

Technical Disclosure Commons

Defensive Publications Series

May 03, 2018

WOPTIMOP - A Cloud Based Intelligent Method Of Automatically Creating And Delivering Workload Optimized Platform Contents

Suprateeka R Hegde
Hewlett Packard Enterprise

Sujoy Saraswati
Hewlett Packard Enterprise

Amrita H S
Hewlett Packard Enterprise

Follow this and additional works at: https://www.tdcommons.org/dpubs_series

Recommended Citation

Hegde, Suprateeka R; Saraswati, Sujoy; and H S, Amrita, "WOPTIMOP - A Cloud Based Intelligent Method Of Automatically Creating And Delivering Workload Optimized Platform Contents", Technical Disclosure Commons, (May 03, 2018)
https://www.tdcommons.org/dpubs_series/1181



This work is licensed under a [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).

This Article is brought to you for free and open access by Technical Disclosure Commons. It has been accepted for inclusion in Defensive Publications Series by an authorized administrator of Technical Disclosure Commons.

WOPTIMOP - A Cloud Based Intelligent Method Of Automatically Creating And Delivering Workload Optimized Platform Contents

Abstract

UNIX/Linux being the most preferred enterprise operating platform, provides state-of-art compiler and developer tool chain for almost all programming languages. And so far, has greatly helped enterprise application developers to optimize their applications for their specific workloads, through various set of options and tunables.

However, the solution is still partial, as only the application is optimized for the workload. The Operating System and other components of the platform, though allows tuning and customization, is common for all. Hence each and every application and workload being characteristically different, cannot benefit much from such operating platforms.

Here we propose a unique innovative solution that provides an Operating Platform completely optimized for a specific enterprise application and workload. The solution is in the form of a cloud based Content Delivery Network (CDN) which internally uses machine learning algorithms to intelligently exercise Profile Based Optimization (PBO)[2] and Runtime Architecture Diagnostics (RTADIAG)[1], and then deliver workload optimized operating platform.

Problem Statement

Performance and security are two of the most critical needs of any enterprise application or service. It is imperative for UNIX/Linux operating platform, being the most preferred in enterprise industry, to provide performance and security to the maximum possible extent.

However, performance being a characteristic of each application, cannot be one for all. In other words, each application, for each different workload, demands customized performance tuning. Research and development in the field of compilers and developer tool chain, have solved this customization problem partially, through Profile Based Optimization (PBO)[2] technique. This technique improves application performance for the workload it handles.

However, the operating system and other components of the platform as a whole, is not optimized for the application and the workload. The Operating Platform is one and common for all enterprise deployments. For instance, Operating Platforms like HP-UX or Linux are available as one release for all; they are not available in customized/tailor-made forms for different consumers in enterprise market.

The invention proposed here solves this problem and provides an intelligent solution for enterprise consumers to get a customized version of the entire Operating Platform that provides performance and security optimized for their application and specific workloads.

Our solution

The invention described here primarily consists of a specialized Content Delivery Network (CDN) through which subscribed enterprise customers can get the platform specific software optimized for their workload. The CDN internally comprised of Machine Learning (ML) based Profile Guided Optimization (PGO/PBO) [4].

Following three key aspects triggered the invention described:

1. Any enterprise solution or service provider would expect the best price-performance ratio from the hardware where they run their workload

2. The workload and its characteristics differ from one enterprise to another. Often, the difference is very significant.

3. Today, almost all enterprise server (hardware or software) makers, provide Operating System and software components, which is common to all -- one release across the world. Options to customize and tune are indeed there, but these are again common to all and in reality, results in very minimal gain.

The invention here describes a CDN system to provide software components that are extremely tailor-made and optimized to specific workload characteristics of each different enterprise. As mentioned above, the ML and PGO engines are the core components involved.

As described in the Diagram-1, WOPTIMOP system is designed and constructed as follows:

- A secure portal that provides web based tools is given to enterprise customers who would like to subscribe to the WOPTIMOP service.
- The request is then received by the WOPTIMOP service provider
- The MLIS server, then sends an initial platform with all software components which are instrumented for profiling.
- The customer then uses the instrumented platform to run their workload. At the end of the exercise, this produces all data required by the WOPTIMOP system and the data is uploaded back to the WOPTIMOP system from the customer's server. This data does not contain any information or data related to customer. This contains data purely specific to platform provider. This makes it easy for the customer to safely upload the data back to WOPTIMOP system.
- The MLIS server, then uses the data produced by instrumented platform, and generates the final platform content that is Profile Guided Optimized. In addition to the PGO, the MLIS server also detects various characteristics of the workload and uses right set of compiler-n-toolchain [2] options to build the optimized platform and software components. This platform is now tailor-made to the characteristics of the customer's workload. The platform in its entirety is profile guided optimized. That means, the shared libraries, the executables from various software packages, the kernel [3] itself, all are PGO'ed.
- The MLIS server then sends a report of all that has been done to the Labs/Developers of the RnD division of WOPTIMOP system, for review. The MLIS server provides options to the RnD engineers to modify decisions taken by the MLIS server, in which case, the MLIS server rebuilds the platform with RnD inputs too.
- Once the final optimized platform is ready, an e-Mail is sent to the customer to notify them about the availability of the "Workload OPTIMized Operating Platform".
- The customer then downloads the platform and installs them on their production server after testing. This provides the best price-performance ratio for the hardware associated with the platform for the specific workload.
- The MLIS server also maintains data on what went wrong and what went good in various cases. The MLIS server continually passes this information through the ML engine to improve itself over a period of time.

Some insights into internal algorithm of WOPTIMOP to provide deeper construction details:

- As an example, when MLIS detects that the workload is memory intensive, the MLIS uses compiler options to optimize memory accesses, reduce cache penalties, etc. Similarly, if MLIS

detects if workload is CPU intensive, it uses compiler options that optimizes loops, uses hardware specific instructions, etc. These are some simple examples of what MLIS does.

- The MLIS also uses the invention of RTADIAG [1] to get runtime diagnostics and improve hard-to-detect aspects.
- Since different customers may have different supported versions of the platform, the MLIS server maintains indexes into code sources corresponding to each released version, so that it can build the platform precisely.
- The ML engine uses algorithms like linear regression, logistic regression, support vector machines (SVM), etc.

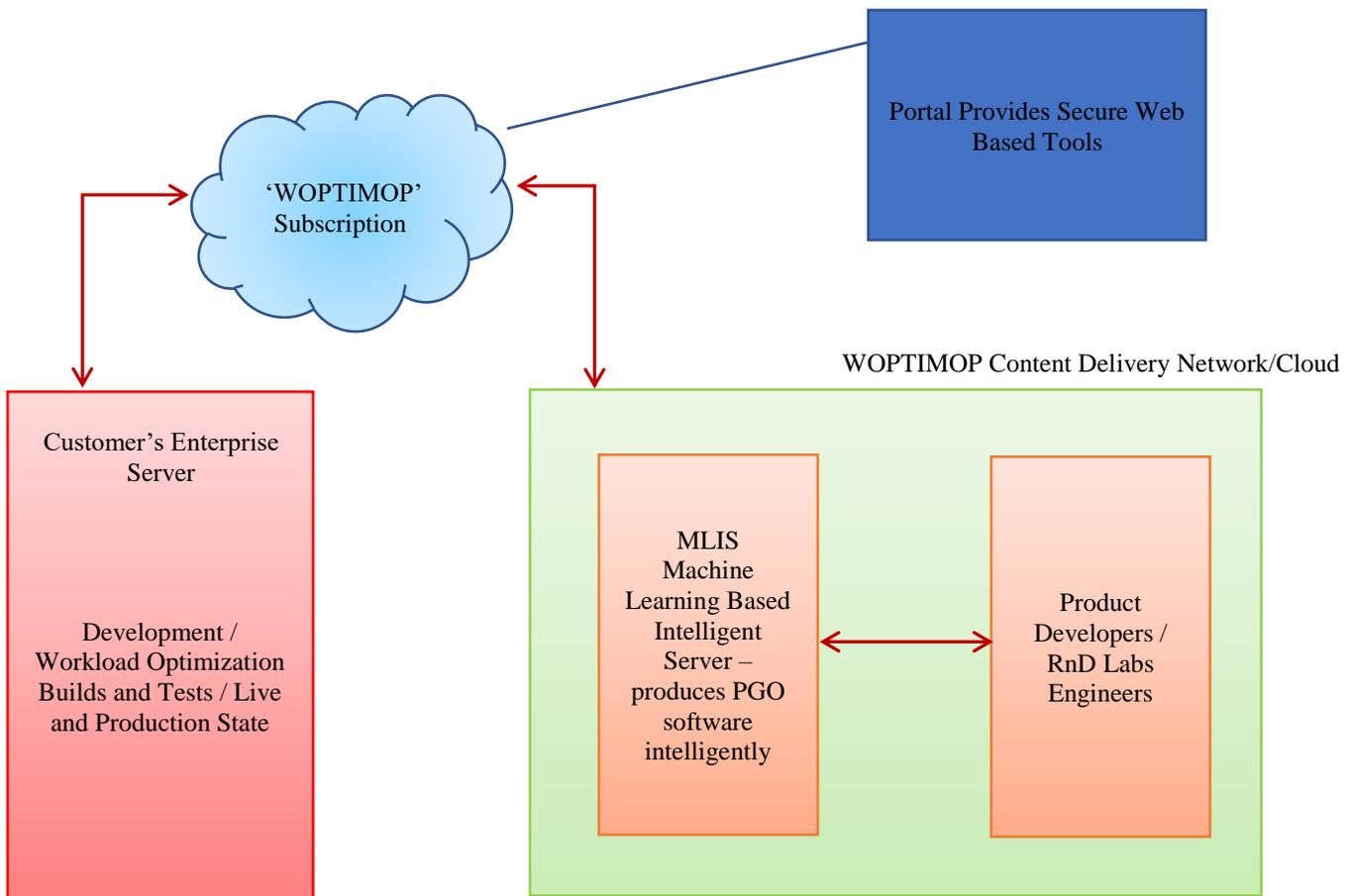


Diagram-1

Prior Solutions

All prior solutions and attempts to customize the operating platform for security and performance optimization of workload, was made through a few operating system kernel flag/tunable and/or configuration settings.

Advantages

In none of the prior solutions, the code of the operating system or components were optimized specially for a specific workload. All operating platforms were/are available as one global release.

There is a significant advantage from invention described here as it enables enterprise application developers to:

1. get a custom operating platform optimized for their specific workload
2. get benefits of machine learning based continuous improvement of the operating platform
3. get security end-to-end and performance optimized for the workload

References

[1] RTADIAG - Runtime Architecture Based Diagnostics

[http://www.tdcommons.org/dpubs_series/491/]

[2] HP-UX PGO

[https://support.hpe.com/hpsc/doc/public/display?docId=emr_na-c02936806]

[3] PGO Kernel Performance Improvement

[http://www.usenix.net/legacy/publications/library/proceedings/sf94/full_papers/partridge.pdf]

[4] Profile Guided Optimization

[https://en.wikipedia.org/wiki/Profile-guided_optimization]

Disclosed by:

Suprateeka R Hegde – Hewlett Packard Enterprise

Sujoy Saraswati – Hewlett Packard Enterprise

Amrita H S – Hewlett Packard Enterprise