

ABSTRAK

Pada era digitalisasi ini *video conference* adalah hal yang umum dalam masyarakat. Teknologi ini sangat membantu manusia dalam menjalankan aktivitasnya sehari-hari. Karena begitu banyaknya *user* yang mengakses, maka tidak jarang juga layanan ini mengalami *downtime*. Salah satu penyebab *downtime* adalah *resource* pada *server* yang telah habis digunakan.

Solusi dalam mengatasi *downtime* adalah dengan menggunakan infrastruktur yang dibangun dengan sebuah *container orchestration tool* Bernama Kubernetes. Kubernetes Cluster akan dijalankan diatas Linode yang berfungsi sebagai *Cloud Service*. Kubernetes menjalankan *service* di dalam *node*. Terdapat komponen terkecil di dalam node yaitu *pod*.

Penelitian ini bertujuan untuk melakukan uji coba layanan menggunakan Kubernetes yang bertugas untuk mengelola *cluster* yang terdapat pada server. *Micro Kubernetes cluster* adalah pengembangan terbaru dari *Kubernetes* yang memiliki ukuran lebih kecil dan lebih cepat dalam mengelola cluster. Dengan *micro Kubernetes cluster* akan dibangun layanan *Video Conference*.

Service yang diimplementasikan adalah WebRTC. *Service* ini akan diuji dan dilakukan perbandingan dengan beberapa parameter. Parameter yang diperhatikan adalah *CPU Usage*, *response code*, *response time*, dan *throughput*. Berdasarkan parameter tersebut, maka didapat hasil penelitian yaitu *service* yang dijalankan di arsitektur monolitik memiliki hasil *response time* yang lebih baik dengan nilai 2668 ms, *response code* sebesar 106, dan *CPU Usage* pada rentang 35 - 68 %. *Service* yang dijalankan pada arsitektur *micro Kubernetes cluster* dengan fitur HPA (*Horizontal Pod Autoscaler*) memiliki *CPU Usage* yang lebih stabil karena adanya *load balancing* yang terjadi dengan mereplikasi *pod* oleh HPA.

Kata Kunci: *Kubernetes, High Availability, Cluster, WebRTC.*