

Rozproszone bazy danych

Protokół dwufazowego zatwierdzania

Plan zajęć

- Cechy transakcji rozproszonej
- Uczestnicy transakcji
- Zatwierdzanie transakcji – 2PC
- Awarie transakcji rozproszonej
- Ręczne odtwarzanie transakcji



Transakcja rozproszona

- Transakcja rozproszona vs zdalna
- Transakcja, której polecenia DML odwołują się do tabel znajdujących się w co najmniej dwóch węzłach rozproszonej bazy danych
- Zbiór lokalnych transakcji
- ACID
- Atomowość zapewniona przez protokół dwufazowego zatwierdzania (two-phase commit)



Aktorzy

- Koordynator globalny (**KG**)
 - inicjator transakcji (korzeń)
- Koordynator lokalny
 - węzeł, któremu podlegają inne węzły
- Uczestnik
 - węzeł z transakcją lokalną
- Węzeł zatwierdzania (**WZ**)
 - wybierany przez SZBD na początku zatwierdzania
 - najważniejszy węzeł w transakcji
 - zawiera status zatwierdzania całej transakcji rozproszonej

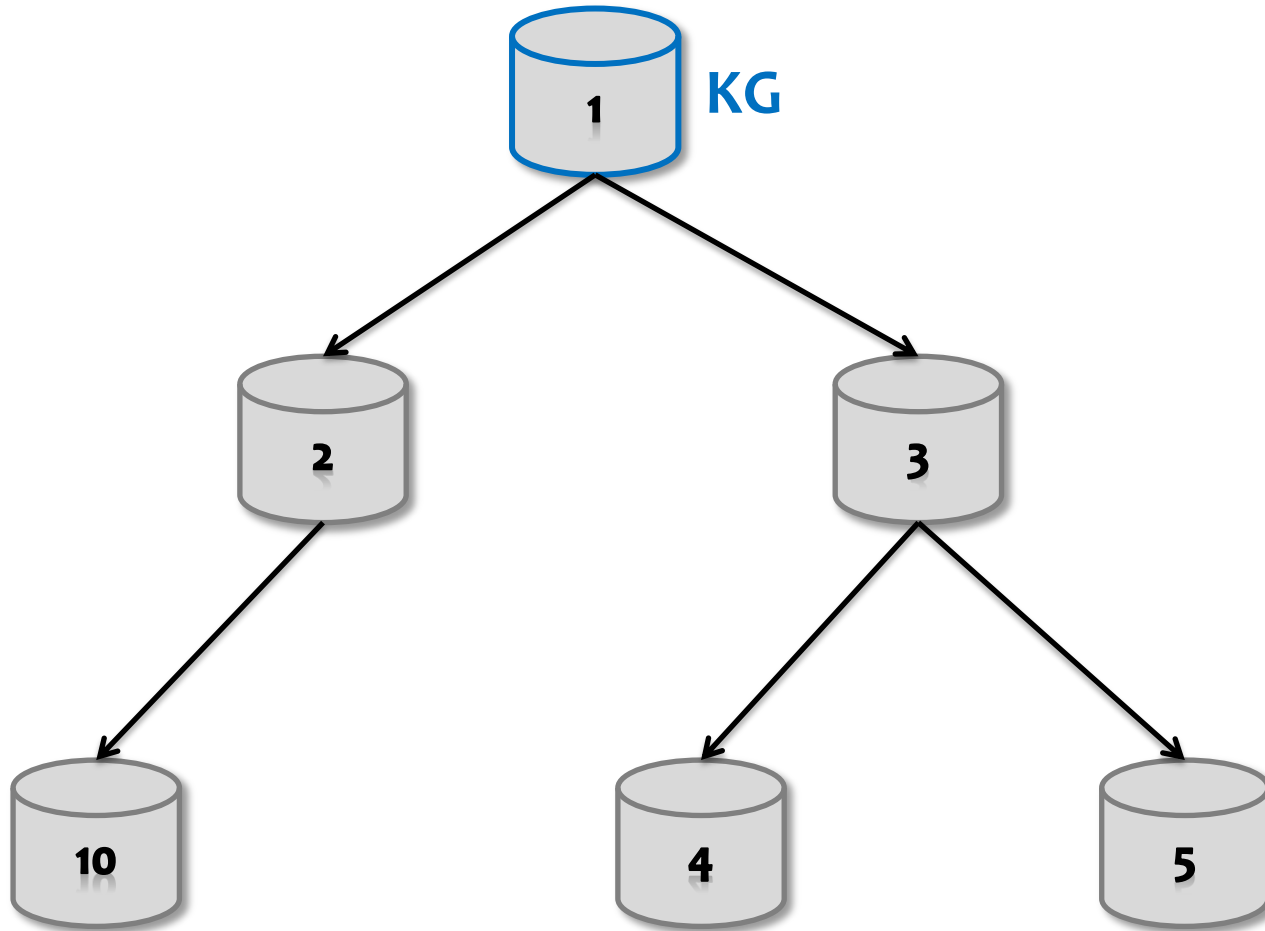


Wybór węzła zatwierdzania

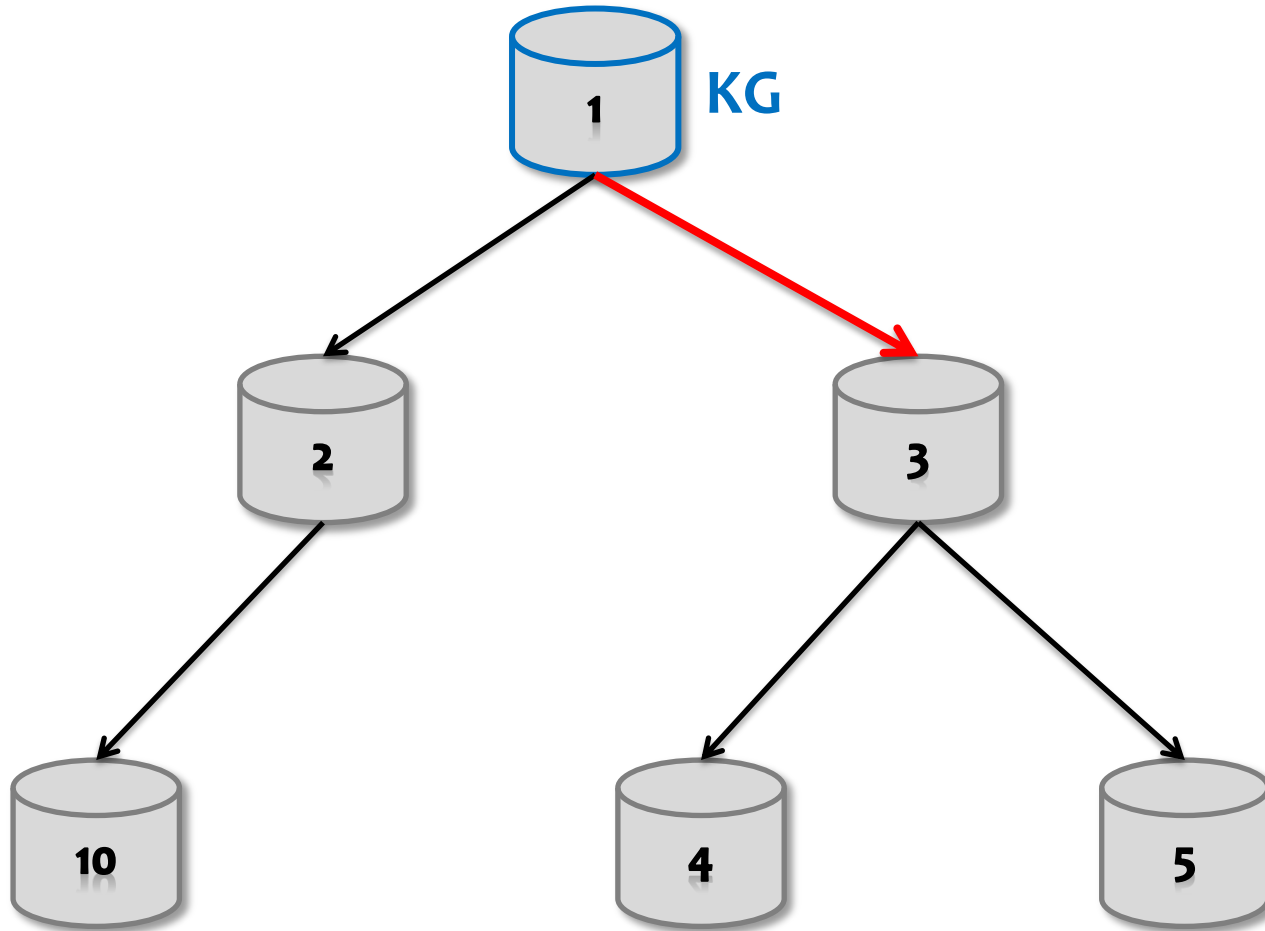
- Parametr COMMIT_POINT_STRENGTH
 - ustawiany przez administratora systemu
 - wartość 0-255
 - odzwierciedla ilość danych krytycznych w węźle oraz jego niezawodność
 - **KG** wybiera **WZ** poszukując w sposób **zachłanny** węzła o najwyższej wartości tego parametru



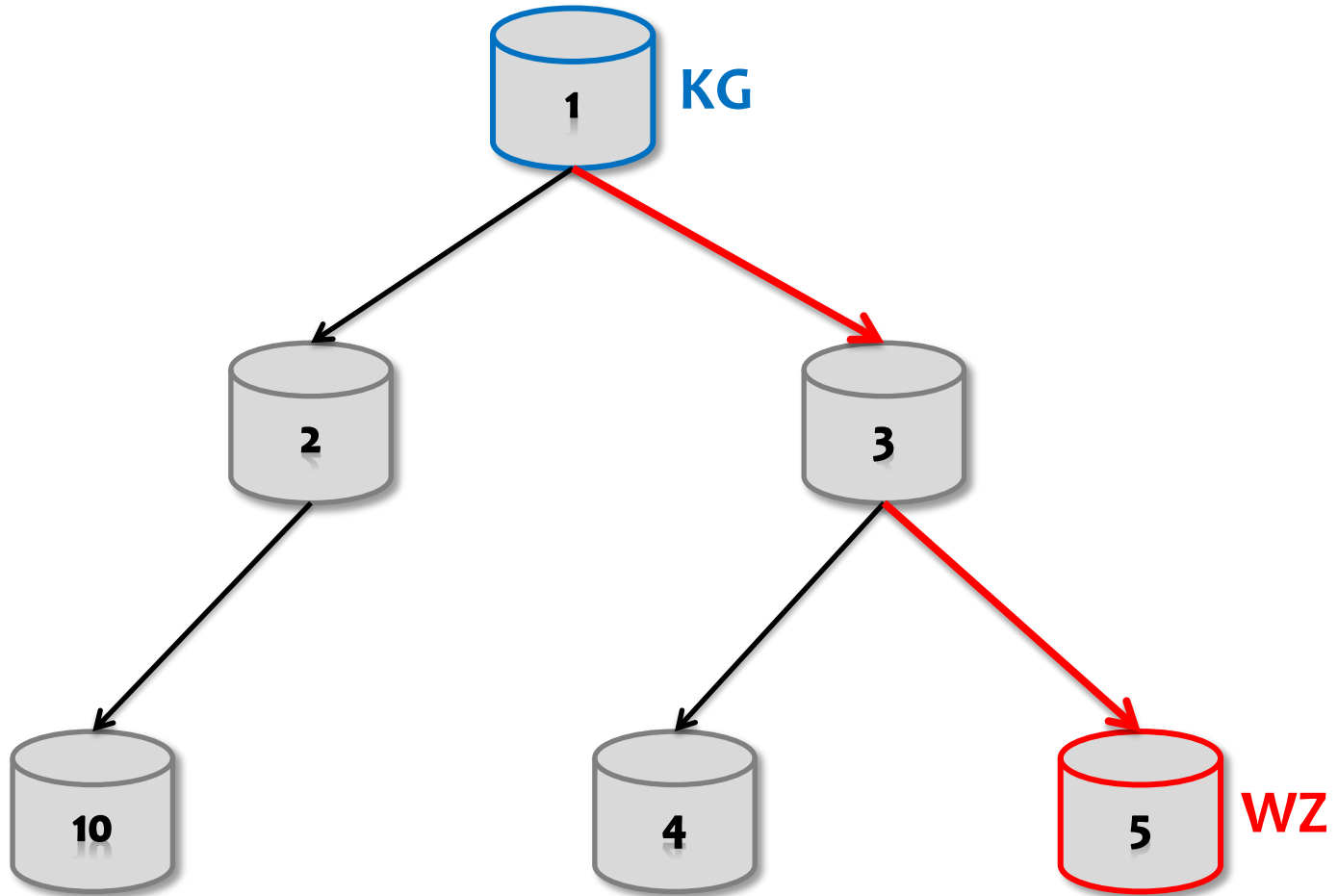
Wybór węzła zatwierdzenia – c.d.



Wybór węzła zatwierdzania – c.d.



Wybór węzła zatwierdzenia – c.d.



Protokół dwufazowego zatwierdzenia

- Fazy realizacji
 - przygotowanie (PREPARE)
 - zatwierdzenie (COMMIT)
 - [zakończenie (FORGET)]



Faza przygotowania – koordynator globalny

1. Wybranie węzła zatwierdzania
2. Wysłanie żądania przygotowania do wszystkich uczestników transakcji, poza węzłem zatwierdzania **WZ**
3. Odebranie komunikatów zwrotnych od uczestników



Faza przygotowania – uczestnik

1. Odebranie żądania przygotowania od **KG**
 2. Przekazanie żądania do podległych uczestników
 3. Odbiór komunikatów zwrotnych od podległych uczestników
 4. Wysłanie komunikatu zwrotnego do **KG**
- Możliwe odpowiedzi uczestnika:
 - **READ-ONLY** – brak modyfikacji
 - **PREPARED** – węzeł oraz wszystkie węzły mu podległe są gotowe to zatwierdzenia
 - **ABORT** – przynajmniej jeden węzeł nie jest gotowy do zatwierdzenia (momentalne wycofanie transakcji lokalnej)



Faza zatwierdzania

1. **KG** odbiera potwierdzenia od uczestników
2. Jeśli wszyscy odpowiedzieli **PREPARED**
 1. **KG** wysyła żądanie zatwierdzenia do **WZ**
 2. **WZ** zatwierdza transakcję i wysyła komunikat do **KG**
 3. **KG** wysyła żądanie zatwierdzenia do pozostałych węzłów
3. Jeśli choć jeden uczestnik odpowiedział **ABORT**
 1. **KG** wysyła żądanie wycofania do **WZ**
 2. **WZ** wycofuje transakcję i wysyła komunikat do **KG**
 3. **KG** wysyła żądanie wycofania do pozostałych węzłów

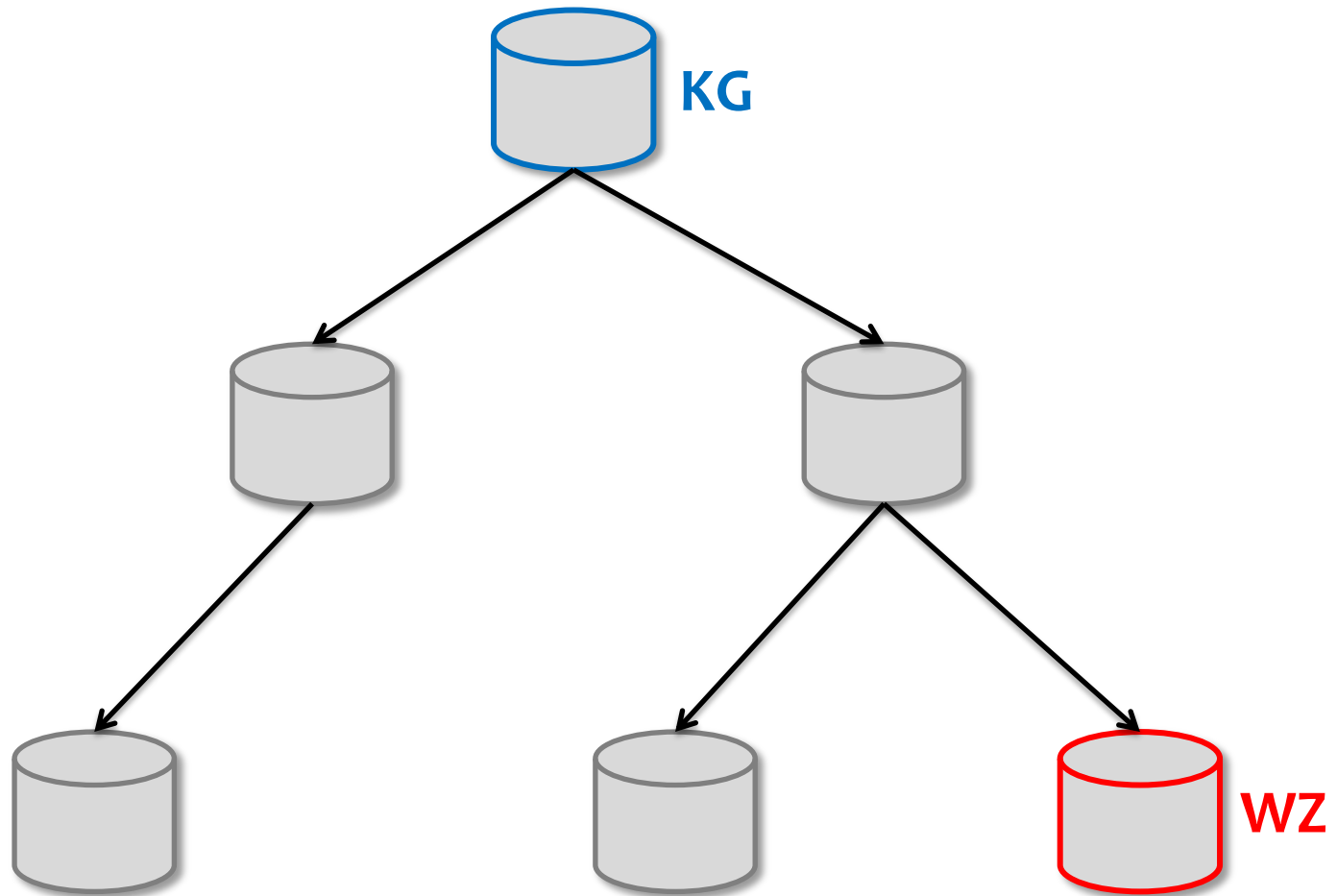


Faza zatwierdzania – uczestnik

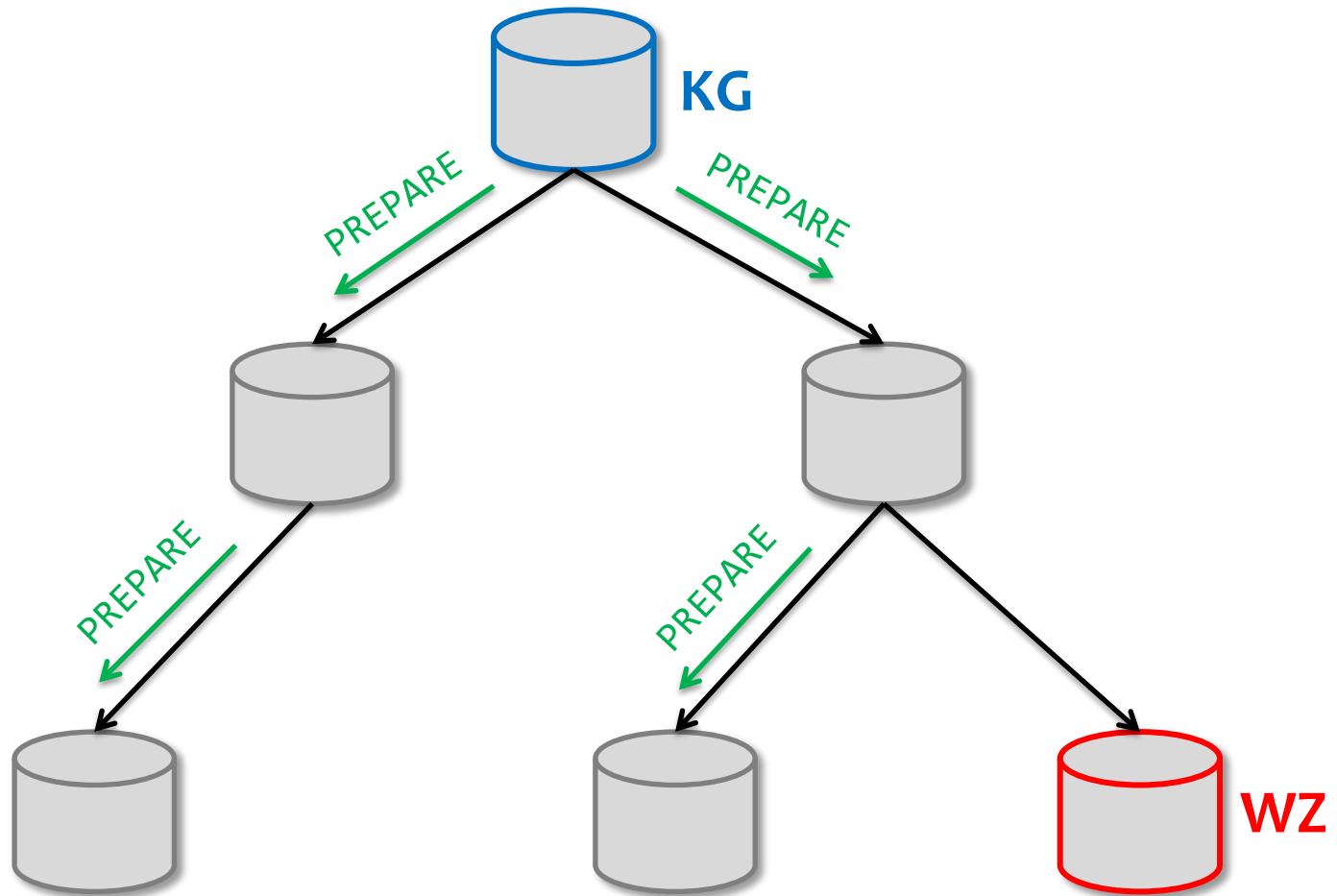
1. Odebranie od **KG** komunikatu zatwierdzenia transakcji
2. Zatwierdzenie lokalnej transakcji
3. **Zwolnienie blokad**



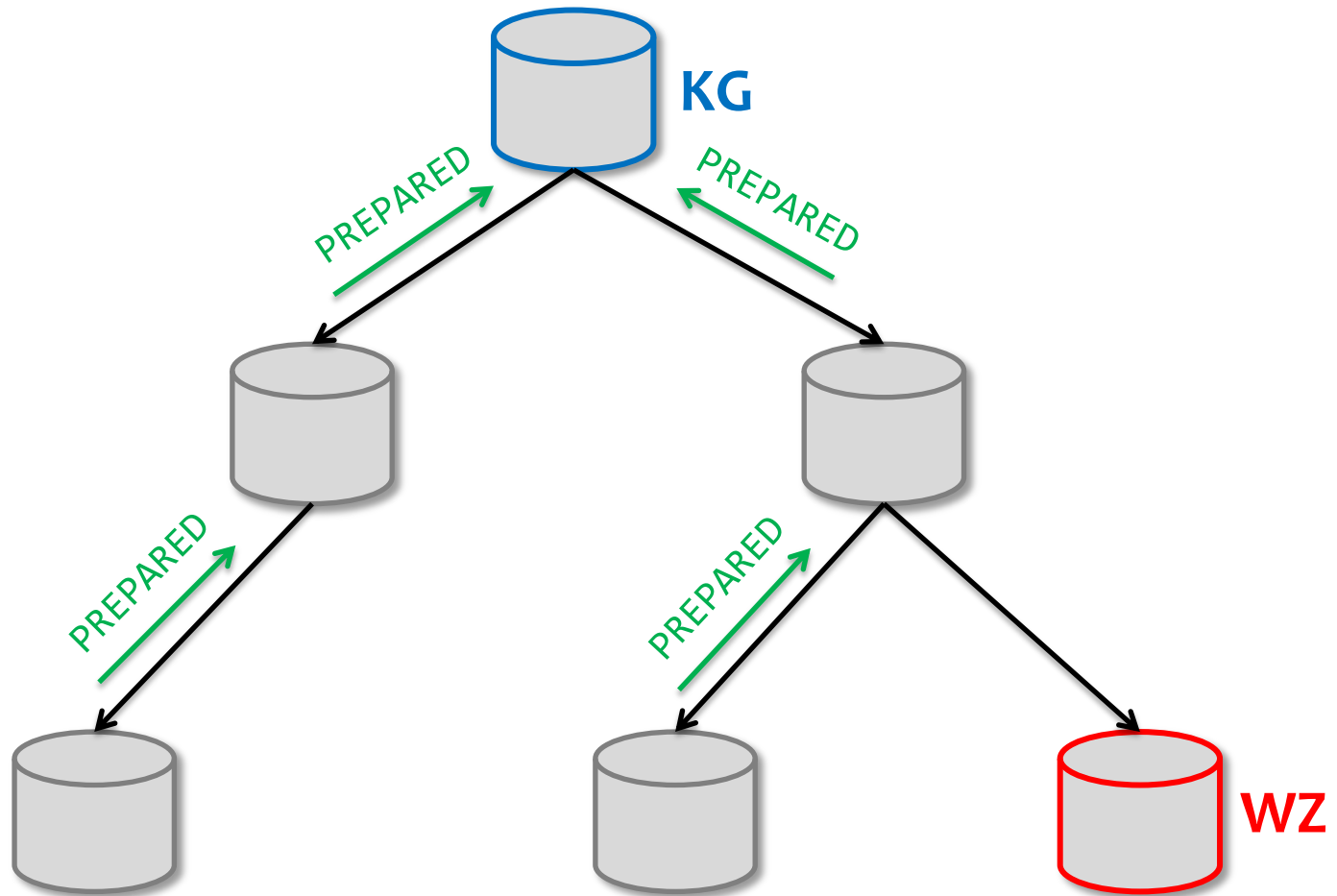
2PC – podsumowanie



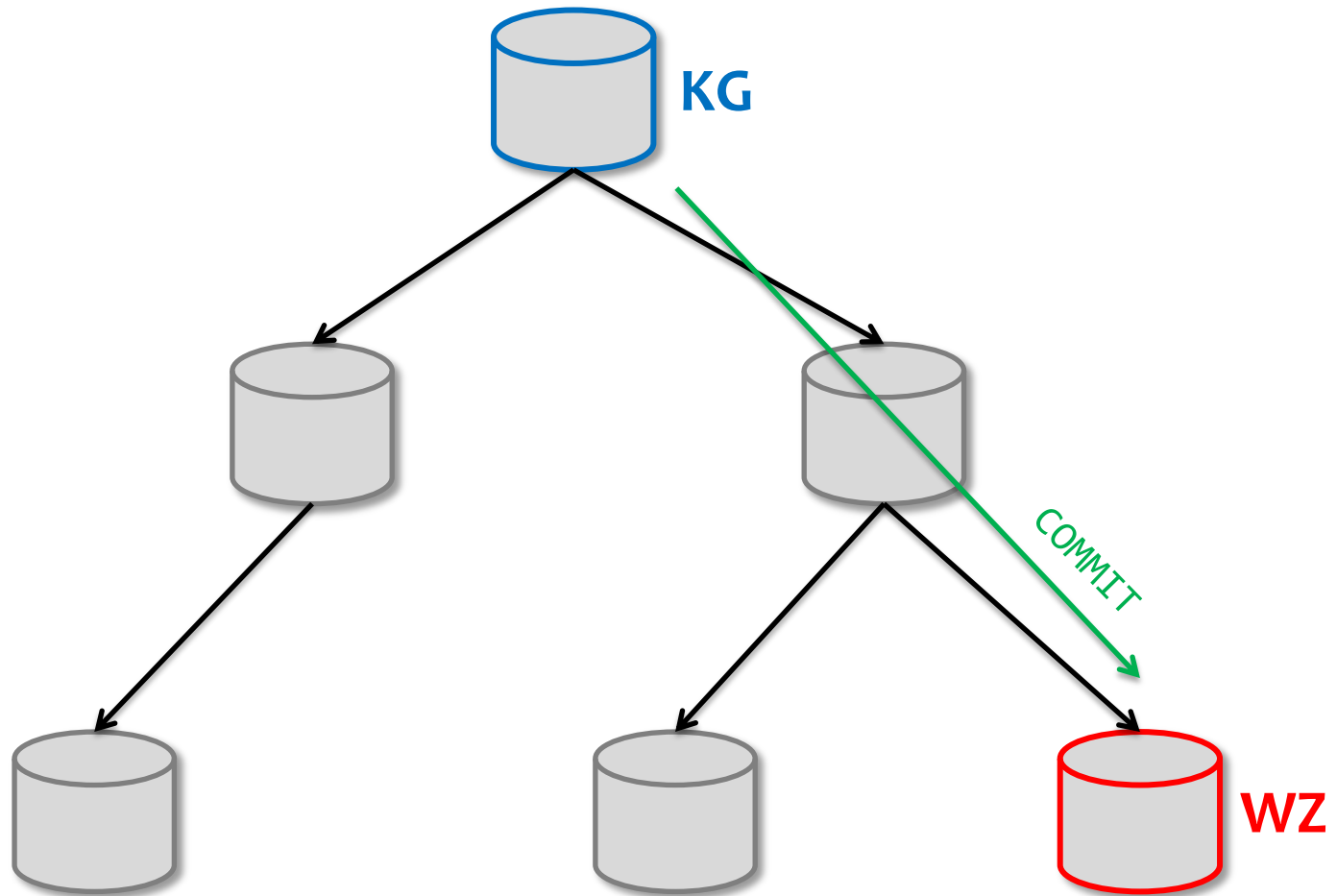
2PC – podsumowanie – faza I



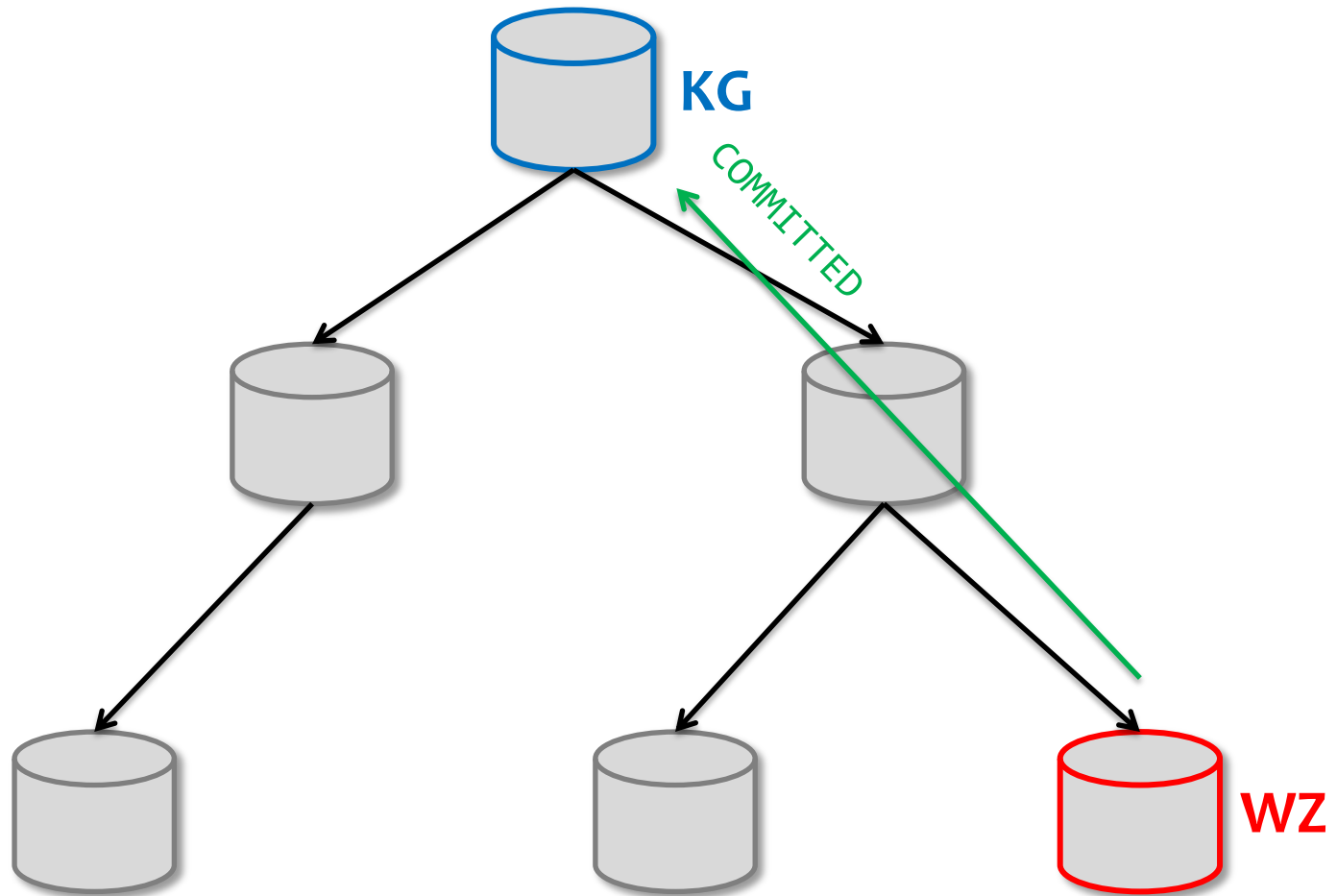
2PC – podsumowanie – faza I



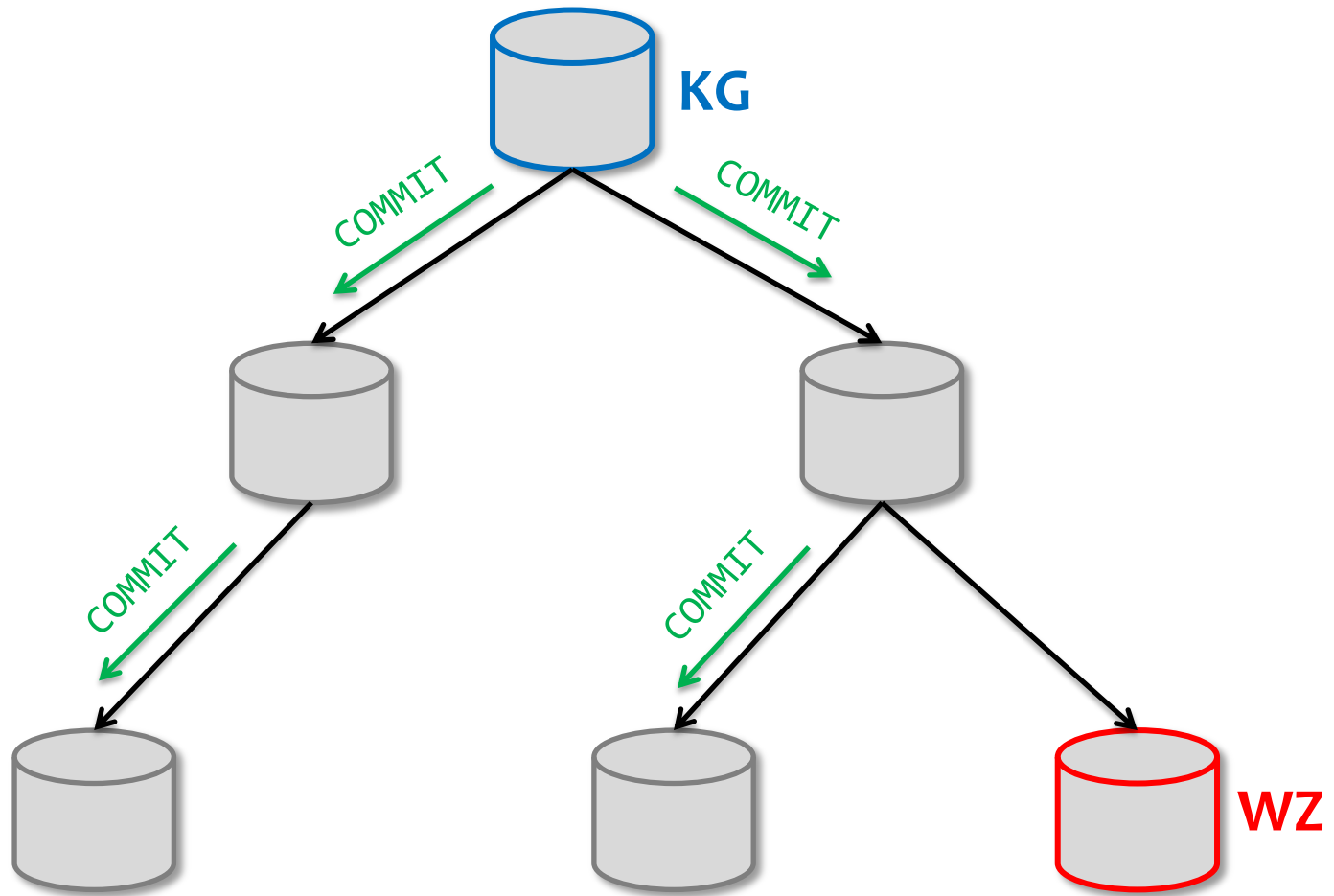
2PC – podsumowanie – faza II



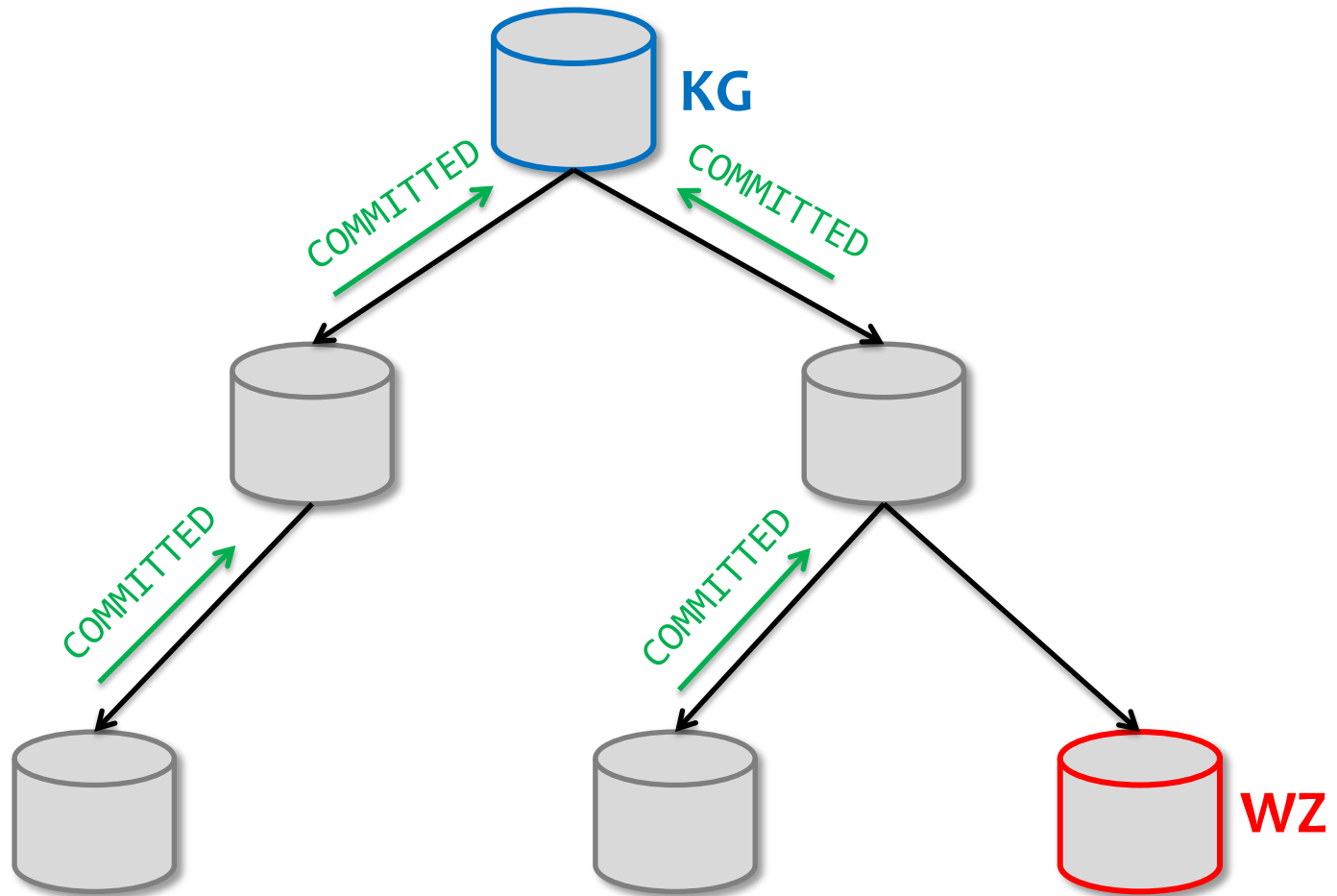
2PC – podsumowanie – faza II



2PC – podsumowanie – faza II



2PC – podsumowanie – faza II



Awarie

- W trakcie zatwierdzania może wystąpić awaria
 - nie wszystkie węzły zatwierdziły/wycofały transakcję
 - nie wszystkie węzły potwierdziły zakończenie operacji
- Transakcja rozproszona w stanie "*in-doubt*"
 - dane są zablokowane
 - **ORA-01591: lock held by in-doubt distributed transaction <id>**
- Automatyczne odtwarzanie transakcji
 - proces RECO
 - wszystkie węzły muszą zatwierdzić lub wycofać zmiany



Ręczne odtwarzanie transakcji

- Stosowane gdy:
 - blokowane dane muszą być natychmiast zwolnione
 - czas usunięcia awarii sprzętowej jest bardzo długi
 - nie działa proces automatycznego odtwarzania



Ręczne odtwarzanie transakcji

- Próbujemy pobrać dane w lokalnej bazie

ORA-01591: lock held by in-doubt distributed transaction 1.21.17

- Analiza zawartości perspektyw słownikowych
 - SYS.DBA_2PC_PENDING
 - SYS.DBA_2PC_NEIGHBORS



Perspektywa SYS.DBA_2PC_PENDING

- LOCAL_TRAN_ID – identyfikator transakcji lokalnej
- GLOBAL_TRAN_ID – identyfikator transakcji rozproszonej
- STATE – stan węzła
 - COLLECTING – czeka na potwierdzenia od węzłów podległych
 - PREPARED – gotowy do zatwierdzenia
 - COMMITTED – zmiany zostały poprawnie zatwierdzone
 - FORCED COMMIT – zmiany zostały zatwierdzone lokalnie
 - FORCED ABORT – zmiany zostały wycofane lokalnie



Perspektywa SYS.DBA_2PC_NEIGHBORS

- **LOCAL_TRAN_ID** – identyfikator transakcji lokalnej
- **IN_OUT** – kierunek komunikacji
 - **IN** – węzeł został wywołany przez daną bazę
 - **OUT** – węzeł wywołuje daną bazę
- **DATABASE** – nazwa bazy danych
- **INTERFACE** – informacja o węźle zatwierdzania
 - **C** – węzeł jest **WZ** lub znajduje się na ścieżce do **WZ**
 - **N** – węzeł nie jest **WZ** ani nie znajduje się na ścieżce do **WZ**



Zwalnianie blokad

- Zatwierdzenie transakcji

```
COMMIT FORCE 'identyfikator.transakcji.lokalnej';
```

- Wycofanie transakcji

```
ROLLBACK FORCE 'identyfikator.transakcji.lokalnej';
```

- Wymagane uprawnienia
 - FORCE ANY TRANSACTION



Symulowanie awarii

1. Crash commit point site after collect
2. Crash non-commit point site after collect
3. Crash before prepare (non-commit point site)
4. Crash after prepare (non-commit point site)
5. Crash commit point site before commit
6. Crash commit point site after commit
7. Crash non-commit point site before commit
8. Crash non-commit point site after commit
9. Crash commit point site before forget
10. Crash non-commit point site before forget

Przykład:

```
COMMIT COMMENT 'ORA-2PC-CRASH-TEST-5' ;
```

