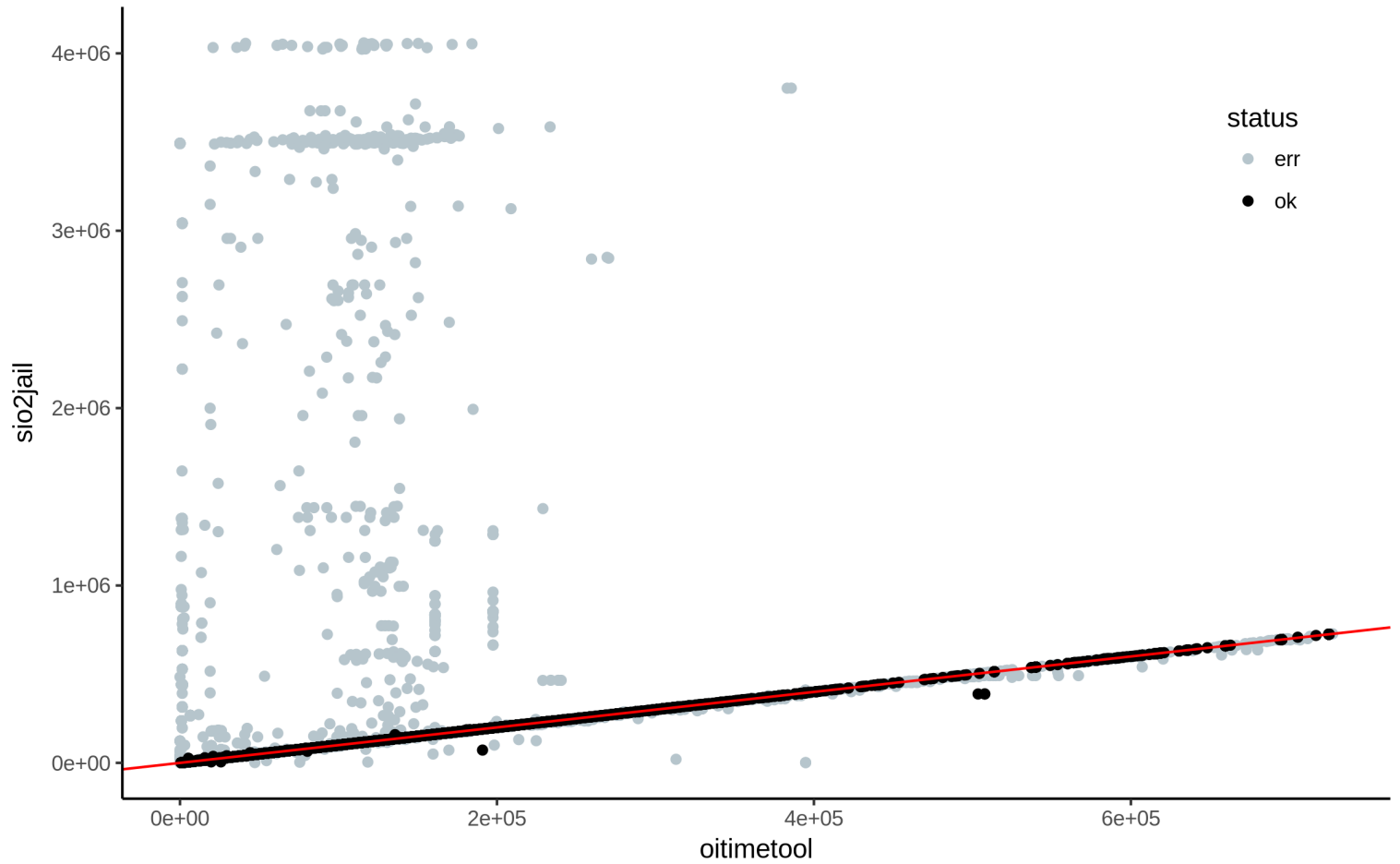


Figure 3: Zużycie pamięci zmierzone sio2jail w zależności od zmierzonego oitimetool

Zużycie pamięci

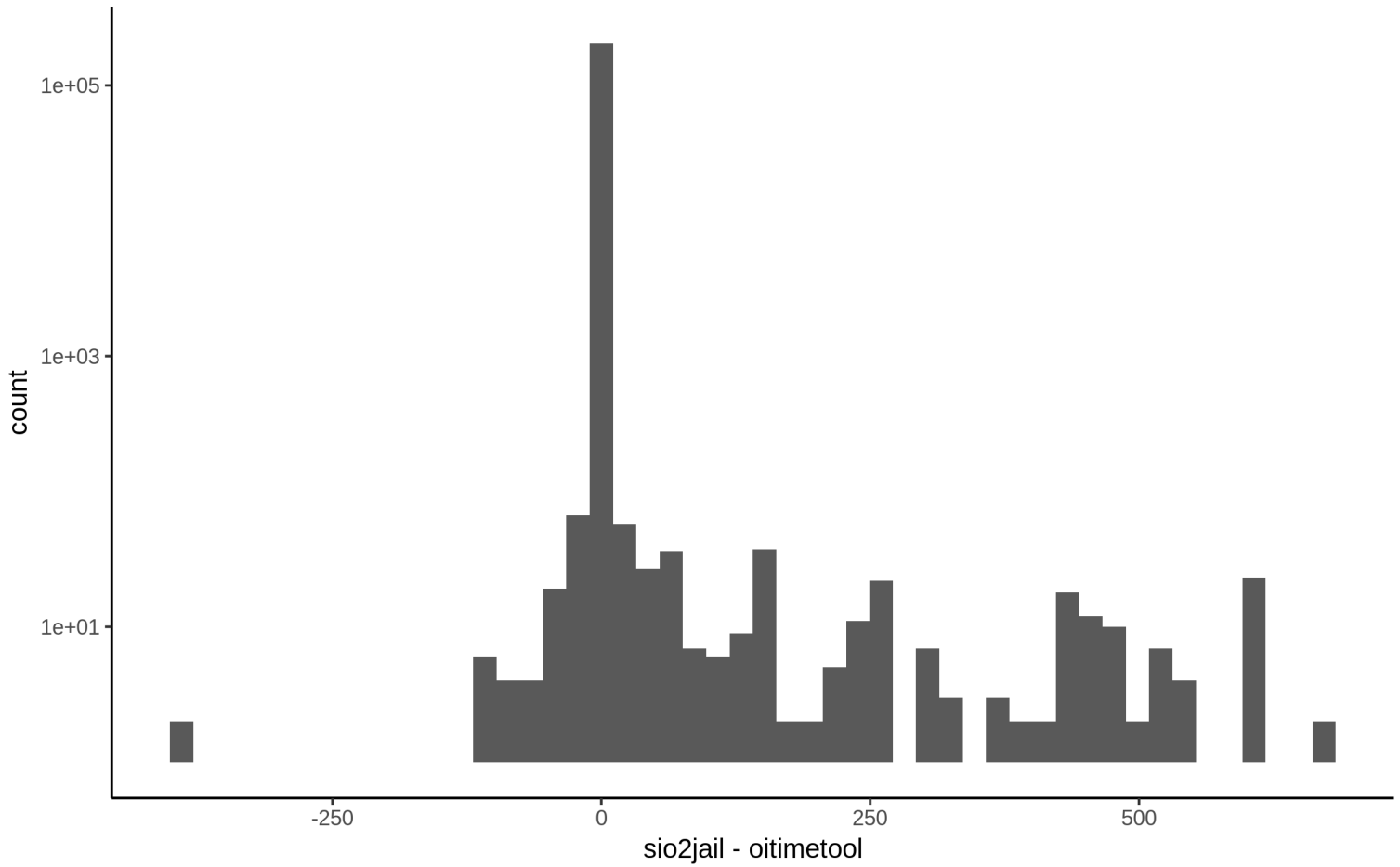


Figure 4: Histogram różnicy zmierzonych zużyć pamięci ($M_{sio2jail} - M_{oitimetool}$)

Rzeczywisty czas sprawdzania

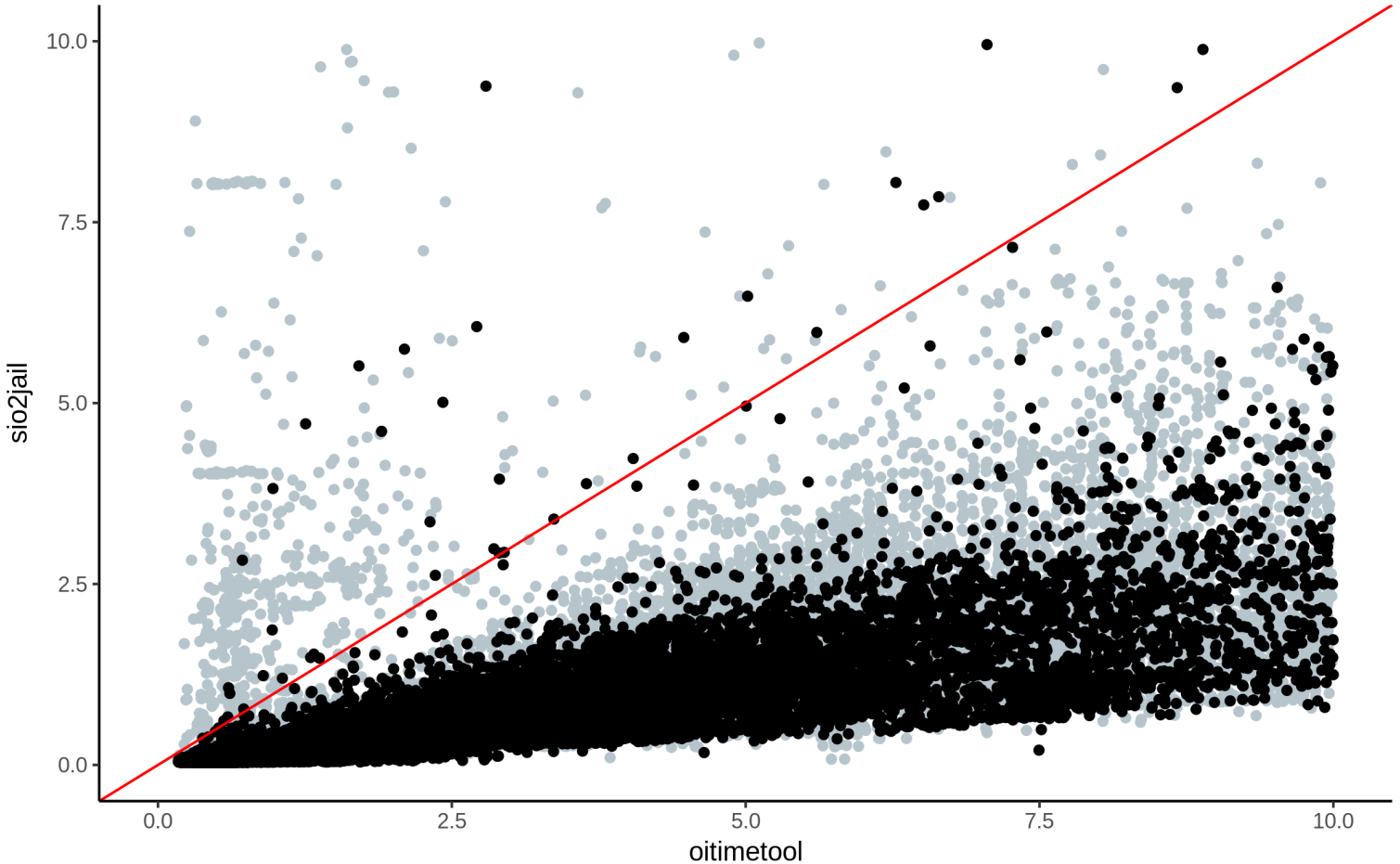
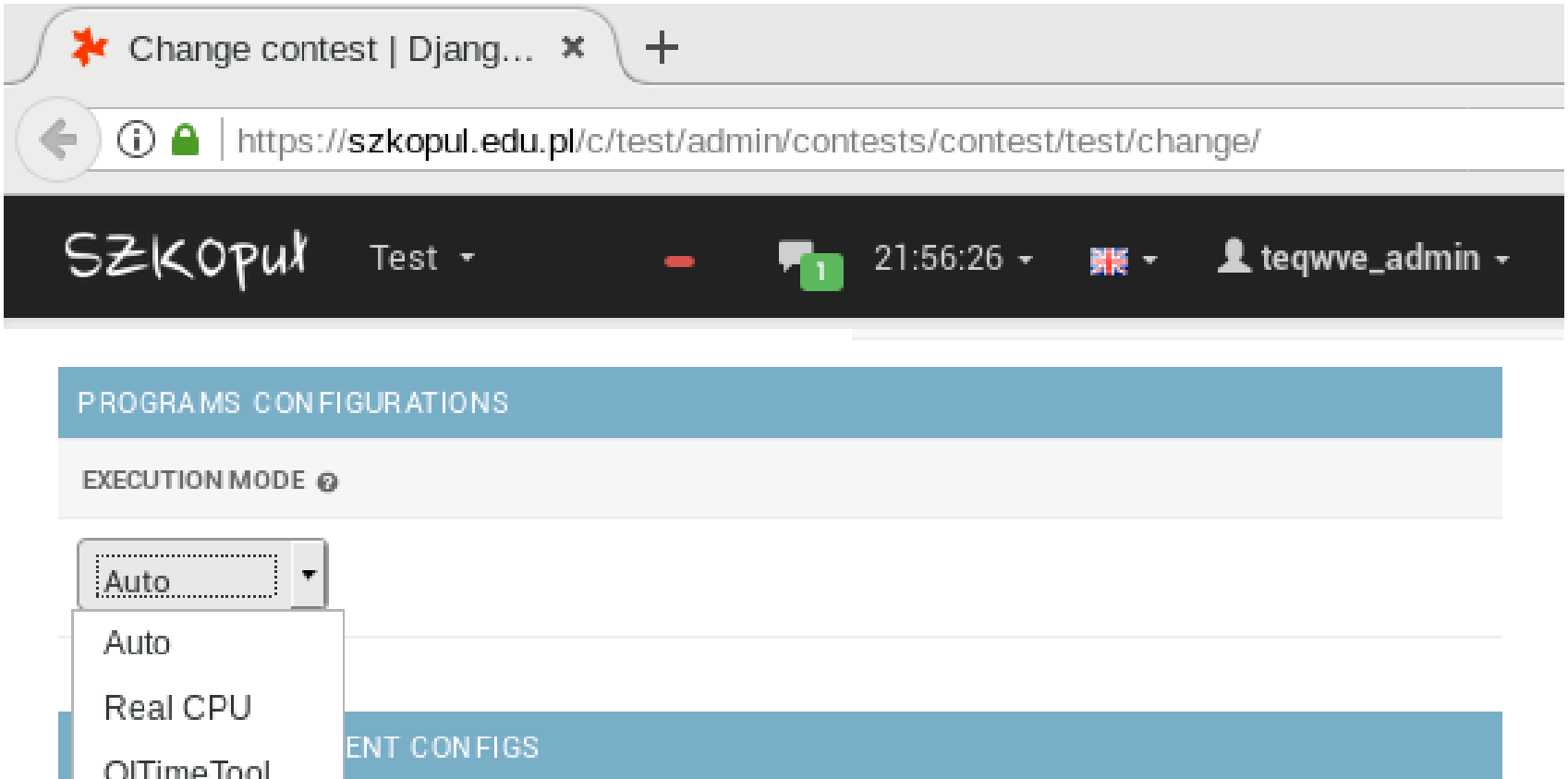


Figure 5: Rzeczywisty czas sprawdzania za pomocą SIO2jail w zależności od czasu sprawdzania za pomocą OITimeTool

Integracja z SIO2

nowy tryb sprawdzania korzystający z SIO2jail

- wdrożony na szkopul.edu.pl, domyślnie używany od maja 2018
- wdrożony na sio2.mimuw.edu.pl, prawdodpobonie zostanie użyty podczas XXVI OI



Przykładowy raport sprawdzania na OITimeTool

		fla	fla2	fla3	flas1	flas2	flas3	flas4	flas5	flab1	flab2	flab3	flab4	flab5	flab6	flab7	flab8	flab9	flab10	flab11	flab12
		INI_OK	INI_OK	INI_OK	INI_ERR	INI_ERR	INI_ERR	INI_ERR	INI_ERR	INI_ERR	INI_ERR	INI_OK	INI_OK	INI_OK	INI_OK	INI_ERR	INI_ERR	INI_OK	INI_OK	INI_ERR	INI_OK
		100	100	100	28	14	51	28	51	0	0	0	0	37	13	0	37	37	37	49	57
Total	95.50s	5.91s	7.35s	3.00s	7.36s	34.71s	9.68s	1.19s	25.47s	0.00s	0.00s	5.79s	5.88s	5.86s	7.37s	7.25s	3.35s	3.36s	3.39s	2.93s	2.94s
0	0.50s	0.00s	0.00s	0.00s	0.03s	0.09s	0.00s	0.00s	0.00s	WA	WA	0.00s	0.00s	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
0a	0.50s	0.00s	0.00s	0.00s	0.03s	0.00s	0.00s	0.00s	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s	WA	0.00s	0.00s	WA	0.00s
1a	0.50s	0.00s	0.00s	0.00s	0.03s	0.00s	0.00s	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
1b	0.50s	0.00s	0.00s	0.00s	0.03s	0.06s	0.00s	0.00s	0.00s	WA	WA	0.00s	0.00s	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
1c	0.50s	0.00s	0.00s	0.00s	0.03s	TLE	0.03s	0.00s	0.02s	WA	WA	WA	WA	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
1d	0.50s	0.00s	0.00s	0.00s	0.03s	TLE	0.01s	0.00s	0.00s	0.00s	WA	WA	0.00s	0.00s	0.00s	0.00s	0.00s	0.00s	0.00s	0.00s	0.00s
1e	0.50s	0.00s	0.00s	0.00s	0.02s	TLE	0.01s	0.00s	0.00s	0.00s	WA	WA	0.00s	0.00s	0.00s	0.00s	0.00s	0.00s	0.00s	0.00s	0.00s
1f	0.50s	0.00s	0.00s	0.00s	0.03s	TLE	0.01s	0.00s	0.00s	WA	0.00s	WA	WA	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
1g	0.50s	0.00s	0.00s	0.00s	0.03s	TLE	0.00s	0.00s	0.00s	WA	WA	0.00s	0.00s	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
1h	0.50s	0.00s	0.00s	0.00s	0.03s	TLE	0.01s	0.00s	0.00s	0.00s	WA	WA	0.00s	WA	0.00s	0.00s	WA	WA	WA	0.00s	0.00s
1i	0.50s	0.00s	0.00s	0.00s	0.03s	TLE	0.00s	0.00s	0.00s	0.00s	WA	WA	0.00s	WA	0.00s	0.00s	WA	WA	WA	0.00s	0.00s
1j	0.50s	0.00s	0.00s	0.00s	0.03s	0.00s	0.00s	0.00s	0.00s	0.00s	WA	WA	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s	0.00s
1k	0.50s	0.00s	0.00s	0.00s	0.03s	0.33s	0.00s	0.00s	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s	WA	0.00s	0.00s	WA	0.00s
1ocen	0.50s	0.00s	0.00s	0.00s	0.03s	0.12s	0.00s	0.00s	0.00s	WA	WA	0.00s	0.00s	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
2a	0.50s	0.00s	0.00s	0.00s	RE	0.00s	0.00s	MLE	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
2b	1.25s	0.00s	0.00s	0.00s	RE	TLE	0.12s	0.05s	0.03s	WA	WA	0.00s	WA	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
2c	1.25s	0.00s	0.00s	0.00s	RE	TLE	0.18s	MLE	0.12s	WA	WA	0.00s	0.00s	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
2d	1.25s	0.00s	0.00s	0.00s	RE	TLE	0.10s	MLE	0.11s	0.00s	WA	WA	0.00s	WA	0.00s	0.00s	WA	WA	WA	0.00s	0.00s
2e	1.25s	0.00s	0.00s	0.00s	RE	TLE	0.13s	RE	0.12s	WA	WA	0.00s	WA	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
2f	0.50s	0.00s	0.00s	0.00s	0.03s	TLE	0.01s	0.00s	0.01s	WA	0.00s	0.00s	WA	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
2g	0.50s	0.00s	0.00s	0.00s	0.03s	TLE	0.01s	0.00s	0.01s	WA	WA	0.00s	WA	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
2h	0.50s	0.00s	0.00s	0.00s	0.03s	TLE	0.01s	0.00s	0.00s	0.00s	WA	WA	0.00s	WA	0.00s	0.00s	WA	WA	WA	0.00s	0.00s
2i	0.50s	0.00s	0.00s	0.00s	RE	TLE	0.00s	0.00s	0.00s	0.00s	WA	WA	0.00s	WA	0.00s	0.00s	WA	WA	WA	0.00s	0.00s
2j	0.50s	0.00s	0.00s	0.00s	0.03s	0.00s	0.00s	0.00s	0.00s	0.00s	WA	WA	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s	0.00s
2k	3.25s	0.00s	0.00s	0.00s	MLE	0.03s	0.80s	RE	0.00s	0.00s	WA	WA	0.00s	WA	0.00s	0.00s	WA	WA	WA	WA	WA
2l	0.50s	0.00s	0.00s	0.00s	0.03s	TLE	0.01s	0.00s	0.07s	0.00s	WA	WA	0.00s	WA	0.00s	0.00s	WA	WA	WA	0.00s	0.00s

Przykładowy raport sprawdzania na SIO2Jail

		fla	fla2	fla3	flas1	flas2	flas3	flas4	flas5	flab1	flab2	flab3	flab4	flab5	flab6	flab7	flab8	flab9	flab10	flab11	flab12
		INI_OK	INI_OK	INI_OK	INI_ERR	INI_ERR	INI_ERR	INI_ERR	INI_ERR	INI_ERR	INI_ERR	INI_OK	INI_OK	INI_OK	INI_OK	INI_ERR	INI_ERR	INI_OK	INI_OK	INI_ERR	INI_OK
		100	100	100	28	14	51	28	51	0	0	0	0	37	13	0	37	37	37	49	57
Total	95.50s	5.91s	7.35s	3.00s	7.25s	33.50s	9.65s	1.19s	25.32s	0.00s	0.00s	5.79s	5.88s	5.86s	7.37s	7.25s	3.35s	3.36s	3.39s	2.93s	2.94s
0	0.50s	0.00s	0.00s	0.00s	0.03s	0.09s	0.00s	0.00s	0.00s	WA	WA	0.00s	0.00s	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
0a	0.50s	0.00s	0.00s	0.00s	0.03s	0.00s	0.00s	0.00s	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s	WA	0.00s	0.00s	WA	0.00s
1a	0.50s	0.00s	0.00s	0.00s	0.03s	0.00s	0.00s	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
1b	0.50s	0.00s	0.00s	0.00s	0.03s	0.06s	0.00s	0.00s	0.00s	WA	WA	0.00s	0.00s	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
1c	0.50s	0.00s	0.00s	0.00s	0.03s	TLE	0.03s	0.00s	0.02s	WA	WA	WA	WA	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
1d	0.50s	0.00s	0.00s	0.00s	0.03s	TLE	0.01s	0.00s	0.00s	0.00s	WA	WA	0.00s	0.00s	0.00s	0.00s	0.00s	0.00s	0.00s	0.00s	0.00s
1e	0.50s	0.00s	0.00s	0.00s	0.02s	TLE	0.01s	0.00s	0.00s	0.00s	WA	WA	0.00s	0.00s	0.00s	0.00s	0.00s	0.00s	0.00s	0.00s	0.00s
1f	0.50s	0.00s	0.00s	0.00s	0.03s	TLE	0.01s	0.00s	0.00s	WA	0.00s	WA	WA	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
1g	0.50s	0.00s	0.00s	0.00s	0.03s	TLE	0.00s	0.00s	0.00s	WA	WA	0.00s	0.00s	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
1h	0.50s	0.00s	0.00s	0.00s	0.03s	TLE	0.01s	0.00s	0.00s	0.00s	WA	WA	0.00s	WA	0.00s	0.00s	WA	WA	WA	0.00s	0.00s
1i	0.50s	0.00s	0.00s	0.00s	0.03s	TLE	0.00s	0.00s	0.00s	0.00s	WA	WA	0.00s	WA	0.00s	0.00s	WA	WA	WA	0.00s	0.00s
1j	0.50s	0.00s	0.00s	0.00s	0.03s	0.00s	0.00s	0.00s	0.00s	0.00s	WA	WA	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s	0.00s
1k	0.50s	0.00s	0.00s	0.00s	0.03s	0.33s	0.00s	0.00s	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s	WA	0.00s	0.00s	WA	0.00s
1ocen	0.50s	0.00s	0.00s	0.00s	0.03s	0.12s	0.00s	0.00s	0.00s	WA	WA	0.00s	0.00s	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
2a	0.50s	0.00s	0.00s	0.00s	RE	0.00s	0.00s	MLE	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
2b	1.25s	0.00s	0.00s	0.00s	RE	TLE	0.12s	0.05s	0.03s	WA	WA	0.00s	WA	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
2c	1.25s	0.00s	0.00s	0.00s	RE	TLE	0.18s	MLE	0.12s	WA	WA	0.00s	0.00s	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
2d	1.25s	0.00s	0.00s	0.00s	RE	TLE	0.10s	MLE	0.11s	0.00s	WA	WA	0.00s	WA	0.00s	0.00s	WA	WA	WA	0.00s	0.00s
2e	1.25s	0.00s	0.00s	0.00s	RE	TLE	0.13s	RE	0.12s	WA	WA	0.00s	WA	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
2f	0.50s	0.00s	0.00s	0.00s	0.03s	TLE	0.01s	0.00s	0.01s	WA	0.00s	0.00s	WA	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
2g	0.50s	0.00s	0.00s	0.00s	0.03s	TLE	0.01s	0.00s	0.01s	WA	WA	0.00s	WA	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s
2h	0.50s	0.00s	0.00s	0.00s	0.03s	TLE	0.01s	0.00s	0.00s	0.00s	WA	WA	0.00s	WA	0.00s	0.00s	WA	WA	WA	0.00s	0.00s
2i	0.50s	0.00s	0.00s	0.00s	RE	TLE	0.00s	0.00s	0.00s	0.00s	WA	WA	0.00s	WA	0.00s	0.00s	WA	WA	WA	0.00s	0.00s
2j	0.50s	0.00s	0.00s	0.00s	0.03s	0.00s	0.00s	0.00s	0.00s	0.00s	WA	WA	0.00s	0.00s	WA	0.00s	0.00s	0.00s	0.00s	0.00s	0.00s
2k	3.25s	0.00s	0.00s	0.00s	RE	0.03s	0.80s	RE	0.00s	0.00s	WA	WA	0.00s	WA	0.00s	0.00s	WA	WA	WA	WA	WA
2l	0.50s	0.00s	0.00s	0.00s	0.03s	TLE	0.01s	0.00s	0.07s	0.00s	WA	WA	0.00s	WA	0.00s	0.00s	WA	WA	WA	0.00s	0.00s

Dalszy rozwój

- Wsparcie dla innych języków
- Wsparcie dla programów wielowątkowych
- Bogarszy format wyjściowy
- Dodatkowe zabezpieczenia

namespaces

Przynależność do namespace-u jest dziedziczona z procesu-rodzica na jego procesy potomne. Proces nie jest w stanie uzyskać dostępu do obiektów spoza swojego namespace-u.

User namespacy umożliwiają definiowanie innych namespace-ów i tworzenie w nich zasobów.

- uprawnienia procesu są definiowane przez flagi `capabilities`
- osobne `capabilities` na każdy user namespace
- utworzenie user namespace daje wewnątrz niego pełne `capabilities` dotyczące utworzonych wewnątrz niego zasobów