

DST, Summer 2017

#### Programming Dynamic Features and Monitoring Distributed Software Systems

#### Hong-Linh Truong Distributed Systems Group, TU Wien

truong@dsg.tuwien.ac.at dsg.tuwien.ac.at/staff/truong @linhsolar

1





#### **OVERVIEW**





#### Things that make a difference

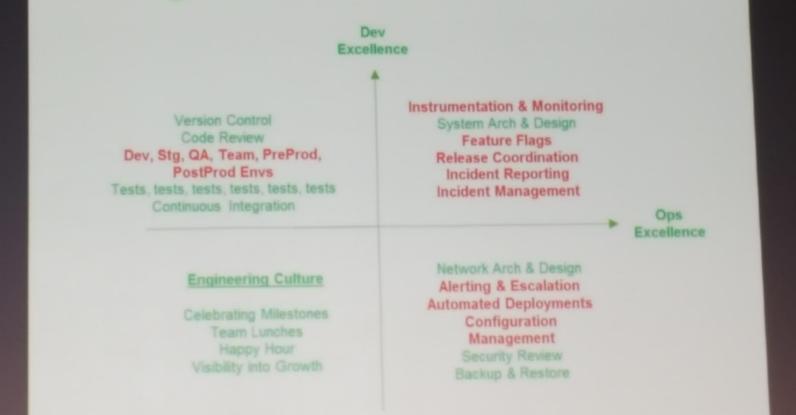
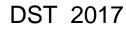


Photo from "Lessons Learned From a Fast Growing Startup", Arul Kumaravel, Vice President, Engineering, Grab at CloudAsia2016, Singapore







"In the context of computer programming, instrumentation refers to an ability to monitor or measure the level of a product's performance, to diagnose errors and to write trace information"

https://en.wikipedia.org/wiki/Instrumentation\_%28computer\_programming%29

"To monitor or monitoring generally means to be aware of the state of a system, to observe a situation for any changes which may occur over time, using a monitor or measuring device of some sort."

https://en.wikipedia.org/wiki/Monitoring

Programming dynamic features enable instrumentation and monitoring



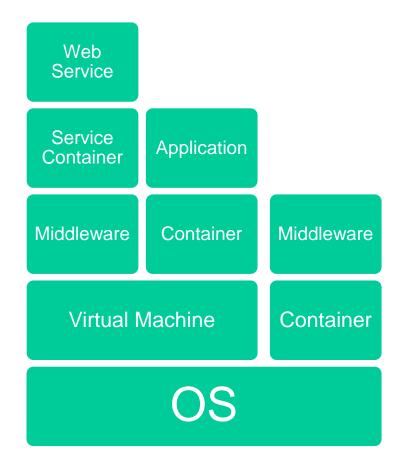
DISTRIBUTED SYSTEMS GR

### Key techniques for today's lecture

- Big picture in full-stack and large-scale monitoring
- Dynamic features
- Code inspection
  - Dynamic loading
  - Reflection
  - Dynamic proxy
- Instrumentation and Program Analysis
- Annotation
- Aspect-oriented Programming



## Full stack monitoring

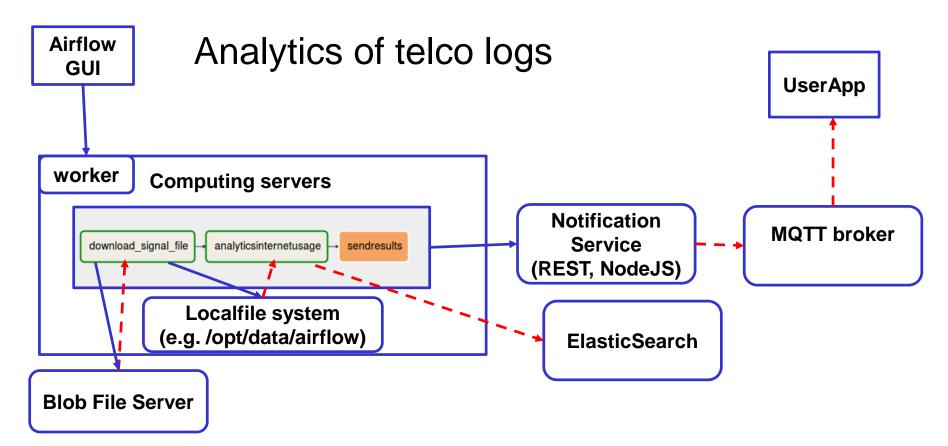


- You might need to monitor from OS to the application components
  - You might own or just rent them
- Artifacts: binary, runtime, source
- Monitoring functions about computation, data and network



DISTRIBUTED SYSTEMS GRO

### Monitoring at the Large-scale



- Many distributed components across various enterprise boundaries
- Events/Measurement collection, Storage, Analytics and Visualization



## **Dynamicity needs (1)**

- Monitoring, performance analysis, tracing, debugging
  - Dependent on specific contexts → static ways are not good due to overhead and flexibility problems
- Common concerns among several classes/objects
  - Do not want to repeat programming effort
- Provide metadata information for lifecycle management
  - Provisioning and runtime configuration

## Dynamic needs (2)

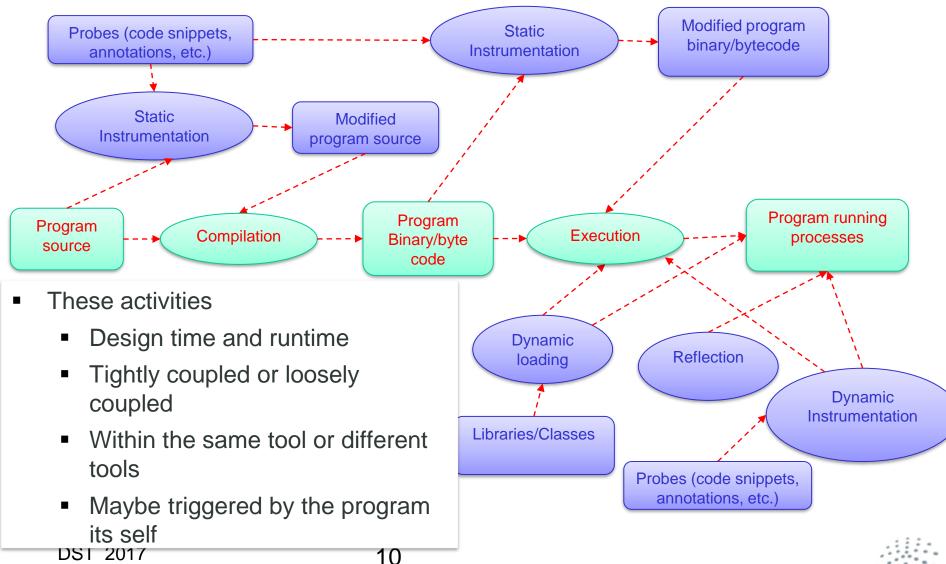
- Provide metadata information for code generation
  - Service interfaces and validation
- Flexible software delivery for some core functions (e.g., patches)
- Enable continuous update without recompiling/redeploying applications
- Extensibility



DISTRIBUTED SYSTEMS GRO



# Main activities for programming dynamic features



DISTRIBUTED SYSTEMS GROUP

#### Grab instrumentation and monitoring

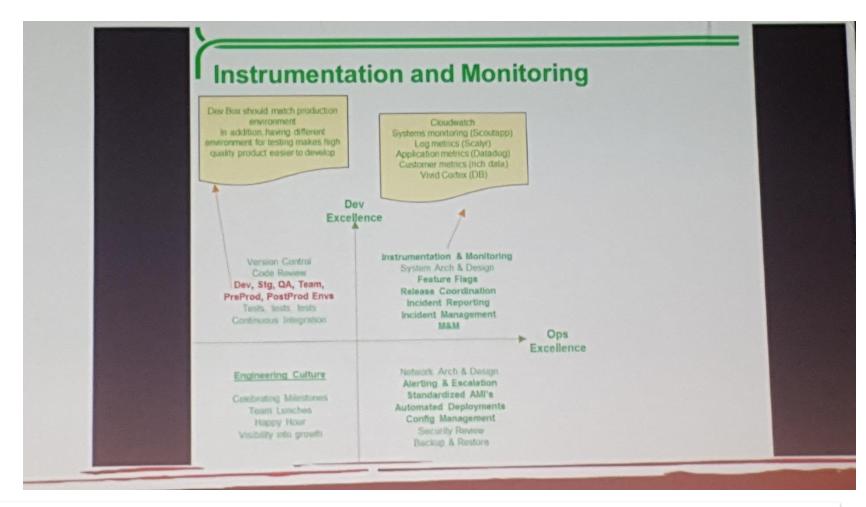
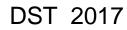


Photo from "Lessons Learned From a Fast Growing Startup", Arul Kumaravel, Vice President, Engineering, Grab at CloudAsia2016, Singapore







#### **CODE INSPECTION**





- How do we know program structures?
- Can we analyze the program structure within program processes during runtime?
- Are we able to examine statically and dynamically linked code?
- What kind of tasks we could do if we know the program structure?





- Code inspection/analysis
  - Analyze different forms of program code at design and runtime
- Source code analysis
- Bytecode and binary code analysis
- Program's running process analysis



## Dynamic loading

- Code can be dynamically loaded into a running program
  - At runtime, libraries are loaded into a program memory
  - Variables, functions, etc. in these libraries can be used by the program
  - Dynamic linking: no libraries are needed at compile or before loading at runtime
- Implementations

- C: void \*dlopen(const char \*filename, int flag);
- Java ClassLoader and System.loadLibrary(String libname)





## Bytecode/Binary inspection

- Read and build program structures from bytecode
  - Example, using javap –c to see Java bytecode
- Tools to process bytecodes
  - Javassist (https://github.com/jbossjavassist/javassist/releases)
  - BCEL (<u>http://commons.apache.org/proper/commons-bcel/</u>)
  - CGLIB (https://github.com/cglib/cglib)
- Cannot see the dynamic code which will be loaded at runtime







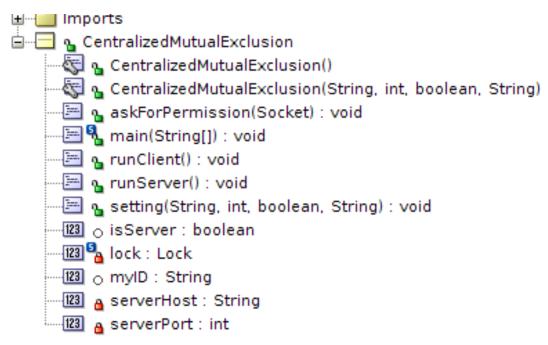
- Allows running program code to introspect its own definition (e.g., iterate methods and fields of a class)
- Additionally allows for creation and invocation of object instances at runtime
- Basic feature in most current object-oriented languages



DISTRIBUTED SYSTEMS GRO

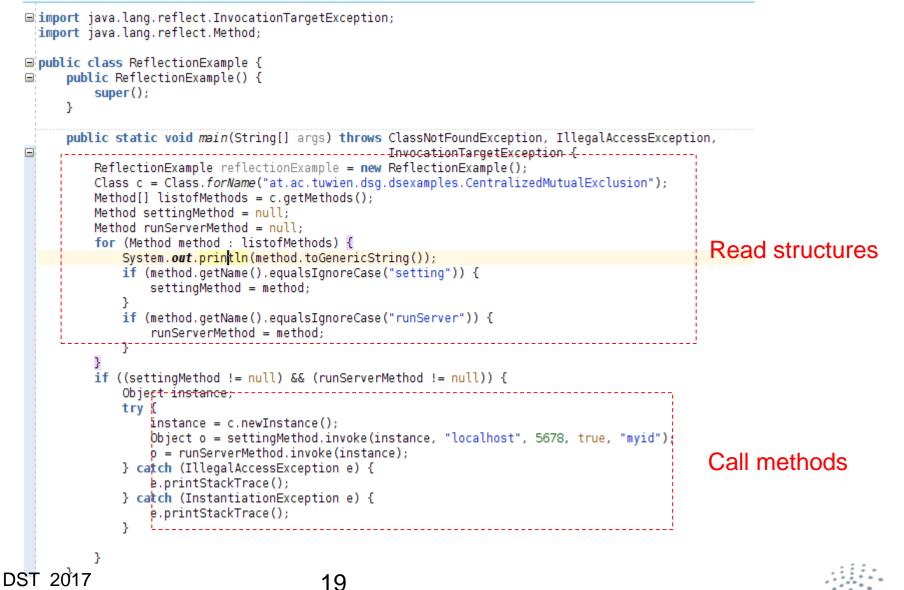
### **Example: Reflection in Java**

#### Source code



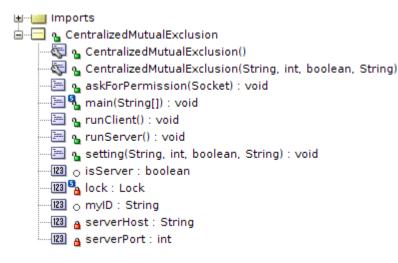


#### **Example: Reflection in Java**



### **Example:** Reflection in Java

#### Source code



#### Inspection output

Ctruong@truong-dsg:~/docs/Dropbox/teaching/ds-resources/dssyn-exs/classes\$ /usr/local/java/bin/java at.ac.tuwien.dsg.dsexamples.Reflec ublic static void at.ac.tuwien.dsg.dsexamples.CentralizedMutualExclusion.main(java.lang.String[]) throws java.io.IOException,java.lang ublic void at.ac.tuwien.dsg.dsexamples.CentralizedMutualExclusion.askForPermission(java.net.Socket) throws java.io.IOException ublic void at.ac.tuwien.dsg.dsexamples.CentralizedMutualExclusion.runClient() throws java.net.UnknownHostException,java.io.IOException ublic void at.ac.tuwien.dsg.dsexamples.CentralizedMutualExclusion.setting(java.lang.String,int,boolean,java.lang.String) throws java. ublic void at.ac.tuwien.dsg.dsexamples.CentralizedMutualExclusion.runServer() throws java.io.IOException ublic final void java.lang.Object.wait(long,int) throws java.lang.InterruptedException ublic final native void java.lang.Object.wait(long) throws java.lang.InterruptedException ublic final void java.lang.Object.wait() throws java.lang.<u>InterruptedException</u> ublic boolean java.lang.Object.equals(java.lang.Object) ublic java.lang.String java.lang.Object.toString() ublic native int java.lang.Object.hashCode() ublic final native java.lang.Class<?> java.lang.Object.getClass() ublic final native void java.lang.Object.notify() ublic final native void java.lang.Object.notifyAll() etting parameter am the centralized server for mutual exclusion





#### Reflection – Advantages and Disadvantages

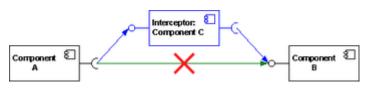
- Advantages:
  - Built-in features in most modern programming languages
  - Allows for flexible applications and design patterns
- Disadvantages:
  - Complex solutions are difficult to write and error-prone (method names in strings, etc.)
  - Performance degradation
  - Security restrictions
  - Reflection is read-only it is not (easily) possible to add methods or change the inheritance hierarchy of an object



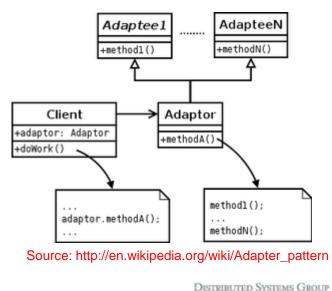
DISTRIBUTED SYSTEMS GRO



- Allow us to implement a proxy class whose interfaces specified at runtime
- Create proxy instance with a set of interfaces
- Forward method invocations on interfaces of the proxy instance to another object

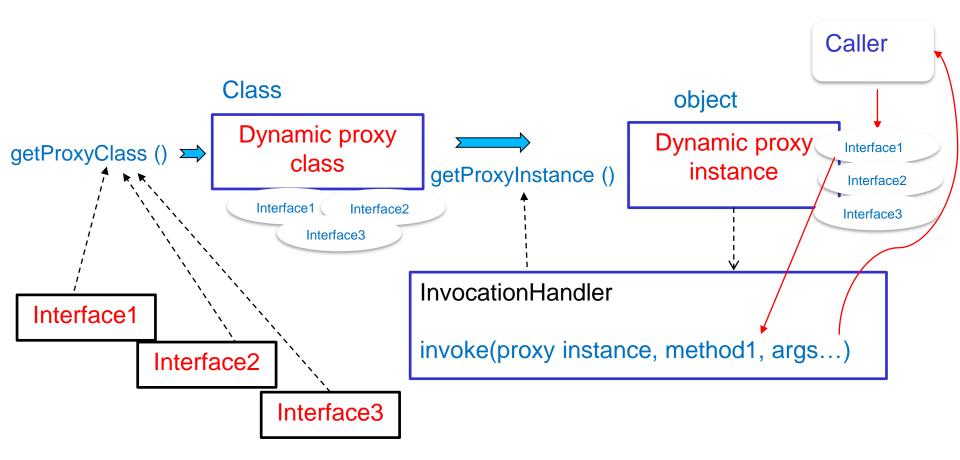


Source: http://en.wikipedia.org/wiki/Interceptor\_pattern

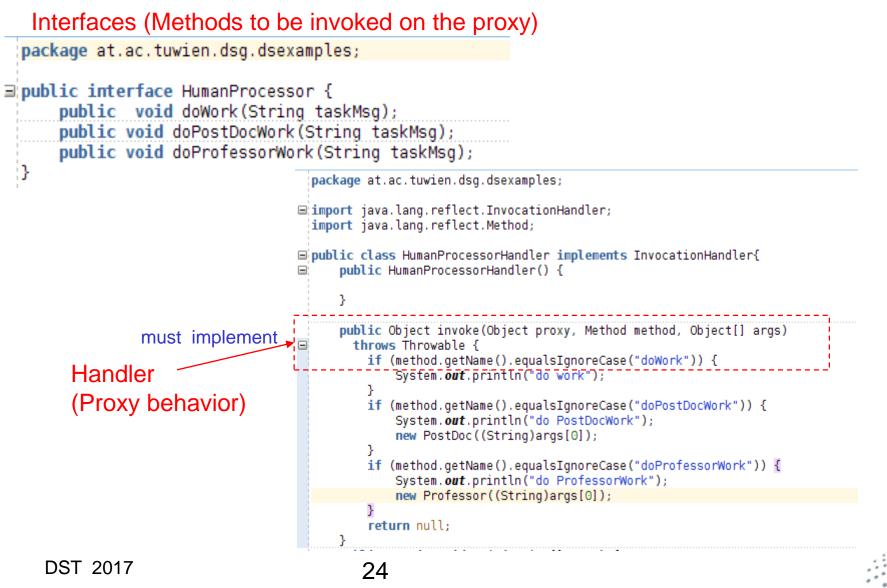




#### **Dynamic Proxy – Conceptual Model**



### **Example of Dynamic Proxy**



DISTRIBUTED SYSTEMS GROUP

### **Example of Dynamic Proxy**

```
import java.lang.reflect.InvocationHandler;
 import java.lang.reflect.InvocationTargetException;
 import java.lang.reflect.Method;
 import java.lang.reflect.Proxy;
public class HumanDynamicProxy {
     public HumanDynamicProxy() {
         super();
     3
     public HumanProcessor getHumanProcessorProxy() throws NoSuchMethodException, InstantiationException,
             IllegalAccessException, InvocationTargetException {
         InvocationHandler hph = new HumanProcessorHandler();
         Class proxyClass =
             Proxy.getProxyClass(HumanProcessor.class.getClassLoader(), new Class[] { HumanProcessor.class });
         return (HumanProcessor)proxyClass.getConstructor(new Class[] { InvocationHandler.class }).newInstance(new Object[] { hph });
         //return (HumanProcessor) Proxy.newProxyInstance(HumanProcessor.class.getClassLoader(),
         //new Class[] {HumanProcessor.class }, new HumanProcessorHandler());
                                                 _____
     public static void main(String[] args) throws InstantiationException, IllegalAccessException,
                                                   InvocationTargetException {
         HumanDynamicProxy humanDynamicProxy = new HumanDynamicProxy();
         HumanProcessor hp = null;
         try {
             hp = humanDynamicProxy.getHumanProcessorProxy();
             hp.doWork("Read the assignment 3");
             hp.doPostDocWork("Read the paper about java dynamics");
             hp.doProfessorWork("Read the news about Wurst");
         } catch (NoSuchMethodException e) {
             e.printStackTrace();
         3
```



#### PROGRAM INSTRUMENTATION AND ANALYSIS



DISTRIBUTED SYSTEMS GROUP

## Program instrumentation

- A process to inspect and insert additional code/meta-data, etc., into a program/process
  - Static and runtime
  - Manual or automatic
- Examples:
  - Source code annotations/directives
  - Byte code modification
  - Dynamic code modification at loading time
  - Process instructions at runtime





# Static versus Dynamic instrumentation

- Dynamic instrumentation
  - Perform the instrumentation during the process running
    - E.g., Dyninst (<u>http://www.dyninst.org</u>)
  - Java support:
    - E.g., Dtrace + Btrace, <u>http://www.oracle.com/in/javaonedevelop/dtrace-j1-</u> <u>sritter-400745-en-in.pdf</u>
    - Java instrumentation API
    - Some works on static + dynamic instrumentation based on dynamic class loading mechanisms
- Static instrumentation
  - Source code, bytecode and binary code levels
- In many cases: a combination of different methods





# Where we can insert instrumented code into the program execution?

- At any join point: a point in the control flow of a program
- Examples:
  - method calls, entry/exit of method body, statements (set/get, assignment, etc.)

If we instrument probes before, after or around these join points, when the program execution reaches these points, the probes will be executed accordingly.

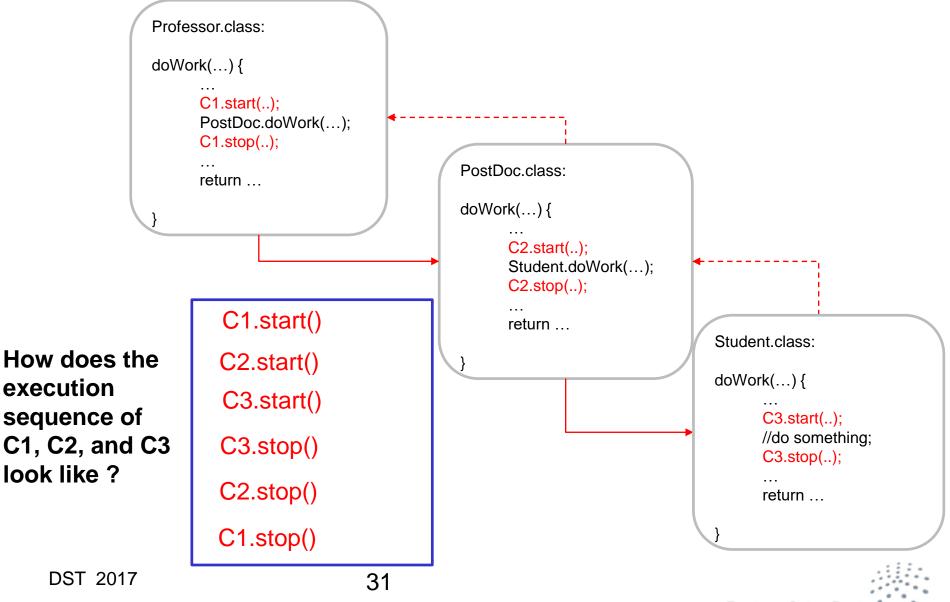




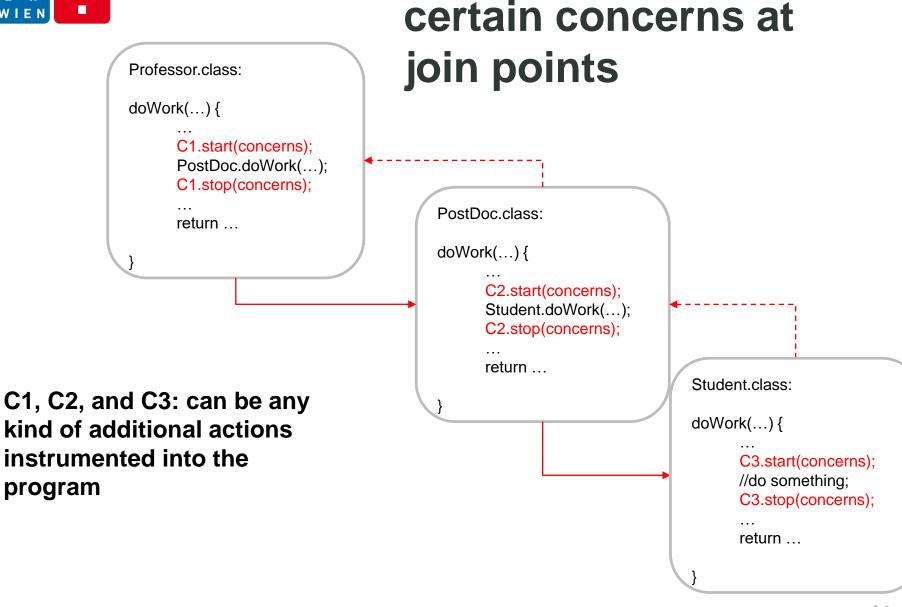
#### **Example: Dynamic call graph** "A call graph is a directed graph that represents calling relationships between subroutines in a Professor.class: computer program." doWork(...) { https://en.wikipedia.org/wiki/Call\_graph C1.start(..); PostDoc.doWork(...); C1.stop(..); PostDoc.class: return ... doWork(...) { C2.start(..); Student.doWork(...); C2.stop(..); return ... Student.class: doWork(...) { C3.start(..); //do something; C3.stop(..); return ... DST 2017 30

#### Example: Dynamic call graph

FN







If we want to deal with



#### **SOURCE CODE ANNOTATION**





- Annotations are added into source code
  - Can be considered as static instrumentation
  - Can be considered as a part of typical programming activities
- Goal: provide additional metadata/instructions
  - For supporting compilation process
  - For code generation at compiling and deployment
  - For runtime processing
  - Etc.
- Very popular in Java/C#/Python, ...





Format

@AnnotationName (....)

- Pre-defined versus user-defined
  - Pre-defined: supporting by runtime systems or some well-known libraries in programming frameworks
  - User-defined: it is up to the developer to define annotations
- Points at which annotations can be added
  - declarations of classes, fields, methods, and other program elements
  - type uses (Java 8, e.g. @NonNull String serverName)





## **Example of EE Annotation Support**

- In most EE programming frameworks
- Some Spring annotations (http://docs.spring.io/spring/docs)
  - @Resource, @PostConstruct and @PreDestroy

```
@IfProfileValue(name="java.vendor", value="Oracle Corporation")
@Test
public void testProcessWhichRunsOnlyOnOracleJvm() {
    // some logic that should run only on Java VMs from Oracle Corporation
}
```

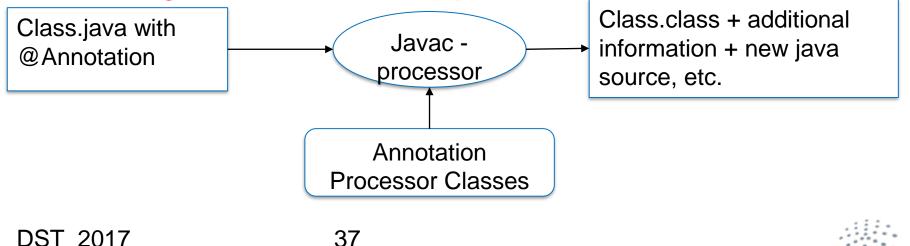
- JAX-RS (https://jax-rs-spec.java.net/)
  - @Path("scu/{scuid}")
  - @Consumes("application/json")
  - @Produces("application/xml")



## Java Annotation Processing

- Parsing source codes
- Reflection APIs also return Annotation Method.class: public Annotation[] getDeclaredAnnotations()
- Reading bytecode to get Annotation

### Processing Model in Java compilation



DISTRIBUTED SYSTEMS GRO

# Example – your case study with New Relic

Check:

https://docs.newrelic.com/docs/agents/javaagent/custom-instrumentation/javainstrumentation-annotation

@Trace

protected void methodWithinTransaction() {
 // work

}



DISTRIBUTED SYSTEMS GROU



### ASPECT-ORIENTED PROGRAMMING

DST 2017



## **Cross-cutting concerns**

- We have some common concerns that across multiple objects/methods
  - Tracing, measuring time, logging, checking security, etc.
- We want to have dynamically programming features to address these concerns



DISTRIBUTED SYSTEMS GRO

# Cross-cutting concerns – when, where and how

- We can use "probes" instrumented into targeting programs → creating hooks
  - Probes specify code for dealing these concerns
  - Probes create addition actions at runtime
- But we need dynamic and flexible way
  - Probes are instrumented when and where we need but they can be replaced!
- How
  - Can we use annotation? Can we use dynamic loading? Bytecode/binary instrumentation? Dynamic instrumentation?



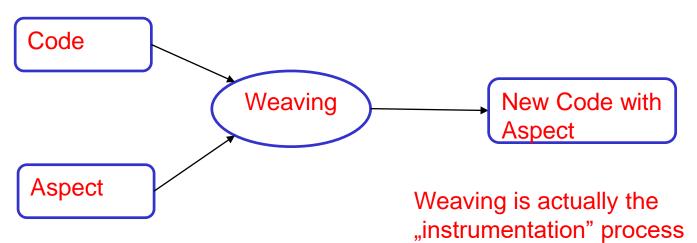
### Aspect-Oriented Programing

- Aspect: common feature in various methods, classes, objects, etc. → crosscutting concern
- Separate from functional concerns and crosscutting concerns
  - In Aspect-Oriented Software Development (AOSD), functional concerns are built in the usual way
  - Cross-cutting concerns are built as independent modules
- Combining these two types of concerns using semi-automatic instrumentation techniques



DISTRIBUTED SYSTEMS GRO

### **Conceptual model**



- Some java implementations
  - AspectJ
    - The standard implementation of AOP in Java
  - SpringAOP







### Join Point

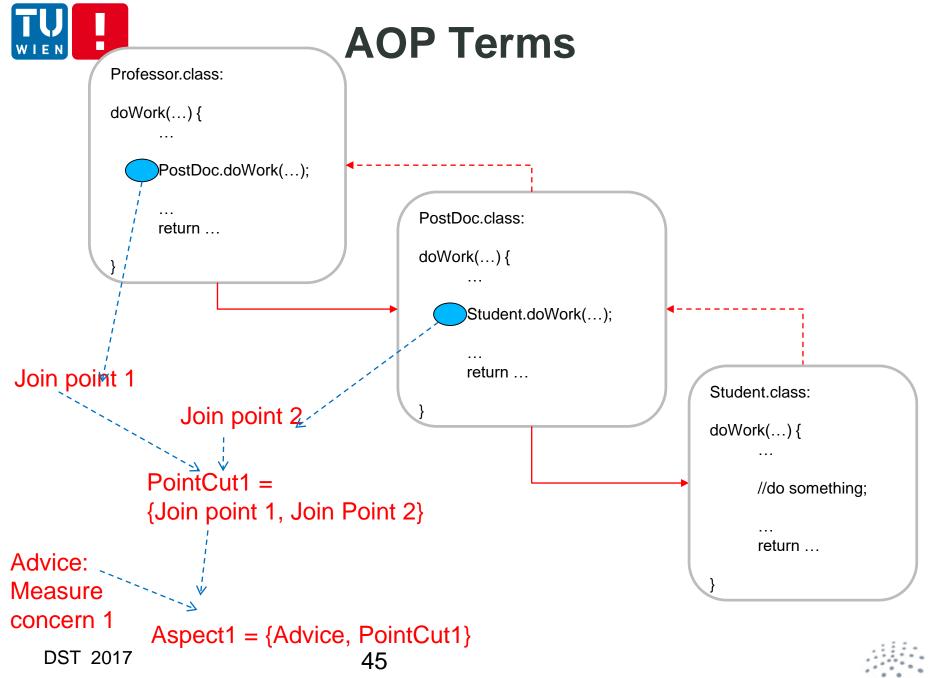
point in the execution

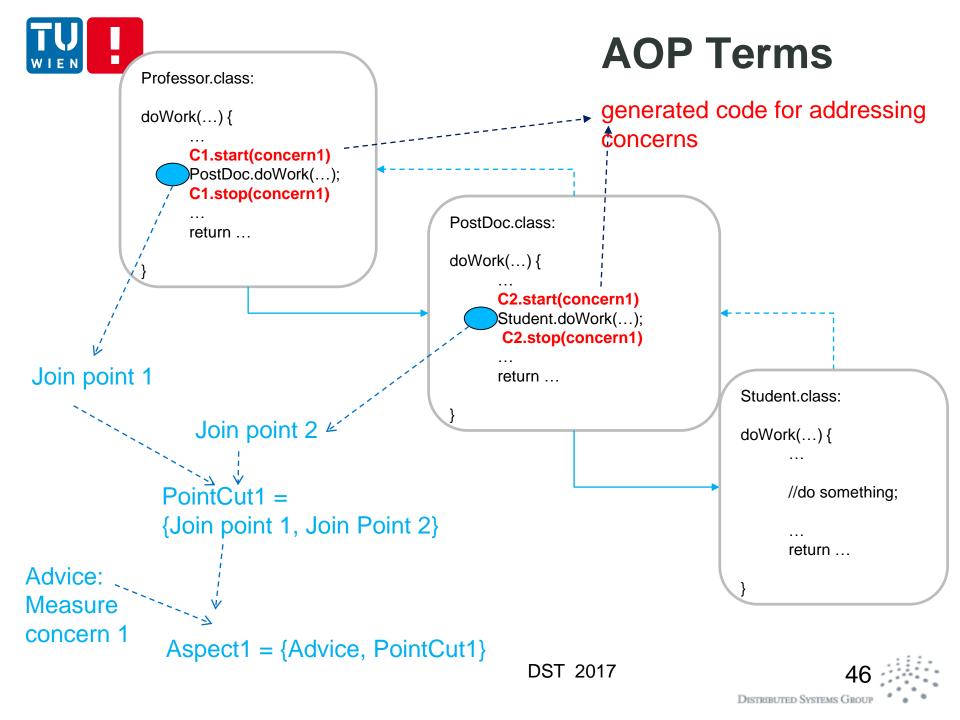
### Pointcut

A set of join points (composed using different operators such as &&, ||, !)

- Advice
  - Additional action that should be executed at join points in a pointcut
- Aspect
  - Cross-cutting type and its implementation (advices + others)







## Main types of Join Points

- Execution: when a method body executes execution(public void doWork(String))
- Call: when a method is called call(void doWork(String))
- Handler: when an exception handler executes

handler(ArrayOutOfBoundsException)





## Main types of Join Points

this: when the current executing object is of the specified type

### this(Student)

- target: when the target object is of the specified type target(Student)
- within: when the executing code within the specified class within(Student)
- withincode: within a method withincode(void doWork())
- set/get: field access/references set(String Student.name)

48





- Avoid to confuse call with execution join points for methods
- Call matches before or after a method is called (i.e., still in the scope of the caller)

E.g., the call join point is the last thing that happens before the method is actually invoked

 Execution matches when the method starts to execute (i.e. already in the scope of the callee)

E.g., the execution join point is the first thing that happens during method invocation



DISTRIBUTED SYSTEMS GRO



- Advice defines code of aspect implementation that is executed at defined points
- Main types of advice

```
before () : methodCall() {
    ...
}
after () : methodCall() {
    ...
}
around () : methodCall() {
    ....
}
```

methodCall is a pointcut





- The process of merging aspects into the program code is called weaving
- Three ways of weaving:
  - Compile-Time Weaving (weave as part of source-tobinary compilation)
  - Binary Weaving (compile normally, then merge binaries in a post-compilation step)
  - Load-Time Weaving (like binary weaving, but done when the class is loaded by the classloader – implemented via Java agent mechanism)



DISTRIBUTED SYSTEMS GRO

## **Compile-Time Weaving**

- Advantages:
  - No startup performance degradation
  - Allows you to store and see the produced source code (good for debugging, and error tracking)
- Disadvantages:
  - Can't weave third-party code (e.g., used libraries)
  - Java: requires usage of special compiler (ajc), which may produce worse regular Java code than your Sun or openjdk compiler



## Binary Weaving

- Advantages:
- No startup performance degradation
- Allows you to store and see the produced binary (eases debugging)
- Can be used with any compiler
- Can weave third-party libraries, as long as you can live with permanently modifying them
- Disadvantages:
- Complex when dealing with system libraries

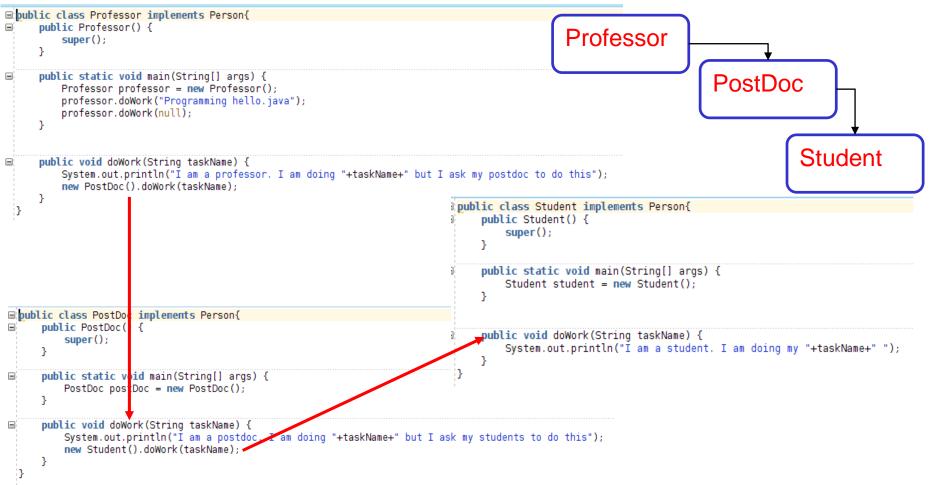




- Advantages:
  - Can be used with any compiler
  - Can sensefully weave pretty much anything, even system libraries
- Disadvantages:
  - Slow startup
  - Hard to debug and understand, as the running code exists nowhere outside the classloader



### Example of AOP with AspectJ



### Example of AOP with AspectJ

```
Professor.class:
//impore juvurneeroveree,
public aspect Tracing {
       private pointcut methodExecution () :
                                                                         public void doWork(String taskName) {
           execution(public void doWork(String ));
                                                                               System.out.println("I am a professor. I am doing
       private pointcut methodCall () :
           call(void doWork(String )) &&within (PostDoc);
                                                                          '+taskName+" but I ask my postdoc to do this");
       private pointcut withinClass () :
                                                                               new PostDoc().doWork(taskName);
           within (Student) && call(void println(String ));
       private pointcut methodParameter (String task) :
  ( 4)
           (call(void doWork(String )) &&args(task)) && within (Professor);
                                                                          public static void main(String[] args) {
       before () : methodExecution() {
                                                                               Professor professor = new Professor();
               System.out.println("START> " + thisJoinPoint);
                                                                               professor.doWork("Programming hello.java"
       3
                                                                                 professor.doWork(null);
       after () : methodExecution() {
               System.out.println("<END " + thisJoinPoint);</pre>
                                                                          PostDoc.class:
       before () : methodCall() {
               System.out.println("CALL> " + thisJoinPoint);
       }
                                                                         public void doWork(String taskName) {
       after () : methodCall() {
                                                                          System.out.println("I am a postdoc. I am doing
               System.out.println("<CALL " + thisJoinPoint);</pre>
                                                                          "+taskName+" but I ask my students to do this");
                                                                               new Student().doWork(taskName);
       before () : withinClass() {
               System.out.println("WITHIN> " + thisJoinPoint);
       3
                                                                          Student.class:
       after () : withinClass() {
               System.out.println("<WITHIN " + thisJoinPoint);</pre>
                                                                          public void doWork(String taskName) {
                                                                               System.out.println("I am a student. I am doing my
       before (String task) : methodParameter(task) {
           if (task ==null) {
                                                                           "+taskName+" ");
               System.out.println("Error!!!");
               System.exit(0);
           }
```

DST 2017

### **Example of AOP with AspectJ**

#### Call gaph tracing information

Professor.doWork()	START> execution(void Professor.doWork(String)) I am a professor. I am doing Programming hello.java but I ask my postdoc to do this START> execution(void PostDoc.doWork(String))
PostDoc.doWork()	I am a postdoc. I am doing Programming hello.java but I ask my students to do this CALL> call(void Student.doWork(String))
Student.doWork()	START> execution(void Student.doWork(String)) WITHIN> call(void java.io.PrintStream.println(String)) I am a student. I am doing my Programming hello.java WITHIN call(void java.io.PrintStream.println(String))
	<pre><end <call="" <end="" call(void="" execution(void="" postdoc.dowork(string))="" pre="" professor.dowork(string))<="" student.dowork(string))=""></end></pre>







- Not all features are supported
  - String AOP only method execution join points
- Using Java annotation or XML
- Java Annotation
  - @Aspect, @Pointcut, @Before, @After,
     @AfterReturning, @Around
- Using XML
  - aop:config, aop:aspect, aop:before, etc.

What is the underlying mechanism?

→ Using dynamic proxy to delegate/process advices



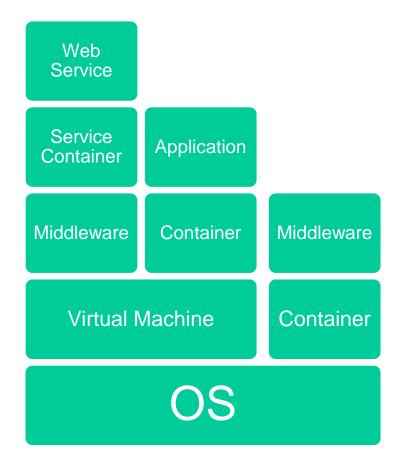


### INSTRUMENTATION AND MONITORING AT SCALE

DISTRIBUTED SYSTEMS GROUP

DST 2017

## Full stack monitoring



- Combine many techniques: instrumentation, API interface, etc.
- Push and pull methods
- Exact measurement and sampling



DISTRIBUTED SYSTEMS GROU

## Scale of systems and of monitoring

- Many monitoring components
- Scalable middleware for relaying monitoring data
  - Various protocols, HTTP, AMQP, MQTT
- Scalable storage: file systems and time series data
- Visualization and other types of big/fast data analytics





### What can we do with messaging, complex event processing (lecture 2) and dynamic features programming (lecture 3)?

Building real-world instrumentation and monitoring for (cloud-based) services → instrumentation and monitoring ecosystem for complex distributed systems





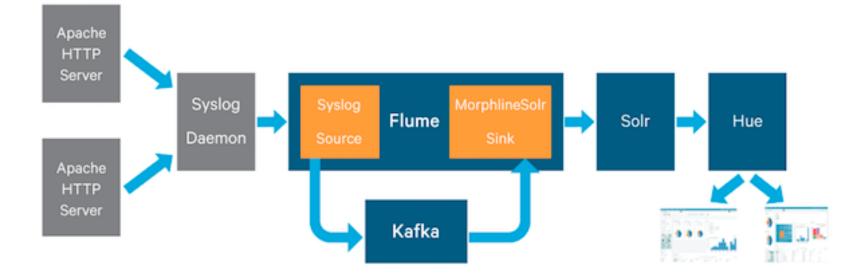
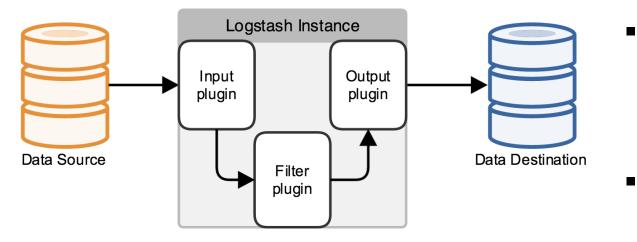


Figure source: https://blog.cloudera.com/blog/2015/02/how-to-do-real-time-log-analytics-with-apache-kafka-cloudera-search-and-hue/







Codecs: stream filters within inputs or outputs that change data representation

E.g.: multilines  $\rightarrow$  a single event

Source: https://www.elastic.co/guide/en/logstash/current/advanced-pipeline.html







#### The Beats Family

All kinds of shippers for all kinds of data.



Heartbeat Uptime Monitoring

#### LIGHTWEIGHT

#### Ship from the Source. Plain and Simple.

Beats are great for gathering data. They sit on your servers and centralize data in Elasticsearch. And if you want more processing muscle, Beats can also ship to Logstash for transformation and parsing.

Windows Event Logs

#### https://www.elastic.co/products/beats

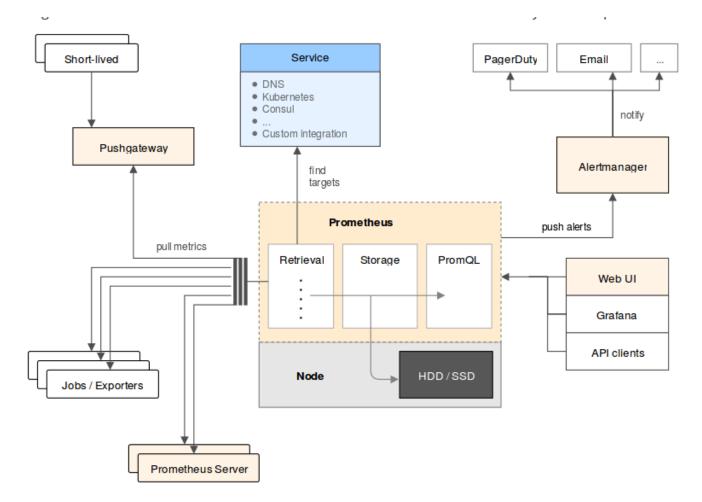








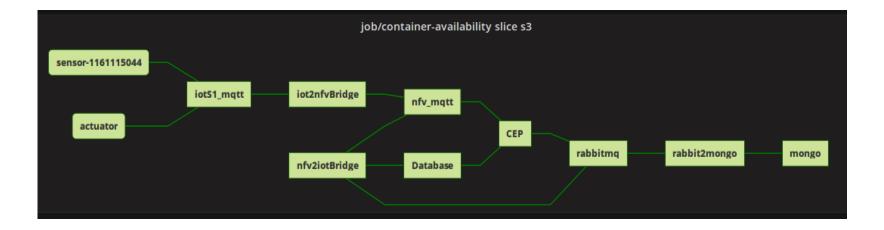
### Promethesus Architecture



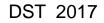
Source: https://prometheus.io/docs/introduction/overview/

DISTRIBUTED SYSTEMS GROUP





#### From Christian Proinger (Master thesis work TU Wien), using Prometheus and Grafana, etc. for Raspberry PI dockers + clouds.





### Build instrumentation and monitoring for your cloud application/services

### Do a real-world test!

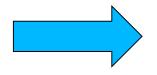
Instrumentation and monitoring (aspects, annotation, etc.)

Your Application

Your Middleware

Your VM/Docker

Logs/traces/metrics



STACKDRIVER MONITORING

Cloud monitoring services

For applications running on Google Cloud Platform and Amazon Web Services VIEW DOCUMENTATION

TRY IT FREE





#### Amazon CloudWatch

Scout 🛩

Amazon CloudWatch is a monitoring service for AWS cloud resources

App Monitoring

DST 2017



Server Monitoring



To design your monitoring and instrumentation solutions together with communication and storage system middleware

Try to analyze existing examples and tools to see the complexity of programming dynamic features and monitoring (not just simple AOP)



DISTRIBUTED SYSTEMS GI



- Many features are generic but very useful for dynamic features programming
  - Required by complex distributed software
- Dynamicity programming can be achieved through different design and runtime activities
- There are different tools for programming dynamic features
- Understanding which instrumentation techniques should be used and what will be instrumented is crucial
- Often we need to combine different techniques



DISTRIBUTED SYSTEMS GR

## **Further materials**

- http://eclipse.org/aspectj/doc/released/progguide
- http://docs.spring.io/spring/docs/current/spring-framework-reference/
- http://docs.oracle.com/javase/tutorial/java/annotations/
- http://commons.apache.org/proper/commons-bcel/
- https://jcp.org/en/jsr/detail?id=160
- <u>http://docs.oracle.com/javase/7/docs/api/java/lang/instrument/package-summary.html</u>
- http://docs.oracle.com/javase/8/docs/technotes/guides/reflection
- http://docs.oracle.com/javase/tutorial/reflect/index.html





### Thanks for your attention

Hong-Linh Truong Distributed Systems Group, TU Wien truong@dsg.tuwien.ac.at http://dsg.tuwien.ac.at/staff/truong

