

vhodnější název

Application for parallel processing

Aktuálně spuštěné úlohy

Přehled výpočetních uzlů

Historie

Přehled 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

Application for parallel processing

Aktuálně spuštěné úlohy

Přehled výpočetních uzlů

Historie

Přehled uložených úloh

Vytvořit novou úlohu

Zpět

Detail 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 běžící kroky

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

Možnosti úlohy

Zobrazit konfiguraci

Pozastavit úlohu

Spustit

Stav úlohy na jednotlivých výpočetních uzlech

Filtrovat

Status:

☐ běží

☐ běží/chyba

☐ chyba

☐ čeká

☐ pozastaveno

Krok:

Zvolte krok

Doba běhu:

od:

0

 dnů

0

 hod

do:

0

 dnů

0

 hod

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

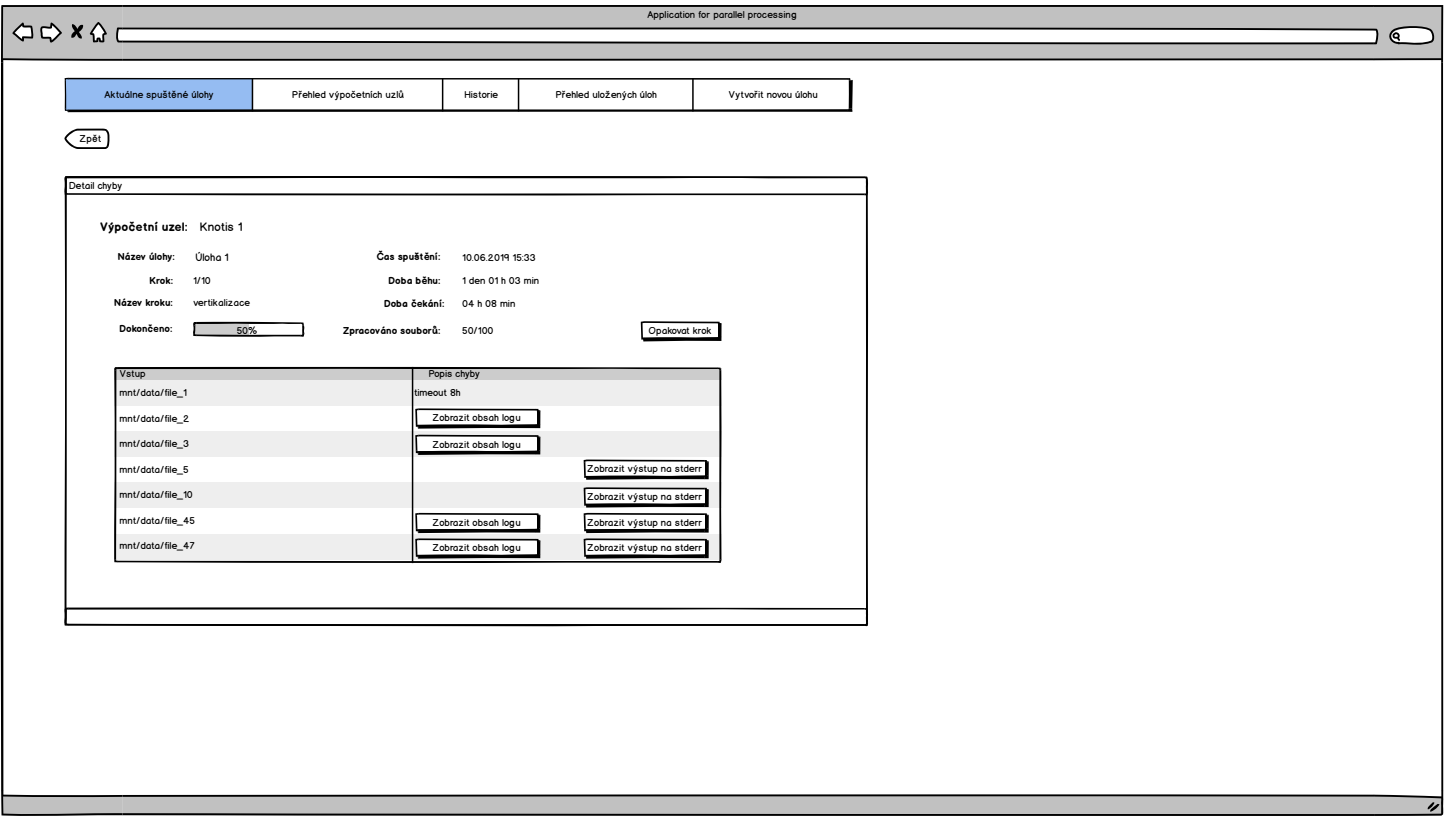
Běžící kroky

- 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



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%

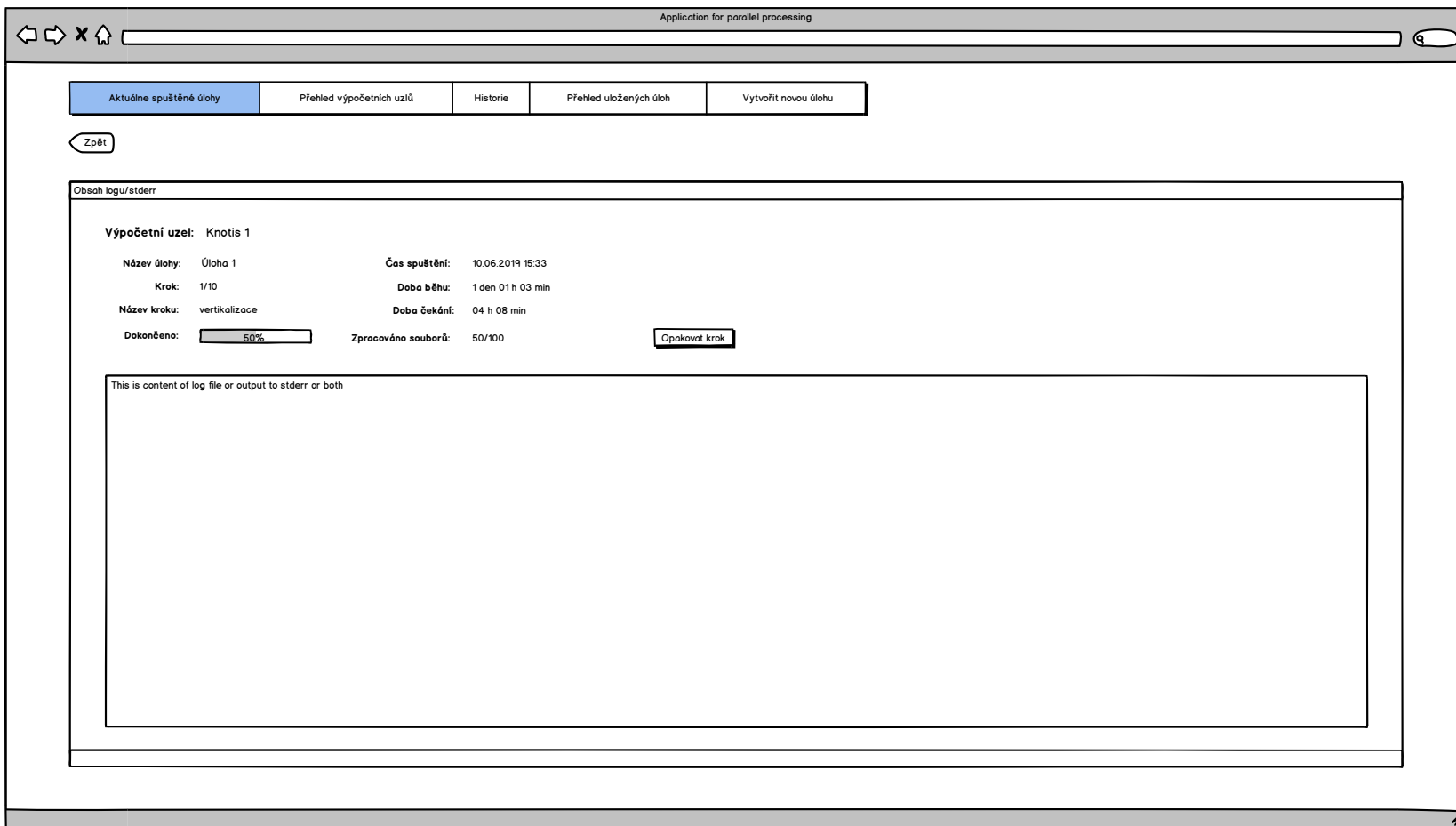
Popis chyby může mít 4 varianty:

- 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.



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)

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.

Application for parallel processing

Aktuálně spuštěné úlohy

Přehled výpočetních uzlů

Historie

Přehled 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

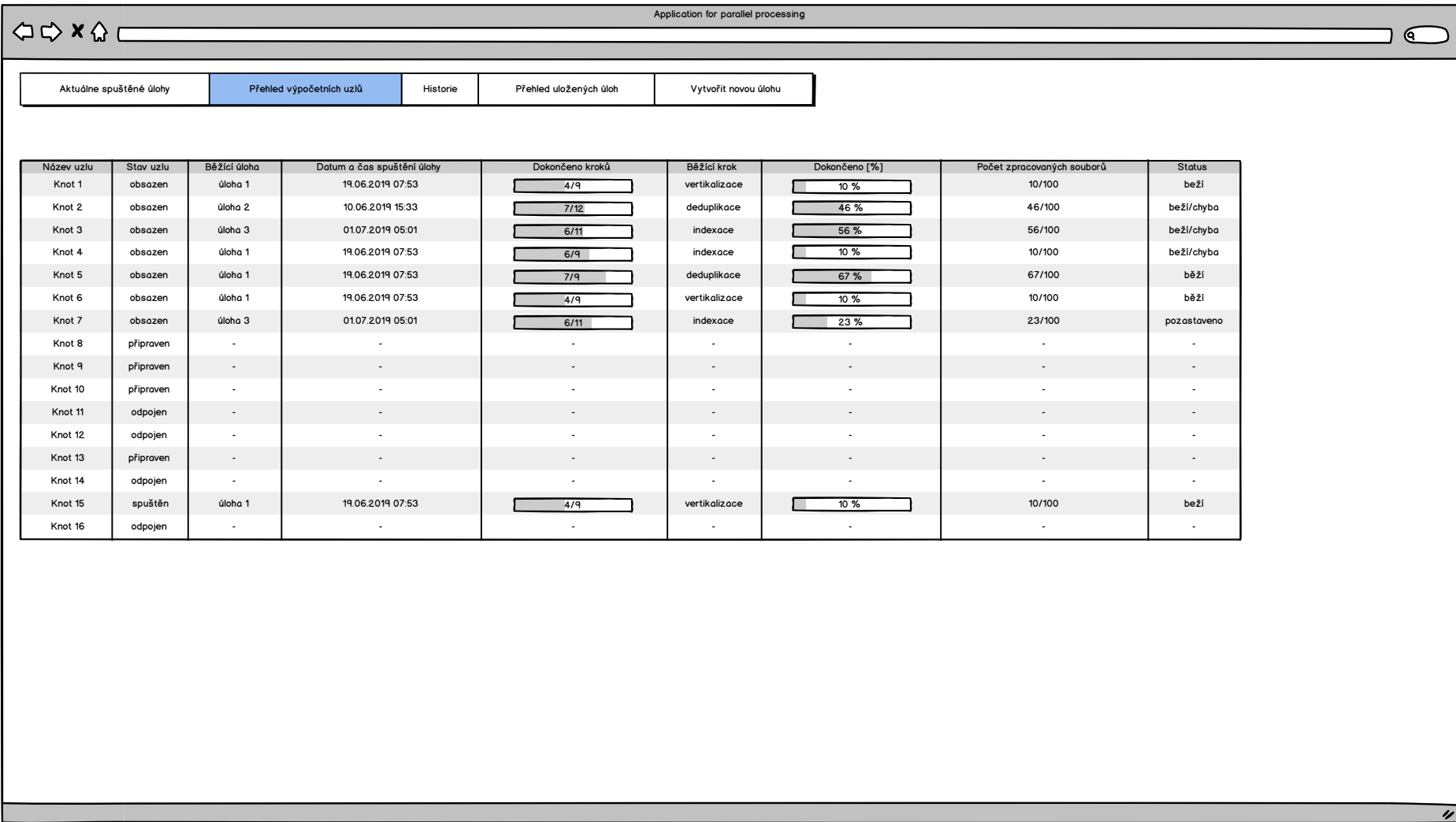
☒ Při chybě vypsát obsah logu /mnt/logs/\$1.err

☒ Nutná synchronizace kroku

druhá varianta - konfigurace úlohy je společně s detailem úlohy na jedné stránce - já osobně bych volil tuto variantu

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



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) ?

Application for parallel processing

Aktuálně spuštěné úlohy

Přehled výpočetních uzlů

Historie

Přehled uložených úloh

Vytvořit novou úlohu

Filtry

Název úlohy:

Čas spuštění:

od:

//

do:

//

Celková doba běhu:

od: dnů

0

 hod

do:

dnů

0

 hod

Vytíženo serverů:

0

 -

100

Zrušit všechny filtry

Seřadit podle:

Data spuštění

Doby běhu

Počet vytížených 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

Zobrazit statistiky

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

Zobrazit statistiky

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

Zobrazit statistiky

<

1

2

3

4

...

>

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

přidat ještě nějaké grafy ?

zaznamenávat ještě nějaké další informace do historie?



Application for parallel processing

Aktuálně spuštěné úlohy

Přehled výpočetních uzlů

Historie

Přehled uložených úloh

Vytvořit novou úlohu

Seřadit podle:

Data spuštění

Názvu

Filtrovat

Název úlohy:

Čas spuštění:

od:

/

/

do:

/

/

Vytíženo serverů:

0

-

100

Zrušit všechny filtry

Název úlohy	Naposledy Spuštěno	Vytíženo serverů
Úloha 14	09.07.2019 23:52	10
Úloha 13	08.07.2019 07:47	12
Úloha 12	30.05.2019 10:42	6
Úloha 11	02.04.2019 08:02	55
Úloha 10	01.04.2019 13:12	10
Úloha 9	30.03.2019 07:47	38
Úloha 8	30.03.2019 10:42	7
Úloha 7	02.03.2019 08:02	55
Úloha 6	01.02.2019 23:22	27
Úloha 5	21.12.2018 07:47	12
Úloha 4	30.11.2018 10:42	8
Úloha 3	10.09.2018 08:02	60
Úloha 2	02.03.2018 23:52	10
Úloha 1	01.01.2018 07:47	12

<

1

2

3

4

...

>

Název úlohy	Naposledy Spuštěno	Vytíženo serverů
Úloha 14	09.07.2017 23:52	11
Úloha 13	08.07.2017 07:47	12
Úloha 12	30.05.2017 10:42	06
Úloha 11	02.04.2017 08:02	55
Úloha 10	01.04.2017 13:12	10
Úloha 9	30.03.2017 07:47	08
Úloha 8	30.03.2017 10:42	70
Úloha 7	02.03.2017 08:02	55
Úloha 6	01.02.2017 23:22	39
Úloha 5	21.12.2016 07:47	19
Úloha 4	30.11.2016 10:42	28
Úloha 3	10.09.2016 08:02	48
Úloha 2	02.03.2016 23:52	17
Úloha 1	01.01.2016 07:47	14

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

Application for parallel processing

Aktuálně spuštěné úlohy

Přehled výpočetních uzlů

Historie

Přehled uložených úloh

Vytvořit novou úlohu

Zpět

ÚLOHA 1

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

Odstranit

Editovat

Spustit úlohu

Exportovat úlohu

odstranit+spustit s alert dialogem

Application for parallel processing

Aktuálně spuštěné úlohy

Přehled výpočetních uzlů

Historie

Přehled uložených úloh

Vytvořit novou úlohu

Zpět

Editace úlohy

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/\$1vert

Regulární výraz: [0-9]+

+ Přidat parametr

Timeout:

8

h

Ověřit:

☐ existence neprázdného výstupního souboru

☒ výstupní složka není prázdná

☒ stderr je prázdný

☒ velikost logu /mnt/data/log

menší než

10

KB

☐ počet výstupních souborů ve složce /mnt/data/out je 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/\$1err

☐ Vypsat obsah stderr

☒ Nutná synchronizace kroku

Uložit

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

Application for parallel processing

Aktuálně spuštěné úlohy

Přehled výpočetních uzlů

Historie

Přehled 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 je 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:

☐ Vypsát obsah logu

☐ Vypsát obsah stderr

☐ Nutná synchronizace kroku

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