

## SR20. Zadania i ćwiczenia (spójność i zwielokrotnianie)

1. Dostęp do obiektu dzielonego Javy można uszeregować, deklarując używane w nim metody jako synchronizowane. Czy to wystarcza do zagwarantowania uszeregowania, gdy taki obiekt zostanie zwielokrotniony?
2. Wyjaśnij własnymi słowami główną przyczynę, która skłania do zajmowania się modelami spójności słabej.
3. Wyjaśnij, jak odbywa się zwielokrotnianie w systemie DNS, i dzięki czemu działa ono tak dobrze.
4. Omawiając modele spójności, często odwoływaliśmy się do umowy między oprogramowaniem a pamięcią danych. Czemu służy taka umowa?
5. Co należałoby zrobić z replikami (dwoma kopiami) na rysunku 7.2 (s. 361 w S&A, plansza 19 w materiałach do wykładu) w celu sfinalizowania wartości w konicie (*conit* – ang. *consistency unit*, jednostka spójności), aby zarówno w kopii *A*, jak i *B* uwidocznił się ten sam wynik?
6. Czy na rys. 7.7 (s. 366 w S&A, plansza 28 w materiałach do wykładu) wartość 001110 jest dozwolona dla pamięci spójnej sekwencyjnie? Wyjaśnij swoją odpowiedź.
7. Często utrzymujemy, że modele spójności słabej dodatkowo obciążają osoby programujące. Do jakiego stopnia teza ta jest rzeczywiście prawdziwa?
8. Czy całkowicie uporządkowane rozsyłanie za pomocą procesu porządkowego i na rzecz spójności w aktywnym zwielokrotnieniu narusza czynnik powiązań końcowych (zob. p. 5.2.2 w wydaniu polskim) w projektowaniu systemu?
9. Jaki rodzaj spójności proponujesz do realizacji giełdy elektronicznej? Uzasadnij swoją odpowiedź.
10. Rozważ osobistą skrzynkę pocztową użytkownika ruchomego, zrealizowaną jako część rozległej, rozproszonej bazy danych. Jaki rodzaj spójności nastawionej na klienta byłby tu najbardziej odpowiedni?
11. Opisz prostą implementację spójności czytania swoich zapisów do wyświetlania właśnie zaktualizowanych stron Sieci.
12. Dla uproszczenia założyliśmy, że w systemie Bayou<sup>1</sup> nie występują konflikty pisanie-pisanie. Jest to oczywiście założenie nierealne. Wyjaśnij, jak może dochodzić do konfliktów.
13. Czy przy korzystaniu z dzierżawy jest konieczne, aby zegary klienta i serwera były ściśle zsynchronizowane?
13. Rozważmy nieblokowany protokół podstawa-zapas, użyty do zapewnienia spójności sekwencyjnej w rozproszonej pamięci danych. Czy taka pamięć danych zawsze umożliwia spójność czytania swoich zapisów?
14. Ogólnie biorąc, do działania aktywnego zwielokrotnienia jest niezbędne, aby wszystkie operacje były wykonywane na każdej kopii w tej samej kolejności. Czy to uporządkowanie jest zawsze konieczne?
15. Jedną z metod realizacji całkowicie uporządkowanego rozsyłania za pomocą procesu porządkowego jest przekazanie operacji najpierw do porządkowego, który przypisuje jej niepowtarzalny numer, po czym ją rozsyła. Wymień dwie inne metody i porównaj wszystkie trzy.
16. Plik jest zwielokrotniony na 10 serwerach. Wylicz kombinacje kworum czytania i kworum pisania, dozwolone w algorytmie głosowania.

<sup>1</sup>Zob. np. <https://people.eecs.berkeley.edu/istoica/classes/cs268/06/notes/20-BFTx2.pdf>, szczególnie plansze 20-49.

17. Rozważmy monitor omówiony w rozdz. 1 [wydanie polskie]. Jeśli w zwielokrotnionym monitorze wątki mogą się blokować, to co powinniśmy zapewnić przy sygnalizowaniu zmiennej warunkowej?
18. W liniowości zakładamy istnienie globalnego zegara. Przy okazji omawiania spójności ścisłej wykazaliśmy jednak, że takie założenie jest nierealne w większości systemów rozproszonych. Czy da się zrealizować liniowość dla pamięci danych rozproszonych fizycznie?
19. Wieloprocessor ma jedną szynę. Czy możemy wobec tego zrealizować w nim pamięć ściśle spójną?
20. Dlaczego ciąg  $W_1(x)a \ R_2(x)NIL \ R_3(x)a$  na rys. 6.5(b) [wydanie polskie, plansza 22 w materiałach wykładowych] jest niedozwolony?
21. W większości implementacji [pilnej] spójności zwalniania w systemach rozproszonej pamięci dzielonej zmienne dzielone są synchronizowane przy zwalnianiu, lecz nie przy nabywaniu. Po co w takim razie w ogóle je nabywać?
22. Czy system Orca [wydanie polskie] oferuje spójność sekwencyjną, czy spójność wejścia? Uzasadnij swoją odpowiedź.
23. Podaj przykład, w którym spójność nastawiona na klienta może łatwo doprowadzić do konfliktów pisanie-pisanie.
24. W tekście nakreśliliśmy schemat zapobiegania zwielokrotnionym wywołaniom, z udziałem nadawcy. W schemacie z udziałem odbiorcy kopia odbiorcza rozpoznaje kopie nadchodzących komunikatów, dotyczące tego samego wywołania. Opisz, jak można zapobiec zwielokrotnionym wywołaniom w schemacie z udziałem odbiorcy.
25. Rozważ leniwe zwielokrotnianie spójne przyczynowo. W jakich dokładnie okolicznościach można usunąć operację z kolejki pisania?
26. [Zadanie dodatkowe]. W tym ćwiczeniu masz zrealizować prosty system, który będzie umożliwiał rozsyłanie wywołań RPC. Zakładamy, że istnieje pewna liczba zwielokrotnionych serwerów, oraz że każdy klient komunikuje się z serwerem za pomocą wywołania RPC. W wypadku zwielokrotnienia klient musi wysłać wywołanie RPC do każdej kopii. Zaprogramuj klienta tak, aby z punktu widzenia aplikacji odnosiło się wrażenie, że jest wysyłane jedno wywołanie RPC. Załóż, że dokonujesz zwielokrotnień w celach efektywnościowych, lecz serwery są podatne na awarie. [Może zamodelujesz to zadanie, wykorzystując maszyny wirtualne w swoim komputerze? [zpl]].