Cybercrime To Cost The World $10.5 Trillion Annually By 2025
https://cybersecurityventures.com/cybercrime-damages-6-trillion-by-2021/

Investing now can save millions

USD 4.45  51% million

The global average cost of a data breach in 2023 was USD 4.45 million, a 15% increase over 3 years.

51% of organizations are planning to increase security investments as a result of a breach, including incident response (IR) planning and testing, employee training, and threat detection and response tools.

https://www.ibm.com/reports/data-breach

Microsoft: 70 percent of all security bugs are memory safety issues
The Cost of Poor Software Quality in the US: A 2022 Report

From Problem to Solutions

HERB KRASNER
MEMBER, ADVISORY BOARD
CONSORTIUM FOR INFORMATION & SOFTWARE QUALITY (CISQ)
WWW.IT-CISQ.ORG
HKRASNER@UTEXAS.EDU
DATE: DECEMBER 15, 2022

https://www.synopsys.com/software-integrity/resources/analyst-reports/cost-poor-quality-software.html
Integrity is an advanced concept, but it’s the answer to the most serious, costly problems that software developers face.
PART I

How things have changed
Java’s Backward Compatibility

- A remarkable success considering age, size of ecosystem, depth of dependency graphs
- Achieved through the Java SE Specification, but also applies to supported JDK APIs
- Standard APIs are only removed by a deprecation process spanning multiple releases (or by Maintenance Reviews)
  - Even then, only done for APIs that are not widely used or have good alternatives
- Platform components may be restricted in a gradual process spanning multiple releases and involving warnings
- Changes reviewed through the CSR process
Java’s Backward Compatibility

- Not just a principle, but **one of Java’s greatest strengths!**
- When companies invest in expensive software development, they want to preserve their investment:
  - Existing code continues working
  - Platform evolves to offer better performance and new functionality as requirements and environments change
... At Least in Theory

- In Java’s first decade, things were only added, rarely removed — started small
- Then Java experienced some years of slow evolution
Using internals

- Not many new APIs were added and some bugs remained unfixed
- Ecosystem reached for JDK internals
  - New functionality
  - Work around bugs
  - Improve performance
Using internals

- Not many new APIs were added and some bugs remained unfixed
- Ecosystem reached for JDK internals
  - New functionality
  - Work around bugs
  - Improve performance
- Still, it was technical debt

In general, writing java programs that rely on sun.* is risky: they are not portable, and the APIs are not supported.
Using internals

- Not many new APIs were added and some bugs remained unfixed
- Ecosystem reached for JDK internals
  - New functionality
  - Work around bugs
  - Improve performance
- Still, it was technical debt
- \textit{Effectively} backward compatible because internals didn’t change much
Using internals

- Not many new APIs were added and some bugs remained unfixed
- Ecosystem reached for JDK internals
  - New functionality
  - Work around bugs
  - Improve performance
- Still, it was technical debt
- *Effectively* backward compatible because internals didn’t change much
What Happened in JDK 9?

- Modules restricted access to internals breaking lots of libraries
- `sun.misc.Unsafe` was removed, breaking more libraries
Nah

• Modules restricted access to internals breaking lots of libraries
  • Modules’ strong encapsulation of internals wasn’t turned on until JDK 16. All runtime access to internals remained as it was in JDK 8 until then
• `sun.misc.Unsafe` was removed, breaking more libraries
  • `sun.misc.Unsafe` is still here, exactly as accessible as ever
What Really Happened?

Java picked up its pace

JDK 9

The goal of this Project was to produce an open-source reference implementation of the Java SE 9 Platform as defined by JEP 379 in the Java Community Process.

JDK 9 reached General Availability on 21 September 2017. Production-ready binaries under the GPL are available from Oracle; binaries from other vendors will follow shortly.

The features and schedule of this release were proposed and tracked via the JEP Process, as amended by the JEP 2.0 proposal.

Features

102: Process API Updates
110: JHTTP 2 Client
143: Improve Contended Locking
158: Uniform JVM Logging
165: Compiler Control
193: Variable Handles
197: Segmented Code Cache
199: Smart Java Compilation, Phase Two
200: The Modulier JDK
201: Modular Source Code
211: Elide Deprecation Warnings on Import Statements
212: Resolve Lin and Dclлинt Warnings
213: Hiding Project C intricate
214: Remove GC Combinations Deprecated in JDK 8
215: Tiered Compilation for Java
216: Move the Import Statements Correctly
217: Annotations Pipeline 2.0
219: Datagram Transport Layer Security (DTLS)
220: Modular Run-Time Images
221: Simplified Socket API
222: Jshell: The Java Shell (Read-Eval-Print Loop)
223: New Version-String Scheme
224: HTTLEJ javac
225: Javadoc Search
226: UTF-8 Property Files
227: Unicode 7.0
228: Add More Diagnostic Commands
229: Create PKCS12 KeyStores by Default
231: Remove Launch-Time JRE Version Selection
232: Improve Secure Application Performance
233: Generate Run-Time Compiler Tests Automatically
235: Test Class-File Attributes Generated by javac
236: Panfix API for Nashorn
237: Linux/Mach64 Port
238: Multi-Release JAR Files
240: Remove the JMI 1.1 protocol
241: Remove the Jaxp Tool
243: Java-Level JVM Compiler Interface
244: TLS Application-Layer Protocol Negotiation Extension
245: Validate JVM Command-Line Flag Arguments
246: Leverage CPU Instructions for GHASH and RSA
247: Compile for Older Platform Versions
248: Migrate G1 the Default Garbage Collector
249: GCSIP Staging for TLS
250: Store Internal Strings in CDS Archives

251: Multi-Resolution Images
252: Use CLDR Locale Data by Default
253: Prepare JavaFX UI Controls & CSS APIs for Modularization
254: Compact Strings
255: Merge Selected Xeroxs 2.11.0 Updates into JAXP
256: BeanInfo Annotations
257: Update JavaFXMedia to Newer Version of GStreamer
258: HarfBuzz Font-Layout Engine
259: Stack-Walking API
260: Encapsulate Most Internal APIs
261: Module System
262: TIFF Image I/O
263: HiDPI Graphics on Windows and Linux
264: Platform Logging API and Service
265: Martin Graphics Renderer
266: More Concurrency Updates
267: Unicode 8.0
268: XML Catalogs
269: Convenience Factory Methods for Collections
270: Reserved Stack Areas for Critical Sections
271: Unified GC Logging
272: Platform-Specific Desktop Features
273: DRBG-Based SecureRandom Implementations
274: Enhanced Method Handles
275: Modular Java Application Packaging
276: Dynamic Linking of Language-Defined Object Models
277: Enhanced Deprecation
278: Additional Tests for Heterogeneous Objects in G1
279: Improve Test-Failure Troubleshooting
280: Indify String Concatenation
281: HotSpot C= + Unit/Test Framework
282: Jlink: The Java Linker
283: Enable GTK 3 on Linux
284: New HotSpot Build System
285: Spin-Wait Hints
287: SHA 3 Hash Algorithms
288: Disable SHA-3 Certificates
289: Deprecate the Applet API
290: Filter Incoming Serialization Data
291: Deprecate the Concurrent Mark Sweep (CMS) Garbage Collector
292: Implement Selected ECMAScript 6 Features in Nashorn
294: Linux/x399x Port
295: Ahead-of-Time Compilation
297: Unified arm32/arm64 Port
298: Remove Demos and Samples
299: Reorganize Documentation
What Changed?

• The JDK is changing more quickly
  • Reaching for internals can no longer work (the tech debt collector has come)
  • But it is also no longer needed as new standard APIs are added
What Changed? (Internal)

- The JDK is changing more quickly
  - Reaching for internals can no longer work (the tech debt collector has come)
  - But it is also no longer needed as new standard APIs are added
- More of the runtime is written in Java
What Changed? (External)

• Java applications primarily run on the server with a wide and deep dependency trees.
  • Security focus has shifted from defending against malicious code to the greater challenge of defending against vulnerabilities in benevolent code
• One notable exception: Supply-chain attacks
• Server applications run in containers; want to “scale to zero”
PART II

What is integrity?
Integrity: The Ability to Promise

- **Invariant**: A property that’s true everywhere in a section of code (entire program)
- **Integrity Invariant**: An invariant that is *guaranteed* to hold by the language/runtime

Example:
- No out-of-bounds access to an array may or may not be (but *should* be) an invariant in a C program; requires a full-code analysis
- `int[] a = new int[10]` establishes an *integrity invariant* in Java that no out-of-bound access can take place; guaranteed by the runtime
Integrity Invariants in Java

- No out-of-bounds array access
- No use-after-free
- No process crash
- No uninitialized data
- Runtime type-safety (String can’t be cast to Socket)
- Relative file paths are stable (no chdir operation)

Integrity invariants are safety properties: something “bad” never happens
Encapsulation: The Mother of Java Integrity?

```java
public final class Even {
    private int x = 0;
    public int value() { return x; }
    public void incrementByTwo() { x += 2; }
    public void decrementByTwo() { x -= 2; }
}
```

- New invariants can be created from an encapsulation invariant (no access rule violations)
- All integrity invariants depend on encapsulation; those on previous slides depend on native VM code being encapsulated from Java code.
The Structure of a Modern Java Program

Application code
The Structure of a Modern Java Program

Application code → Libraries
The Structure of a Modern Java Program
Encapsulation: The Mother of Java Integrity?

Any 4\textsuperscript{th}-level dependency could violate the invariant:

- Deep reflection: `setAccessible`
- `sun.misc.Unsafe`
- JNI
- Dynamically load an agent and either redefine the methods (or, if class is not yet loaded, transform the field to `public`)

Impossible to establish any integrity invariant in Java if any of these is in play. Invariance requires full-code analysis, same as buffer overflow in C
Encapsulation: The Mother of Java Integrity?

That’s why it matters that more of the runtime is being written in Java:

- JIT is written in Java: Java code could globally disable array bounds checking by encroaching on the JIT’s encapsulation
- Thread scheduling and monitors written in Java: Java code could globally disable the JMM by encroaching on the the implementation of the thread scheduler or the implementation of monitors
But surely they wouldn’t, would they?
PART III

The importance of integrity
The Importance of Integrity

- Evolution
- Security
- Performance
Integrity & Evolution

- Freedom to fearlessly change internals not subject to backward compatibility
- Why do libraries reach for internals
  - New functionality
  - Work around bugs
  - Improve performance
- No longer works in an age of faster-paced evolution
- No longer needed in an age of faster-paced evolution
- May be justified for a library individually, but over the wide and deep dependency ecosystem it leads to a *tragedy of the commons* that demands regulation
- Technical debt is secretly foisted on client applications
public final class Session {
    private boolean superuser
        = authorizeSuperuser();

    public void sensitiveOperation() {
        if (superuser) doSensitiveOperation();
        else throw new UnauthorizedException();
    }
    private void doSensitiveOperation() { ... }
    private boolean authorizeSuperuser() { ... }
}
public final class Session {
    private boolean superuser = authorizeSuperuser();

    public void sensitiveOperation() {
        if (superuser) doSensitiveOperation();
        else throw new UnauthorizedException();
    }

    private void doSensitiveOperation() { ... }
    private boolean authorizeSuperuser() { ... }
}
Integrity & Security

- *Application*, uses *Session*. It also employs library *GoodSerializer* to deserialize JSON. *GoodSerializer* employs library *NeutralEncapsulationBreaker* to instantiate objects w/o constructor and assign private fields.

- A bug in *GoodSerializer*’s input sanitation means an attacker could send an input to get it to set the private field *superuser*.

Who’s at fault?

- *Application* is innocent and doing its best
- *GoodSerializer* is well-intentioned, but bugs happen
- *NeutralEncapsulationBreaker* has no vulnerability

- Any library that can break encapsulation, and any library that uses that library, becomes part of the attack surface area of any code that relies on encapsulation for integrity; we’re back to full-code analysis.
Integrity & Security

- Integrity is not a security mechanism, but no robust security mechanism can be created without it
- Offers “bulkheads” that compartmentalize the blast radius of a vulnerability
Integrity & Security: Aren’t we doomed, anyway?

- A Java library could write to the class files in the file system
- Plus: Spectre, Rowhammer etc.
- **Integrity of components is best enforced by their owner**
  - File system: OS
  - CPU cache: CPU
- Layers can cooperate, but each is in charge of its own integrity
- Java can and must enforce the integrity of the things it owns — Java code and objects — but shouldn’t (and can’t reliably) do more
Integrity & Performance

- Constant folding: Can a `final` field be constant folded? No, may be reassigned with deep reflection, JNI, or `Unsafe`.

- “Tree shaking”: Can a private method unused in a class be removed by a Condenser (that must preserve program meaning)? No, may be invoked with deep reflection or JNI (yes, we could try relying on speculation, but it makes some things much more complicated)

- Mechanical, meaning preserving transformations require `absolute` certainty
But surely they wouldn’t, would they?

- *Tragedy of the Commons*: A library author may feel *individually* justified.
- *Unintentionality*: A vulnerability in library X can unintentionally make library Y use its superpowers for bad.
- *Certainty*: Mechanical transformations require absolute certainty.
PART IV  (and last)

Integrity by Default
Integrity by Default

- Disabling the integrity of an invariant has a global effect
- A library (4th-level dependency) or a framework must not make a global decision

**Integrity by Default:** Every exception to integrity must be explicitly acknowledged by the application in a centralized program configuration

A centralized application configuration is an auditable record of integrity exceptions and accepted risks

The final say on module boundaries and privileges is given to the application
Strong Encapsulation: The Mother of Java Integrity!

**Strong encapsulation:** The encapsulation offered by Java’s access control cannot be broken by code in a different module *by any means* unless:

- The declaring module explicitly grants some other module the permission to do so in `module-info` or programmatically (`java.lang.Module`)
- Code passes on its privileges to other code with a `MethodHandles.Lookup` capability object
- The application redraws the map of encapsulation boundaries with `--add-opens/--add-exports` flags.
- The application grants “superpowers” (JNI, agent) to some/all code
Outdated Libraries

- **--add-opens/exports** are not a “JDK 8 compatibility mode”. A program with many such flags for technical-debt reasons is a program that’s about to break.
- They’re “landmine markers” — keep you alive while you clear the landmines
- They add no burden because landmines must be found to be cleared
- Remember: The program will break even if we required no flags. Short of stopping Java’s evolution, there’s nothing we can do about that
- Plus, ensure that no new uses of internals can be added unnoticed; there’s incompatibility pain only *once* more and only one fix: stop using internals
- Without strong encapsulation by default, 8->9 migration pains would have continued forever and ever and ever (and no other integrity benefits, either)
- There are worse fates than an exception
- If a library is not updated to not require flags that’s a red flag that it’s improperly maintained
Outdated Libraries

• In the real world, companies don’t have the resources to fix technical debt
• That’s absolutely true
• It’s also true that in the real world some countries don’t have resources to fix bridges and make sure buildings are up to code
• And in the real world bridges collapse and buildings burn down
• There are consequences to risk whether we must take it or not
• Ignoring risk doesn’t make it go away; best to know where it is

• Tip: **Add a comment/git message explaining why each flag is needed**
Supporting Old JDK Versions

- Our advice:
  - Develop at the tip, and only for some recent-enough JDK version
  - Largely freeze old library versions.
  - Backport only security patches and serious bugs — not much work
- That’s what we’ve done at Oracle with the JDK since we introduced the LTS service
Operating beyond encapsulation boundaries

- Unit tests
  - Build tools and testing frameworks should *automatically* emit `--add-exports`, `--add-opens`, and `--patch-module` for the module under test, as appropriate
  - For mocking, use agents loaded at startup
- Frameworks
  - Should not use `add-opens` flags; use `MethodHandles.Lookup`
    ```java
    static { AcmeFramework.grantAccess(MethodHandles.lookup()); }
    ```
- APM tools — Use agents loaded at startup
- Serialization (A common cause of vulnerabilities)
  - We have a vision for encapsulation-respecting serialization
  - Until then try serializing only records, collections and other classes with well-known construction.
Even with a memory safe language, memory management is not entirely memory safe. Most memory safe languages recognize that software sometimes needs to perform an unsafe memory management function to accomplish certain tasks. As a result, classes or functions are available that are recognized as non-memory safe and allow the programmer to perform a potentially unsafe memory management task. Some languages require anything memory unsafe to be explicitly annotated as such to make the programmer and any reviewers of the program aware that it is unsafe. Memory safe languages can also use libraries written in non-memory safe languages and thus can contain unsafe memory functionality. Although these ways of including memory unsafe mechanisms subvert the inherent memory safety, they help to localize where memory problems could exist, allowing for extra scrutiny on those sections of code.
Where We Want to Be

The NSA information sheet continues:

For languages with an extreme level of inherent protection, considerable work may be needed to simply get the program to compile due to the checks and protections.

Not in Java!

• Java should be the safest mainstream programming language in the world
• Exceptions to integrity are tracked and localized in an auditable configuration
• The tax on those who don’t care is not large esp. if they use the classpath (no “localization”)}
Why not *opt-in* to integrity?

- Most programs can remain under full integrity due to recent work
  - Enjoy portability and other benefits *much* more easily than ever
  - The minority that don’t *will be inconvenienced*, but not much
- Simple enough to be on by default and reduce attack surface area
- Tighter regulation (integrity) = lower entropy (fewer possible programs)
  - For new programs it’s easy (and better) to start with low entropy
  - Old program need to expend energy to reach low entropy once
A Gradual Yet Resolute Path Forward

- Deep reflection restricted since JDK 16
- Dynamically loaded agents will be restricted (non-SE/optional)
- JNI will be restricted (optional component)
- Unsafe will be removed (non-SE)
- FFM will be restricted (starts out restricted)

As always, we’ll emit warnings & give ecosystem time to adapt

Reminder 1: All command line options can be placed in shared, full or partial, configuration “@files”

Reminder 2: jlink is flexible and widely applicable (more than some seem to think)
A Gradual Yet Resolute Path Forward

- **JEP draft: Integrity and Strong Encapsulation**
- **JEP 451: Prepare to Disallow the Dynamic Loading of Agents**
- **JEP draft: Prepare to Restrict The Use of JNI**
- More to follow