

A screenshot of a web browser window. The title bar at the top reads "Application for parallel processing". The browser's address bar is empty. The main content area of the browser displays a login page. At the top of the page, the text "Aplikace pro řízení paralelního zpracování dat" is centered. Below this, there are two input fields: the first is labeled "Login" and the second is labeled "Heslo". A button labeled "Přihlásit se" is positioned below the "Heslo" field. The browser's status bar at the bottom is empty.

Aktuálně spuštěné úlohy

Application for parallel processing

Aktuálně spuštěné úlohy

Přehled výpočetních uzlů

Historie dokončených úloh

Možnosti uložených úloh

Vytvořit novou úlohu

Seřadit podle:

Data spuštění

Doby běhu

Počet vytížených serverů

Filtrovat

Název úlohy:

Čas spuštění:

od:

/

/

do:

/

/

Doba běhu:

od:

dnů

0

hod

do:

dnů

0

hod

Status:

☐ běží

☐ běží/chyba

☐ chyba

☐ čeká

☐ pozastaveno

Vytíženo serverů:

0

-

100

Zrušit všechny filtry

ÚLOHA 1

Název úlohy: Úloha 1

Čas spuštění: 10.06.2019 15:33

Úspěšně dokončené kroky:

6/10

Poslední úspěšně dokončený krok: vertikalizace

Status: chyba

Čas běhu: 2 dny 03 hod 05 min

Běží na: 5/10 serverů

Detail úlohy

ÚLOHA 2

Název úlohy: Úloha 2

Čas spuštění: 29.06.2019 09:01

Úspěšně dokončené kroky:

7/10

Poslední úspěšně dokončený krok: deduplikace

Status: běží

Čas běhu: 6 dní 05 hod 50 min

Běží na: 12/12 serverů

Detail úlohy

ÚLOHA 3

Název úlohy: Úloha 3

Čas spuštění: 10.07.2019 19:41

Úspěšně dokončené kroky:

3/10

Poslední úspěšně dokončený krok: indexace

Status: pozastaveno

Čas běhu: 11 dní 05 hod 50 min

Běží na: 12/12 serverů

Detail úlohy

<

1

2

3

4

...

>

seřadit podle:

- data spuštění
- počet serverů, na kterých úloha aktuálně běží
- doby běhu

....

vyhledávání s napovídáním názvů úloh

Detail úlohy

Application for parallel processing

Aktuálně spuštěné úlohy

Přehled výpočetních uzlů

Historie dokončených úloh

Možnosti uložených úloh

Vytvořit novou úlohu

Zpět

Detail běžící úlohy

Název úlohy: Úloha 1

Čas spuštění: 10.6.2019 15:33

Úspěšně dokončené kroky: 6/10

Poslední úspěšně dokončený krok: vertiklizace

Status: chyba

Uplynulý čas: 2 dny 3 hod 5 min

Běží na: 5/10 serverů

Aktuálně rozpracované kroky

Název kroku	Krok	Dokončeno [%]	Zpracováno/Celkem souborů	Status
vertikalizace	2/10	98%	950/1000	chyba
indexace	3/10	60%	598/1000	běží
deduplikace	4/10	10%(1/10)	98/100	běží
filtrace	5/10	8%	75/1000	běží

Možnosti úlohy

Zobrazit konfiguraci

Pozastavit úlohu

Spustit úlohu

Stav úlohy na jednotlivých výpočetních uzlech

Filtrovat

Status:

☐ běží

☐ běží/chyba

☐ chyba

☐ čeká

☐ pozastaveno

☐ dokončeno

Krok:

Zvolte krok

Doba běhu:

od:

0

 dnů

0

 h

do:

0

 dnů

0

 h

Název výpočetního uzlu	Krok	Název kroku	Doba běhu	Doba čekání	Dokončeno [%]	Zpracováno/Celkem souborů	Status	Možnosti
Knots 1	2/10	vertikalizace	1 den 01 h 03 min	1 den 02 hod 02 min	55%	50/100	chyba	<div>Detail chyby</div> <div>Pozastavit</div> <div>Opakovat</div>
Knots 2	3/10	indexace	2 dny 03 h 05 min	00 h 00 min	43%	43/100	běží	<div>Pozastavit</div>
Knots 3	4/10	indexace	2 dny 03 h 05 min	00 h 00 min	4%	9/100	běží/chyba	<div>Detail chyby</div> <div>Pozastavit</div> <div>Opakovat</div>
Knots 4	4/10	deduplikace	2 dny 01 h 05 min	02 h 00 min	0%	-	běží	<div>Pozastavit</div>
Knots 5	4/10	indexace	1 den 23 h 05 min	04 h 00 min	34%	36/100	chyba	<div>Detail chyby</div> <div>Pozastavit</div> <div>Opakovat</div>
Knots 6	4/10	indexace	12 h 01 min	1 den 15 h 04 min	95%	10/100	pozastaveno	<div>Spustit</div>
Knots 7	4/10	deduplikace	2 dny 03 h 00 min	00 h 05 min	0%	-	čeká	<div>Pozastavit</div>
Knots 8	4/10	deduplikace	1 den 01 h 05 min	1 den 02 h 00 min	0%	-	pozastaveno	<div>Spustit</div>
Knots 9	4/10	deduplikace	2 dny 03 h 05 min	00 h 00 min	0%	-	běží	<div>Pozastavit</div>
Knots 10	5/10	filtrace	02 h 07 min	00 h 00 min	75%	75/100	běží	<div>Pozastavit</div>
Knots 11	5/10	filtrace	03 h 06 min	00 h 00 min	81%	81/100	běží	<div>Pozastavit</div>

První varianta - po kliknutí na zobrazit konfiguraci se zobrazí konfigurace dané úlohy na samostatné straně

Dokončeno / odpovídá počtu zpracovaných souborů v procentech. V případě, že bude souborem jen jedna složka/soubor, bude počet procent boolovský, tj. 0|100%

doba běhu = jak dlouho daný krok běžel celkem na daném výpočetním uzlu.
Doba čekání = jak dlouho čekal krok na přidělení prostředků, byl pozastaven nebo byl v chybovém stavu = jak dlouho od spuštění úlohy neběžel

Běžící kroky

- Dokončeno - odpovídá počtu celkově zpracovaných souborů na všech serverech NEBO v případě, že krok nepracuje se složkou souborů ale pouze s jedním souborem/složkou, odpovídá procento dokončení: (počtu výpočetních uzlů, na kterých je krok dokončen/celkový počet výpočetních uzlů)*100.
- Sumarizovaný status - odpovídá sumarizovanému statusu všech serverů, na kterých úloha aktuálně běží

Řazení

u obou tabulek bude možnost řadit podle jednotlivých položek v tabulce

- v tabulce výpočetních uzlů bude možné vybrat jednotlivé statusy

Detail chyby

Application for parallel processing

Aktuálně spuštěné úlohy

Přehled výpočetních uzlů

Historie dokončených úloh

Možnosti uložených úloh

Vytvořit novou úlohu

Zpět

Detail chyby

Výpočetní uzel: Knotis 1

Název úlohy: Úloha 1

Krok: 1/10

Název kroku: vertikalizace

Dokončeno:

50%

Čas spuštění: 10.06.2019 15:33

Doba běhu: 1 den 01 h 03 min

Doba čekání: 04 h 08 min

Zpracováno souborů: 50/100

Opakovat krok

Vstup	Popis chyby
mni\data/file_1	timeout 8h
mni\data/file_2	<div>Zobrazit obsah logu</div>
mni\data/file_3	<div>Zobrazit obsah logu</div>
mni\data/file_5	<div><div>Zobrazit výstup na stderr</div></div>
mni\data/file_10	<div><div>Zobrazit výstup na stderr</div></div>
mni\data/file_45	<div><div>Zobrazit obsah logu</div><div>Zobrazit výstup na stderr</div></div>
mni\data/file_47	<div><div>Zobrazit obsah logu</div><div>Zobrazit výstup na stderr</div></div>

Tabulka je potřebná -> nutno zobrazit všechny soubory, ve kterých nastala chyba

Dokončeno / odpovídá počtu zpracovaných souborů v procentech. V případě, že bude souborem jen jedna složka/soubor, bude počet procent boolovský, tj. 0-100%

Popis chyby může mít 4 varianty:

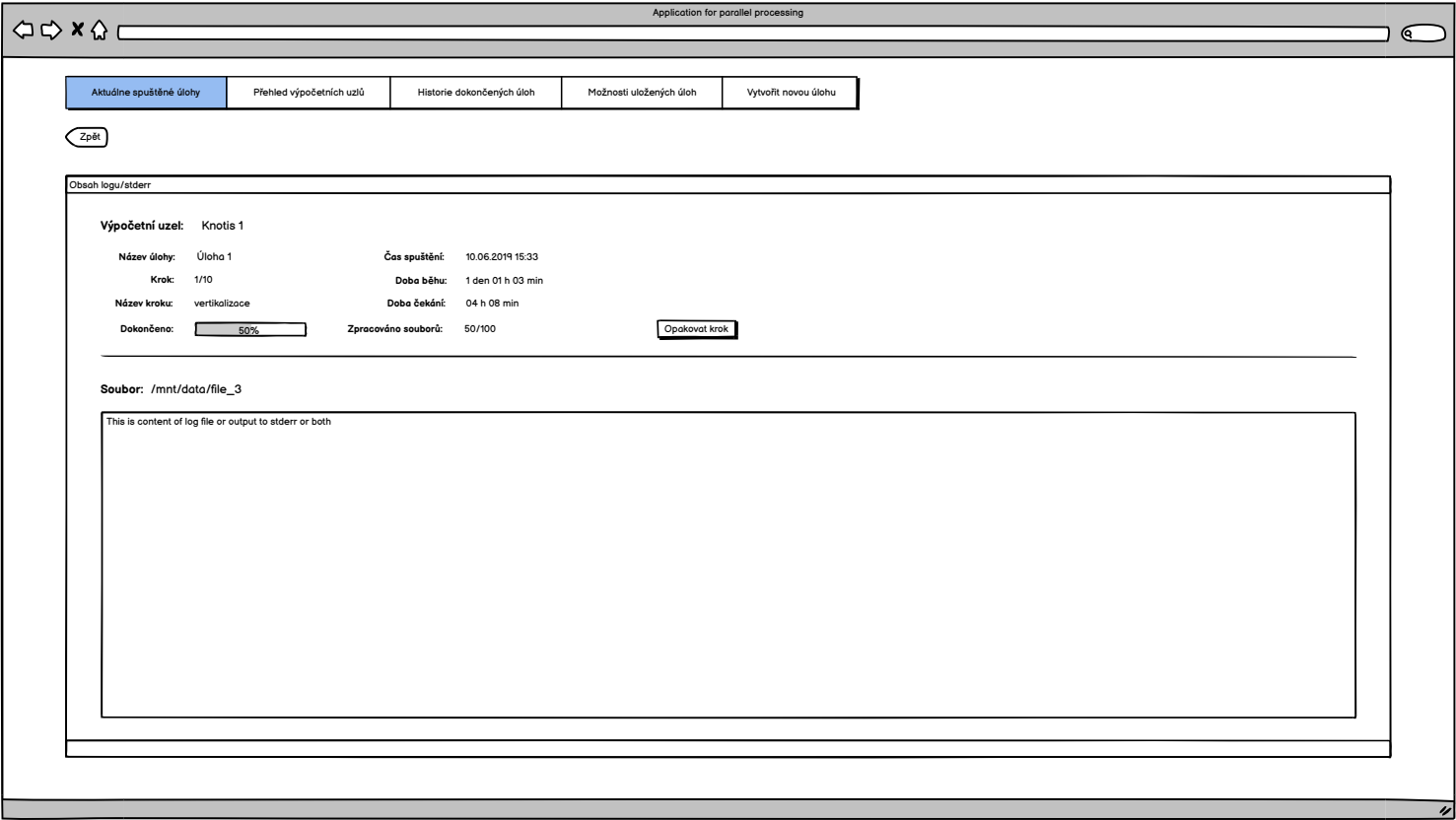
- 1) vlastní popis chyby
- 2) chyba vypsaná na stderr
- 3) výpis chyby z error log file
- 4) kombinace 2) a 3)

doba běhu = jak dlouho daný krok běžel celkem na daném výpočetním uzlu.
Doba čekání = jak dlouho čekal krok na přidělení prostředků, byl pozastaven nebo byl v chybovém stavu = jak dlouho od spuštění úlohy neběžel

opakovat krok - odstraní všechny chyby a spustí daný krok znovu - v případě, že je vstupem jedna složka/soubor, spustí celý krok znovu s danou složkou/souborem jinak jen s nezpracovanými soubory.

Tabulka zobrazuje různé možnosti zobrazení, nejedná se o reálná data - není možné reálně kombinovat obsah logu|stderr|obojí, jelikož toto je jednotné pro daný krok.

Zobrazení obsahu logu/stderr



Tabulka je potřebná -> nutno zobrazit všechny soubory, ve kterých nastala chyba

Dokončeno / odpovídá počtu zpracovaných souborů v procentech. V případě, že bude souborem jen jedna složka/soubor, bude počet procent boolovský, tj. 0-100%

Popis chyby může mít 4 varianty:

- 1) vlastní popis chyby
- 2) chyba vypsána na stderr
- 3) výpis chyby z error log file
- 4) kombinace 2) a 3)

dobu běhu = jak dlouho daný krok běžel celkem na daném výpočetním uzlu.
Doba čekání = jak dlouho čekal krok na přidělení prostředků, byl pozastaven nebo byl v chybovém stavu = jak dlouho od spuštění úlohy neběžel

opakovat krok - odstraní všechny chyby a spustí daný krok znovu - v případě, že je vstupem jedna složka/soubor, spustí celý krok znovu s danou složkou/souborem jinak jen s nezpracovanými soubory.

Konfigurace úlohy

Application for parallel processing

Aktuálně spuštěné úlohy

Přehled výpočetních uzlů

Historie dokončených úloh

Možnosti uložených úloh

Vytvořit novou úlohu

Zpět

Konfigurace úlohy

Název úlohy: Úloha 1

Čas spuštění: 10.06.2019 15:33

Úspěšně dokončené kroky: 6/10

Poslední úspěšně dokončený krok: vertikalizace

Status: chyba

Uplynulý čas: 2 dny 3 hod 5 min

Běží na: 5/10 serverů

1. krok

2. krok

3. krok

+ Přidat další krok

Název kroku: vertikalizace

Program: python3 /mnt/programs/vert.py

Typ vstupu: složka Prefix: --input Cesta: /mnt/data/in Regulární výraz: [0-9]+

Typ výstupu: složka Prefix: --output Cesta: /mnt/data/out/\$1.vert

Parametry: -t target out mnt/data/out/

Timeout: 8 hodin

Ověření:

☒ výstupní složka není prázdná

☒ stderr je prázdný

☒ velikost logu

☒ Pro každou vstupní složku existuje odpovídající neprázdný výstupní soubor nebo složka

☒ Při chybě vypsát obsah logu /mnt/logs/\$1.err

☒ Nutná synchronizace kroku

Dokončeno / odpovídá počtu zpracovaných souborů v procentech. V případě, že bude souborem jen jedna složka/soubor, bude počet procent boolovský, tj. 0-100%

doba běhu = jak dlouho daný krok běžel celkem na daném výpočetním uzlu.
Doba čekání = jak dlouho čekal krok na přidělení prostředků, byl pozastaven nebo byl v chybovém stavu = jak dlouho od spuštění úlohy neběžel

Přehled výpočetních uzlů

↩️↪️❌🏠

Application for parallel processing

🔍

Aktuálně spuštěné úlohy

Přehled výpočetních uzlů

Historie dokončených úloh

Možnosti uložených úloh

Vytvořit novou úlohu

Název uzlu	Stav uzlu	Běžící úloha	Datum a čas spuštění úlohy	Dokončeno kroků	Běžící krok	Dokončeno [%]	Počet zpracovaných souborů	Status úlohy	Využití za			
									hodinu	den	týden	měsíc
Knot 1	obsazen	úloha 1	19.06.2019 07:53	4/9	vertikalizace	10 %	10/100	beží	67 %			
Knot 2	obsazen	úloha 2	10.06.2019 15:33	7/12	deduplikace	46 %	46/100	beží/chyba	36 %			
Knot 3	obsazen	úloha 3	01.07.2019 05:01	6/11	indexace	56 %	56/100	beží/chyba	46 %			
Knot 4	obsazen	úloha 1	19.06.2019 07:53	6/9	indexace	10 %	10/100	beží/chyba	100 %			
Knot 5	obsazen	úloha 1	19.06.2019 07:53	7/9	deduplikace	67 %	67/100	běží	74 %			
Knot 6	obsazen	úloha 1	19.06.2019 07:53	4/9	vertikalizace	10 %	10/100	běží	76 %			
Knot 7	obsazen	úloha 3	01.07.2019 05:01	6/11	indexace	23 %	23/100	pozastaveno	57 %			
Knot 8	připraven	-	-	-	-	-	-	-	81 %			
Knot 9	připraven	-	-	-	-	-	-	-	12 %			
Knot 10	připraven	-	-	-	-	-	-	-	24 %			
Knot 11	odpojen	-	-	-	-	-	-	-	67 %			
Knot 12	odpojen	-	-	-	-	-	-	-	10 %			
Knot 13	připraven	-	-	-	-	-	-	-	33 %			
Knot 14	odpojen	-	-	-	-	-	-	-	61 %			
Knot 15	obsazen	úloha 1	19.06.2019 07:53	4/9	vertikalizace	10 %	10/100	beží	93 %			

Možné stavy uzlu a úlohy:

1. obsazen

a) běží

b) běží/chyba

c) čeká (nezobrazuje se)
2. volný

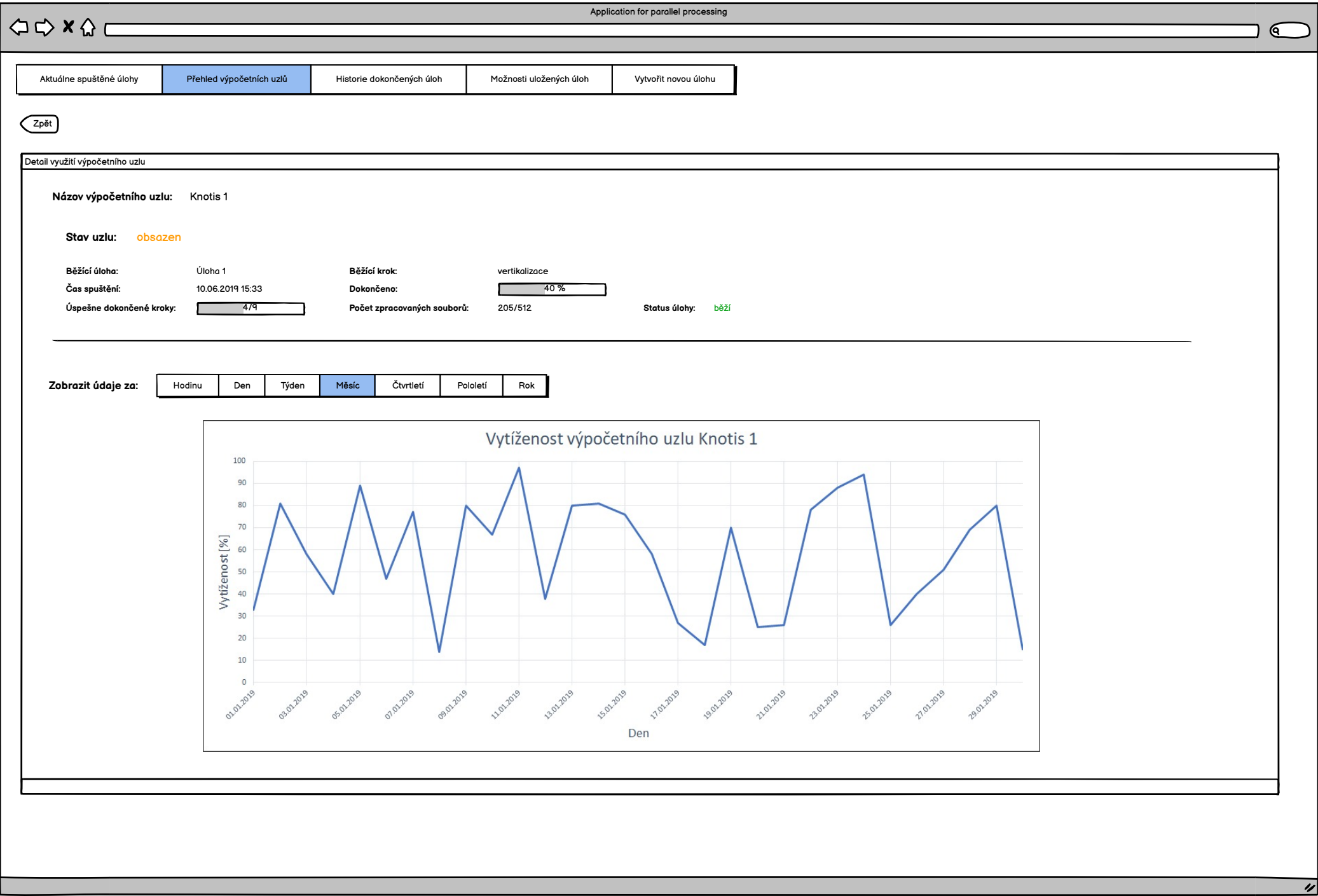
a) pozastaveno (nezobrazuje se)

b) chyba (nezobrazuje se)
3. odpojen

- řazení úloh do front na základě data spuštění úlohy -> při opětovném spuštění úlohy (po odstranění chyby) pozastavit aktuálně běžící úlohu (v případě, že je aktuálně běžící úloha spuštěna později jak úloha, u které byla chyba odstraněna) a spustit úlohu, u které byla chyba odstraněna nebo ji nechat doběhnout a zařadit ji do fronty na první místo (v případě, že je nejstarší ze všech čekajících úloh) ?

po kliknutí na progress bar "využití za" se zobrazí "Detail využití úlohy".

Detail využití úloh



Historie dokončených úloh

Application for parallel processing

↶ ↷ ✕ 🏠

Q

Aktuálně spuštěné úlohy

Přehled výpočetních uzlů

Historie dokončených úloh

Možnosti uložených úloh

Vytvořit novou úlohu

Filtry

Název úlohy:

Čas spuštění:

od:

/

/

📅

do:

/

/

📅

Celková doba běhu: od: dnů

0

▼

 hod

do: dnů

0

▼

 hod

Vytiženo serverů:

0

-

100

Zrušit všechny filtry

Seřadit podle:

Data spuštění

Doby běhu

Počet vytižených serverů

ÚLOHA 1

Název úlohy: Úloha 1

Čas spuštění:10.06.2019 15:33

Čas dokončení:30.06.2019 01:47

Spuštěno na:10 uzlech

Celková doba trvání:5 dní 02 hod 42 min

Celková doba výpočtu:38 dní 20 hod 05 min

Celková doba čekání:8 dní 01 hod 14 min

Podrobnosti běhu úlohy

ÚLOHA 2

Název úlohy: Úloha 2

Čas spuštění:18.07.2019 04:17

Čas dokončení:22.07.2019 13:21

Spuštěno na:40 uzlech

Celková doba trvání:5 dní 02 hod 42 min

Celková doba výpočtu:28 dní 00 hod 55 min

Celková doba čekání:05 hod 03 min

Podrobnosti běhu úlohy

ÚLOHA 3

Název úlohy: Úloha 3

Čas spuštění:08.05.2019 07:11

Čas dokončení:02.07.2019 17:00

Spuštěno na:45 uzlech

Celková doba trvání:20 dní 12 hod 12 min

Celková doba výpočtu:16 dní 10 hod 15 min

Celková doba čekání:0 dní 02 hod 43 min

Podrobnosti běhu úlohy

<

1

2

3

4

...

>

jak se má chovat využití "CPU" pro více jader ?

Podrobnosti běhu úlohy

Application for parallel processing

Aktuálně spuštěné úlohy

Přehled výpočetních uzlů

Historie dokončených úloh

Možnosti uložených úloh

Vytvořit novou úlohu

Zpět

Detail dokončené úlohy

Název úlohy: Úloha 1

Čas spuštění: 10.6.2019 15:33

Čas dokončení: 30.6.2019 01:47

Spuštěno na: 10 uzlech

Celková doba trvání: 5 dní 02 hod 42 min

Celková doba výpočtu: 38 dní 20 hod 05 min

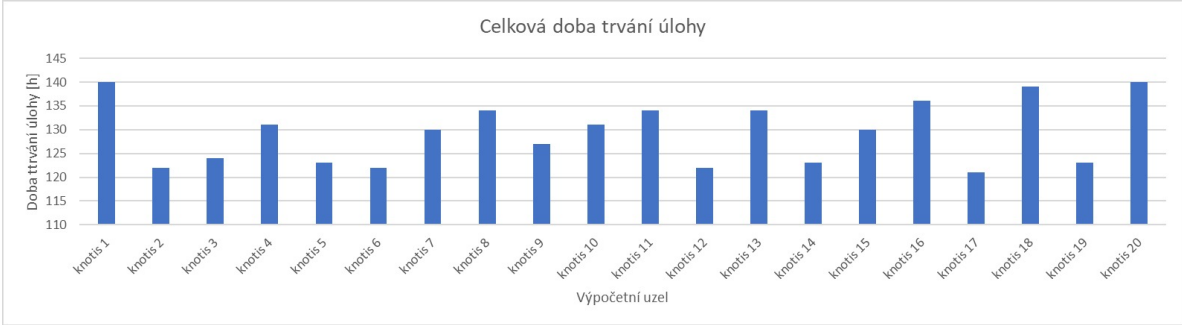
Celková doba čekání: 8 dní 01 hod 14 min

Možnosti

Zobrazit konfiguraci

Odstranit

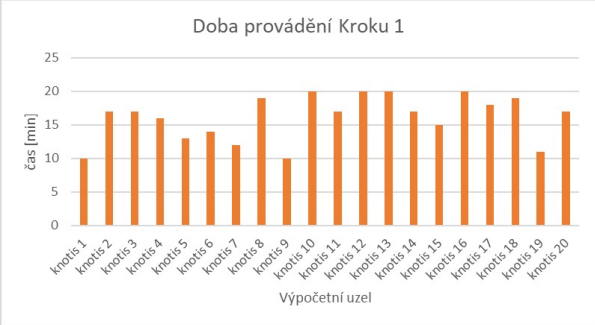
Celková doba trvání úlohy



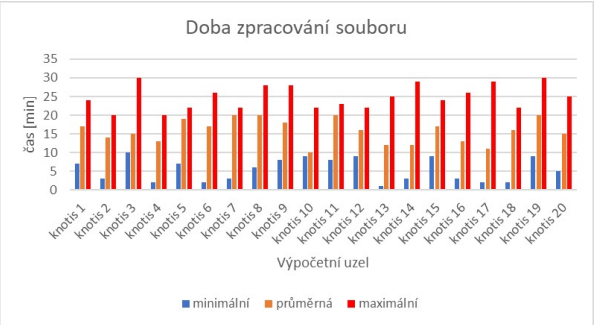
Statistiky provádění úlohy

KROK 1

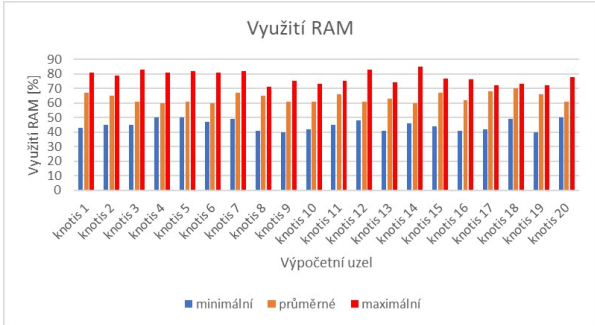
Doba provádění Kroku 1



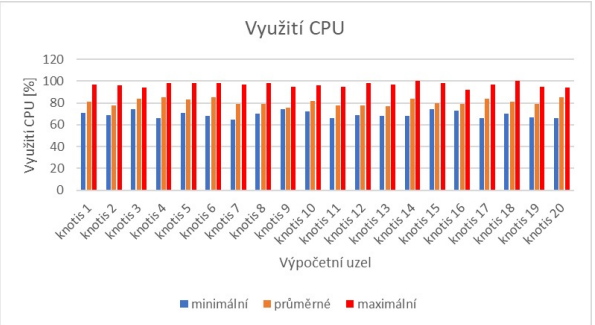
Doba zpracování souboru



Využití RAM



Využití CPU



KROK 2

KROK 3

KROK 4

jak se má chovat využití "CPU" pro více jader ?
Bude od 0 do N%, kde N je 100*počet jader

celková doba čekání = celková doba čekání na prostředky, setrvávání v chybovém stavu, pozastavení úlohy.

odstranit - odstraní záznam o statistikách dané úlohy (ne úlohu samotnou)

po kliknutí na "Detail úlohy" se zobrazí "Detail uložené úlohy"

Uložené úlohy

Application for parallel processing

Aktuálně spuštěné úlohy

Přehled výpočetních uzlů

Historie dokončených úloh

Možnosti uložených úloh

Vytvořit novou úlohu

Filtrovat

Název úlohy:

Čas spuštění:

od:

/

/

do:

/

/

Vytíženo serverů:

0

-

100

Zrušit všechny filtry

Seřadit podle:

Data spuštění

Doby běhu

Počet vytížených serverů

ÚLOHA 1

Název úlohy:

Úloha 1

Status:

Není spuštěna

Naposledy spuštěno:

-

Vytíženo serverů:

40

Doba trvání:

-

Počet kroků úlohy:

5

Detail úlohy

ÚLOHA 2

Název úlohy:

Úloha 2

Status:

Není spuštěna

Naposledy spuštěno:

-

Vytíženo serverů:

40

Doba trvání:

-

Počet kroků úlohy:

5

Detail úlohy

ÚLOHA 3

Název úlohy:

Úloha 3

Status:

Chyba

Naposledy spuštěno:

05.12.2020 12:27:50

Vytíženo serverů:

30

Doba trvání:

5 dní 5 hodin 37 minut

Počet kroků úlohy:

12

Detail úlohy

<

1

2

3

4

...

>

po kliknutí na řádek tabulky se zobrazí "Detail uložené úlohy".

Zobrazují se uložené i spuštěné úlohy (ty co byly spuštěny víckrát se zobrazují pouze 1x).

Detail uložené úlohy

Application for parallel processing

Aktuálně spuštěné úlohy

Přehled výpočetních uzlů

Historie dokončených úloh

Možnosti uložených úloh

Vytvořit novou úlohu

Zpět

Detail úlohy

Název úlohy: 1. Úloha

Výpočetní uzly :

knot01

knot02

Specifikace jednotlivých kroků úlohy

1. krok

2. krok

3. krok

+ Přidat další krok

Název kroku: vertikalizace

Program: python3 /mnt/programs/vert.py

Typ vstupu: složka

Prefix: --input

Cesta: /mnt/data/in

Regulární výraz: [0-9]+

Typ výstupu: složka

Prefix: --output

Cesta: /mnt/data/out/\$1.vert

Parametry: -t target out mnt/data/out/

Ověření:

☒ výstupní složka není prázdná

☒ stderr je prázdný

☒ velikost logu

☒ Pro každou vstupní složku existuje odpovídající neprázdný výstupní soubor nebo složka

☒ Při chybě vypsát obsah logu /mnt/logs/\$1.err

☒ Nutná synchronizace kroku

Vytvořit úlohu založenou na této úloze

Odstranit úlohu

Spustit úlohu

Exportovat úlohu

odstranit+spustit s alert dialogem

odstranit úlohu - odstraní i všechny statistiky běhu úlohy!

Vytvoření úlohy založené na jiné úloze

↩️ ⏪ ⏩ ⏹️ 🏠

Application for parallel processing

🔍

Aktuálně spuštěné úlohy

Přehled výpočetních uzlů

Historie dokončených úloh

Možnosti uložených úloh

Vytvořit novou úlohu

⏪ Zpět

Název úlohy:

1 Úloha

Konfigurace úlohy:

☒ Vlastní

☐ Načíst uloženou konfiguraci úlohy

☐ Importovat konfiguraci úlohy

Načíst

Vybrat soubor

Nahrát

Servery:

knot01

knot02

>><<

knot03

knot04

Specifikace jednotlivých kroků úlohy:

1. krok

2. krok

3. krok

+ Přidat další krok

Název kroku:

vertikalizace

Program:

python3 /mnt/programs/vert.py

Vstup:

složka složek

Prefix:

--input

Cesta:

/mnt/data/in/

Počet procesů:

1

Výstup:

složka souborů

Prefix:

--output

Cesta:

/mnt/data/out/\$2.vert

Regulární výraz:

(*)\\(.*\\)(.*)

+ Přidat parametr

Timeout:

8

h

Ověření:

☐ existence neprázdného výstupního souboru

☒ výstupní složka není prázdná

☒ stderr je prázdný

☒ velikost logu /mnt/data/logs/\$2.log menší než 10 kB

☐ počet výstupních souborů ve složce /mnt/data/out roven počtu vstupních souborů

☐ počet souborů ve vstupní složce je roven počtu souborů ve výstupní složce

☐ Pro každý vstupní soubor existuje odpovídající neprázdný výstupní soubor nebo složka

☒ Pro každou vstupní složku existuje odpovídající neprázdný výstupní soubor nebo složka

Při chybě:

☒ Vypsat obsah logu /mnt/logs/\$1.err

☐ Vypsat obsah stderr

☒ Nutná synchronizace kroku

Odstranit krok

Uložit

Uložit a spustit

Ve vstupu nemá smysl \$1, ne? Naopak v logu ano.

Aplikuji reg. výraz na název vstupního souboru a dostanu nějaké stringy (typ vstupu "soubory ve složce")

A ty stringy můžu využít pro:

- * cestu a název výstupního souboru
- * cestu a název logu
- * nějaké další parametry

Nová úloha

Application for parallel processing

Aktuálně spuštěné úlohy

Přehled výpočetních uzlů

Historie dokončených úloh

Možnosti uložených úloh

Vytvořit novou úlohu

Název úlohy:

Konfigurace úlohy:

Vlastní

Načíst uloženou konfiguraci úlohy

Importovat konfiguraci úlohy

Načíst

Vybrat soubor

Nahrát

Servery:

knot01
knot02

>
>>
<
<<

knot03
knot04

Specifikace jednotlivých kroků úlohy:

1. krok

2. krok

3. krok

+ Přidat další krok

Název kroku:

Program:

Vstup:

typ vstupu

Prefix:

Cesta:

Počet procesů:

1

Výstup:

typ výstupu

Prefix:

Cesta:

Regulární výraz:

+ Přidat parametr

Timeout:

h

Ověření

existence neprázdného výstupního souboru

výstupní složka není prázdná

stderr je prázdný

velikost logu

menší než

kB

počet výstupních souborů ve složce

roven počtu vstupních souborů

počet souborů ve vstupní složce je roven počtu souborů ve výstupní složce

Pro každý vstupní soubor existuje odpovídající neprázdný výstupní soubor nebo složka

Pro každou vstupní složku existuje odpovídající neprázdný výstupní soubor nebo složka

Při chybě:

Vypsat obsah logu

Vypsat obsah stderr

Nutná synchronizace kroku

Odstranit krok

Uložit

Uložit a spustit

načíst konfiguraci = načtu konfiguraci nějaké jiné úlohy - musím ji uložit však jako novou (jinou) úlohu, u které mohu některé parametry změnit jejich přepsáním.

Jednotlivé možnosti ověření se nebudou disablovat, ale dynamicky měnit podle vybrané kombinace typu vstupu a výstupu