



Evropská unie
Evropský sociální fond
Operační program Zaměstnanost



DRG
RESTART

INSTALAČNÍ PŘÍRUČKA A MANUÁL PRO UŽIVATELE NÁSTROJE CZ-DRG GROUPER

Metodický materiál projektu DRG Restart

Zpracoval autorský kolektiv pod vedením: T. Pavlík, M. Bartůňková, P. Klika, J. Linda, L. Dušek

Autoři: P. Klika, K. Kupčák, T. Pavlík, Z. Bortlíček, M. Uher, J. Klika

Verze CZ-DRG: 7.0

Verze dokumentu: 2

Datum: 31.8. 2024



Obsah

Účel nástroje.....	3
Předpoklady pro spuštění programu.....	3
Instalace	4
Nastavení pro spuštění v prostředí Linux	4
Spustitelné prostředí Javy	4
Nastavení dostupné paměti ve spouštěcím souboru	4
Instalace JDBC knihovny pro připojení do databáze	5
Vytvoření databázových tabulek	5
Konfigurace grouperu.....	6
Parametry procesu	6
Nastavení specifických parametrů pro JDBC.....	7
Spuštění grouperu	9
Parametry procesu	9
Příklady spuštění grouperu.....	10
Ukázka spuštění grouperu	11
Logování	11
Spuštění jako služba	11
Testování rychlosti grouperu – verze 1.0 z 05/2018	13
Scénář „velká pojišťovna“	14
Scénář „malá pojišťovna“	14
Scénář „velká nemocnice“	14
Scénář „malá nemocnice“	15
Testování rychlosti grouperu – verze 2.0 z 08/2019	15
Scénář „malý počet záznamů“	16
Scénář „větší počet záznamů“	16
Scénář „jediný záznam“	17
Scénář „spuštění jako služba“	17



Účel nástroje

Program slouží pro klasifikaci vstupní věty grouperu dle pravidel klasifikačního systému CZ-DRG. Vstupem grouperu je soubor odpovídající sestavené vstupní datové větě grouperu (popsané v samostatném dokumentu *DATOVÉ ROZHRANÍ NÁSTROJE CZ-DRG GROUPER*). Výstupem je soubor obsahující záznamy ve formátu výstupní datové věty grouperu (v témže datovém rozhraní), které byly klasifikovány do DRG skupiny systému CZ-DRG.

Nástroj implementuje klasifikační systém popsáný v samostatném dokumentu *DEFINIČNÍ MANUÁL KLASIFIKAČNÍHO SYSTÉMU CZ-DRG*, včetně výpočtu skóre závažnosti vedlejších diagnóz hospitalizačního případu a jednotlivých klasifikačních pravidel. Proces výpočtu DRG klasifikace je popsán v samostatném dokumentu *TECHNICKÁ DOKUMENTACE NÁSTROJE CZ-DRG GROUPER*.

Předpoklady pro spuštění programu

Nástroj grouper je CLI (řádková) aplikace implementovaná v jazyku Java. Grouper je tedy multiplatformní, přenositelný, snadno spravovatelný a bezpečný (za předpokladu pravidelné aktualizace systému, JRE a vlastního nástroje grouper).

Základní požadavky na systém jsou shrnuty v následující tabulce:

Tabulka 1 Požadavky na systém pro běh grouperu

Parametr / vlastnost	Hodnota	Poznámka
HW	Standardní HW umožňující běh aplikací v Javě, 2 GB volné RAM, 50 MB volného HDD pro instalaci + minimálně 2x velikost vstupních dat (v souboru CSV), alespoň 2 vlákna	Optimální parametry dle velikosti zpracovávaných dat
Operační systém	Standardní OS umožňující běh aplikací v Javě, Windows/UNIX/Linux	Ekvivalentní IR-DRG grouperu
Java	JRE (Java Runtime Environment) v aktuální a podporované verzi - Java 8 (poslední free verze je 8u202), Java 11 nebo OpenJDK 11	Ekvivalentní IR-DRG grouperu
Připojení k DB (volitelné)	Standardní JDBC knihovna pro použitou DB	Libovolná DB podporující JDBC rozhraní, např. MS SQL Server, Oracle, DB2, PostgreSQL, ...
Spuštění aplikace	Příkazová řádka OS (např. cmd, bash)	



Instalace

Program je distribuován ve formě balíku (zip), který obsahuje programové soubory (jar), spouštěcí skripty (sh/bat), konfigurační soubor a soubor s vytvořeným modelem pravidel klasifikačního systému CZ-DRG. Tento archiv je třeba rozbalit do samostatného adresáře v cestě, ke které má uživatel právo k přístupu na čtení i zápis. Součástí archivu je i testovací soubor se vzorovými sestavenými hospitalizačními případy pro otestování klasifikačního nástroje.

Název balíku ve formátu: **02__CZ-DRG_<verze_programu>_Grouper.zip**

např.: **02__CZ-DRG_v7_r0_Grouper.zip**

Seznam složek, které jsou součástí instalace:

- **db_install** – vzorové DDL skripty pro vytvoření databázových tabulek pro načítání a ukládání dat z/do databáze pro vybrané databázové servery
- **input** – složka pro ukládání vstupních CSV souborů dle DR grouperu, obsahuje testovací soubor *vstupni_veta.csv*
- **lib** – obsahuje doplňující programové balíky nástroje bez knihoven JDBC pro připojení do databáze
- **log** – prázdná složka, slouží pro uložení logového souboru log.txt (vytvoření této složky umožňuje zvýšit zabezpečení dat při rutinním provozu – hlavní složka grouperu nemusí mít právo zápisu a u jednotlivých podsložek se může nastavit přístup pro zápis jen pro nezbytné složky, tzn. Input, log a output)
- **model** – obsahuje aktuální klasifikační model s pravidly klasifikačního systému CZ-DRG
- **output** – prázdná složka, slouží k uložení výstupního souboru
- **settings** – obsahuje vzorový konfigurační soubor nástroje

Nastavení pro spuštění v prostředí Linux

Spouštěcím skriptům (*.sh) je nutné přidat oprávnění pro spuštění. Obecně lze provést např. pomocí následujícího příkazu spuštěného v hlavním adresáři grouperu (povolí spuštění nástroje všem uživatelům):

```
chmod a+x *.sh
```

Spustitelné prostředí Javy

Od verze CZ-DRG 5.0 je podporováno spustitelné prostředí Javy (JRE) ve verzi 11 a zpětně i verzi 8 a to jak na oficiální distribuci Oracle, tak implementacích zdarma (OpenJDK). Distribuce Oracle těchto verzí již nejsou podporovány za bezplatných podmínek pro jiné než osobní použití. Doporučujeme tedy použití některé z volně dostupných distribucí JRE 11 nebo 8.

Nastavení dostupné paměti ve spouštěcím souboru

Parametr -Xmx určuje maximální hodnotu alokované paměti pro data JVM (Java Virtual Machine). Hodnotu lze nastavit úpravou hlavního spouštěcího bat/sh souboru – *grouper.bat*, resp. *grouper.sh*. Například hodnota -Xmx2G udává maximum možné paměti na 2 GB. Minimálně je nutné nastavit



hodnotu 1500M = 1500 MB. Parametr je možné nastavit dle HW možností zařízení, na němž je program spuštěn, obvykle se nastavuje na polovinu až dvě třetiny dostupné operační paměti.

Parametr -Xms určuje iniciální hodnotu alokované paměti. Je doporučeno nastavit jeho hodnotu shodnou s parametrem -Xmx:

```
java -Xmx2G -Xms2G ...
```

Současně s tímto parametrem je nutné adekvátně upravit i maximální počet záznamů současně zpracovávaných v paměti (viz část [KONFIGURACE GROUPERU](#)). V případě některých databází je také třeba omezit počet záznamů, které se v jednom kroku načítají z databáze – viz část [NASTAVENÍ SPECIFICKÝCH PARAMETRŮ PRO JDBC](#). V případě, kdy tyto parametry nejsou nastaveny odpovídajícími hodnotami, může dojít k pádu aplikace v důsledku nedostatku paměti. V takovém případě je nutné buď navýšit velikost dostupné paměti pro aplikaci nebo omezit počet najednou zpracovávaných (resp. načítaných) záznamů. Ukázky spuštění grouperu s různými hodnotami tohoto parametru jsou v kapitole [TESTOVÁNÍ RYCHLOSTI GROUPERU](#).

Instalace JDBC knihovny pro připojení do databáze

V případě čtení resp. zápisu dat z/do databáze je třeba instalovat do operačního systému příslušný JDBC ovladač odpovídající použitému typu databáze. V případě požadavku na využití některé databáze použijte ovladač dostupný na stránkách výrobce nebo dodavatele příslušné databáze. Můžete tento balík nahrát do standardní složky **lib** anebo do jiné cesty, která bude při spuštění nástroje dostupná. Cestu k této knihovně je potřeba doplnit do spouštěcího souboru sh/bat do parametru -cp (class search path):

```
java ... -cp lib/* ...
```

Kompletní příkaz pro spuštění programu uvedený v souboru grouper.bat a grouper.sh potom může vypadat např. takto:

```
java -Xmx2000M -Xms2000M -cp lib/*;CZ-DRG-Grouper-7.0.jar  
cz.drg.clasificador.Main
```

Vytvoření databázových tabulek

Pro využití možnosti číst vstupní data z databáze a zapisovat tam výsledek je třeba v databázi vytvořit tabulky nebo pohledy, které odpovídají dokumentu **DATOVÉ ROZHRANÍ NÁSTROJE CZ-DRG GROUPER**.

Vytvořené tabulky musejí mít shodné názvy, pořadí sloupců a odpovídající datové typy (text, číslo, datum). Přesné datové typy jsou specifické pro konkrétní databázi. Vzorové příklady tabulek pro vybrané databáze jsou součástí distribuce grouperu ve složce **db_install**. V případě využití jiné databáze může být nutné provést drobné úpravy datových typů, případně další syntaxe v definici tabulek.

Názvy tabulek a jejich umístění je možno konfigurovat v nastavení grouperu (viz následující část [KONFIGURACE GROUPERU](#)).



Konfigurace grouperu

Proces grouperu může být parametrizován pomocí řady konfiguračních parametrů, které jsou uloženy v souboru **settings/programSettings.xml**. Jde o XML dokument, který je členěn do několika částí a popisuje základní chování nástroje, jeho vstupy a výstupy. Hodnoty parametrů se zapisují mezi zahajovací a ukončovací „tag“ příslušného elementu dle XML syntaxe:

```
<nazevElementu>konfiguracni_hodnota</nazevElementu>
```

Tedy např.:

```
<defaultInput>input/vstupni_veta.csv</defaultInput>
```

Parametry procesu

Grouper je možno konfigurovat pomocí parametrů uvedených v následujících třech tabulkách. První tabulka obsahuje parametry, které ovlivňují základní nastavení programu a možnosti načítání a ukládání dat ze/do souboru. V konfiguračním souboru se zadávají přímo pod kořenový element `<program>`.

Tabulka 2 Seznam základních konfiguračních parametrů nástroje grouper

Název parametru	Popis
defaultInput	Definuje cestu k vstupnímu souboru pro nástroj (v případě, že nebyl zadán v argumentech programu). Používá se pouze, pokud se data pro grouper načítají ze souboru.
defaultPmmlInput	Definuje cestu k souboru s definicí klasifikačního modelu grouperu. Tuto cestu není možné zadat jiným způsobem než v tomto souboru.
defaultOutput	Definuje defaultní cestu k výstupnímu souboru pro nástroj (v případě, že nebyl zadán v argumentech programu). Používá se pouze, pokud se data z grouperu ukládají do souboru.
defaultCharset	Definuje znakovou sadu používanou jak pro čtení, tak pro zápis dat. Dle definice DR musí být nastaven na hodnotu ISO-8859-2.
defaultDelimiter	Definuje oddělovač jednotlivých buněk dat při čtení a zápisu z/do souboru. Dle definice DR musí být nastaven na hodnotu \t.
evaluationBatchSize	Počet najednou zpracovávaných záznamů vstupní datové věty grouperu. Velikost tohoto parametru má vliv na rychlost a na využití operační paměti. Počet zpracovávaných záznamů v jednom bloku musí korespondovat s velikostí dostupné paměti pro grouper (viz část NASTAVENÍ DOSTUPNÉ PAMĚTI VE SPOUŠTĚCÍM SOUBORU).

Druhá tabulka obsahuje parametry, které určují vlastnosti pro připojení do databáze. Zadávají se do bloků uzavřených v hierarchii elementů `<IOClasses>` a `<IOClass>`, ve kterých je umístěn element `<forType>` s hodnotou `database`. Lze nastavit samostatně vlastnosti připojení do databáze pro čtení a pro zápis grouperu. K tomu slouží element `<forArg>`, který může nabývat hodnot `input` nebo `output`.



Tabulka 3 Parametry pro připojení grouperu k databázi

Název parametru	Popis
databaseConnection	JDBC url pro připojení do vybrané databáze (schématu) – viz část NASTAVENÍ SPECIFICKÝCH PARAMETRŮ PRO JDBC – pro čtení vstupních dat z DB
databaseTargetTable	Název databázové tabulky pro načtení nebo uložení dat do/z programu (dle hodnoty elementu <i>forArg</i>)
databaseUser	Uživatelské jméno pro připojení do database
databasePassword	Heslo pro připojení do database

Poslední tabulka definuje jeden parametr, který umožňuje zadání Java třídy pro připojení k databázi. Umožňuje spustit grouper s libovolným JDBC ovladačem, který bude doplněn do adresáře **lib/** grouperu. Zadává se dovnitř elementu `<databaseDrivers>` pod kořenovým elementem `<program>`.

Tabulka 4 Seznam dostupných ovladačů pro připojení grouperu k databázi

Název parametru	Popis
driverClass	Název třídy s ovladačem pro danou databázi, například pro připojení do databáze PostgreSQL je to <i>org.postgresql.Driver</i> nebo <i>com.microsoft.sqlserver.jdbc.SQLServerDriver</i> pro připojení do databáze MS SQL – pro čtení vstupních dat z DB

Nastavení specifických parametrů pro JDBC

Základní nastavení připojení do databáze se provádí pomocí dvou konfiguračních parametrů - **databaseConnection** a **driverClass**. Parametr **databaseConnection** obsahuje zejména specifikaci typu databáze, cestu k serveru a číslo portu, případně název databáze. V závislosti na typu připojované databáze může být nutné doplnit specifické parametry – například pro definici schématu (*searchpath*, *currentSchema*) nebo počtu přenášených záznamů (např. *defaultRowFetchSize*). Parametry jsou obvykle oddělovány středníkem (;) nebo znakem ampersand (&) a mezi názvem a hodnotou parametru je znak = (rovnítko).

Příklad vyplnění parametru **databaseConnection** pro napojení do databáze PostgreSQL:

```
jdbc:postgresql://localhost:5432/db_name?currentSchema=schema_name&  
defaultRowFetchSize=1000
```

Např. v tomto případě byl omezen počet záznamů načítaných v jednom bloku parametrem *defaultRowFetchSize*, aby nedošlo k vyčerpání dostupné paměti nástroje grouper.

Další ukázka zobrazuje vzorové nastavení parametru pro připojení do databáze MS SQL Server:

```
jdbc:sqlserver://localhost:1433;databaseName=db_name;  
authentication=SqlPassword;trustServerCertificate=true
```



Evropská unie
Evropský sociální fond
Operační program Zaměstnanost



Přesný popis jednotlivých parametrů pro připojení ke konkrétní databázi najdete v dokumentaci každé konkrétní databáze.



Spuštění grouperu

Grouper se spouští z příkazové řádky s využitím spustitelného prostředí jazyka Java. V kořenové složce nástroje grouper jsou za tímto účelem vytvořeny spouštěcí skripty, které umožňují snadné spuštění nástroje pro vybrané varianty vzorové konfigurace (určené pomocí konfiguračních parametrů –input a –output).

- Linux

`./grouper.sh`

- Windows

`grouper.bat`

Parametry procesu

Program může při spuštění na vstup dostat několik parametrů. Zadávají se v následujícím tvaru za název spouštěcího skriptu:

`[-arg [paramArg1 [paramArg2]] [-arg [paramArg1 [paramArg2]]]]`

Výraz uzavřený v hranatých závorkách [...] značí nepovinné vyplnění. Jako argumenty mohou být zadány následující hodnoty:

Tabulka 5 Parametry zadávané při spuštění grouperu

arg	paramArg1	paramArg2	Popis
input	csv	cesta_k_souboru.csv	Určuje načítání dat ze souboru a případně zadává cestu k němu. Pokud zde cesta není specifikována, načítá se z místa určeného v konfiguraci grouperu.
input	database	-	Určuje načítání dat z databáze. Specifikace připojení je určena v konfiguraci grouperu.
output	csv	cesta_k_souboru.csv	Určuje ukládání dat do souboru a případně zadává cestu k němu. Pokud zde cesta není specifikována, načítá se z místa určeného v konfiguraci grouperu.
output	database	-	Určuje ukládání dat do databáze. Specifikace připojení je určena v konfiguraci grouperu.
configFile	cesta_k_souboru.xml	-	Umožňuje zadat alternativní konfigurační soubor, např. pro jednotlivé uživatele. Pokud parametr není uveden, konfigurace se načítá ze standardního konfiguračního souboru settings/programSettings.xml.



service	číslo portu	-	Nástroj je také možné spustit jako službu, která poslouchá na konkrétním portu a průběžně zpracovává jednotlivé nemocniční případy.
help	-	-	Vypisuje základní možnosti parametrů při spuštění grouperu.

Pokud nejsou žádné argumenty zadány, použije se čtení a ukládání dat ze/do souborů určených v konfiguraci grouperu.

Příklady spuštění grouperu

Následující ukázky ukazují různé možnosti spuštění grouperu (zde pod OS Windows).

- Načtení i uložení dat z/do databáze určené konfigurací grouperu

grouper.bat -input database -output database

nebo

grouper_db.bat

- Načtení i uložení dat z/do souborů, určených v konfiguraci grouperu

grouper.bat -input csv -output csv

nebo

grouper_file.bat

- Načtení dat ze souboru zadaného na příkazové řádce, uložení do databáze určené v konfiguraci grouperu, načtení konfigurace z alternativního konfiguračního souboru

grouper.bat -input csv input/test.csv -output database

-configFile settings_novak/programSettings.xml



Ukázka spuštění grouperu

Ukázka spuštění nástroje v příkazové řádce je uvedena na následujícím obrázku.

```
C:\Windows\System32\cmd.exe
Microsoft Windows [Version 10.0.17763.557]
(c) 2018 Microsoft Corporation. Všechna práva vyhrazena.

c:\Users\klikaj\Nextcloud\sdilena-oddeleni-datove-centrum\UZIS\GROUPER\grouper\czdrg\project\Sestavene\CZ-DRG-Grouper-2.0
>@grouper.bat -input csv -output csv
CZ-DRG Grouper 2.0
(C) 2019, UZIS CR
Batch of size 206
Total of 206 entries.

Input entries..... 206
Output entries..... 206
Valid entries..... 194
Error entries..... 12
No clasification..... 0
Evaluated after 13037ms

Klasifikace vseh zaznamu byla uspesne ukoncena, zpravu a zaznamy o pripadnych chybach najdete v souboru log.txt.
Press any key to continue . . .

c:\Users\klikaj\Nextcloud\sdilena-oddeleni-datove-centrum\UZIS\GROUPER\grouper\czdrg\project\Sestavene\CZ-DRG-Grouper-2.0
>
```

Obrázek 1 Ukázka spuštění CZ-DRG Grouperu nad testovacími daty v příkazové řádce Windows

Logování

Záznamy ze spuštění grouperu a případné chybové hlášky jsou zaznamenány do souboru **log.txt**, umístěného v kořenovém adresáři grouperu:

```
18.06.2019
12:07
Batch of size 206
Total of 206 entries.

Input entries..... 206
Output entries..... 206
Valid entries..... 194
Error entries..... 12
No clasification..... 0
Evaluated after 13037ms

18.06.2019
12:08
```

Spuštění jako služba

Nástroj je také možné spustit jako službu, která umožní rychlé spouštění klasifikace jednotlivých záznamů, vložených do databáze. Spuštěná služba naslouchá na zadaném portu a očekává na něm HTTP POST request s identifikátorem hospitalizačního případu (ID_PRIPADU) nebo identifikátorem spuštění grouperu (RUN_ID) v jeho těle. Pro připojení k databázi se vstupními a výstupními záznamy je



používané nastavení sekce pro čtení a zápis z/do databáze v konfiguračním souboru. Čtení a ukládání záznamů do souboru není v tomto režimu podporováno.

Službu lze spustit například následujícím příkazem:

```
„java -cp lib/*;CZ-DRG-Grouper-7.0.jar cz.drg.clasificator.Main –service 6666“
```

Ten spustí grouper na daném počítači na portu 6666.

Pro vytvoření HTTP POST requestu lze použít příkazovou řádku (Linux nebo Windows) s následujícím příkazem:

curl -i -X POST -d 1 http://localhost:6666/

Zde je “1” id hospitalizačního případu, které bude vyhledáno v tabulce vstupní datové věty grouperu (dle konfigurace nástroje) v parametru ID_PRIPADU a u kterého se má provést klasifikace. Nově tento parametr může identifikovat konkrétní spuštění nástrojů pre-grouper a grouper, kdy budou zpracovány pouze záznamy ze vstupní datové věty, které mají uvedenou hodnotu ve sloupci RUN_ID. Zadané záznamy musí již v tabulce existovat (mohou v ní být i další záznamy, které ale nebudou klasifikovány). Grouper hledá zadaný identifikátor v obou sloupcích tabulky, ID_PRIPADU i RUN_ID a provede klasifikaci všech záznamů, u kterých najde shodu.

Text “http://localhost:6666/” je URL adresa, kam se má požadavek odeslat, tedy adresa, kde běží spuštěná aplikace grouperu. Odpovědí na tento požadavek je buď stavový HTTP kód 200, který reprezentuje průběh bez chyb, anebo kód 404, reprezentující chybu v průběhu klasifikace. V případě úspěšné klasifikace je záznam uložen do tabulky výstupní datové věty grouperu dle konfigurace nástroje.

Více informací o protokolu HTTP a jeho metodách lze naléznout na následující adrese:

https://www.w3schools.com/tags/ref_httpmethods.asp



Testování rychlosti grouperu – verze 1.0 z 05/2018

V rámci vývoje grouperu byla provedena řada testovacích běhů nad testovacími daty Národního registru hrazených zdravotních služeb (NRHZS). Byla vybrána data dvou pojišťoven a dvou nemocnic, vždy jedna velká a jedna malá. Byly vybírány doklady za jeden kalendářní rok (podle roku dávky). Vstupní věta grouperu byla sestavena pomocí nástroje CZ-DRG Pre-grouper.

Testováno bylo na virtuálním serveru s procesory Intel(R) Xeon(R) CPU E5-2620 v4 @ 2.10GHz, který měl k dispozici až 24 GB RAM a 8 jader. Operační systém je Ubuntu Linux 16.04 a virtualizační platforma KVM. Testovací konfigurace byla nastavována na různé hodnoty dostupných prostředků, aby byla ověřena použitelnost nástroje pro různé účely. Na serveru běžela v danou chvíli pouze testovaná úloha, ale výsledky mohou být ovlivněny nastavením virtualizace nebo systémovými procesy operačních systémů. Data byla vždy načítána ze souboru.

Následující sekce zahrnují souhrn údajů o vybraných testovacích datech a výsledky z testování nad vybranými konfiguracemi. Na základě těchto výsledků je možné nastavit optimální konfiguraci pro jednotlivé uživatele tohoto nástroje a odhadnout časovou náročnost klasifikace hospitalizačních případů.



Scénář „velká pojišťovna“

Tabulka 6 Zpracovávaná data scénáře "velká pojišťovna"

Ukazatel	Hodnota
Počet sestavených HP	1 197 612

Tabulka 7 Výsledky testování scénáře "velká pojišťovna"

Počet jader	Dostupná paměť	Počet v jednom bloku zpracovávaných záznamů	Čas trvání (min)
8	16 GB	250 000	29,5
8	4 GB	75 000	29,9
8	2 GB	10 000	27,3
4	2 GB	10 000	43,3

Scénář „malá pojišťovna“

Tabulka 8 Zpracovávaná data scénáře "malá pojišťovna"

Ukazatel	Hodnota
Počet sestavených HP	81 080

Tabulka 9 Výsledky testování scénáře "malá pojišťovna"

Počet jader	Dostupná paměť	Počet v jednom bloku zpracovávaných záznamů	Čas trvání (s)
8	6 GB	100 000	220
8	2 GB	10 000	255
8	1 750 MB	1 000	236
4	1 750 MB	1 000	442

Scénář „velká nemocnice“

Tabulka 10 Zpracovávaná data scénáře "velká nemocnice"

Ukazatel	Hodnota
Počet sestavených HP	70 693



Tabulka 11 Výsledky testování scénáře "velká nemocnice"

Počet jader	Dostupná paměť	Počet v jednom bloku zpracovávaných záznamů	Čas trvání (s)
8	8 GB	100 000	190
8	2 GB	10 000	207
8	1 750 MB	1 000	198
4	1 750 MB	1 000	224

Scénář „malá nemocnice“

Tabulka 12 Zpracovávaná data scénáře "malá nemocnice"

Ukazatel	Hodnota
Počet sestavených HP	6 343

Tabulka 13 Výsledky testování scénáře "malá nemocnice"

Počet jader	Dostupná paměť	Počet v jednom bloku zpracovávaných záznamů	Čas trvání (s)
8	2 GB	10 000	116
8	1,75 GB	1 000	122
8	1 500 MB	500	138
4	1 500 MB	500	168

Testování rychlosti grouperu – verze 2.0 z 08/2019

V rámci vývoje grouperu byly provedeny některé úpravy programu, především:

- na konec do výstupní tabulky byla doplněna nová položka DRG_KAT – DRG klasifikace kategorie
- možnost využití database Oracle a Informix
- možnost spuštění programu jako služby (výpočet jednoho hospitalizačního případu trvá obvykle jen zlomek vteřiny)
- zrychlení programu především v části úvodního načítání xml klasifikačního modelu

Test rychlosti klasifikace původní a nové verze programu byl proveden na PC Intel® Core™ i5-5300U CPU @ 2.30 GHz, 8 GB RAM, Windows 10 Pro, vstup i výstup byl z csv souboru. Výjimkou je poslední scénář „spuštění jako služba“, který klasifikuje jediný záznam v databázi pomocí grouperu, spuštěného jako služba.



Scénář „malý počet záznamů“

Tabulka 14 Zpracovávaná data scénáře "malý počet záznamů"

Ukazatel	Hodnota
Počet sestavených HP	206

Tabulka 15 Výsledky testování scénáře "malý počet záznamů – původní program"

Počet jader	Dostupná paměť	Počet v jednom bloku zpracovávaných záznamů	Čas trvání (s)
4	4 GB	1 000	138

Tabulka 16 Výsledky testování scénáře "malý počet záznamů – nový program"

Počet jader	Dostupná paměť	Počet v jednom bloku zpracovávaných záznamů	Čas trvání (s)
4	4 GB	1 000	11

Výsledky dokládají řádové zrychlení výpočtu klasifikace u malého počtu záznamů (dáno zkrácení doby zahájení klasifikace).

Scénář „větší počet záznamů“

Tabulka 17 Zpracovávaná data scénáře "větší počet záznamů"

Ukazatel	Hodnota
Počet sestavených HP	20000

Tabulka 18 Výsledky testování scénáře "větší počet záznamů – původní program"

Počet jader	Dostupná paměť	Počet v jednom bloku zpracovávaných záznamů	Čas trvání (s)
4	4 GB	1 000	201

Tabulka 19 Výsledky testování scénáře "větší počet záznamů – nový program"

Počet jader	Dostupná paměť	Počet v jednom bloku zpracovávaných záznamů	Čas trvání (s)



4	4 GB	1 000	93
---	------	-------	----

Výsledky dokládají významné zrychlení u většího počtu záznamů, jehož vliv se však s rostoucím počtem klasifikovaných záznamů snižuje.

Scénář „jediný záznam“

Tabulka 20 Zpracovávaná data scénáře "jediný záznam"

Ukazatel	Hodnota
Počet sestavených HP	1

Tabulka 21 Výsledky testování scénáře "jediný záznam – původní program"

Počet jader	Dostupná paměť	Počet v jednom bloku zpracovávaných záznamů	Čas trvání (s)
4	4 GB	1	130

Tabulka 22 Výsledky testování scénáře "jediný záznam – nový program"

Počet jader	Dostupná paměť	Počet v jednom bloku zpracovávaných záznamů	Čas trvání (s)
4	4 GB	1	10

Mezi klasifikací jednoho záznamu a menšího počtu záznamů není u nového grouperu významný rozdíl. Většinu doby trvá inicializace program před zahájením vlastní klasifikace.

Scénář „spuštění jako služba“

Tabulka 23 Zpracovávaná data scénáře "spuštění jako služba"

Ukazatel	Hodnota
Počet sestavených HP	1

Tabulka 24 Výsledky testování scénáře "jediný záznam – nový program"

Počet jader	Dostupná paměť	Počet v jednom bloku zpracovávaných záznamů	Čas trvání (s)



4	4 GB	1	0,1
---	------	---	------------

Čas pro provedení klasifikace jednoho záznamu je u varianty spuštění grouperu jako služba zanedbatelný, řádově setiny až desetiny sekund.