CLAIMS

What is claimed is:

1. A computer system comprising:

a processing unit operatively coupled to memory;

the processing unit operatively coupled with an artificial intelligence (AI) platform to orchestrate a multi-objective optimization of a cloud application, the AI platform comprising:

an input manager to receive a set of two or more key performance indicators (KPIs) associated with the application, each KPI in the set of KPIs associated with a KPI value;

a trial manager to leverage a machine learning (ML) based surrogate function learning model in combination with an acquisition function to conduct an adaptive trial execution of the application with one or more parameters associated with one or more topological constraints, wherein the one or more topological constraints include a hierarchical arrangement of micro-services;

the trial manager to observe output from each trial execution and to compute a pareto surface of the observed output;

an optimization manager to explore the computed pareto surface and select an optimal application operating point corresponding to configuration of the one or more parameters; and

the optimization manager to execute the application at the selected optimal operating point.

2. The computer system of claim 1, wherein the ML based surrogate function learning model maps each configuration of the one or more parameters to the two or more KPI values.

3. The computer system of claim 2, wherein the ML based surrogate function learning model is implemented using a neural network, Gaussian Progress regression, kernel regression techniques, or a combination thereof.

4. The computer system of claim 1, wherein the acquisition function selects a second configuration of one or more second parameters for a second trial execution of the application.

5. The computer system of claim 4, where the acquisition function is implemented as an expected improvement function, maximum probability of improvement function, upper confidence bound function, or a combination thereof.

6. The computer system of claim 1, wherein the set of two or more KPI values comprises latency, error rate, throughput, or a combination thereof, and wherein the one or more parameters associated with the one or more topological constraints comprises timeout, retry, replication count, or a combination thereof.

7. The computer system of claim 6, wherein the hierarchical arrangement of micro-services and the one or more parameters are set with respect to a hierarchy, including ensuring that the one or more parameters guarantees a query from a parent micro-service in the hierarchy is active as long as one or more queries from a child micro-service is active.

8. A computer program product to orchestrate a multi objective optimization of cloud applications, the computer program product comprising a computer readable storage medium having program code embodied therewith, the program code executable by a processor to:

receive a set of two or more key performance indicators (KPIs) associated with an application, each KPI in the set of KPIs associated with a KPI value;

leverage a machine learning (ML) based surrogate function learning model in combination with an acquisition function to conduct an adaptive trial execution of the application with one or more received parameters associated with one or more topological constraints, wherein the one or more topological constraints include a hierarchical arrangement of micro-services;

observe output from each trial execution and compute a pareto surface of the observed output;

explore the computed pareto surface and select an optimal application operating point corresponding to configuration of the one or more parameters; and

execute the application at the selected optimal operating point.

9. The computer program product of claim 8, wherein the ML based surrogate function learning model maps each configuration of the one or more parameters to the two or more KPI values.

10. The computer program product of claim 9, wherein the ML based surrogate function learning model is implemented using a neural network, Gaussian Progress regression, kernel regression techniques, or learning combination thereof.

11. The computer program product of claim 8, wherein the acquisition function selects a second configuration of one or more second parameters for a second trial execution of the application.

12. The computer program product of claim 11, where the acquisition function is implemented as an expected improvement function, maximum probability of improvement function, upper confidence bound function, or a combination thereof.

13. The computer program product of claim 8, wherein the set of two or more KPI values comprises latency, error rate, throughput, or a combination thereof, and wherein the one or more parameters associated with the one or more topological constraints comprises timeout, retry, replication count, or a combination thereof.

14. The computer program product of claim 13, wherein the hierarchical arrangement of micro-services and the one or more parameters are set with respect to a hierarchy, including ensuring that the one or more parameters guarantees a query from a parent micro-service in the hierarchy is active as long as one or more queries from a child micro-service is active.

15. A computer implemented method comprising:

receiving a set of two or more key performance indicators (KPIs) associated with an application, each KPI in the set of KPIs associated with a KPI value;

leveraging a machine learning (ML) based surrogate function learning model in combination with an acquisition function to conduct an adaptive trial execution of the application with one or more received parameters associated with one or more topological constraints, wherein the one or more topological constraints include a hierarchical arrangement of micro-services;

observing output from each trial execution and computing a pareto surface of the observed output;

exploring the computed pareto surface and selecting an optimal application operating point corresponding to configuration of the one or more parameters; and

executing the application at the selected optimal operating point.

16. The method of claim 15, wherein the ML based surrogate function learning model maps each configuration of the one or more parameters to the two or more KPI values.

17. The method of claim 16, wherein the ML based surrogate function learning model is implemented using a neural network, Gaussian Progress regression, kernel regression techniques, or a combination thereof.

18. The method of claim 15, wherein the acquisition function selects a second configuration of one or more second parameters for a second trial execution of the application.

19. The method of claim 18, where the acquisition function is implemented as an expected improvement function, maximum probability of improvement function, upper confidence bound function, or a combination thereof.

20. The method of claim 15, wherein the set of two or more KPI values comprises latency, error rate, throughput, or a combination thereof, and wherein the one or more parameters associated with the one or more topological constraints comprises timeout, retry, replication count, or a combination thereof.