基于SBOM的开源社区软件供应链安全解决方案

目录：

开源软件行业洞察分析：使用无处不在、面临诸多不可控风险

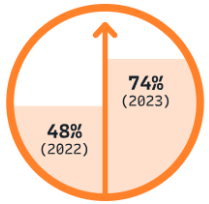
OpenEuler

17个行业1067个代码库：
96%的代码库中包含开源代码，77%的代码库直接使用开源代码

96%
of the total codebases contained open source

77%
of all code in the total codebases originated from open source

高危漏洞呈上升趋势



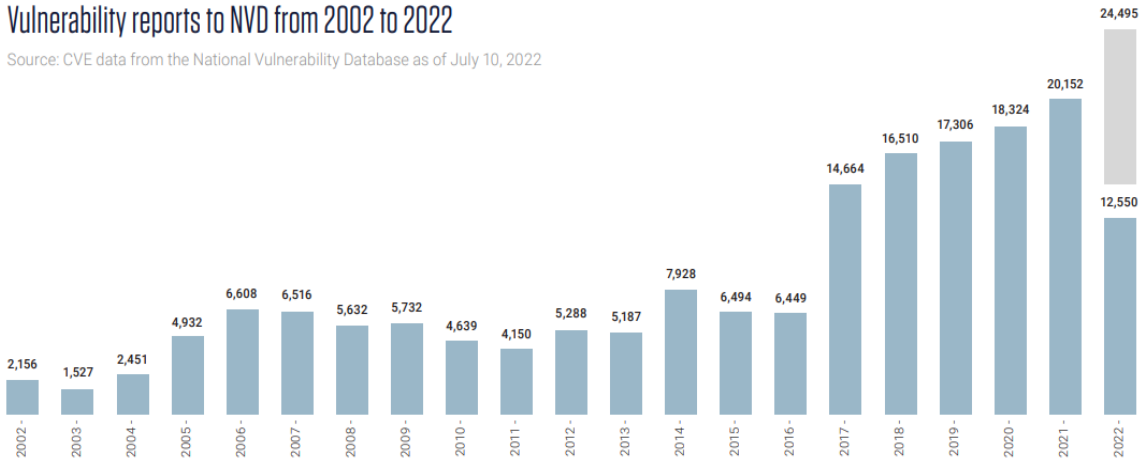
54%
increase in codebases containing high-risk vulnerabilities in the past year

49%开源软件未使用新版本

49%
of codebases that had risk assessments contained open source that had no new development in the last two years

《 2024 Open Source Security and Risk Analysis Report 》

NVD漏洞越来越多、CVE每月呈增长趋势



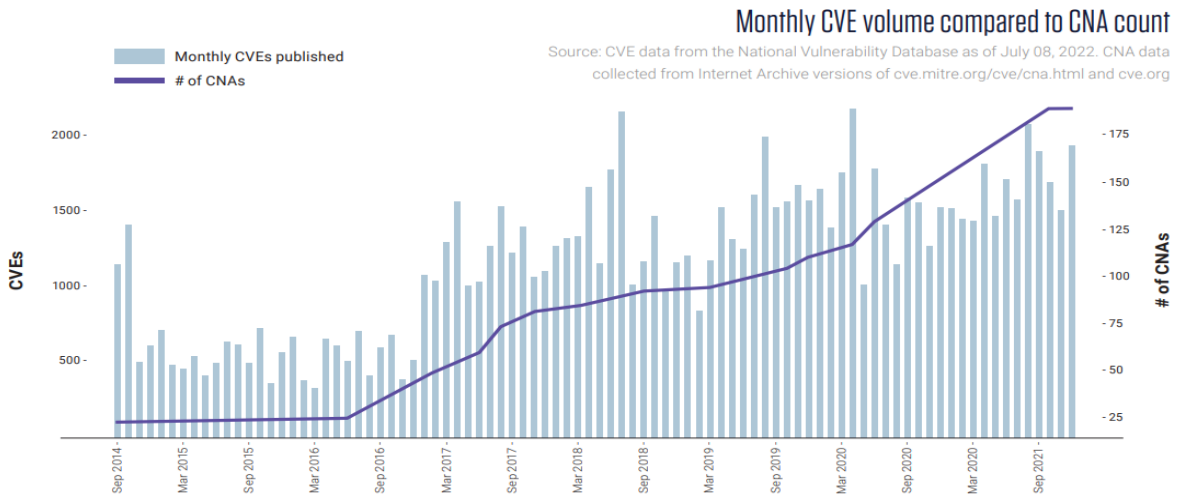


Figure 2. The number of monthly CVEs has grown with the number of CVE Number Authorities (CNAs).

开源软件供应链重大安全事件：受影响面广、让人措手不及

OpenEuler

重大漏洞	时间线	描述
Ripple 20	2019年	物联网安全风暴、JSOF安全研究人员在Treck TCP/IP软件库中发现19个0day安全漏洞，影响**70多家厂商**的全球**十亿台联网设备**
Amnesia 33	2020年	Forescout在4个开源TCP/IP 协议栈中发现33个0day安全漏洞，影响超过**150家厂商的数百万智能和工业产品和设备**
Apache Log4j2	2021年	被曝存在严重高危的远程代码执行漏洞，攻击难度低、影响人群广泛，被称为**“史诗级”高危漏洞;** 披露一个月多时间，CVE-2021-44228漏洞为起始点，Apache Log4j 总计爆发8个漏洞（7个高危；软件在官方发布漏洞修复补丁后依旧被黑客多次绕过 影响全球**6万+开源软件，70%以上企业业务系统、几乎使用Java所有互联网公司**
开源软件“围剿”俄罗斯	2022年	**30+开源项目加入了对俄罗斯的抵制**，开源软件的伦理道德和安全性正在遭遇前所未有的挑战其中甚至包括亚马逊（AWS Terraform modules）和Oracle等科技巨头的项目，也不乏MongoDB、pnpm、es5-ext、Drupal、RedisDesktopManager等流行项目， 例如：**Vue CLI的依赖项node-ipc包以反战为名进行供应链投毒**，而被投毒的 node-ipc 包在 npm 上**每周下载量超百万+**
xz maintainer投毒攻击	2024年	xz是用于压缩/解压缩文件的一款压缩库，微软的安全研究员Andres Freund调查程序性能下降时发现了该漏洞，xz漏洞从5.6.0版本开始存在**恶意后门**，该后门存在于XZ Utils的5.6.0和5.6.1版本中，攻击者可能利用这一漏洞在受影响的系统上**绕过SSH的认证获得未授权访问权限，执行任意代码**。

OpenEuler

灵魂拷问：

1、如何判断我是否受影响?

在大规模软件产品中要分清楚 **我依赖了谁？ 谁依赖了我**？ 建立软件的正反向依赖关系全链路可追溯

2、如何检测和修复软件供应链攻击?

前提是需要建立现代化的**DevSecOps软件工程体系**、从依赖分析→License分析→漏洞分析等均需要从人工排查到工程自动化

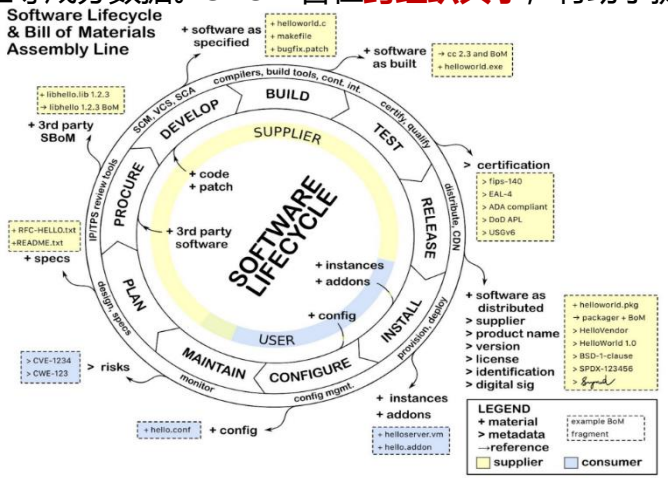
目录：

SBOM简介：致力于软件安全供应链透明的数据底座、跨组织共享

OpenEuler

SBOM是什么？

SBOM是一种正式标准化的、机器可读的元数据，它唯一地标识软件组件及其内容；也可能包括版权和许可证等成分数据。SBOM旨在**跨组织共享**，有助于提供软件供应链成分清单与透明度



SBOM数据标准: SPDX、CycloneDX、SWID

SBOM标准	SPDX	SWID	CycloneDX
组织	Linux Fundation	ISO & IEC	OWASP
标准化	ISO/IEC 5962: Under Development	ISO/IEC 19770-2	No plan
当前版本	2.2.1	2	1.3
软件标识	Package SPDX ID, Hash	SWID	GAV, PURL, CPE, SWID, Hash
内容	Document Creation Information Package Information File Information Snippet Information Other Licensing Information Relationships Annotations	SWID Entity Payload Link rel	Meta Components Pidgrees Dependencies Compositions Vulnerabilities Signatures
工具	SPDX tools: 20	SWID tools: 12	CycloneDX Tools: 67
主要应用场景	合规场景；漏洞追溯等	软件标识：安装/发现/移除/补丁	场景分析：漏洞追溯/合规/等18种场景
转换工具	SwiftBOM, DecoderRing		

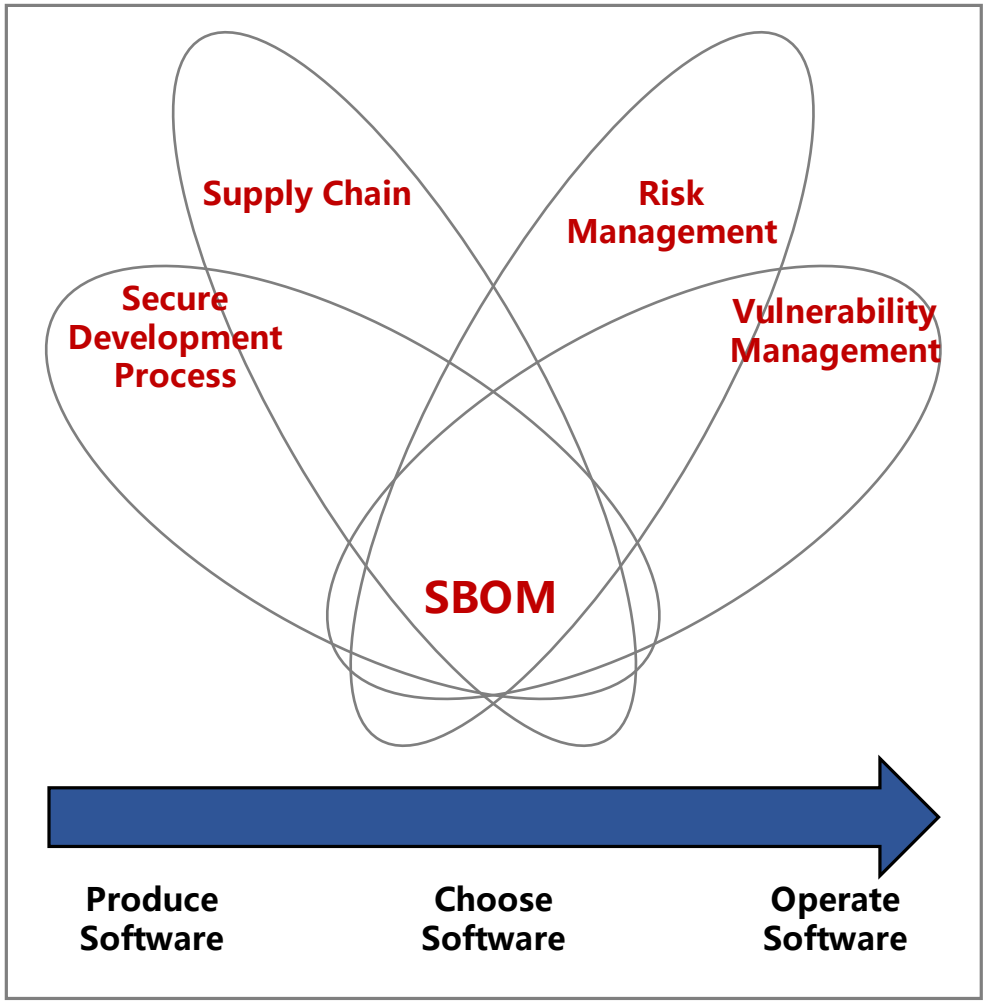
- **标准组织**
  - 主要推动组织：**NTIA**，2018/6/6起，启动软件组件透明度计划Software Component Transparency
    - 目标：定义SBOM（软件物料清单）基础格式达成行业共识
    - 四个工作组：框架组，实践组，格式与工具组，医疗健康POC
  - 主要协议组织
    - **Linux Fundation**: SPDX
    - **ISO/IEC**: SWID 19770-2
    - **OWASP**: CycloneDX
    - **CISQ** & **OMG**: 3T-SBOM
  - 相关标准
    - **OpenChain** ISO 5230：软件供应链管理需要SBOM

- **公司行业**
  - NTIA 2021年6月征集SBOM意见，包括协议，应用场景等。包括**Google**，**微软**，**爱立信**在内的**88家**公司和政府部门提供了积极的反馈
  - **美国白宫2021年5月份发布第14028号行政令**，要求美国商务部在60天内协同美国国家通信和信息管理局 (NTIA) 发布软件物料清单 (SBOM) 的“最小元素”
  - 包括**Snyk**，**SonaType**，**Fossology**在内的众多安全软件已逐步开始支持生成SBOM或者使用SBOM作为漏洞分析的输入
  - 荷兰**NCSC**：在2021年1月发布《Using the SBOM for Enhancing Cybersecurity》推荐基于CycloneDX，Dependency Track等整体方案
  - **ENISA**：在2020/11/9发布《Guidelines For Securing The Internet Of Things》建议使用针对物联网设备的SBOM，并建议使用Dependency Track工具来管理底层软件依赖关系等

SBOM的消费场景与用途

场景：软件供应链安全管理，安全漏洞管理、安全应急响应，高可信安全应用管理

用途：能帮助软件生产商、购买者和运营商更高效地识别软件成分、排查License风险/合规风险/安全漏洞影响风险、履行义务声明等



Benefits	Produce	Choose	Operate
Cost	Less unplanned, unscheduled work	A more accurate total cost of ownership	More efficient administration
Security Risk	Avoid known vulnerabilities	Easier due diligence	Faster identification and resolution. Know if and where specific software is affected.
License Risk	Quantify and manage licenses and associated risk	Easier due diligence	More efficient, accurate response to license claims
Compliance Risk	Easier risk evaluation. Identify compliance requirements earlier in lifecycle	More accurate due diligence, catch issues earlier in lifecycle	Streamlined process
High Assurance	Make assertions about artifacts, sources, and processes used	Make informed, attack-resistant choices about components	Validate claims under changing and adversarial conditions

SBOM最小集字段定义：

美国国家电信和信息管理局(National Telecommunications and Information Administration)发布SBOM最小集的定义：

数据字段是关于必须捕获和维护每个组件的基础数据，以便在整个软件供应链中跟踪组件、并基于此扩展License和漏洞库等其他数据地段

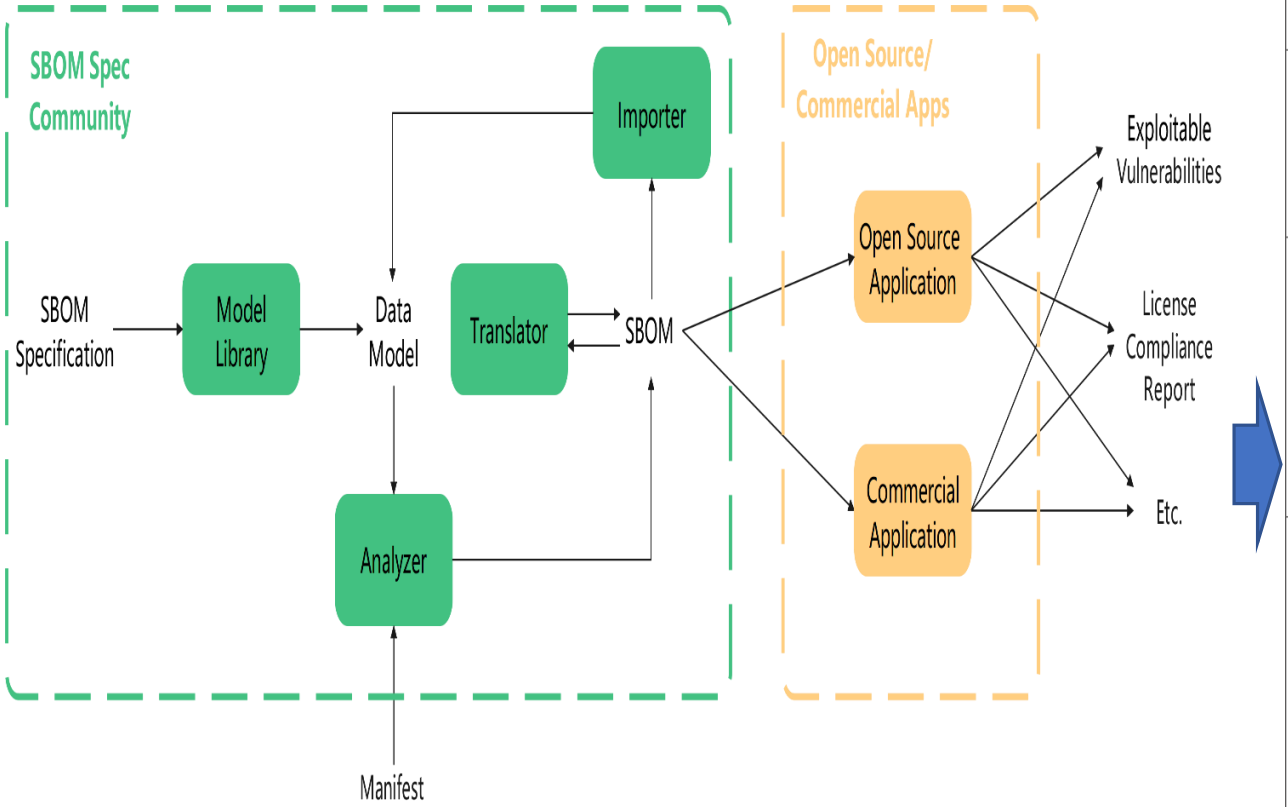
数据字段	描述
供应商名称	创建、定义和标识组件的实体的名称。
组件名称	分配给原始供应商定义的软件单元的名称。
组件的版本	供应商用来指定软件从先前标识的版本发生变化的标识符。
其他唯一标识符	用于标识组件或用作相关数据库的查找键的其他标识符。
依赖关系	表征上游组件 X 包含在软件 Y 中的关系。
SBOM 数据的作者	为此组件创建 SBOM 数据的实体的名称。
时间戳	记录 SBOM 数据组装的日期和时间。
推荐的数据	
组件哈希	组件的唯一哈希，以帮助允许列表或拒绝列表。
生命周期阶段	SDLC 中捕获 SBOM 数据的获取的阶段。
.....	

参考： https://www.ntia.doc.gov/blog/2021/ntia-releases-minimum-elements-software-bill-materials

SBOM技术生态：不仅仅是数据标准、还需开源&商业生态链配套

OpenEuler

- **SBOM标准的设计是起点**：标准的推广离不开一个活跃的**生态**。SBOM的一大特征是**机器可读**，因此，SBOM标准社区需要提供完善的基础设施，以供开发者和厂商方便地根据SBOM标准完成SBOM的**生成**、**消费**、**转换**
- **SBOM标准生态离不开实际的高阶应用**：（例如漏洞感知、license合规）支持，没有高阶应用等消费场景支持的SBOM标准只不过是空中楼阁



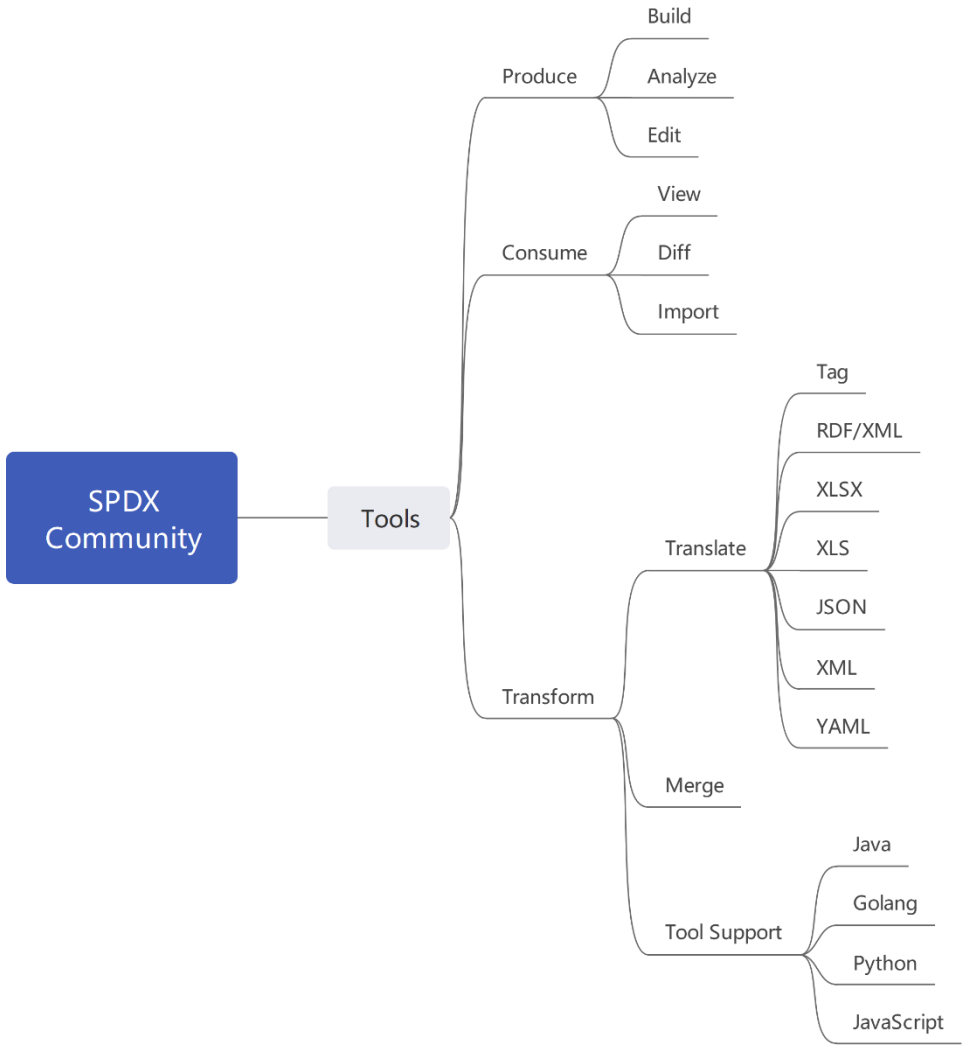
SBOM工具分类

分类	子类	描述	以SPDX为例
生成	构建	构建软件制品时，SBOM随之自动生成，并包含构建信息	https://github.com/spdx/spdx-maven-plugin https://github.com/spdx/spdx-build-tool
	分析	分析源码或二进制文件等以生成SBOM	https://github.com/spdx/spdx-sbom-generator
	编辑	协助人工编辑SBOM数据	https://github.com/spdx/spdx-online-tools
消费	查看	能够理解以人类可读的形式（图、表、文字等）展现的内容。用于支持决策和业务流程	https://github.com/spdx/tools-java https://github.com/spdx/spdx-online-tools
	比较	能够比较多个SBOM并清楚地找出差异（例如，比较同一软件的两个不同版本）	https://github.com/spdx/tools-java https://github.com/spdx/spdx-online-tools
	导入	能够发掘、检索以及导入SBOM，以便进一步分析与处理	https://github.com/spdx/tools-java https://github.com/spdx/tools-golang https://github.com/spdx/tools-python https://github.com/spdx/spdx-tools-js https://github.com/spdx/spdx-online-tools
变换	转换	在保有原始内容的前提下，将SBOM从一种格式转换为另一种格式	https://github.com/spdx/tools-java https://github.com/spdx/tools-golang https://github.com/spdx/tools-python https://github.com/spdx/spdx-tools-js https://github.com/spdx/spdx-online-tools https://github.com/spdx/spdx-to-osv
	合并	合并多个SBOM及其他数据	https://github.com/spdx/tools-java
	工具支持	通过API、对象模型、库或其他引用源来支持工具使用SBOM	https://github.com/spdx/spdx-3-model https://github.com/spdx/Spdx-Java-Library https://github.com/spdx/tools-java https://github.com/spdx/tools-golang https://github.com/spdx/tools-python https://github.com/spdx/spdx-tools-js https://github.com/spdx/spdx-online-tools

SBOM-->SPDX生态洞察：

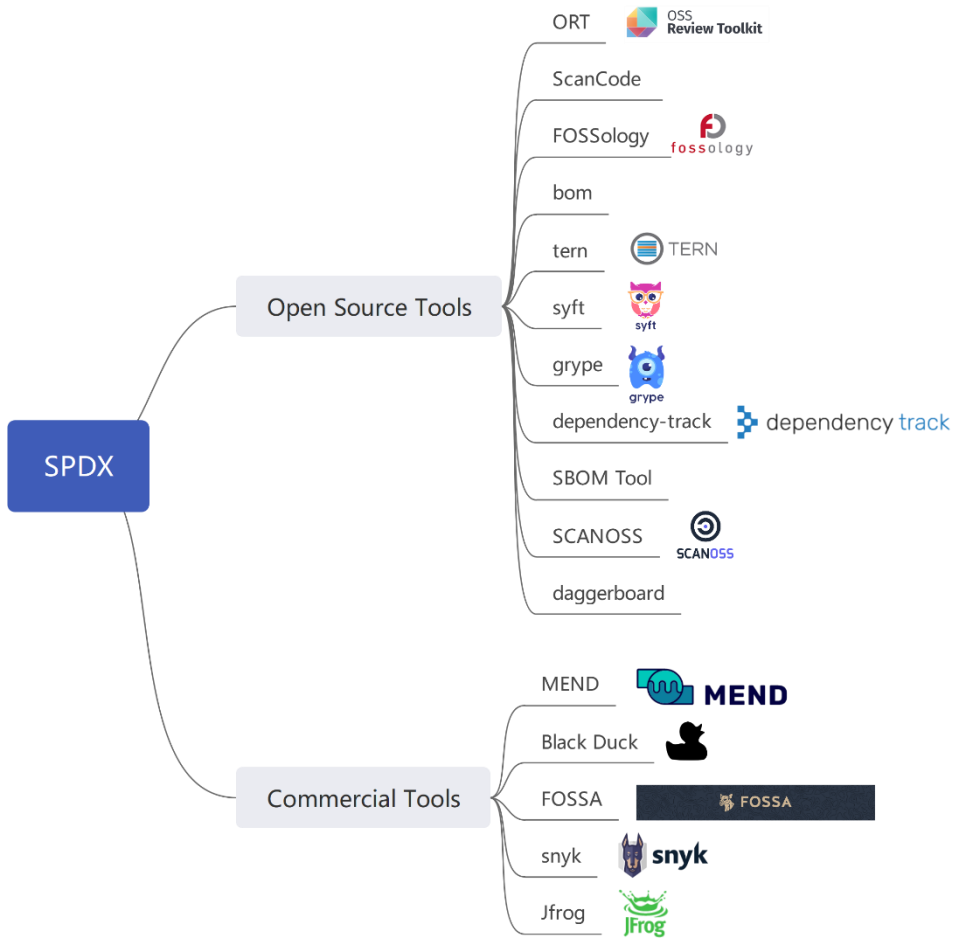
SPDX社区提供了多语言类库和工具：

能够完成SBOM生成、消费、转换等任务、语言类库当前支持Java、Golang、Python、JavaScript等主流编程语言



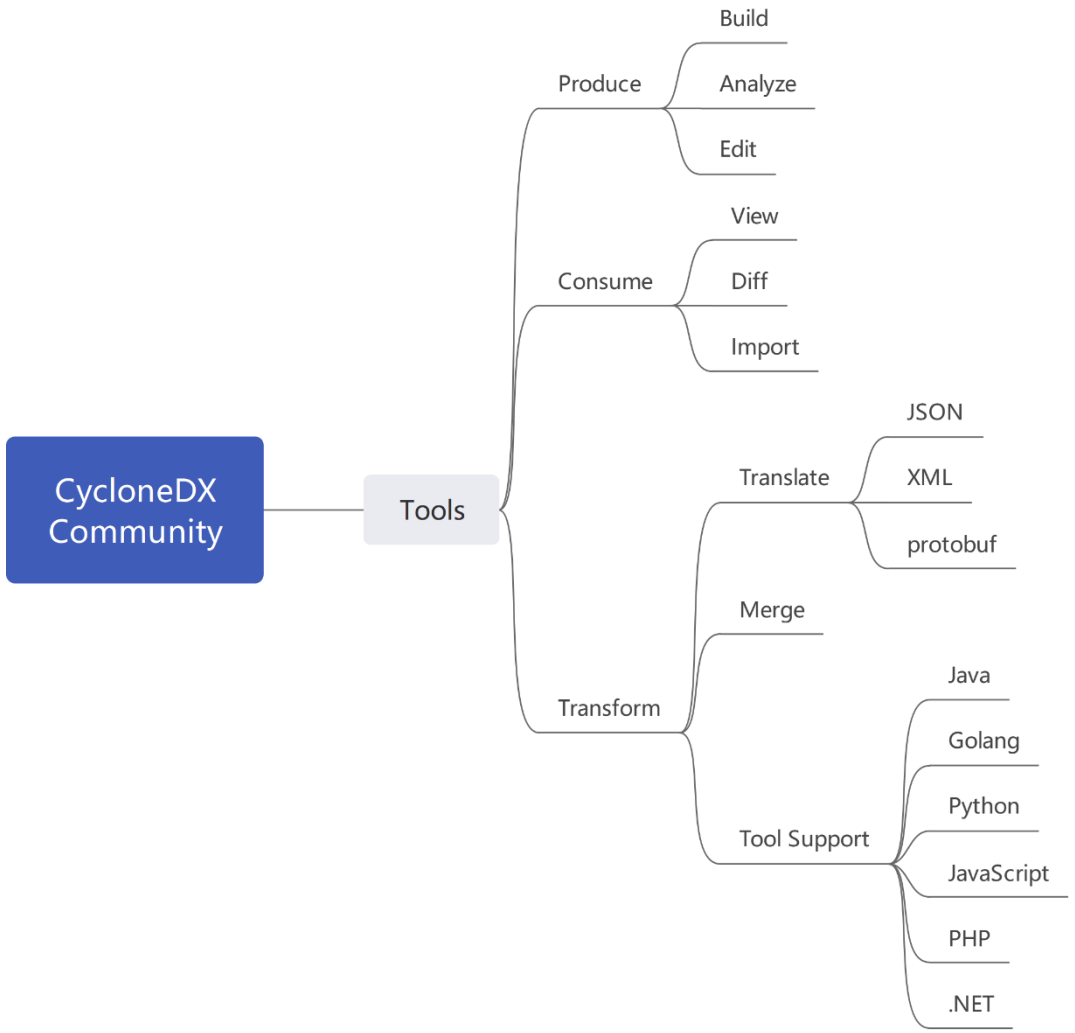
众多主流开源/商业工具及解决方案：

支持生成、消费SPDX SBOM，license合规、漏洞管理等



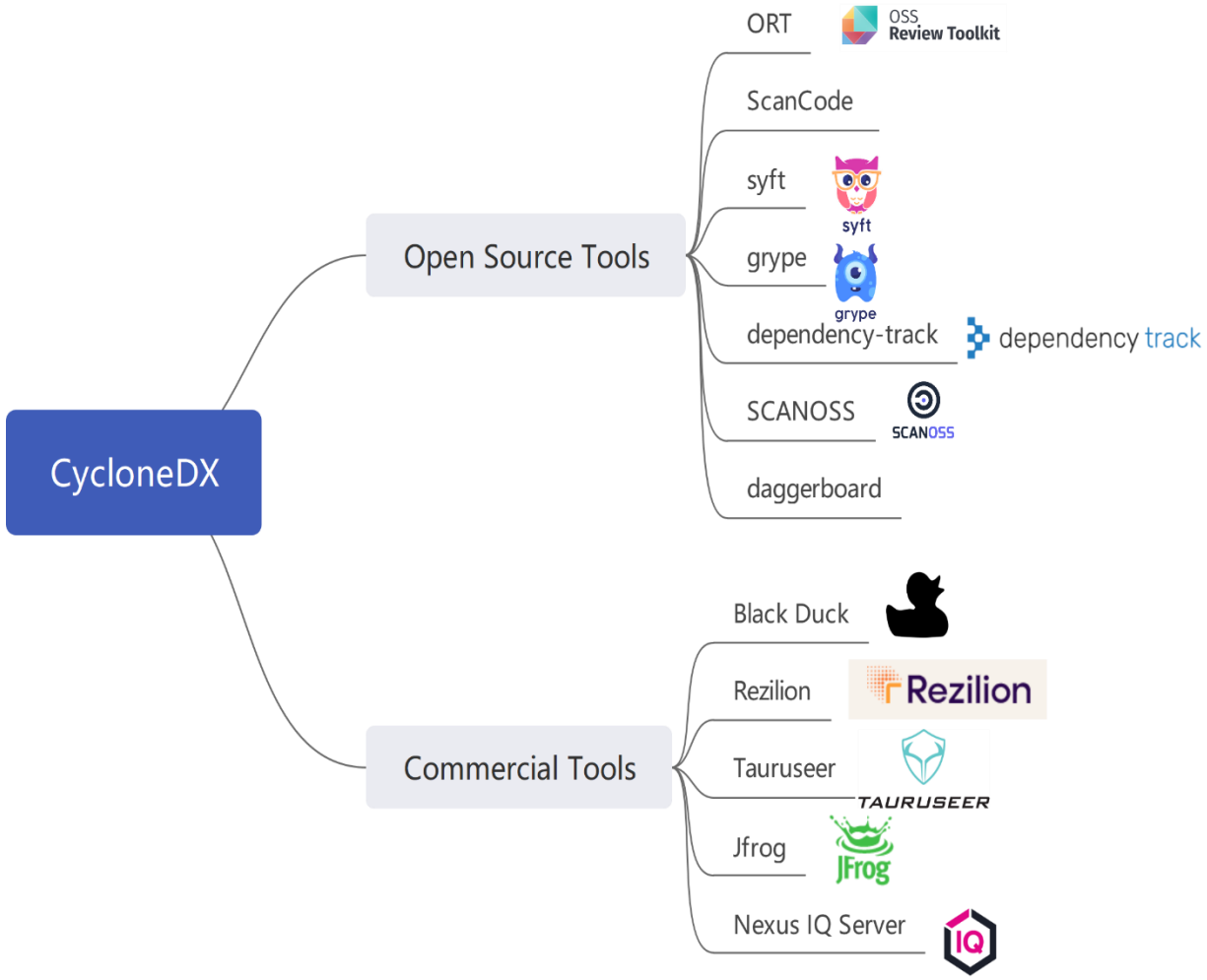
SBOM--> CycloneDX生态洞察：

CycloneDX社区同样提供了多语言类库和工具：
能够完成SBOM生成、消费、转换等任务、语言类库当前支持Java、Golang、Python、JavaScript、PHP、.NET等主流编程语言



众多主流开源/商业工具及解决方案：

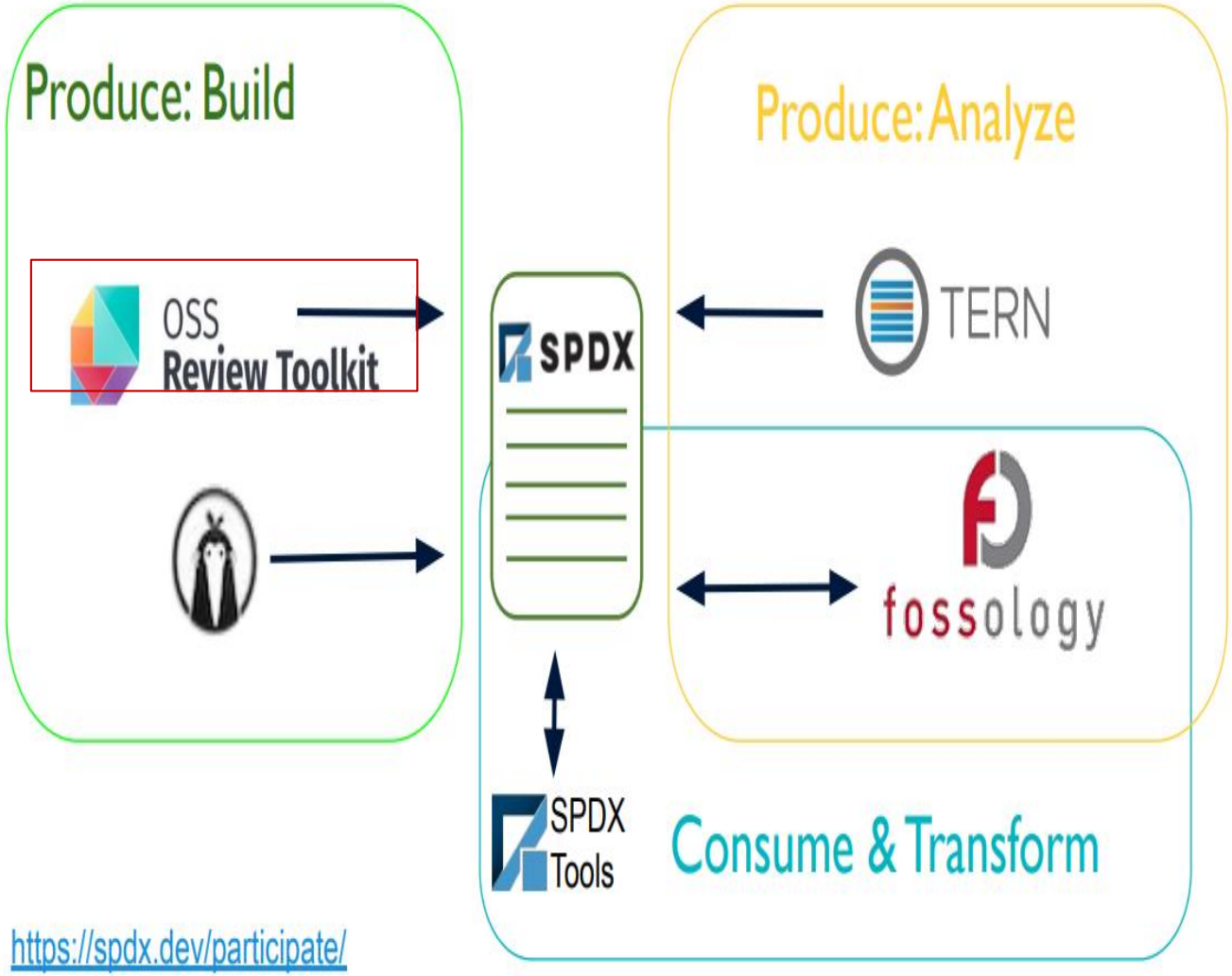
支持生成、消费SPDX SBOM，license合规、漏洞管理等



洞察Linux基金会：基于SBOM开源软件供应链安全解决方案

OpenEuler

ORT(oss-review-toolkit)： 基于SBOM提供完整解决方案而且对主流第三方商业工具，开源工具的集成：VulnerableCode，ScanCode，SCANOSS ，Fossid， Sonatype
工具参考： https://github.com/oss-review-toolkit/ort



- *Analyzer* - determines the dependencies of projects and their metadata, abstracting which package managers or build systems are actually being used.

- *Downloader* - fetches all source code of the projects and their dependencies, abstracting which Version Control System (VCS) or other means are used to retrieve the source code.

- *Scanner* - uses configured source code scanners to detect license / copyright findings, abstracting the type of scanner.

- *Advisor* - retrieves security advisories for used dependencies from configured vulnerability data services.

- *Evaluator* - evaluates license / copyright findings against customizable policy rules and license classifications.

- *Reporter* - presents results in various formats such as visual reports, Open Source notices or Bill-Of-Materials (BOMs) to easily identify dependencies, licenses, copyrights or policy rule violations.

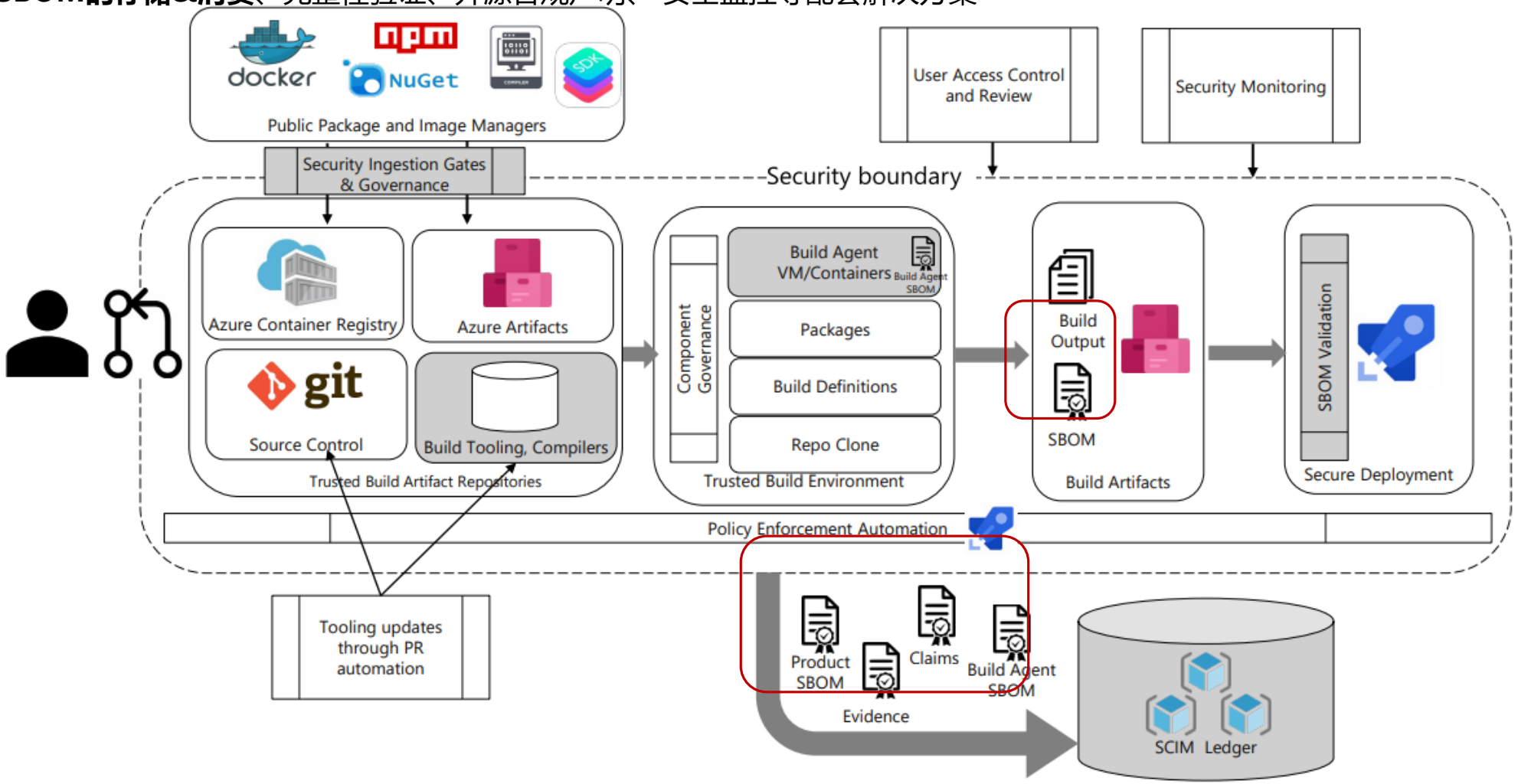
- *Notifier* - sends result notifications via different channels (like emails and / or JIRA tickets).

洞察微软：基于SBOM开源软件供应链安全解决方案

1、**SBOM Generator：**基于SPDX、跨平台、全场景语言、包管理、操作系统的支持（NPM, NuGet, PyPI, CocoaPods, Maven, Golang, Rust Crates, RubyGems, containers (and their Linux packages), Gradle, Ivy, GitHub public repositories)

工具参考：https://github.com/microsoft/sbom-tool

2、基于SBOM的存储&消费、完整性验证、开源合规声明、 安全监控等配套解决方案



目录：

openEuler开源社区SBOM解决方案应用架构全景图

OpenEuler

1、作业层：围绕社区开发者作业流、任何发布二进制能自动生成SBOM、自动提交Issue

2、服务层：提供原子化服务解耦如SBOM生成&格式转化工具套件、License服务、漏洞排查服务、漏洞感知、开源片段引用扫描服务等

3、数据层：提供SBOM数据库存储、核心License数据库、漏洞数据库、开源片段数据库等数据资产

社区作业

openEuler
开发者社区提交PR→PR门禁检查

CI/CD发布流水线

Release与维护

服务层

软件成分分析SCA（SBOM化）：应用市场(Marketplace)

社区SBOM应用服务

SBOM生成工具

License合规分析应用

SBOM各种格式导出服务

软件正反向依赖链路

基于SBOM漏洞分析应用

基于SBOM的开源声明及义务履行

数据层

SBOM元数据库

制品仓库数据

元数据中心数据库

社区信息数据库（貂蝉）

License数据库
(ScanCode...etc)

软件信息画像数据库
(官网地址、软件成熟度等)

统一漏洞信息数据库（CVE-MGR）

公开开源漏洞库
(ossIndex,NVD,OSV, Vulnerablecode)

补充第三方商业漏洞库Vtopia

开源片段引用数据库

开源&第三方服务
SCANOSS/FossID.etc

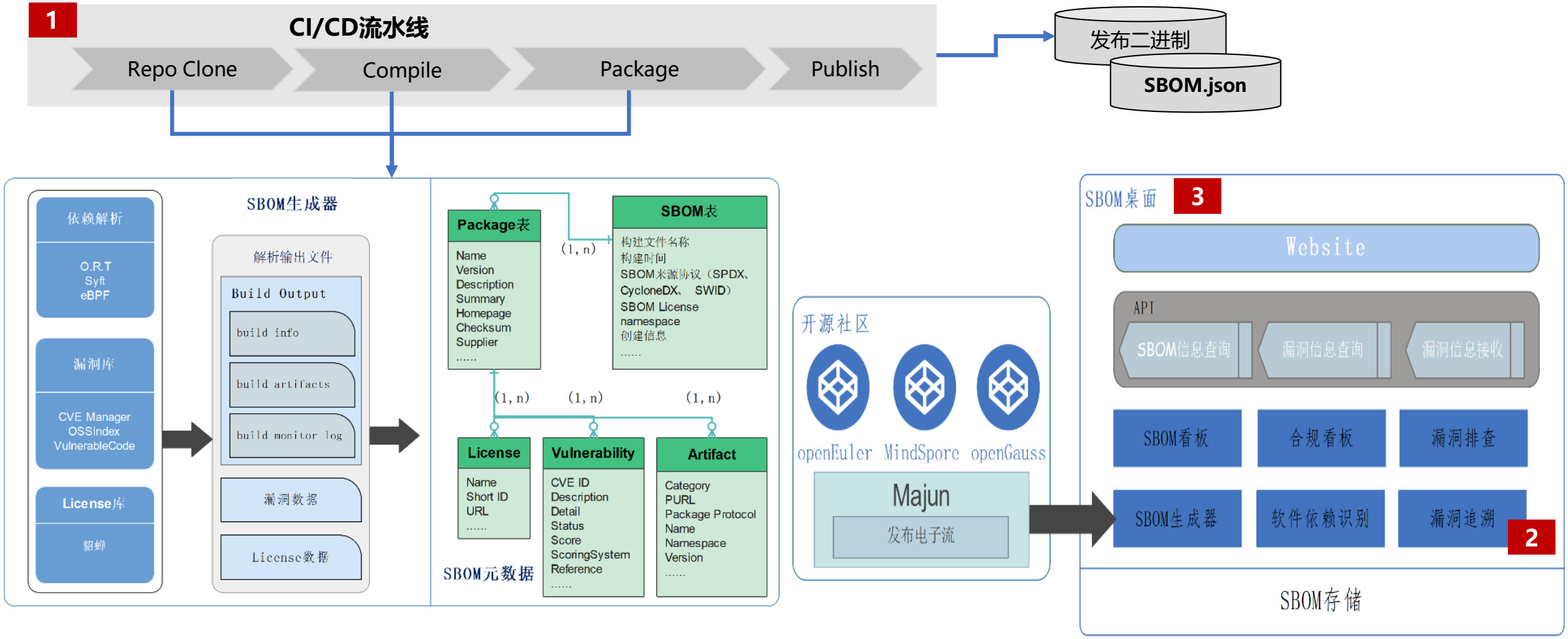
SBOM的数据流向图：生产、存储、消费

OpenEuler

1、数据生产&存储：SBOM基于CI/CD流水线自动生成、随制品一起归档存储、实现发布二进制与SBOM关联

2、数据消费： 基于SBOM的License合规、漏洞排除与感知

3、社区门户：依托于各社区门户网站对全量SBOM信息进行管理、并与源码仓交互、进行issue提交、漏洞修复、开源合规履行义务声明



SBOM数据组装与颗粒度：面向软件包的一包一SBOM、一包一PURL

OpenEuler

对外可见的发布包（一包一SBOM）			
Everything ISO Img	openEuler-22.03-LTS-everything-x86_64-dvd.iso	15.6 GiB	2022-Apr-01 08:41
Docker img	openEuler-docker.x86_64.tar.xz	44.7 MiB	2022-Apr-01 08:41
raspi_img	openEuler-22.03-LTS-raspi-aarch64.img.xz	234.8 MiB	2022-Apr-01 07:58
virtual_machine_img	openEuler-22.03-LTS-x86_64.qcow2.xz	385.9 MiB	2022-Apr-01 08:15
RPM Package	log4j12-1.2.17-25.oe2203.noarch.rpm	462.3 KiB	2022-Apr-01 07:53

SBOM依赖细分

- **Packages**：例如ISO里面包含有多少个RPM包
- **Modules**：单个RPM包有那些自己的内部组件。例如hive.rpm→hive-cli.jar
- **依赖**：单个RPM包里面打入了哪些传递性依赖，例如：hive.rpm→log4j-core.jar
- **运行时依赖**：RPM包安装部署依赖的RPM，例如：hive.rpm→mysql5-server.rpm

一包一PURL原则

遵从业界PURL specification、数据标准一样、数据格式一样，实现在不同SBOM标准、不同工具解决方案的软件包PURL的唯一性

- **包管理器颗粒度：** RPM,Conan,Maven,NPM,Cargo,Docker,Gem，例如pkg:maven/org.apache.hive.hcatalog/hcatalog-core@3.1.2
- **源码颗粒度**：github,github,gitee源码引入依赖场景，样例pkg:github/google/flatbuffers@2.0.0

openEuler SBOM披露字段(SPDX 2.2)

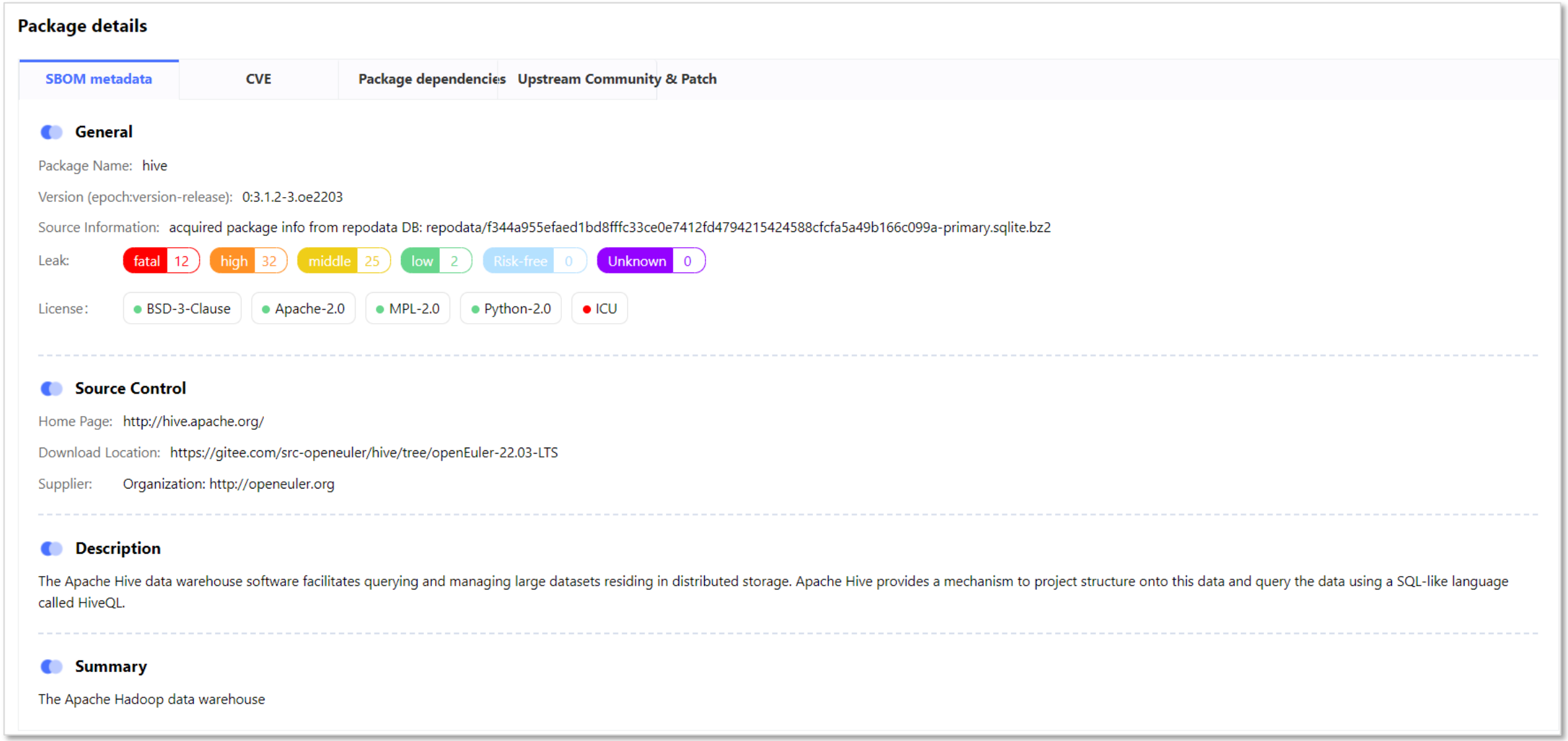
覆盖7个基本字段，且更全面地给出扩展license和copyright等字段的值；

openEuler				
	SPDXID			f3a4ac8e-33ec-45f3-b8d8-a6b83893a213
	spdxVersion			SPDX-2.2
	creationInfo			
SBOM时间戳		created		2023-07-31T08:45:17.8597553Z
SBOM数据作者		creators		Organization: Anchore, Inc", "Tool: syft-[not provided]
	licenseListVersion			3.17
	name			C-\\Users\\admin\\Desktop\\openEuler-22.03-LTS-SP2-everything-x86_64-dvd.iso
	dataLicense			CC0-1.0
	documentNamespace			https://anchore.com/syft/file/C-%5CUsers%5Cadmin%5CDesktop%5CopenEuler-22.03-LTS-SP2-everything-x86_64-dvd.iso-90832dc2-a415-4b21-803a-ac0c7795c91c
	packages			
		SPDXID		SPDXRef-rpm-texlive-thesis-titlepage-fhac-doc-svn15878.0.1
		checksums		
			algorithm	SHA256
			checksumValue	4ec690ac89af14168c07af825ddf81ad97760c21b19ff2bbeea8d284691d8065
		copyrightText		null
		description		Documentation for thesis-titlepage-fhac
		downloadLocation		https://gitee.com/src-openeuler/texlive-split-x/tree/openEuler-22.03-LTS-SP2
组件其他唯一标识		externalRefs		
			referenceCategory	PACKAGE_MANAGER
			referenceType	purl
			referenceLocator	pkg:rpm/texlive-thesis-titlepage-fhac-doc@svn15878.0.1-24.oe2203sp2?arch=noarch&epoch=8&upstream=texlive-split-x-2018-24.oe2203sp2.src.rpm
			referenceCategory	SOURCE_MANAGER
			referenceType	url
			referenceLocator	https://tug.org/svn/texlive
		filesAnalyzed		FALSE
		homepage		http://tug.org/texlive/
		licenseConcluded		Artistic 2.0 and GPLv2 and GPLv2+ and LGPLv2+ and LPPL-1.3a and LPPL-1.3c and MIT and Public Domain and UCD and Utopia
		licenseDeclared		Artistic 2.0 and GPLv2 and GPLv2+ and LGPLv2+ and LPPL and MIT and Public Domain and UCD and Utopia
组件名称		name		texlive-thesis-titlepage-fhac-doc
		sourceInfo		acquired package info from repodata DB: repodata/f29318f9b0cd6387921a68c63ae38f00ec6f4532797108b889ab6d84d0af18a2-primary.sqlite.bz2
		summary		Documentation for thesis-titlepage-fhac
组件供应商名称		supplier		Organization: http://openeuler.org
组件版本		versionInfo		8:svn15878.0.1-24.oe2203sp2
组件依赖关系	relationships			
		spdxElementId		SPDXRef-rpm-xorg-x11-xauth-help-1.1.2
		relationshipType		DEPENDS_ON
		relatedSpdxElement		SPDXRef-rpm-man-db-2.11.0

目录：

SBOM元数据信息

SBOM元数据信息包括：源码信息,、软件License、漏洞信息、依赖关系等.



软件包依赖信息全景

展示软件包、其自身组件、可传递依赖项和运行时依赖项的全景视图，实现前向和后向依赖项查询。

Package details

SBOM metadata | CVE | Package dependencies | Upstream Community & Patch

Package Info

请输入namespace或name，按enter键搜索

numbering	The package name	Package managen	name	version	qualifiers
1	hive	rpm	hive	3.1.2-3.oe22...	arch: noarch epoch: 0 upstream: hiv...

Module Info

请输入namespace或name，按enter键搜索

numbering	The package na	Package manag	namespace	name	version
1	hive	maven	org.apac...	hcatalog-core	3.1.2
2	hive	maven	org.apac...	hcatalog-pig-ada...	3.1.2
3	hive	maven	org.apac...	hcatalog-server-e...	3.1.2
4	hive	maven	org.apac...	Hive-Accumulo H...	3.1.2
5	hive	maven	org.apac...	hive-beeline	3.1.2
6	hive	maven	org.apac...	hive-classification	3.1.2

Dependency

请输入namespace或name，按enter键搜索

numbering	The package name	Package manager	namespace	name	version
1	hive	maven	org.apach...	Accumulation-core	1.7.3
2	hive	maven	org.apach...	Accumulation-Fate	1.7.3
3	hive	maven	org.apach...	Accumulation-Start	1.7.3

Runtime Dependency

请输入namespace或name，按enter键搜索

numberi	The package nam	Package mana	name	version	qualifiers
1	java-11-ope...	rpm	java-11-openjdk-h...	11.0.13...	arch: x86_64 epoch: 1 upstream: ja...
2	mysql5-server	rpm	mysql5-server	5.7.34-...	arch: x86_64 epoch: 0 upstream: m...

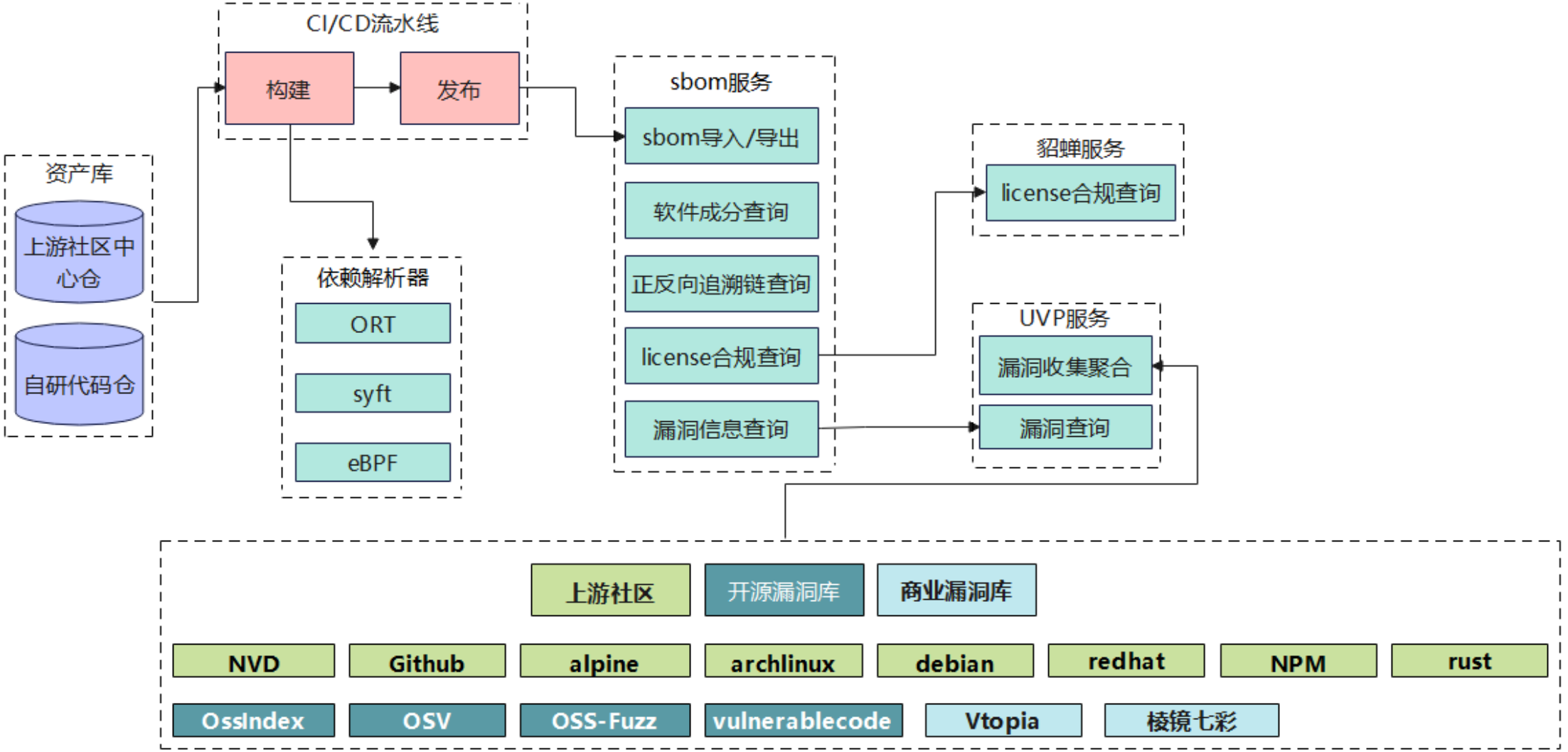
漏洞影响范围跟踪

OpenEuler

通过反向依赖信息可视化漏洞影响的完整路径。

SBOM工具及应用服务介绍

- openEuler提供能力导入，包括依赖解析器、sbom服务、UVP服务等组件，OSV基于这些工具建立自己的SBOM基础设施；
- OSV伙伴披露的SBOM数据字段建议先和openEuler保持一致，应用于漏洞分析服务、合规检测服务等场景；




OpenEuler