

## Presentation Abstract

# Detection of Layered Bottlenecks in Microservices

TATSUSHI INAGAKI<sup>1,a)</sup> YOHEI UEDA<sup>1,b)</sup> MORIYOSHI OHARA<sup>1,c)</sup> TATSUHIRO CHIBA<sup>1,d)</sup> QI ZHANG<sup>2,e)</sup>

Presented: July 21, 2021

This presentation proposes a method to detect layered bottlenecks in microservices. A bottleneck in a service is a resource which limits the maximum performance of the service. Bottlenecks in software resources such as threads, locks, and channels, and hardware resources such as processors and disks, form a layered structure since a single request to the service can use multiple software resources simultaneously. Detection of such layered bottlenecks should be agile since development and operation of an application are integrated and agile. We detect layered bottlenecks in microservices by constructing a weighted dependency graph of microservices. A node represents a microservice and is weighted by the number of requests calculated from custom performance metrics. A link represents dependency between microservices which is extracted from connectivity between microservices. We detect layered bottlenecks in standard microservice benchmarks deployed on a container orchestration platform by applying the proposed method. We demonstrate that mitigating the bottlenecks improves the performance of the benchmarks.

---

This is the abstract of an unrefereed presentation, and it should not preclude subsequent publication.

<sup>1</sup> IBM Research – Tokyo, IBM Japan, Ltd., Chuo, Tokyo 103–8510, Japan

<sup>2</sup> IBM Research –Thomas J. Watson Research Center, Yorktown Heights, NY 10598

a) e29253@jp.ibm.com

b) yohei@jp.ibm.com

c) ohara@jp.ibm.com

d) chiba@jp.ibm.com

e) Q.Zhang@ibm.com