
 skodapetr Update documentation ✓

 History

 2 contributors



 227 lines (163 sloc) | 12.2 KB

...

# Konfiguračná príručka a pokyny pre diagnostiku - Národný katalog otvorených dat

Čtenář této příručky by měl být obeznámen s:

- [Kubernetes](#)
- [OCI](#)
- [Inštalčná príručka a pokyny na inštaláciu \(úvodnú/opakovanú\)](#)
- [Aplikační příručka](#)
- [Uživatelská příručka](#)

## 1. ROZSAH PLATNOSTI A ÚČEL

Tato konfiguračná příručka slouží k popisu Národního katalogu otevřených dat (NKOD), části projektu OD2.0.

## 2. DEFINÍCIA POJMOV A SKRATIEK

**NKOD**

Národní katalog otevřených dat

**OCI**

Oracle Cloud Infrastructure

**LP-ETL**

LinkedPipes ETL

**OCPU**

The Oracle CPU

## 3. POŽIADAVKY NA ZÁKLADNÉ HW A SW PROSTREDIE

---

Požadavky na běhové prostředí jsou specifikovány v [Inštalačná příručka a pokyny na inštaláciu \(úvodnú/opakovanú\)](#).

## 4. POSTUP KONFIGURÁCIE

---

Většina potřebné konfigurace je prováděna v rámci instalace popsané v [Inštalačná příručka a pokyny na inštaláciu \(úvodnú/opakovanú\)](#). Pro většinu konfigurace platí, že po její úpravě je nutné provést přenasazení Deploymentu v Kubernetes. Toho je možné dosáhnout příkazem:

```
kubectl rollout restart deploy {deployment-name} --namespace=nodc
```

V případě úpravy existující konfigurace (ConfigMap) nebo tajemství (Secret) není možné znovu využít příkazu vytvoření. Řešením je vytvořit zdroj nanečisto a pak ho načíst do Kubernetes. Například pro konfigurační soubor `./configuration/nginx.conf` může příkaz vypadat následovně:

```
kubectl create configmap nodc-nginx --namespace=nodc --from-  
file=website.nginx=./configuration/nginx.conf --dry-run=client -o yaml |  
kubectl apply -f -
```

### 4.1 Konfigurácia na strane servera (INF)

Tato sekce popisuje konfiguraci NKOD na straně serveru, která je nezávislá na prostředí OCI.

### 4.1.1 SPARQL Dotazy

Uživatel si může v klientské části aplikace komponenty Webové stránky zvolit z předpřipravených SPARQL dotazů (viz. [Aplikační příručka](#)) Tyto dotazy jsou uloženy v konfiguraci komponenty Webové stránky v repositáři [NKOD-SW](#) v souboru `components/website/www/src/sparql-queries.yaml` . V případě změny této konfigurace dojde k automatickému spuštění sestavení a publikace Docker image na GitHubu pomocí [GitHub actions](#). Jakmile je akce dokončena je třeba provést přenasazení komponenty Webové stránky v Kubernetes pomocí příkazu:

```
kubectl rollout restart deploy nodc-website-deployment --namespace=nodc
```

### 4.1.2 Obnova či získání HTTPS certifikátů

Prvotní získání HTTPS certifikátů je popsáno v [Inštalačná příručka a pokyny na instaláciu \(úvodnú/opakovanú\)](#) sekce 4.1.10 Získání HTTPS certifikátu. Platnost certifikátu vystaveného Let's Encrypt jsou tři měsíce.

Při prodlužování certifikátu je třeba provést odstávku systému, respektive komponenty Webové stránky. Následně je třeba spustit Job který provede obnovu certifikátu. V posledním kroku je pak třeba opět pustit služby Webové stránky. Toto je možné provést pomocí příkazů:

```
kubectl delete --namespace=nodc deployment/nodc-website-deployment
kubectl apply -f ./k8s/certificate-manager/update-certificate.yaml
kubectl apply -f ./k8s/website/website.yaml
```

Po ověření úspěšného dobehnutí je vhodné Job smazat:

```
kubectl delete job update-certificate.yaml --namespace=nodc
```

Alternativně je možné deployment nemazat ale upravit počet instancí pomocí příkazu:

```
kubectl scale deploy nodc-website-deployment --replicas=0 --namespace=nodc
```

## 4.2 Konfigurácia a postup pre Cloud riešenie

Tato sekce se věnuje konfiguraci, která je specifická pro prostředí OCI. Konfigurace související s nasazením do clusteru je nastavena při zavádění definic do Kubernetes. Jedná se například o *{Availability Domain}*, *{Compartment}*, či vytvoření disků. Neb je tuto konfiguraci třeba provést při instalaci aplikace, je popsána v [Inštalčná príručka a pokyny na inštaláciu \(úvodnú/opakovanú\)](#). Většina konfigurace je v zásadě fixní a není očekáváno, že by jí bylo třeba měnit.

Konfigurací kterou by mohlo být potřeba měnit je využití výkonnostní třídy *Block Storage*. Tato konfigurace ovlivňuje dvě komponenty:

- RDF úložiště Graph
- Metadatového procesoru - LinkedPipes ETL

První komponenta je nasazena dle definice ze soubor `k8s/graphdb/graphdb.yaml` v rámci *Deployment* `nodc-graphdb-deployment`. V případě druhé komponenty LinkedPipes ETL se jedná o její komponenty `executor` a `executor-monitor`. Ty jsou definovány v souborech `k8s/linkedpipes/executor.yaml` v rámci *Deployment* `nodc-executor-deployment`.

Obě tyto komponenty používají ve výchozím nastavení *Block storage* třídu `oci-block-balanced` definovanou v souboru `k8s/oci/oci-block-storage-balanced.yaml`. V případě nedostatečného výkonu je možné zvážit změnu na třídu `oci-block-high` definovanou v souboru `k8s/oci/oci-block-storage-high.yaml`. Nedostatek výkonu se může projevit zvýšenou dobou harvestování nebo odpovídání na SPARQL dotazy. Před tímto zásahem je ovšem také vhodné zvážit zvýšení počtu OCPU.

Změnu je možné provést v souborech `k8s/graphdb/graphdb-pvc.yaml` a `k8s/linkedpipes/executor-pvc.yaml`. Součástí této změny je ztráta starých *Block storage* úložišť a tedy i uložených dat. Tato data budou znovu vytvořena v rámci harvestování.

## 4.3 Konfigurácia na strane klienta (INF)

Klientskou částí aplikace je webová stránka běžící v prohlížeči uživatele (Návštěvník bez znalosti SPARQL, Návštěvník se znalostí SPARQL).

### 4.3.1 Vnitřní konfigurace pro Yasgui

Klientská část aplikace je dostupná skrze komponentu Webové stránky. Jedná se o knihovnu [Yasgui](#) pro dotazování nad SPARQL endpointem. Součástí této dokumentace je například určení HTTP metody pro vykonání SPARQL dotazu, nebo SPARQL endpointu.

Tatko konfigurace není za normálních okolností dostupná uživateli. Nicméně jakmile uživatel jednou zobrazí stránku Yasgui si svojí konfiguraci uloží do session storage prohlížeče. Tato konfigurace je použita při dalším načtení stránky.

Při změně této konfigurace, ve zdrojovém kódu komponenty Webové stránky, může být třeba toto konfiguraci smazat. Alternativně je možné implementovat mechanismus který smazání provede automaticky při detekci změny.

## 4.4 Popis rozhraní IS (APL)

Rozhraní IS je popsáno v dokumentech [Aplikační příručka](#) a [Uživatelská příručka](#).

### 4.1.7 Konfigurácia programu

Většina konfigurace se provádí skrze objekty v Kubernetes. V případě integrace s OCI jde přímo o úpravu objektů v případě konfigurace a tajemství pak o úpravu konfiguračních souborů. Konfigurace SPARQL dotazů se provádí skrze úpravu souboru v repozitáři [NKOD-SW](#).

### 4.1.8 Možné chyby programu

Selháním harvestačních pipeline v LinkedPipes ETL je možné z důvodu chybné konfigurace či nedostupnosti používaných služeb.

Ostatní chyby související s konfigurací jsou popsány v 5. CHYBOVÉ STAVY PRI KONFIGURÁCIÍ.

### 4.2.2 Význam polí konfigurovanej oblasti

Význam konfiguračních záznamů je popsán v konfiguračních souborech.

### 4.2.3 Dôvod a predpoklady konfigurácie oblasti

Hlavní částí konfigurace jsou pipelines pro LinkedPipes ETL. Ty určují proces harvestace, jaké kvalitativní atributy jsou kontrolovány a jak jsou data publikována.

Další oblastí konfigurace jsou hodnoty závislé na prostředí OCI. Jedná se o konfiguraci popsanou v 4.2 Konfigurácia a postup pre Cloud riešenie. Důvodem této konfigurace je nutnost nastavit odpovídající identifikátory pro integrace Kubernetes s OCI zdroji.

Konfigurace server popsaná v 4.1 Konfigurácia na strane servera (INF) pak slouží pro nastavení zabezpečení a vstupně výstupním rozhraním. Změny v této oblasti lze očekávat při změně: domény, zdrojů externích dat jako jsou pipelines nebo registrační záznamy, SPARQL dotazy. Pravidelným důvodem pro změnu části konfigurace je prodloužení platnosti HTTPs certifikátu.

Na straně klienta, konfigurace popsána v 4.3 Konfigurácia na strane klienta (INF), by neměl být důvod ke změně.

## 5. CHYBOVÉ STAVY PRI KONFIGURÁCIÍ

---

Chyby způsobené špatnou konfigurací se mohou projevit:

- Selháním harvestačních pipeline v LinkedPipes ETL
- Neúspěšným pokusem o nahrání definice do Kubernetes
- Nefunkčním Pod v Kubernetes

Selhání harvestačních pipeline může být kromě externího důvodu způsobeno také chybnou konfigurací. V takovém případě je postup podobný jako v případě 4.1.8 Možné chyby programu.

V případě pokusu o nahrání nevalidního souboru do Kubernetes je zobrazena chybová hláška, která popisuje problém. Neb je náprava takové chyby otázkou znalosti Kubernetes tak se jí tento dokument nevěnuje.

V některých případech může dojít k úspěšnému načtení definice do Kubernetes, nicméně vytvořený object nepracuje jak by měl. V případě Persistent Volume Claim může být důvodem nedostupnost, nebo nepřipravenost Persistent Volume. Dalším může být selhání spuštění Pod. Zde může být příčinou nedostupný Docker image pro spuštění, či chybový stav spouštěné aplikace. První případ může být způsobem nedostupností externí služby, nebo neplatnými autorizačními údaji. Ve druhém případě se může jednat o chybu konfiguraci. Například pro komponentu Webové stránky může dojít k nefunkčnosti v případě absence HTTPs certifikátu. Základem pro diagnostiku zdrojů jsou příkazy:

```
kubectl describe --namespace=nodc {object name}
kubectl logs {pod name}
```

První příkaz ukazuje zprávy z životního cyklu zdroje a chyby před spuštěním Docker kontejneru. Druhý příkaz zpřístupňuje standardní výstup a je ho tedy možné použít ke čtení logů.

V případě problému s Pod doporučujeme získat chybový výstup a následně zrušit Deployment který Pod definuje. Kubernetes se totiž bude snažit Pod opakovaně vytvářet. Díky konfiguraci stahování Docker image pro každé vytvoření může snadno dojít k překročení počtu stažení Docker image z externího repository.

## 6. UMIESTNENIE ZAKLADNEJ DOKUMENTÁCIE

---

<https://github.com/datova-kancelaria/nkod-dokumentacia>

## 7. SÚVISIACA DOKUMENTÁCIA

---

- [Aplikační příručka](#)
- [Inštalčná príručka a pokyny na inštaláciu \(úvodnú/opakovanú\)](#)
- [Uživatelská příručka](#)

## 8. ZOZNAM INŠTALAČNÝCH MÉDIÍ

---

Instalace NKOD specifického kódu se provádí z [NKOD-SW](#) repositáře. Definice pipeline pro LinkedPipes ETL je uložena v [NKOD-PIPELINE](#) repositáři.

[Give feedback](#)