



NSO 6.0 からの新機能 とそれによる問題の解決

Akira Iwamoto

2/28/2024

自己紹介

- 岩本 彰 (akiwamot@cisco.com)
- Technical Leader / Cisco Systems TAC

- -2011 (他エンタープライズ企業)
 - Catalyst 6500, 7200VXR 等レガシールータ、スイッチ
- 2011-2015
 - CRS / ASR9000 シリーズルータ (IOS-XR)
- 2015-current
 - SP Automation (NSO / ESC / EPNM / Crosswork)

今日の紹介

- 追加された機能のリスト紹介
- HA (5分)
- Concurrency Model (5分)
- その他 (5分)

6.0 からの新機能

新機能 (6.0)

機能	内容
Concurrency Control Model	複数のトランザクションを同時に実行（トランザクション中 Critical Section は同時実行しない）
Lockless dry-run	commit dry-run 実施時に、Transaction Lock取得が不要
Transaction Lock Time Violation Alarm	トランザクションロックの時間が指定した時間より長くなると、アラームを上げる（デフォルトはオフ）
Metric Framework	Insight ManagerがWebUIに登場。Commit Queue や North Bound の状態について、1/5/15分毎の統計情報で表示
JSON-RPC Progress Trace	JSON-RPC で Progress Trace の出力をSubscribe、commit 動作についての情報を取得
Biased free API	昨今の風潮から、Master/Slave ではなく Primary/Secondary と呼び方が変更
Enhanced /services/check-sync	全サービスを対象として、フィルタしながらまとめて check-sync

新機能 (6.1)

機能	内容
HA Raft	新アーキテクチャを使ったHA
SSO and Package Authentication	シングルサインオン (SAML2.0) サポート (6.1.3) (6.1 / 6.1.2 ではリミテーションがあり)
Progress Trace Enhancement	Progress-Trace に span-id の導入し、トランザクションの追跡がSpan単位で可能となる
Compliance Reporting	コンプライアンスレポート機能の機能強化、結果をSQLite ファイルへ出力が可能となり、他システムから再利用が可能
Unified rollback commit API	rollback-files action が、既存の NETCONF/RESTCONF/MAAPI に追加し、CLI / JSON-RPCでも使用可能
Devices platform smart license	接続先機器の smart-license 状態を、NSO側でキャッシュ
HA primary listen	HA ステータスがprimary となった際に、指定した ip/port で Listen する。netconf, restconf, http(s), json-rpc, cli/ssh. 外部LBによるTCPポート監視で、適切に必要なノードが使用される
Package Add	NEDパッケージの新規追加時、フルリロード (packages reload) をせずに単純追加が可能。運用環境において作業時間の短縮
New counters	種々の動作履歴カウンタ (トランザクションがabort した回数など) /metric/sysadmin/counter/* /metric/log/enabled (ログファイルに記録する場合)
CDB Compaction Improvements	自動で実行される CDB Compactionのタイミングについて、種々のパラメータを指定が可能 (ファイルサイズ、トランザクション数等)

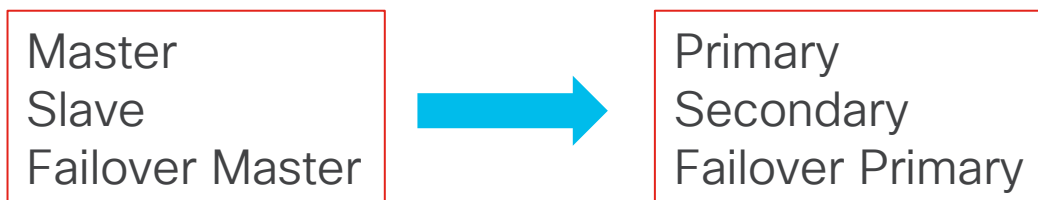
新機能 (6.2)

機能	内容
CDB+transaction log improvement	CDBに格納されたトランザクションログについて、操作エラー時に詳細を記録する transaction-error-log を追加
Service/Device timestamp	サービスインスタンスやデバイスの更新日時の記録
Automatic self component	ncs:self コンポーネントが自動的に追加されるため、Nanoサービスの設計時に明示的に記述不要
CDB improvements	CDB Upgrade の高速化と、処理時のログ出力見直しによる Debuggability の向上
SSO Package Authentication	LDAP/TACACS+, 2FA を認証で使用可能に
Selective Rollback	rollback-files apply-roollback-file アクションに、exclude オプションの追加
Progress Trace Java and Python API	Progress Traceを使用するための、Java と Pythonの 新設 API
HA Logging improvements	デバッグビリティ。HA の状態遷移のログを追加。
Device Group Authorization	Device Groupに対する、NACMルールの設定が可能
JSON-RPC get_leafref_values pagination	get_leafref_values メソッドに、offset パラメータが追加。全件取得ではなく、ページ単位として取得可能に。
NACM action input	Action の入力パラメータに対する NACM 設定。特定のユーザのみがパラメータを使用可能。(sync-from dry-run のみ使用可能、など)

HA (High Availability)

ノードの呼び方

- Master / Slave は Primary / Secondary といった呼び方に置き換え
- ドキュメントと製品内で、まだ呼称が混在している場合がある
- <https://www.cisco.com/c/en/us/about/social-justice/inclusive-language-policy.html>



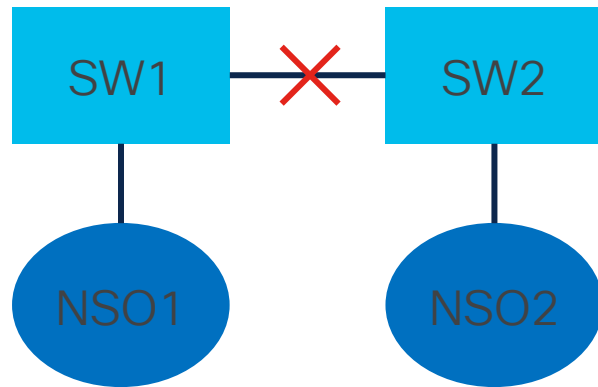
本ドキュメントでは、ノードの呼称については 1号機、2号機、
役割については、Primary / Secondary と呼びます

HA の問題点

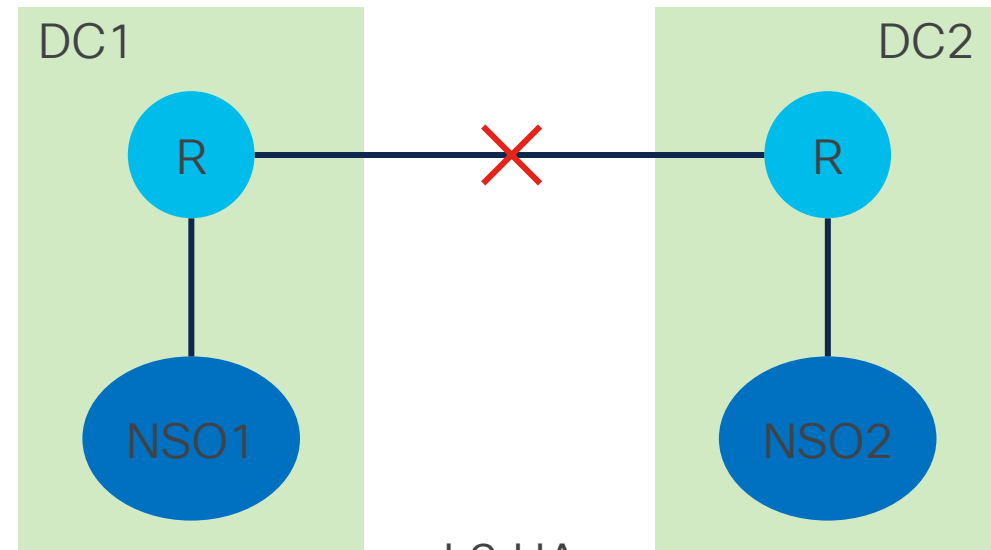
- 既存のHA 構成における問題点

- スプリットブレイン

- 2つのノード間がネットワーク的にダウンした場合に、もう一方の状態が分からないため、結果 Active-Active となる
- 誤って2つのNSOへ別の変更がされる可能性がある



L2 HA



L3 HA

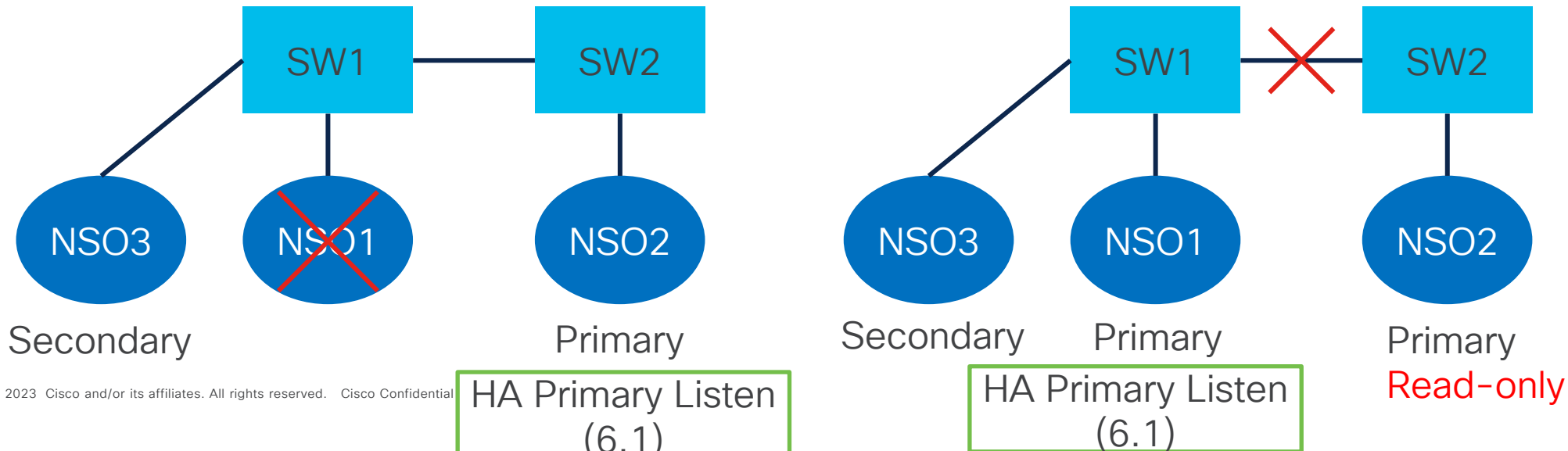
Split Brain 問題の解決策 – Rule Based Algorithm (5.7)

- Failover 実施後、2号機が Primary (Failover-Primary) となった際、Read-only モードとなる
 - (1) Secondaryが join すると解除
 - (2) 手動解除
- Perpetual Secondary を設置
 - Perpetual Secondary が、Primary となった2号機に join し、自動的に解除

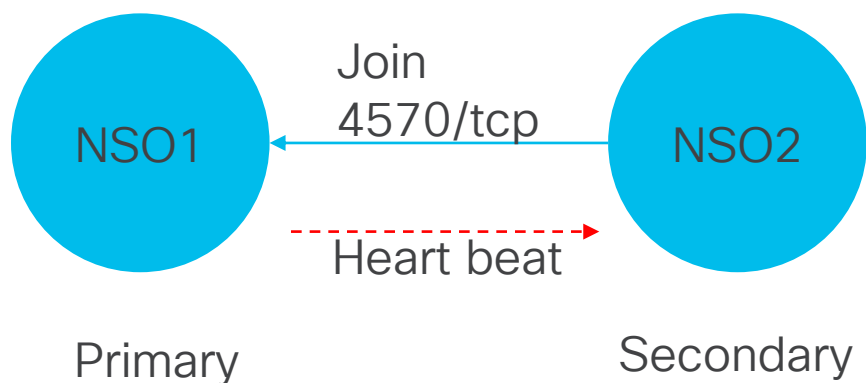
設定 (デフォルトで有効)

high-availability settings consensus enabled

high-availability settings consensus algorithm rule-based

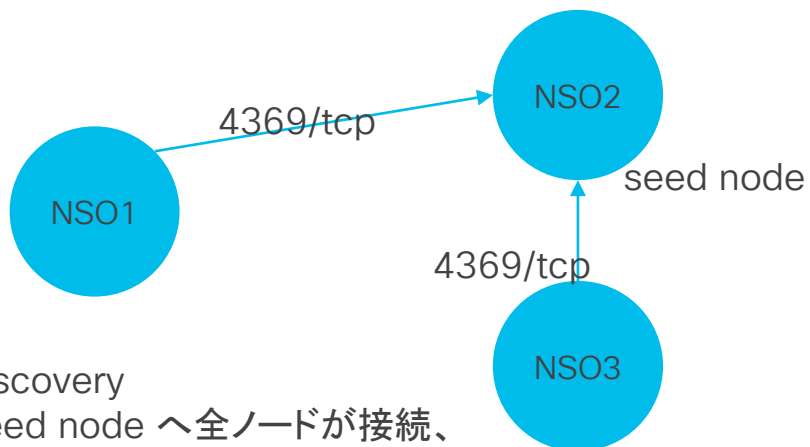


HA のアーキテクチャ (旧来)

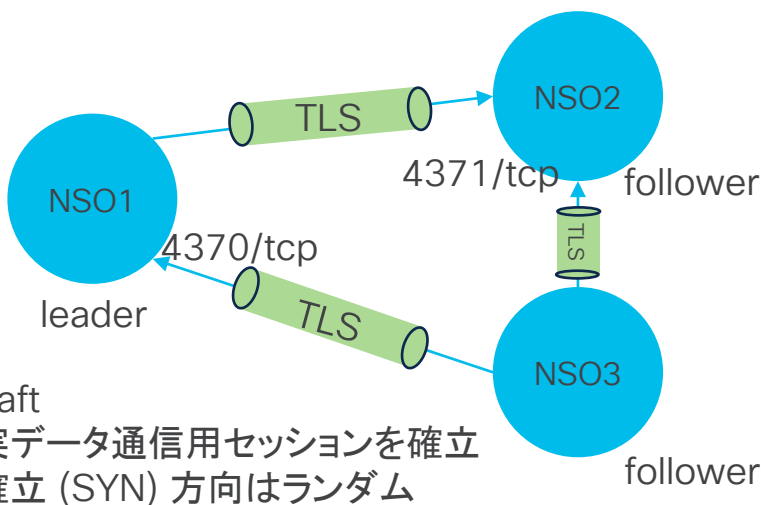


- NSO1 がサーバとなり、port 4570 をlisten
- NSO2 が NSO1 の 4570 へ TCP 接続
- その接続上で、Heart beat や CDB Sync を実施
- 1-way HA
 - 1号機に障害、Failoverされ2号機がPrimaryとして動作
 - 1号機をSecondaryとして復旧後、2号機の障害時には自動Failoverしない
 - Failover後には、速やかにRecoveryを実施する必要がある
- TCP接続上のメッセージ交換は平文

HA のアーキテクチャ (HA Raft)



- Seed node
 - Discovery フェーズにおいて、ノード発見のために使用される
 - ncs.conf で指定
 - サービス提供用NSOと兼務でOK
- Leader node
 - ノードの中から選ばれる
- Follower node
 - Leader 以外のノード



- NSO 起動時
 - NSO1 を seed ノードと決定。NSO1 で他のノードを Discovery
- create-cluster の実行
 - 3ノード間で leader/followerを決定する
- その接続上で、Heart beat や CDB Sync を実施
- Failover後も、ダウンしたノードが復旧した場合はProtectされる
- TCP接続上のメッセージ交換はTLSを使用して暗号化

今後のHA方法の選定

- high-availability
- ha-raft

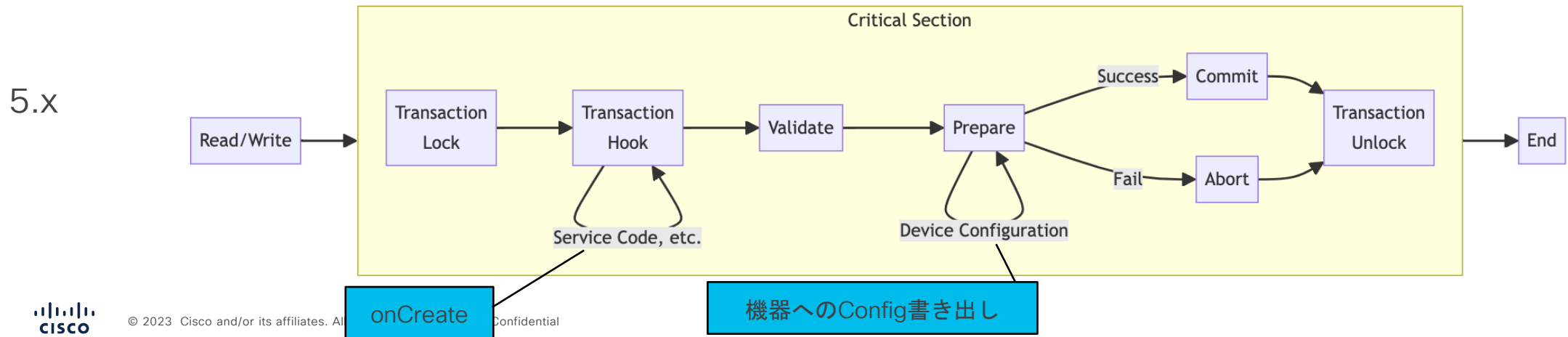
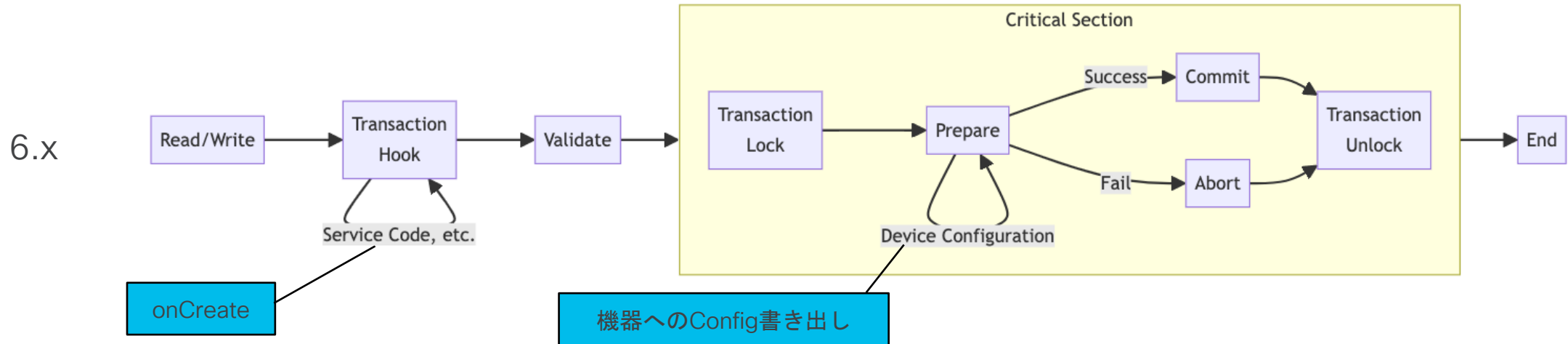
- どちらもサポートされ、ユーザは選択できます
- L3 BGP による Virtual IP を使うためには、tailf-hcc が必要です

	NSO 5.7, 5.8	6 and up
tailf-HCC 4.X	tailf-hcc-4.6.x	n/a
tailf-HCC 5.X	tailf-hcc-5.x	tailf-hcc-5.x

Concurrency Model

Concurrency Model (6.0)

- <https://qiita.com/akiraiwamoto/items/4ca9fbb7095c095ffce1>



他の新機能

Faster CDB Upgrades (6.2)

- CDBのアップグレードを行う内部コードが刷新され、ラボでの計測では最大8倍の速さで完了
- CDB Upgrade
 - NSO のアップグレード
 - NED, サービス・パッケージのアップグレード
- CDBはYangのスキーマ情報に基づいて、内部ツリー構造を構成することでデータを保存
- アップグレード実行時には、スキーマの変更や構造の変更が行われるため、一定の時間が必要
- 内部ツリー構造や、アップグレード時のアルゴリズム変更により、それが高速化

Service/Device timestamp (6.2)

```
admin@ncs# show link t2
```

TYPE	NAME	BACK TRACK	GOAL	STATE	STATUS	WHEN	ref	POST ACTION STATUS
self	self	false	-	init	reached	2024-01-31T16:44:13	-	-
				ready	reached	2024-01-31T16:44:13	-	-
vlan-link	ex1-eth0-ex2-eth0	false	-	init	reached	2024-01-31T16:44:13	-	-
				dev-setup	reached	2024-01-31T16:44:13	-	-
				ready	reached	2024-01-31T16:44:13	-	-

```
modified devices [ ex1 ex2 ]
```

```
directly-modified devices [ ex1 ex2 ]
```

```
plan-location /link[name='t2']
```

```
created 2024-01-31T16:44:12.960235+00:00
```

```
last-modified 2024-01-31T16:44:12.967658+00:00
```

```
last-run 2024-01-31T16:44:13.030749+00:00
```

```
admin@ncs#
```

```
admin@ncs# show devices device ex0
```

```
devices device ex0
```

```
last-modified 2024-01-31T16:43:56.333349+00:00
```

```
last-connected 2024-02-01T04:06:28.859463+00:00
```

```
last-in-sync 2024-01-31T16:43:56.388648+00:00
```

```
...
```

- 最後にサービスが作成・更新・deploy/re-deployされた日時
- 最後にデバイスが作成・変更・接続・Syncされた日時

カウンタ

```
admin@ncs# show metric sysadmin counter transaction
```

NAME	COMMIT	TOTAL TIME	SERVICE TIME	VALIDATION TIME	GLOBAL	GLOBAL	ABORTS	CONFLICTS	RETRIES
					LOCK TIME	LOCK WAIT TIME			
operational	6	0.043s	0s	0s	0s	0s	0	-	-
running	3	0.954s	0s	0.202s	0.037s	0.008s	0	0	0

```
admin@ncs# show metric sysadmin counter device
metric sysadmin counter device connect 1
metric sysadmin counter device connect-failed 0
metric sysadmin counter device sync-from 1
metric sysadmin counter device sync-to 0
metric sysadmin counter device out-of-sync 0
admin@ncs#
```

Devices platform smart license

```
admin@ncs> show devices device c0 platform smart-license
platform smart-license enabled
platform smart-license registration registration-state UNREGISTERED
platform smart-license registration smart-account "Smart Account Name"
platform smart-license registration virtual-account "Virtual Account 1"
platform smart-license utility disabled
platform smart-license transport transport-type Callhome
platform smart-license authorization authorization-state "OUT OF COMPLIANCE"
```

接続先 cisco 機器上の、Smart License状況を sync-from 実施時に取得し、oper data として保持。
使用可否はNEDや接続先機器に依る

- Added support for populating tailf-ncs-devices-cisco-smart-license.yang upon connect.
Note: NSO 6.1 or newer only
(CISCOIOS-2587 / ENG-32798)

用語

Span

- トランザクション中に起こる動作の固まり
- Spanの中にSpanがある場合もある
- Progress Trace管理上の概念

トランザクションフェーズ

```
admin@ncs(config)# devices device xr0 config hostname def
admin@ncs(config-config)# commit | details debug
applying transaction for running datastore usid=50 tid=1143 trace-id=c3f3cc96-2d2b-4d9e-b43d-c07567fc8802
2024-02-04T18:19:29.869 waiting to apply... ok (0.000 s)
entering validate phase
2024-02-04T18:19:29.870 creating rollback checkpoint... ok (0.000 s)
2024-02-04T18:19:29.870 creating rollback file... ok (0.005 s)
...
2024-02-04T18:19:29.885 taking transaction lock... ok (0.000 s)
2024-02-04T18:19:29.885 holding transaction lock...
2024-02-04T18:19:29.885 check for read-write conflicts... ok (0.000 s)
leaving validate phase (0.015 s)
entering write-start phase
...
leaving write-start phase (0.003 s)
entering prepare phase
2024-02-04T18:19:29.889 cdb: prepare
2024-02-04T18:19:29.891 device-manager: prepare
2024-02-04T18:19:29.895 device xr0: calculating southbound diff... ok (0.013 s)
2024-02-04T18:19:29.909 device xr0: connect...
2024-02-04T18:19:29.909 device xr0: send NED connect
2024-02-04T18:19:30.349 device xr0: package cisco-iosxr-cli-7.49: connecting... ok (0.095 s)
2024-02-04T18:19:30.807 device xr0: send NED is-alive
2024-02-04T18:19:30.812 device xr0: connect: ok (0.902 s)
...
2024-02-04T18:19:30.963 device xr0: prepare...
2024-02-04T18:19:30.963 device xr0: send NED prepare
2024-02-04T18:19:30.967 device xr0: package cisco-iosxr-cli-7.49: transforming output... ok (0.001 s)
2024-02-04T18:19:30.969 device xr0: package cisco-iosxr-cli-7.49: sending config... ok (0.013 s)
2024-02-04T18:19:30.983 device xr0: prepare: ok (0.020 s)
2024-02-04T18:19:30.984 device xr0: commit...
2024-02-04T18:19:30.984 device xr0: send NED commit
2024-02-04T18:19:30.985 device xr0: package cisco-iosxr-cli-7.49: committing config... ok (0.045 s)
2024-02-04T18:19:31.031 device xr0: commit: ok (0.047 s)
leaving prepare phase (1.141 s)
entering commit phase
...
2024-02-04T18:19:31.035 device xr0: persist...
2024-02-04T18:19:31.035 device xr0: send NED persist
2024-02-04T18:19:31.036 device xr0: persist: ok (0.000 s)
2024-02-04T18:19:31.036 device xr0: push configuration: ok (0.223 s)
2024-02-04T18:19:31.036 device xr0: get-trans-id...
2024-02-04T18:19:31.036 device xr0: send NED get-trans-id
...
2024-02-04T18:19:31.169 device xr0: package cisco-iosxr-cli-7.49: reading config: ok (0.132 s)
2024-02-04T18:19:31.170 device xr0: get-trans-id: ok (0.134 s)
2024-02-04T18:19:31.178 holding transaction lock: ok (1.292 s)
leaving commit phase (0.146 s)
applying transaction for running datastore usid=50 tid=1143 trace-id=c3f3cc96-2d2b-4d9e-b43d-c07567fc8802 (1.308 s)
Commit complete.
admin@ncs(config-config)#
```