

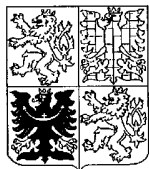
PŘIHLÁŠKA VYNÁLEZU

zveřejněná podle § 31 zákona č. 527/1990 Sb.

(21) Číslo dokumentu:

1999 -4221

(19)
ČESKÁ
REPUBLIKA



ÚŘAD
PRŮMYSLOVÉHO
VLASTNICTVÍ

(22) Přihlášeno: **27.05.1998**

(32) Datum podání prioritní přihlášky: **29.05.1997**

(31) Číslo prioritní přihlášky: **1997/865156**

(33) Země priority: **US**

(40) Datum zveřejnění přihlášky vynálezu: **11.10.2000**
(Věstník č. 10/2000)

(86) PCT číslo: **PCT/GB98/01548**

(87) PCT číslo zveřejnění: **WO98/54648**

(13) Druh dokumentu: **A3**

(51) Int. Cl. ⁷:

G 06 F 11/14

(71) Přihlašovatel:
INTERNATIONAL BUSINESS MACHINES
CORPORATION, Armonk, NY, US;

(72) Původce:
Dias Daniel Manuel, Mahopac, NY, US;
Jhingran Anant Deep, Elmsford, NY, US;
King Richard Pervin, Thornwood, NY, US;
Polyzois Christos, Chatham, NJ, US;

(74) Zástupce:
Kalenský Petr JUDr., Hálkova 2, Praha 2, 12000;

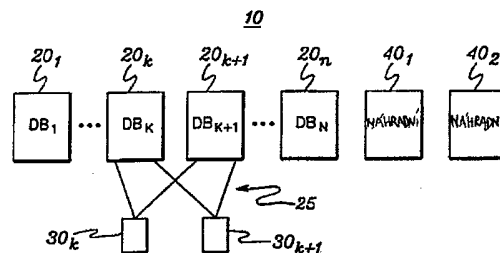
(54) Název přihlášky vynálezu:

**Způsob a systém pro obnovu v rozděleném
databázovém systému bez sdílených prostředků
pomocí virtuálních sdílených disků**

(57) Anotace:

Je popsán způsob a systém pro obnovu ze selhání procesorového uzlu v rozděleném databázovém systému bez sdílených prostředků. Databázový systém (10) může obsahovat dvojici procesorových uzlů (20_k , 20_{k+1}), které mají k sobě zdvojeně připojené úložné zařízení (30_k , 30_{k+1}). Na prvním procesorovém uzlu (20_k) ze dvojice procesorových uzlů (20_k , 20_{k+1}) běží první instance databáze DB_k , která přistupuje k první části dat na úložném zařízení (30_k , 30_{k+1}) před selháním. Po zjištění selhání je zajištěn přístup k první části dat na úložném zařízení (30_k , 30_{k+1}), třetímu náhradnímu procesorovému uzlu (40_1) prostřednictvím druhého procesorového uzlu (20_{k+1}) ze dvojice procesorových uzlů. Na třetím procesorovém uzlu (40_1) běží náhradní instance databáze instance pro první instanci databáze DB_k , která běžela na prvním procesorovém uzlu (20_k) před jeho selháním. Náhradní instance databáze přistupuje k první části dat na úložném zařízení (30_k , 30_{k+1}) prostřednictvím druhého procesorového uzlu (20_{k+1}) tím k obnově ze selhání prvního procesorového uzlu (20_k). Přístup k první části dat může obsahovat použití utility virtuálního sdíleného disku, která má serverovou část na druhém procesorovém uzlu (20_{k+1}) a

klientskou část na třetím procesorovém uzlu (40_1).



ZPŮSOB A SYSTÉM PRO OBNOVU V ROZDĚLENÉM DATABÁZOVÉM SYSTÉMU BEZ SDÍLENÝCH PROSTŘEDKŮ POMOCÍ VIRTUÁLNÍCH SDÍLENÝCH DISKŮ

Oblast techniky

Vynález se týká počítačových databázových systémů. Podrobněji se tento vynález týká obnovy ze selhání procesorového uzlu v databázovém systému bez sdílených prostředků.

Dosavadní stav techniky

Moderní počítačové systémy často obsahují více samostatných procesorů nebo uzlů, které jsou spolu propojeny prostřednictvím komunikační sítě. V takových systémech se často ukládá a zpracovává velké množství informací. Navíc kromě vybavení pro zpracování má každý uzel typicky digitální úložná zařízení (např. magnetické disky) k ukládání informací. Informace jsou často uspořádány jako databáze, která zabírá dostupný úložný prostor na různých uzlech v systému.

Použité techniky k uspořádání požadovaného prostoru databáze a přístup k databázi v počítačovém systému s více uzly závisí na požadavcích na určitý systém. Jisté požadavky jsou však běžné pro většinu systémů. Všechna data v databázi by měla být dostupná pro přístup z libovolného uzlu v systému. Míra režie úložného prostoru a režie zpracování se musí udržovat na minimu, aby se systému umožnil účinný

provoz systému a strategie ukládání/přístupu musí být obecně odolná vůči selhání, které nastane na libovolném uzlu.

V moderních systémech se používají dvě obecné techniky pro ukládání databází nebo dělení na části. První, sdílení dat, obsahuje zajištění fyzického přístupu ke všem diskům z každého uzlu v systému. Aby se však udržela souvislost databáze, je nutné globální uzamykání nebo seznamy změn, aby se zajistilo, že žádné dva uzly nezmění různým způsobem část databáze.

Druhá technika ukládání dat zahrnuje fyzické rozdělení dat a distribuci výsledných částí odpovědným nebo vlasťicím uzlům v systému, které budou odpovědné za transakce týkající se jejich vlastních odpovídajících částí.

Tato architektura bez sdílených prostředků "shared nothing" vyžaduje další režii komunikace, aby bylo možné zajistit přístup ke všem datům ze všech uzlů. Požadující uzel musí vydat databázové požadavky vlasťicímu uzlu. Vlasťicí uzel pak buďto: (i) provede požadovaný databázový požadavek týkající se jeho odpovídající části (tj., dodávka funkce) nebo (i.i) přenese samotná data požadujícímu uzlu (tj.; I/O dodávka).

Problém přístupu bez sdílených prostředků je možnost selhání na kterémkoli uzlu a z toho vyplývající neschopnost tohoto uzlu přijmout nebo zpracovat databázové požadavky týkající se jeho části.

V současnosti jsou známy dvě hlavní metody obnovy selhání uzlu v systému databáze bez sdílených prostředků: (i) asynchronní replikace, kde se posílají asynchronně do

repliky aktualizace dat (viz např. "An Efficient Scheme for Providing High Availability," A. Bhide, A. Goyal, H. Hsiao a A. Jhingran; SIGMOD '92, str. 236-245, který je zde zahrnutý odkazem); a (ii) obnova na přidruženém uzlu, ke kterému jsou disky uzlu, který selhal, připojeny zdvojeným spojem. Zdvojeně připojené diskové jednotky do přidružených procesorových uzlů jsou v současném stavu techniky známy a obsahují fyzické spojení mezi jedním diskem a více nežli jedním procesorovým uzlem. V jednom režimu zdvojeného propojení je aktivní pouze jeden uzel a přistupuje k disku kdykoli. V jiném režimu zdvojeného propojení je umožněno oběma uzlům přistupovat k disku současně a jsou opatřeny protokoly k zamezení/rozřešování konfliktů, aby se zabránilo porušení dat.

Hlavní výhodou tohoto způsobu (i) je, že se může obnovit jak ze selhání disku, tak uzlu, hlavními nevýhodami tohoto způsobu však je, že data jsou zrcadlena, takže spotřebují dvojnásobek diskového prostoru a režie týkající se normálního bezchybného provozu kvůli šíření dat do repliky. Hlavní výhodou způsobu (ii) je, že během normálního provozu není žádná režie, hlavní nevýhoda však je, že po selhání je na přidružený uzel uvalena dvojnásobná zátěž a to může vést k poloviční propustnosti celého seskupení, protože hledání dotazů nebo volání transakčních funkcí přidruženého uzlu k uzlu, který selhal, budou úzkým hrdlem pro celé seskupení.

Tudíž se vyžaduje technika obnovy ze selhání procesorového uzlu v databázovém systému bez sdílených prostředků, který nezpůsobuje značnou režii zpracování během normálního provozu ani režii úložného prostoru kvůli úplné replikaci dat.

Podstata vynálezu

Technika obnovy selhání procesorového uzlu je zajištěna tímto vynálezem, který se z jednoho hlediska týká způsobu a systému obnovy ze selhání prvního procesorového uzlu v databázovém systému, který má více procesorových uzlů. První instance databáze běží na prvním procesorovém uzlu před selháním. První procesorový uzel a druhý procesorový uzel mají k sobě společně připojené první úložné zařízení pro ukládání prvních dat pro první instanci databáze. Po zjištění selhání prvního procesorového uzlu je zajištěn přístup k prvním datům třetímu procesorovému uzlu prostřednictvím druhého procesorového uzlu. První instance databáze pak běží na třetím procesorovém uzlu, který přistupuje k prvním datům na prvním úložném zařízení prostřednictvím druhého procesorového uzlu. Je tudíž zajištěna obnova ze selhání prvního procesorového uzlu.

V upraveném provedení se první data kopírují z prvního úložného zařízení do druhého úložného zařízení připojeného ke třetímu procesorovému uzlu. Za běhu první instance databáze na třetím procesorovém uzlu se mohou další aktualizace první instance databáze zrcadlit na první úložné zařízení a kopírovaná data na druhé úložné zařízení. Po novém spuštění prvního procesorového uzlu může být první procesorový uzel vyhrazen jako náhradní procesorový uzel v systému kvůli následujícím selháním uzlů.

První úložné zařízení může obsahovat dvě úložná zařízení, přičemž každé je zdvojeně připojené k prvnímu a druhému procesorovému uzlu.

Přístup k prvním datům prostřednictvím druhého procesorového uzlu přednostně obsahuje použití utility virtuálního sdíleného disku, která má serverovou část na druhém procesorovém uzlu a klientskou část na třetím procesorovém uzlu.

Protože na druhém procesorovém uzlu může také běžet jeho vlastní druhá instance databáze, může být zajištěn přístup ke druhým datům pro druhou instanci databáze čtvrtému procesorovému uzlu prostřednictvím druhého procesorového uzlu. V tom případě může pak druhá instance databáze běžet na čtvrtém procesorovém uzlu přístupem ke druhým datům na prvním úložném zařízení prostřednictvím druhého procesorového uzlu. Na druhém procesorovém uzlu by se tudíž vyžadovala pouze podpora serverové části a zpracování instance databáze lze úplně odlehčit z druhého procesorového uzlu na třetí a čtvrtý procesorový uzel, na kterých běží příslušné klientské části utility virtuálního sdíleného disku.

Jsou zde zveřejněny další provedení a úpravy těchto technik, včetně obnovy první instance databáze na druhém procesorovém uzlu, během které je proveden pokus znovu spustit první procesorový uzel. Pokud pokus skončí úspěšným spuštěním, spustí se znovu první instance databáze na prvním procesorovém uzlu. Pokud neskončí pokus úspěšně, spustí se instance databáze na druhém procesorovém uzlu nebo náhradním procesorovém uzlu jak bylo popsáno výše.

Tento vynález tudíž zajišťuje techniku účinné obnovy v databázovém systému bez sdílených prostředků, která s sebou nenese podstatnou režii zpracování během normálního provozu,

ani režii úložného prostoru kvůli úplné replikaci dat.

Přehled obrázků na výkresech

Vynález bude blíže vysvětlen prostřednictvím konkrétních příkladů provedení znázorněných na výkresech, na kterých představuje

obr. 1 databázový systém, který má více procesorových uzlů, dva náhradní procesorové uzly a úložná zařízení připojená k alespoň některým z procesorových uzlů;

obr. 2 první provedení tohoto vynálezu, kde se po selhání uzlu spustí instance databáze na jednom z náhradních procesorových uzlů a přistupuje k datům prostřednictvím utility virtuálního sdíleného disku na procesorovém uzlu, ke kterému je připojené úložné zařízení, které má v sobě data pro instanci databáze;

obr. 3 blokové schéma kroků obnovy následující na selháním jednoho z uzlů;

obr. 4 upravené provedení tohoto vynálezu, kde se spustí na dvou příslušných náhradních procesorových jednotkách dvě instance databáze, přičemž každá přistupuje k virtuálnímu sdílenému diskovému serveru na jiném procesorovém uzlu, ke kterému je připojeno úložné zařízení obsahující data pro tyto dvě instance databáze;

obr. 5 další upravené provedení tohoto vynálezu, ve kterém je vytvořena kopie dat z úložných zařízení dat na předtím náhradních procesorových uzlech kvůli podpoře obnovy z budoucích selhání uzlů;

obr. 6 blokové schéma ukazující ještě další upravené provedení tohoto vynálezu, kde je pokus o nové zavedení systému uzlu, který selhal, doprovázen současnou obnovou instance databáze.

Příklady provedení vynálezu

S odkazem na obr. 1 je ukázán databázový systém 10, který má skupinu databázových procesorových uzlů 20₁, ..., 20_n, přičemž na každém normálně běží příslušná instance databáze DB₁; ..., DB_N. Vhodná síť (není ukázána) zajišťuje komunikaci mezi uzly. Disky 30_k a 30_{k+1} jsou zdvojeně propojené k přidruženým procesorovým uzlům 20_k a 20_{k+1} pomocí spojení 25. Zde je ukázána vzorová zdvojená implementace, ale disky mohou být vlastně obecně připojeny vícenásobně k více procesorovým uzlům. Obr. 1 tudíž ukazuje uzel 20_k na kterém běží instance DB_k databáze a jeho přidružený uzel 20_{k+1}, na kterém běží instance DB_{k+1} databáze.

Za normálního provozu jsou disky logicky rozdělené mezi přidružené uzly tak, aby jeden uzel logicky vlastnil jednu podskupinu zdvojeně připojených disků a přidružený uzel zbytek. Mohou existovat řádově desítky, možná až stovky databázových procesorových uzlů. Navíc je v systému nakonfigurována skupina náhradních procesorových uzlů 40₁, 40₂. V systému může existovat alespoň jeden náhradní uzel,

přednostně dva a eventuálně více náhradních uzlů.

Technika tohoto vynálezu obnovy ze selhání procesorového uzlu je znázorněná na obr. 2. Tento obrázek znázorňuje případ, kdy uzel 20_{k+1} , na kterém běžela instance DB_{k+1} databáze, selže. Způsob podle předchozího stavu techniky (ii) (popsaný výše) by obnovil instanci DB_{k+1} databáze na přidruženém uzlu 20_k , takže by po selhání na uzlu 20_k běžela jak instance databáze DB_k , tak instance DB_{k+1} . Jak bylo popsáno výše, mohlo by toto vést k dvojnásobné zátěži na uzlu 20_k s následnou ztrátou výkonu celého systému. Obr. 2 znázorňuje zveřejněnou techniku k vyřešení tohoto problému: Po selhání se spustí instance DB_{k+1} databáze na samostatném náhradním uzlu 40_1 . Tato instance databáze stále potřebuje přistupovat ke stejným diskům, které jí byly logicky přiřazené před selháním. Jak je znázorněno na obr. 2, po selhání se disky, které byly v logické části uzlu 20_{k+1} , který selhal, znovu nakonfigurují s přístupem prostřednictvím přidruženého uzlu 20_k komunikační cestou 50 vhodnou komunikační sítí (není ukázána). Tento přístup je zajištěn například utilitou obnovitelných virtuálních sdílených disků (RVSD) zveřejněnou v patentové přihlášce číslo 08/332157, nazvanou "Application-Transparent Recovery for Virtual Shared Disks". Při selhání uzlu se RVSD transparentně přepne kvůli zajištění přístupu k diskům 30_k a 30_{k+1} prostřednictvím přidruženého uzlu 20_k , z kteréhokoli uzlu v systému. Po obnově RVSD se znovu spustí instance DB_{k+1} databáze, která běžela na uzlu, který selhal, na jednom ze záložních uzlů 40_1 . Instance DB_{k+1} logicky vlastní stejné disky a přistupuje k části databáze selhané instance vytvářením požadavků na čtení/zápis disku prostřednictvím klientské části RVSD na uzlu 40_1 serverové části na uzlu 20_k . RVSD utilita transparentně předává požadavky uzlu 20_k a

získává příslušná data.

S použitím techniky znázorněné na obr. 2 má uzel 20_k zátěž odpovídající zátěži instance DB_k databáze a zátěži VSD serveru podporující instanci DB_{k+1} na uzlu 40_1 . Tato zátěž by byla nižší, nežli zátěž při obnově plné instance DB_1 na uzlu 20_k po selhání. Tudiž s touto volbou bude propustnost po selhání o něco málo snížena kvůli zdvojeným povinnostem uzlu 20_k , ale tato propustnost je značně větší, nežli ve výše uvedených případech podle předchozího stavu techniky.

Obr. 3 je blokové schéma kroků nutných k implementaci techniky obnovy z obr. 2. Po selhání uzlu 20_{k+1} v kroku 100, se zvolí náhradní uzel 40_1 , na kterém se má spustit instance DB_{k+1} v kroku 110. Za předpokladu, že disk 30_{k+1} nese části týkající se instance DB_{k+1} , provede uzel 20_k VSD převzetí disku 30_{k+1} v kroku 120. Příslušná klientská část VSD se nakonfiguruje na uzlu 40_1 v kroku 130. V kroku 140 jsou dále informovány všechny ostatní uzly (prostřednictvím aktualizace příslušných tabulek v systému) o tom, že instance DB_{k+1} nyní poběží na uzlu 40_1 .

Všechny příslušné požadavky na instanci DB_{k+1} databáze budou tudiž směřovány do uzlu 40_1 . Nakonec v kroku 150 se spustí instance DB_{k+1} na uzlu 40_1 .

Upravené provedení tohoto vynálezu je znázorněno na obr. 4. Obr. 4 ukazuje instanci databáze DB_k , která se také znovu spouští na jiném náhradním uzlu 40_2 se vzdáleným VSD přístupem ke svým datům cestou 60 uzlem 20_k . Spuštěním obou instancí DB_k a DB_{k+1} databází na náhradních uzlech 40_2 a 40_1 se po zátěži na uzlu 20_k požaduje pouze zpracovávat VSD přístupy od obou těchto instancí. Měření ukazují, že s touto

konfigurací je zatížení VSD na uzlu 20_k pravděpodobně nižší, než zatížení při normálním provozu. Dále je propustnost sekvenčního přístupu prostřednictvím VSD velmi blízká propustnosti sekvenčního přístupu lokálního disku a propustnost při náhodném čtení může také zajistit VSD. Tato konfigurace tudíž povede k velice podobnému výkonu po selhání uzlu 20_{k+1} jako je normální výkon. Avšak nepříjemným důsledkem může být, že přesun instance DB_k databáze může způsobit pád této instance a její nové spuštění na náhradním uzlu. Vliv tohoto závisí na pracovním zatížení. Co se týče podpory rozhodování, ovlivní selhání uzlu 20_{k+1} pravděpodobně většinu, pokud ne všechny, běžící dotazy; ukončení a nové spuštění DB_k bude tudíž také pravděpodobně přípustné. Pro OLTP bude tato volba záviset na části pracovní zátěže ovlivněné selháním uzlu 20_k oproti případu kdy jsou oba uzly 20_k a 20_{k+1} ukončeny při selhání.

Jeden možný problém s touto technikou je ošetřování nového sjednocování po návratu uzlu 20_{k+1}. V nejjednodušším případě mohl uzel 20_{k+1} selhat kvůli zhroucení operačního systému a pouhé nové spuštění by jej mohlo obnovit. Ideálně by bylo žádoucí obnovit systém do konfigurace, která může ošetřit další selhání (tj., režim, který má vyhrazen dostatek náhradních uzlů). Jedna alternativa je přesunout instance databází DB_k a DB_{k+1} zpět na uzly 20_k a 20_{k+1} v uvedeném pořadí. To však typicky vyžaduje ukončení instancí databáze a pak jejich nové spuštění na původních uzlech. Rozšíření techniky zde zveřejněné, které ošetřuje nové začlenění bez ukončování instancí databází zahrnuje kopírování dat z disků 30_k a 30_{k+1} do zdvojeně připojených (65) disků 70₁ a 70₂ na náhradních uzlech. To lze provést současně s provozem databáze po změně konfigurace po selhání. Přístupem prostřednictvím VSD lze data zrcadlit do

zdvojeně připojených disků na náhradních uzlech a jakékoli další aktualizace disků se musí zrcadlit jak na uzlu 20_k tak na předtím náhradních uzlech 40₁ a 40₂. Odborníci snadno ocení, že toho lze dosáhnout příslušnou synchronizací. Uzly 20_k a 20_{k+1} mohou být poté navrženy jako náhradní uzly v systému pro obnovu z budoucích selhání jiných uzlů.

Jak bylo uvedeno výše, může pouhé nové zavedení stačit k obnově uzlu 20_{k+1}, který selhal, do pracovního stavu. V takovém případě může být žádoucí zamezit naprosto převzetí. To však vyžaduje odklad rozhodnutí o převzetí až po dokončení nového spuštění uzlu, který selhal, a po zavedení systému, což patřičně prodlouží dobu obnovy. Následující techniku znázorněnou na blokovém schématu obr. 6 lze použít k překrytí akcí obnovy s pokusem o nové zavedení systému na uzlu, který selhal. Jakmile uzel 20_k selže, (krok 200), převezme jeho přidružený uzel 20_k jeho disky a zahájí obnovu, tj. provede obnovu systému souborů a obnovu instance databáze založenou na protokolu (krok 210). Během této doby obnovy se může uzel 20_{k+1}, který selhal, pokusit o nové zavedení systému (krok 220). Pokud bude úspěšný, (rozhodnutí 230, "A"), převezme znovu řízení svých původních disků a znovu spustí lokálně instanci databáze (krok 250). Pokud se mu nepodaří znovu zavést systém, (rozhodnutí 230, "N"), spustí se instance databáze na přidruženém uzlu 20_k nebo na náhradním uzlu, jak bylo popsáno výše (krok 240). Ve všech případech je nové spuštění instance databáze okamžité, protože obnova disků byla již provedena přidruženým uzlem 20_k.

Techniky tohoto vynálezu jsou použitelné na databázové systémy a zejména na libovolně rozdělené databázové systémy (bez sdílených prostředků).

Tento vynález může být zahrnut do výrobku (např. jednoho nebo více produktů počítačových programů), který má například počítačem použitelné médium. Médium má v sobě zahrnuty například počítačem čitelné prostředky programového kódu kvůli zajištění a usnadnění mechanismů tohoto vynálezu. Výrobek může být zahrnut jako část počítačového systému nebo se může prodávat zvlášť.

I když byl vynález ukázán a popsán zejména s odkazem na své upřednostňované provedení, odborníci rozumí, že v něm lze provádět různé změny ve formě a podrobnostech bez odchýlení od ducha a rozsahu vynálezu.

Zastupuje:

Dr. Petr Kalenský v.r.



STÁTNÍ A VYKONÁVACÍ KANCELARIE
VÝKONNÝ ZÁSTUPCE PŘEDSÍDKY
A PŘEDNÍ
120 00 Praha 2, Hájkova 2
Česká republika

PATENTOVÉ NÁROKY

1. Způsob pro obnovu ze selhání prvního procesorového uzlu v databázovém systému, který má více procesorových uzlů, **vyznačující se tím**, že obsahuje:

spuštění první instance databáze na prvním procesorovém uzlu, přičemž první procesorový uzel a druhý procesorový uzel má k sobě běžně připojeno alespoň jedno první úložné zařízení pro ukládání prvních dat pro první instanci databáze;

zjištění selhání prvního procesorového uzlu;

zajištění třetímu procesorovému uzlu přístupu k prvním datům na alespoň jednom úložném zařízení prostřednictvím druhého procesorového uzlu; a

spuštění první instance databáze na třetím procesorovém uzlu, obsahující přístup k prvním datům na alespoň jednom prvním úložném zařízení prostřednictvím druhého procesorového uzlu, a tím obnovení ze selhání prvního procesorového uzlu.

2. Způsob podle nároku 1, **vyznačující se tím**, že dále obsahuje:

zkopírování prvních dat z alespoň jednoho prvního úložného zařízení na alespoň jedno druhé úložné zařízení připojené ke třetímu procesorovému uzlu;

kde spuštění první instance databáze na třetím procesorovém uzlu obsahuje zrcadlení následných aktualizací první instance databáze do prvních dat na alespoň jednom prvním úložném zařízení a zkopírovaných prvních dat na alespoň jednom druhém úložném zařízení.

3. Způsob podle nároku 1 nebo nároku 2, **vyznačující se tím**, že dále obsahuje po novém spuštění prvního procesorového uzlu:

navržení prvního procesorového uzlu jako prvního náhradního procesorového uzlu v databázovém systému.

4. Způsob podle kteréhokoli z nároků 1 až 3, **vyznačující se tím**, že alespoň jedno první úložné zařízení obsahuje dvě úložná zařízení, přičemž každé je zdvojeně připojené k prvnímu a druhému procesorovému uzlu.

5. Způsob podle kteréhokoli z nároků 1 až 3, **vyznačující se tím**, že alespoň jedno první úložné zařízení obsahuje více úložných zařízení, přičemž každé je vícenásobně připojené k prvnímu, druhému a jiným procesorovým uzlům v databázovém systému.

6. Způsob podle kteréhokoli z předcházejících nároků, **vyznačující se tím**, že první data obsahují část rozdělené databáze bez sdílených prostředků, nacházející se v databázovém systému.

7. Způsob podle kteréhokoli z předchozích nároků, **vyznačující se tím**, že zajištění třetímu procesorovému uzlu přístupu k prvním datům obsahuje použití utility virtuálního sdíleného disku, která má serverovou

část na druhém procesorovém uzlu a klientskou část na třetím procesorovém uzlu.

8. Způsob podle kteréhokoli z předchozích nároků, **vyznačující se tím**, že třetí procesorový uzel je vyhrazený náhradní procesorový uzel v databázovém systému.

9. Způsob podle kteréhokoli z předchozích nároků, **vyznačující se tím**, že dále obsahuje:

před zjištěním selhání prvního procesorového uzlu, spuštění druhé instance databáze na druhém procesorovém uzlu, přičemž alespoň jedno první úložné zařízení má v sobě uloženo druhá data pro druhou instanci databáze; a

po zjištění selhání prvního procesorového uzlu:

zajištění čtvrtému procesorovému uzlu přístupu ke druhým datům na alespoň jednom prvním úložném zařízení prostřednictvím druhého procesorového uzlu a

spuštění druhé instance databáze na čtvrtém procesorovém uzlu obsahujícím přístup ke druhým datům na alespoň jednom prvním úložném zařízení prostřednictvím druhého procesorového uzlu.

10. Způsob podle nároku 9, **vyznačující se tím**, že dále obsahuje:

zkopírování prvních dat a druhých dat z alespoň jednoho prvního úložného zařízení alespoň na jedno druhé úložné zařízení normálně připojené ke třetímu a čtvrtému procesorovému uzlu;

kde spuštění první instance databáze na třetím procesorovém uzlu obsahuje zrcadlení následných aktualizací první instance databáze do prvních dat na alespoň jednom prvním úložném zařízení a do zkopírovaných prvních dat na alespoň jednom druhém úložném zařízení; a

kde spuštění druhé instance databáze na čtvrtém procesorovém uzlu obsahuje zrcadlení následujících aktualizací do druhé instance databáze do druhých dat na alespoň jednom prvním úložném zařízení a do zkopírovaných druhých dat na alespoň jednom druhém úložném zařízení.

11. Způsob podle nároku 9 nebo 10, **vyznačující se tím**, že dále obsahuje po novém spuštění prvního procesorového uzlu:

určení prvního procesorového uzlu jako prvního náhradního procesorového uzlu v databázovém systému; a

navržení druhého procesorového uzlu jako druhého náhradního procesorového uzlu v databázovém systému.

12. Způsob podle jednoho z nároků 9 až 11, **vyznačující se tím**, že první a druhá data každá obsahují příslušné části rozdělené databáze bez sdílených prostředků umístěné v databázovém systému.

13. Způsob podle kteréhokoli z nároků 9 až 12, kde zajištění čtvrtému procesorovému uzlu přístupu ke druhým datům obsahuje použití utility virtuálního sdíleného disku, který má serverovou část na druhém procesorovém uzlu a první klientskou část na čtvrtém procesorovém uzlu.

25.11.99

14. Způsob podle nároku 13, **vyznačující se tím**, že zajištění třetímu procesorovému uzlu přístupu k prvním datům obsahuje použití utility virtuálního sdíleného disku, který má serverovou část na druhém procesorovém uzlu a druhou klientskou část na třetím procesorovém uzlu.

15. Způsob podle nároku 10, **vyznačující se tím**, že třetí procesorový uzel a čtvrtý procesorový uzel jsou vyhrazené náhradní procesorové uzly v databázovém systému.

16. Způsob obnovy ze selhání prvního procesorového uzlu v databázovém systému, který má více procesorových uzlů, **vyznačující se tím**, že obsahuje:

spuštění první instance databáze na prvním procesorovém uzlu, přičemž první procesorový uzel a druhý procesorový uzel má k sobě běžně připojeno alespoň jedno úložné zařízení pro ukládání prvních dat pro první instanci databáze;

zjištění selhání prvního procesorového uzlu;

provedení obnovy první instance databáze na druhém procesorovém uzlu, obsahujícím přístup k prvním datům na alespoň jednom úložném zařízení prostřednictvím druhého procesorového uzlu;

během provádění obnovy databáze, pokus o nové spuštění prvního procesorového uzlu; a

pokud je výsledkem pokusu úspěšné nové spuštění prvního procesorového uzlu, pak nové spuštění první instance

databáze na prvním procesorovém uzlu obsahující přístup k prvním datům alespoň na jednom úložném zařízení prostřednictvím prvního procesorového uzlu, nebo

pokud výsledek uvedeného pokusu není úspěšné nové spuštění prvního procesorového uzlu, pak spuštění první instance databáze na druhém procesorovém uzlu, obsahujícím přístup k prvním datům na alespoň jednom úložném zařízení prostřednictvím druhého procesorového uzlu, tím obnovení ze selhání prvního procesorového uzlu.

17. Způsob podle nároku 16, **vyznačující se tím**, že alespoň jedno úložné zařízení obsahuje dvě úložná zařízení, přičemž každé je zdvojeně připojené k prvnímu a druhému procesorovému uzlu.

18. Způsob podle nároku 16, **vyznačující se tím**, že alespoň jedno úložné zařízení obsahuje více úložných zařízení, přičemž každé je vícenásobně připojené k prvnímu, druhému a jiným procesorovým uzlům v databázovém systému.

19. Způsob podle kteréhokoli z nároků 16 až 18, **vyznačující se tím**, že první data obsahují část rozdělené databáze bez sdílených prostředků nacházející se v databázovém systému.

20. Způsob podle nároku 1, **vyznačující se tím**, že obsahuje:

provedení obnovy databáze první instance databáze na druhém procesorovém uzlu, obsahujícím přístup k prvním datům na alespoň jednom úložném zařízení prostřednictvím druhého

procesorového uzlu;

během provádění obnovy databáze, provedení pokusu o nové spuštění prvního procesorového uzlu; a

pokud je výsledek pokusu úspěšný, nové spuštění prvního procesorového uzlu, pak nové spuštění první instance databáze na prvním procesorovém uzlu, obsahujícím přístup k prvním datům na alespoň jednom úložném zařízení prostřednictvím prvního procesorového uzlu, nebo

pokud není výsledek pokusu úspěšné nové spuštění prvního procesorového uzlu, pak provedení kroků:

zajištění třetímu procesorovému uzlu přístupu k prvním datům na alespoň jednom úložném zařízení prostřednictvím druhého procesorového uzlu; a spuštění první instance databáze na třetím procesorovém uzlu, obsahujícím přístup k prvním datům na alespoň jednom úložném zařízení prostřednictvím druhého procesorového uzlu.

21. V rozděleném databázovém systému bez sdílených prostředků, způsob podle nároku 1 pro obnovu ze selhání prvního procesorového uzlu ze dvojice procesorových uzlů, které k sobě mají zdvojeně připojené alespoň jedno úložné zařízení, přičemž na prvním procesorovém uzlu běží první instance databáze a přistupuje k první části dat na alespoň jednom úložném zařízení před selháním, přičemž způsob , **se vyznačuje tím**, že obsahuje:

zajištění třetímu procesorovému uzlu přístupu k první části dat na alespoň jednom úložném zařízení prostřednictvím druhého procesorového uzlu z dvojice procesorových uzlů; a

spuštění na třetím procesorovém uzlu první instance náhradní databáze pro první instanci databáze, která běžela na prvním procesorovém uzlu před jeho selháním, obsahujícím přístup k první části dat na alespoň jednom úložném zařízení prostřednictvím druhého procesorového uzlu, a tím obnovení ze selhání prvního procesorového uzlu.

22. Systém pro obnovu ze selhání prvního procesorového uzlu v databázovém systému, **vyznačující se tím**, že má více procesorových uzlů, obsahujících:

prostředky pro spuštění první instance databáze na prvním procesorovém uzlu, přičemž první procesorový uzel a druhý procesorový uzel má k sobě běžně připojeno alespoň jedno první úložní zařízení pro ukládání prvních dat pro první instanci databáze;

prostředky pro zjišťování selhání prvního procesorového uzlu;

prostředky pro zajištění třetímu procesorovému uzlu přístupu k prvním datům na alespoň jednom úložném zařízení prostřednictvím druhého procesorového uzlu; a

prostředky pro spuštění první instance databáze na třetím procesorovém uzlu, obsahujícím prostředky pro přístup k prvním datům na alespoň jednom prvním úložném zařízení prostřednictvím druhého procesorového uzlu, a tím obnovení ze selhání prvního procesorového uzlu.

23. Systém podle nároku 22, **vyznačující se tím**, že alespoň jedno první úložné zařízení obsahuje dvě

úložná zařízení, přičemž každé je zdvojeně připojené k prvnímu a druhému procesorovému uzlu.

24. Systém podle nároku 22, **vyznačující se tím**, že alespoň jedno první úložné zařízení obsahuje více úložných zařízení, přičemž každé je vícenásobně připojeno k prvnímu, druhému a jiným procesorovým uzlům v databázovém systému.

25. Systém podle kteréhokoli z nároků 22 až 24, **vyznačující se tím**, že první data obsahují část rozdělené databáze bez sdílených prostředků nacházející se v databázovém systému.

26. Systém podle kteréhokoli z nároků 22 až 25, **vyznačující se tím**, že uvedené prostředky pro zajištění třetímu procesorovému uzlu přístupu k prvním datům obsahují utilitu virtuálního sdíleného disku, která má serverovou část na druhém procesorovém uzlu a klientskou část na třetím procesorovém uzlu.

27. Systém podle kteréhokoli z nároků 22 až 26, **vyznačující se tím**, že dále obsahuje:

prostředky pro spuštění druhé instance databáze na druhém procesorovém uzlu před zjištěním selhání prvního procesorového uzlu, přičemž alespoň jedno první úložné zařízení má v sobě uložená druhá data pro druhou instanci databáze;

prostředky pro zajištění čtvrtému procesorovému uzlu přístupu ke druhým datům na alespoň jednom prvním úložném zařízení prostřednictvím druhého procesorového uzlu po

zjištění selhání prvního procesorového uzlu; a

prostředky pro spuštění druhé instance databáze na čtvrtém procesorovém uzlu obsahujícím prostředky pro přístup ke druhým datům na alespoň jednom prvním úložném zařízení prostřednictvím druhého procesorového uzlu.

28. V rozděleném databázovém systému bez sdílených prostředků, systém podle nároku 22 pro obnovu ze selhání prvního procesorového uzlu ze dvojice procesorových uzlů, které mají k sobě zdvojeně připojeno alespoň jedno úložné zařízení, přičemž na prvním procesorovém uzlu běží první instance databáze a přistupuje k první části dat na alespoň jednom úložném zařízení před selháním, přičemž systém **se vyznačuje tím**, že obsahuje:

prostředky k zajištění třetímu procesorovému uzlu přístupu k první části dat na alespoň jednom úložném zařízení prostřednictvím druhého procesorového uzlu ze dvojice procesorových uzlů; a

prostředky pro spuštění na třetím procesorovém uzlu první náhradní instance databáze pro první instanci databáze, která běžela na prvním procesorovém uzlu před jeho selháním, obsahujícím prostředky pro přístup k první části dat na alespoň jednom úložném zařízení prostřednictvím druhého procesorového uzlu, a tím obnovení ze selhání prvního procesorového uzlu.

29. Výrobek obsahující počítačem použitelné médium, které má v sobě počítačem čitelné prostředky kódu pro obnovu ze selhání prvního procesorového uzlu v databázovém systému, který má více procesorových uzlů, **vyznačující se**

tím, že prostředky počítačem čitelného programového kódu ve výrobku obsahují:

počítačem čitelné prostředky programového kódu pro spuštění první instance databáze na prvním procesorovém uzlu, přičemž první procesorový uzel a druhý procesorový uzel k sobě mají běžně připojeno alespoň jedno první úložné zařízení pro ukládání prvních dat pro první instanci databáze;

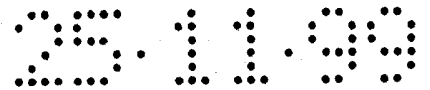
počítačem čitelné prostředky programového kódu pro zjišťování selhání prvního procesorového uzlu;

počítačem čitelné prostředky programového kódu pro zajištění třetímu procesorovému uzlu přístupu k prvním datům na alespoň jednom úložném zařízení prostřednictvím druhého procesorového uzlu; a

počítačem čitelné prostředky programového kódu pro spuštění první instance databáze na třetím procesorovém uzlu, obsahující prostředky kódu pro přístup k prvním datům na alespoň jednom prvním úložném zařízení prostřednictvím druhého procesorového uzlu, a tím obnovení ze selhání prvního procesorového uzlu.

30. Výrobek podle nároku 29, **vyznačující se tím**, že první data obsahují část rozdělené databáze bez sdílených prostředků nacházející se v databázovém systému.

31. Výrobek podle nároku 29 nebo 30, **vyznačující se tím**, že prostředky kódu pro zajišťování třetímu procesorovému uzlu přístupu k prvním datům, obsahují utilitu virtuálního sdíleného disku, která má serverovou část na



druhém procesorovém uzlu a klientskou část na třetím procesorovém uzlu.

32. Výrobek podle kteréhokoli z nároků 29 až 31, **vyznačující se tím**, že dále obsahují:

počítačem čitelné prostředky programového kódu pro spuštění druhé instance databáze na druhém procesorovém uzlu před zjištěním selhání prvního procesorového uzlu, přičemž alespoň jedno první úložné zařízení má v sobě uložená druhá data pro druhou instanci databáze;

počítačem čitelné prostředky programového kódu pro zajištění čtvrtému procesorovému uzlu přístupu ke druhým datům na alespoň jednom prvním úložném zařízení prostřednictvím druhého procesorového uzlu po zjištění selhání prvního procesorového uzlu; a

počítačem čitelné prostředky programového kódu pro spuštění druhé instance databáze na čtvrtém procesorovém uzlu, obsahujícím prostředky kódu pro přístup ke druhým datům na alespoň jednom prvním úložném zařízení prostřednictvím druhého procesorového uzlu.

33. Výrobek podle nároku 29, **vyznačující se tím**, že obsahuje počítačem použitelné médium, které má v sobě počítačem čitelné prostředky programového kódu pro obnovu v rozděleném databázovém systému bez sdílených prostředků ze selhání prvního procesorového uzlu ze dvojice procesorových uzlů, které mají k sobě zdvojeně připojené alespoň jedno úložné zařízení, přičemž na prvním procesorovém uzlu běží první instance databáze a přistupuje k první části dat na alespoň jednom úložném zařízení před

25.11.99

selháním, přičemž počítačem čitelné prostředky programového kódu ve výrobku obsahují:

počítačem čitelné prostředky programového kódu pro zajištění třetímu procesorovému uzlu přístupu k první části dat na alespoň jednom úložném zařízení prostřednictvím druhého procesorového uzlu z dvojice procesorových uzlů; a

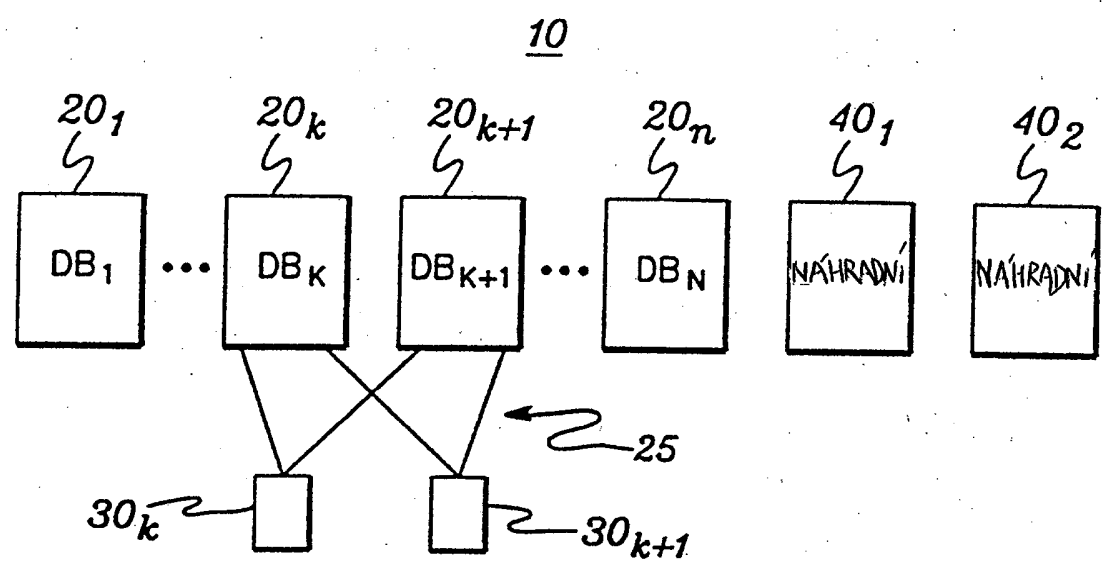
počítačem čitelné prostředky programového kódu pro spuštění na třetím procesorovém uzlu první náhradní instance databáze pro první instanci databáze, která běžela na prvním procesorovém uzlu před jeho selháním, obsahujícím prostředky kódu pro přístup k první části dat na alespoň jednom úložném zařízení prostřednictvím druhého procesorového uzlu, a tím obnovení ze selhání prvního procesorového uzlu.

Zastupuje:

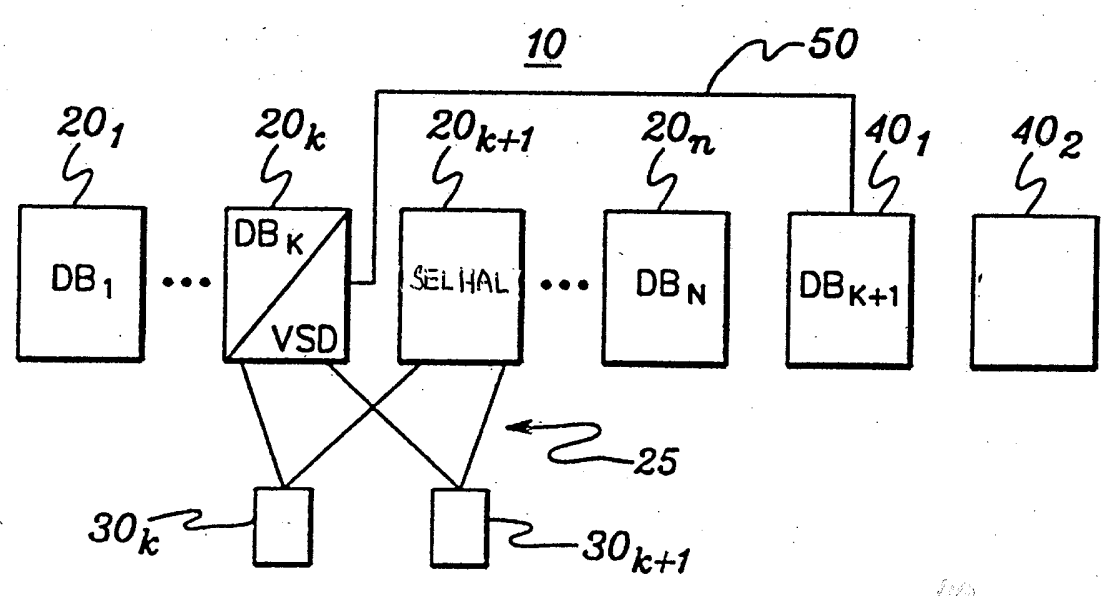
Dr. Petr Kalenský v.r.



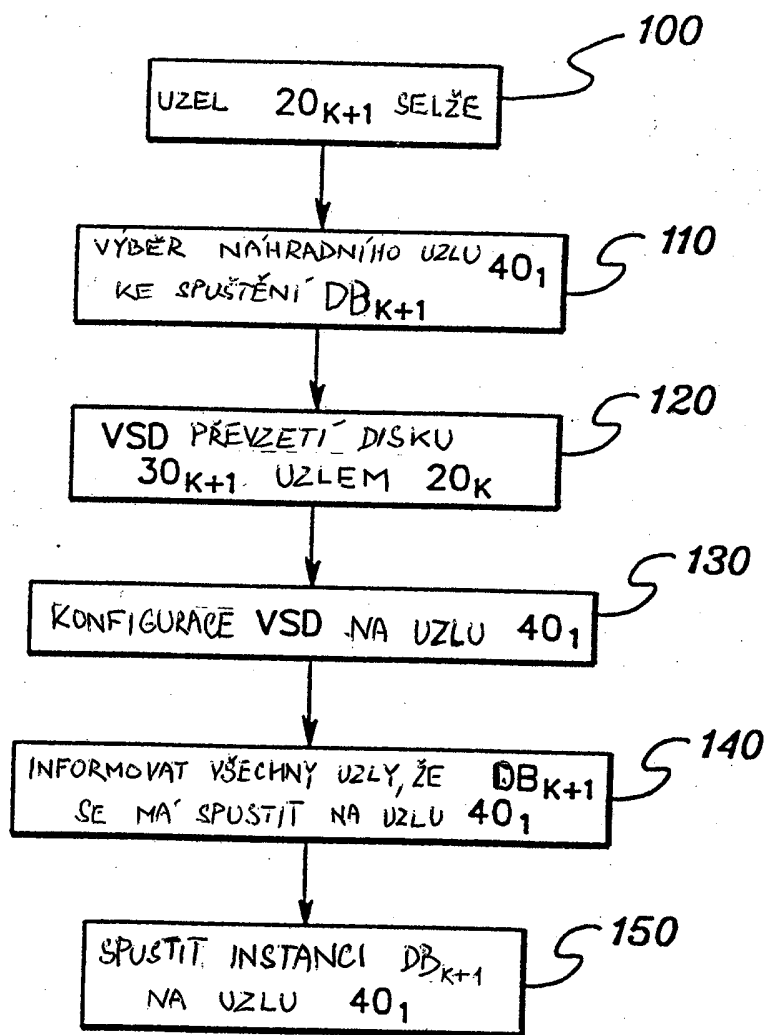
ADVOKÁTNÍ KANCELÁŘ
DR. PETR KALENSKÝ
A PARTNERŮ
120 00 Praha 2, Hájkova 2
Česká republika



OBR. 1

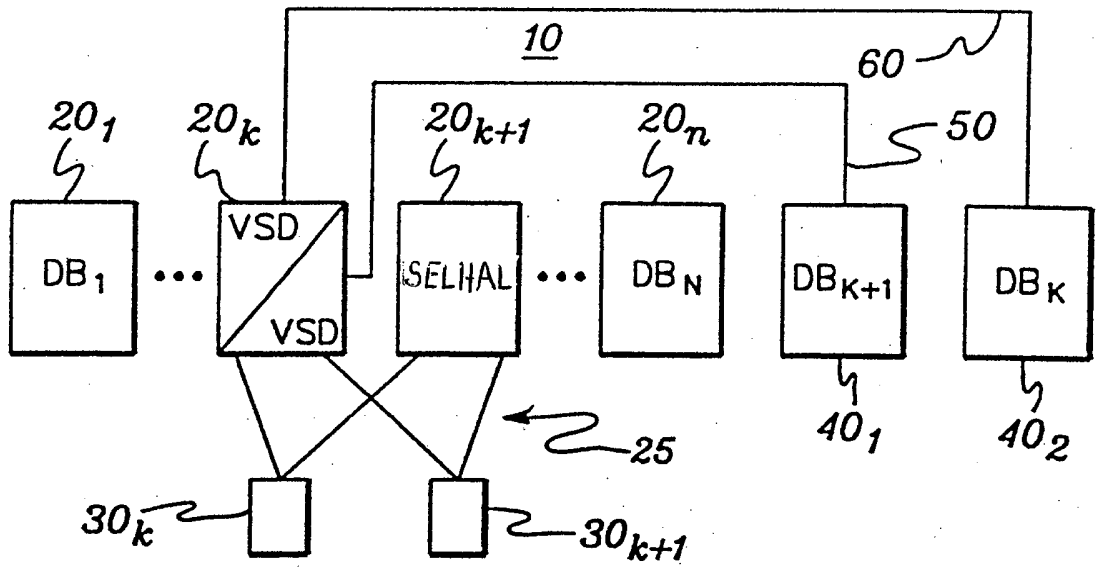


OBR. 2

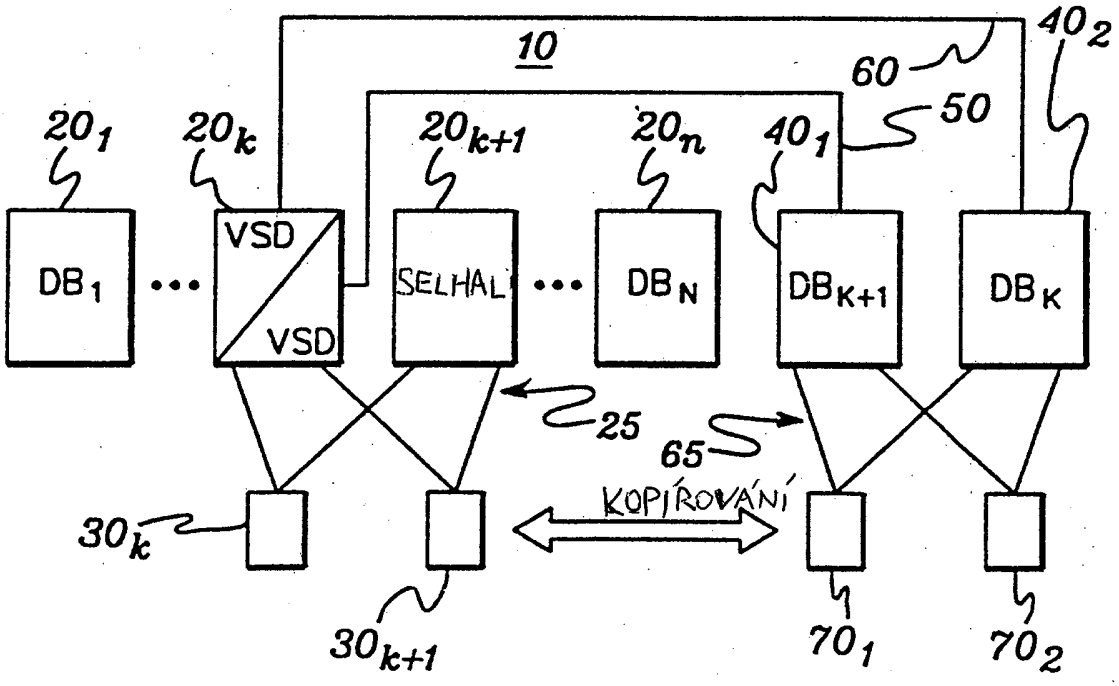


OBR . 3

PETR KALENSKÝ
ATTORNEY AT LAW



OBR. 4



OBR. 5

