RFC-233: GPU Capabilities

Draft

8 Pages

Abstract

More and more GPU hardware is used to accelerate computations in applications compared to running on a CPU. In order to know whether or not an OSGi framework has access to GPU hardware and a GPU accelerated bundle can be deployed, this RFC provides a namespace with capabilities and requirements.

0 Document Information

0.1 License

DISTRIBUTION AND FEEDBACK LICENSE, Version 2.0

The OSGi Alliance hereby grants you a limited copyright license to copy and display this document (the “Distribution”) in any medium without fee or royalty. This Distribution license is exclusively for the purpose of reviewing and providing feedback to the OSGi Alliance. You agree not to modify the Distribution in any way and further agree to not participate in any way in the making of derivative works thereof, other than as a necessary result of reviewing and providing feedback to the Distribution. You also agree to cause this notice, along with the accompanying consent, to be included on all copies (or portions thereof) of the Distribution. The OSGi Alliance also grants you a perpetual, non-exclusive, worldwide, fully paid-up, royalty free, limited license (without the right to sublicense) under any applicable copyrights, to create and/or distribute an implementation of the Distribution that: (i) fully implements the Distribution including all its required interfaces and functionality; (ii) does not modify, subset, superset or otherwise extend the OSGi Name Space, or include any public or protected packages, classes, Java interfaces, fields or methods within the OSGi Name Space other than those required and authorized by the Distribution. An implementation that does not satisfy limitations (i)-(ii) is not considered an implementation of the Distribution, does not receive the benefits of this license, and must not be described as an implementation of the Distribution. "OSGi Name Space" shall mean the public class or interface declarations whose names begin with "org.osgi" or any recognized successors or replacements thereof. The OSGi Alliance expressly reserves all rights not granted pursuant to these limited copyright licenses including termination of the license at will at any time.

EXCEPT FOR THE LIMITED COPYRIGHT LICENSES GRANTED ABOVE, THE OSGi ALLIANCE DOES NOT GRANT, EITHER EXPRESSLY OR IMPLIEDLY, A LICENSE TO ANY INTELLECTUAL PROPERTY IT, OR ANY THIRD PARTIES, OWN OR CONTROL. Title to the copyright in the Distribution will at all times remain with the OSGi Alliance. The example companies, organizations, products, domain names, e-mail addresses, logos, people, places, and events depicted therein are fictitious. No association with any real company, organization, product, domain name, email address, logo, person, place, or event is intended or should be inferred.

THE DISTRIBUTION IS PROVIDED "AS IS," AND THE OSGi ALLIANCE (INCLUDING ANY THIRD PARTIES THAT HAVE CONTRIBUTED TO THE DISTRIBUTION) MAKES NO REPRESENTATIONS OR WARRANTIES, EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, NON-INFRINGEMENT, OR TITLE; THAT THE CONTENTS OF THE DISTRIBUTION ARE SUITABLE FOR ANY PURPOSE; NOR THAT THE IMPLEMENTATION OF SUCH CONTENTS WILL NOT INFRINGE ANY THIRD PARTY PATENTS, COPYRIGHTS, TRADEMARKS OR OTHER RIGHTS.
NEITHER THE OSGi ALLIANCE NOR ANY THIRD PARTY WILL BE LIABLE FOR ANY DIRECT, INDIRECT, SPECIAL, INCIDENTAL OR CONSEQUENTIAL DAMAGES ARISING OUT OF OR RELATING TO ANY USE OR DISTRIBUTION OF THE DISTRIBUTION.

Implementation of certain elements of this Distribution may be subject to third party intellectual property rights, including without limitation, patent rights (such a third party may or may not be a member of the OSGi Alliance). The OSGi Alliance is not responsible and shall not be held responsible in any manner for identifying or failing to identify any or all such third party intellectual property rights.

The Distribution is a draft. As a result, the final product may change substantially by the time of final publication, and you are cautioned against relying on the content of this Distribution. You are encouraged to update any implementation of the Distribution if and when such Distribution becomes a final specification.

The OSGi Alliance is willing to receive input, suggestions and other feedback (“Feedback”) on the Distribution. By providing such Feedback to the OSGi Alliance, you grant to the OSGi Alliance and all its Members a non-exclusive, non-transferable,

worldwide, perpetual, irrevocable, royalty-free copyright license to copy, publish, license, modify, sublicense or otherwise distribute and exploit your Feedback for any purpose. Likewise, if incorporation of your Feedback would cause an implementation of the Distribution, including as it may be modified, amended, or published at any point in the future ("Future Specification"), to necessarily infringe a patent or patent application that you own or control, you hereby commit to grant to all implementers of such Distribution or Future Specification an irrevocable, worldwide, sublicenseable, royalty free license under such patent or patent application to make, have made, use, sell, offer for sale, import and export products or services that implement such Distribution or Future Specification. You warrant that (a) to the best of your knowledge you have the right to provide this Feedback, and if you are providing Feedback on behalf of a company, you have the rights to provide Feedback on behalf of your company; (b) the Feedback is not confidential to you and does not violate the copyright or trade secret interests of another; and (c) to the best of your knowledge, use of the Feedback would not cause an implementation of the Distribution or a Future Specification to necessarily infringe any third-party patent or patent application known to you. You also acknowledge that the OSGi Alliance is not required to incorporate your Feedback into any version of the Distribution or a Future Specification.

I HEREBY ACKNOWLEDGE AND AGREE TO THE TERMS AND CONDITIONS DELINEATED ABOVE.

0.2 Trademarks

OSGi™ is a trademark, registered trademark, or service mark of the OSGi Alliance in the US and other countries. Java is a trademark, registered trademark, or service mark of Oracle Corporation in the US and other countries. All other trademarks, registered trademarks, or service marks used in this document are the property of their respective owners and are hereby recognized.

0.3 Feedback

This document can be downloaded from the OSGi Alliance design repository at https://github.com/osgi/design The public can provide feedback about this document by opening a bug at https://www.osgi.org/bugzilla/.

0.4 Table of Contents

0.5 Terminology and Document Conventions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY" and "OPTIONAL" in this document are to be interpreted as described in 10.1.

```
Source code is shown in this typeface.
```

0.6 Revision History

The last named individual in this history is currently responsible for this document.

Revision	Date	Comments
Initial	March 13 2017	Initial version of the RFC
0.1	June 2017	Updates with feedback from Montpellier F2F

1 Introduction

More and more GPU hardware is used to accelerate computations in applications compared to running on a CPU. In an OSGi framework one might have multiple candidate bundles providing the same service and/or packages, out of which the resolver has to figure out which ones resolve on the framework. In order to know whether or not an OSGi framework has access to GPU hardware and a GPU accelerated bundle can be deployed, this RFC provides a namespace with capabilities and requirements.

2 Application Domain

Besides the straightforward applications of GPUs in video and audio processing applications, GPU acceleration is also used more and more in the domain of machine learning and big data processing.

3 Problem Description

In an OSGi framework one might have multiple candidate bundles providing the same service and/or packages, out of which the resolver has to figure out which ones resolve on the framework. For example, when using native code, the resolver will only resolve bundles containing native code builds for your system based on the osgi.native requirements and capabilities. However, when having native code that contains code to be executed on GPU hardware, e.g. using CUDA or OpenCL APIs, there is currently no standard way to figure out which GPU hardware and API is available on the system. This RFC provides a namespace with capabilities and requirements that allow implementors to declare GPU capabilities of the hardware, as well as GPU requirements for bundles that contain native code compiled for specific GPUs.

4 Requirements

[GPU0001] The solution must declare a namespace for GPU related requirements and capabilities.

[GPU0002] The presence of a GPU has to be announced by means of an OSGi capability.

[GPU0004] Bundles can declare requirements on GPU attributes based on the native code they contained

[GPU0005] GPU requirement and capabilities might contain GPU vendor specific attributes

5 Technical Solution

To support the declaration of GPU capabilities and requirements, we introduce the osgi.native.gpu namespace to be added in Chapter 135 – Common namespaces.

To announce the presence of a GPU, a bundle can poll the hardware and provide a GPU capability:

```
Provide-Capability: osgi.native.gpu;

  osgi.native.gpu.vendor=”NVIDIA”;

  osgi.native.gpu.model=”GTX 980”;

  osgi.native.gpu.memory=”1024”

  com.nvidia.cuda.version:List<Version>=”2.0,2.1,2.2,2.3,3.0,3.1,3.2,4.0,4.1,4.2,5.0
,5.5,6.0,6.5,7.0,7.5,8.0”

  com.nvidia.computecapability:Version=”7.0”
```

A bundle containing native GPU code, can provide a requirement on the GPUs where it can run on (i.e. for which GPUs and APIs it was compiled):

```
Require-Capability: osgi.native.gpu;

filter:=(&(osgi.native.gpu.memory>=2048)(com.nvidia.cuda.version =”7.5”))
```

6 Data Transfer Objects

7 Javadoc

8 Considered Alternatives

9 Security Considerations

No new security considerations are required by this design

10 Document Support

10.1 References

[1]. Bradner, S., Key words for use in RFCs to Indicate Requirement Levels, RFC2119, March 1997.

[2]. Software Requirements & Specifications. Michael Jackson. ISBN 0-201-87712-0

10.2 Author's Address

Name	Tim Verbelen
Company	Ghent University - imec
Address	
Voice	
e-mail	tim.verbelen@ugent.be

10.3 Acronyms and Abbreviations

10.4 End of Document