Automating creation of Software Bills of Materials: Generating SPDX documents for CMake and Zephyr

Steve Winslow

steve@swinslow.net

github.com/swinslow/slides

ZEPHYR

RTOS for resource-constrained devices, >200 boards

- CMake to manage builds
- west as meta-tool for builds and other actions

https://zephyrproject.org

Zephyr™

GOAL

Create a Zephyr bill-of-materials at build time

Express metadata about:

- file hashes
- licenses
- relationships

for sources as well as build artifacts

OBJECTIVES

make it fully automated

don't leverage external knowledge sources

don't rewrite existing build systems

(also don't know much about them)

make it Zephyr-agnostic

PROOF OF CONCEPT

cmake-spdx

github.com/swinslow/cmake-spdx/

SPDX

An open standard for communicating software bill of material information, including:

- components
- licenses
- copyrights
- security references

https://spdx.dev

SPDX

CMAKE FILE-BASED API

1. Create query directory
2. Add zero-byte query file
3. Run CMake
4. CMake outputs build metadata

https://cmake.org/cmake/help/latest/manual/cmake-file-api.7.html

CMAKE-SPDX PROCESS (1/2)

1. Enable CMake file-based API
2. Run CMake
3. Run Zephyr build
4. Walk through sources and create SPDX document
    - collect file hashes
    - look for `SPDX-License-Identifier`

CMAKE-SPDX PROCESS (2/2)

5. Walk through build artifacts and create SPDX document
6. Parse CMake JSON responses
   - codemodel
   - projects, targets, directories
7. Add relationships to SPDX build document
   - `GENERATED_FROM`
   - `STATIC_LINK`

SPDX DOCUMENT: SOURCES

```
SPDXVersion: SPDX-2.2
DataLicense: CC0-1.0
SPDXID: SPDXRef-DOCUMENT
DocumentName: sources
DocumentNamespace: http://spdx.org/spdxdocs/zephyr-ba8741fd-214a-42bf-bfd9-8c39c67292b6/sources
Creator: Tool: cmake-spdx
Created: 2021-01-07T14:06:12Z

##### Package: my-blinky sources

PackageName: my-blinky sources
SPDXID: SPDXRef-my-blinky
PackageDownloadLocation: NOASSERTION
FilesAnalyzed: true
PackageVerificationCode: 8dbb6f496c10462b5d92672a95a1858fe96d0323
PackageLicenseConcluded: Apache-2.0
PackageLicenseInfoFromFiles: Apache-2.0
PackageLicenseDeclared: NOASSERTION
PackageCopyrightText: NOASSERTION

Relationship: SPDXRef-DOCUMENT DESCRIBES SPDXRef-my-blinky

FileName: ./CMakeLists.txt
SPDXID: SPDXRef-File-CMakeLists.txt
FileChecksum: SHA1: 4d7260c43039d5c3e9caa69105334655b7bf40c6
FileChecksum: SHA256: 692898e82cc4c4b30c266cdfd97ad1b3977f0d6f9b8746e630d2af36ed18f204
LicenseConcluded: Apache-2.0
LicenseInfoInFile: Apache-2.0
FileCopyrightText: NOASSERTION

FileName: ./build/.ninja_deps
SPDXID: SPDXRef-File-.ninja-deps
FileChecksum: SHA1: d7cd689b284b8271fea0340a9a7d0fab8770e7ac
FileChecksum: SHA256: dfd6e041109dbb38ca17abc709dc3a8bf836f4393ad835e2e62c7cd29bee5074
LicenseConcluded: NOASSERTION
LicenseInfoInFile: NONE
```

SPDX DOCUMENT: BUILD ARTIFACTS

```
SPDXVersion: SPDX-2.2
DataLicense: CC0-1.0
SPDXID: SPDXRef-DOCUMENT
DocumentName: build
DocumentNamespace: http://spdx.org/spdxdocs/zephyr-ba8741fd-214a-42bf-bfd9-8c39c67292b6/build
Creator: Tool: cmake-spdx
Created: 2021-01-07T14:06:12Z
ExternalDocumentRef: DocumentRef-sources http://spdx.org/spdxdocs/zephyr-ba8741fd-214a-42bf-bfd9-8c3
SHA256:09d110a7ff29405dee4f858c81c1f0b098545824fbb540f953bc7b0e52bb3897

##### Package: build

PackageName: build
SPDXID: SPDXRef-build
PackageDownloadLocation: NOASSERTION
FilesAnalyzed: true
PackageVerificationCode: a55f7692969f9a5c9fd693288feba610da88302a
PackageLicenseConcluded: Apache-2.0
PackageLicenseInfoFromFiles: Apache-2.0
PackageLicenseDeclared: NOASSERTION
PackageCopyrightText: NOASSERTION

Relationship: SPDXRef-DOCUMENT DESCRIBES SPDXRef-build

FileName: ./.ninja_deps
SPDXID: SPDXRef-File-.ninja-deps
FileChecksum: SHA1: 8e4e616984c9ab0fd6e9c4df22e2f2d439bc21d2
FileChecksum: SHA256: 71748c43e71cc687e92ff625364ebd0a19b2475b4146c55bf987d2bc2ea913a3
LicenseConcluded: NOASSERTION
LicenseInfoInFile: NONE
FileCopyrightText: NOASSERTION
```

SPDX DOCUMENT: RELATIONSHIPS

```
Relationship: SPDXRef-File-libzephyr.a GENERATED_FROM DocumentRef-sources:SPDXRef-File-validate-enabled-instances.c
Relationship: SPDXRef-File-libzephyr.a GENERATED_FROM DocumentRef-sources:SPDXRef-File-power.c-10
Relationship: SPDXRef-File-libzephyr.a GENERATED_FROM DocumentRef-sources:SPDXRef-File-policy-residency.c
Relationship: SPDXRef-File-libzephyr.a GENERATED_FROM DocumentRef-sources:SPDXRef-File-uart-console.c
Relationship: SPDXRef-File-libzephyr.a GENERATED_FROM DocumentRef-sources:SPDXRef-File-clock-control-nrf.c
Relationship: SPDXRef-File-libzephyr.a GENERATED_FROM DocumentRef-sources:SPDXRef-File-sys-clock-init.c
Relationship: SPDXRef-File-libzephyr.a GENERATED_FROM DocumentRef-sources:SPDXRef-File-nrf-rtc-timer.c
Relationship: SPDXRef-File-zephyr.elf GENERATED_FROM DocumentRef-sources:SPDXRef-File-empty-file.c
Relationship: SPDXRef-File-zephyr.elf GENERATED_FROM SPDXRef-File-isr-tables.c
Relationship: SPDXRef-File-zephyr.elf STATIC_LINK SPDXRef-File-libapp.a
Relationship: SPDXRef-File-zephyr.elf STATIC_LINK SPDXRef-File-libzephyr.a
Relationship: SPDXRef-File-zephyr.elf STATIC_LINK SPDXRef-File-libisr-tables.a
Relationship: SPDXRef-File-zephyr.elf STATIC_LINK SPDXRef-File-libarch--common.a
Relationship: SPDXRef-File-zephyr.elf STATIC_LINK SPDXRef-File-libarch--arm--core--aarch32.a
Relationship: SPDXRef-File-zephyr.elf STATIC_LINK SPDXRef-File-libarch--arm--core--aarch32--cortex-m.a
Relationship: SPDXRef-File-zephyr.elf STATIC_LINK SPDXRef-File-libarch--arm--core--aarch32--cortex-m--mpu.a
Relationship: SPDXRef-File-zephyr.elf STATIC_LINK SPDXRef-File-liblib--libc--minimal.a
```

INTERESTING FINDINGS (1/2)

- CMake API data on build artifacts is incomplete
- Found some invalid license IDs in Zephyr code
- Graphviz output to visualize CMake target relationships

INTERESTING FINDINGS (2/2)

NEXT STEPS

Contributing to Zephyr, under review:

zephyrproject-rtos/zephyr: #31065

Auto-conclude binary licenses from source IDs

Cleanup handling directories, license IDs

next-steps.md

TAKEAWAYS

Software bills of materials can be created:

- automatically
- at build time
- without rewriting build systems (sometimes)
- without external knowledge sources
- in a standardized SBOM format

Start small, improve over time

RESOURCES

- cmake-spdx:
  - example CMake file-based API replies
  - running cmake-spdx
  - cmake-spdx internals
- SPDX:
  - specification dev repo
  - past specifications
  - short-form license IDs

THANK YOU

slides: github.com/swinslow/slides