



# Exotanium DOE SBIR Phase I Results Summary

March 2022

*Changing the World's Energy Future*

Taylor Reo Ashbocker, Nathan Lee Woodruff, Christopher S Ritter



**DISCLAIMER**

This information was prepared as an account of work sponsored by an agency of the U.S. Government. Neither the U.S. Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness, of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. References herein to any specific commercial product, process, or service by trade name, trade mark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the U.S. Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the U.S. Government or any agency thereof.

# **Exotanium DOE SBIR Phase I Results Summary**

**Taylor Reo Ashbocker, Nathan Lee Woodruff, Christopher S Ritter**

**March 2022**

**Idaho National Laboratory  
Idaho Falls, Idaho 83415**

**<http://www.inl.gov>**

**Prepared for the  
U.S. Department of Energy  
Under DOE Idaho Operations Office  
Contract DE-AC07-05ID14517**



# Exotanium DOE SBIR Phase I Results Summary

March 2022

Taylor Ashbocker  
Nathan Woodruff  
Christopher Ritter  
*Exotanium Inc.*



*INL is a U.S. Department of Energy National Laboratory  
operated by Battelle Energy Alliance, LLC*

**DISCLAIMER**

This information was prepared as an account of work sponsored by an agency of the U.S. Government. Neither the U.S. Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness, of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. References herein to any specific commercial product, process, or service by trade name, trade mark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the U.S. Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the U.S. Government or any agency thereof.

## CONTENTS

1.	INTRODUCTION.....	1
2.	EXPERIMENT AND RESULTS.....	1
3.	NEXT STEPS.....	2

## ACRONYMS

AWS	Amazon Web Services
GPU	graphics processing unit
HPC	high-performance computing
INL	Idaho National Laboratory
PCI	peripheral component interconnect
SOCA	Scale-Out Compute on AWS
VM	virtual machine

# Exotanium DOE SBIR Phase I Results Summary

## 1. INTRODUCTION

Exotanium has developed novel technologies that dynamically allocate idle or over-sized application containers and virtual machines (VMs) to take advantage of deeply discounted server space. The technology addresses the cost-saving needs of its customer segments by leveraging spot market discounts (X-Spot), and further, consolidating idle workloads and over-sized application containers and VMs (X-Consolidate).

Exotanium developed a strategic platform for these technologies and demonstrated results in partnership with cloud architects at Idaho National Laboratory (INL).

The first product, X-Spot, spawns application containers on discounted VMs and orchestrates containers between such instances based on availability and price. Spot instances can be reclaimed by the cloud provider at any time with very short notice; so traditionally, containers can only use these instances for workloads that tolerate interruption or failure. X-Spot dramatically improves the spot-instance usability for critical applications, however, with comparable reliability and cost reductions upwards of 70% compared to regular on-demand, or native, instances.

In a demonstration for INL, these cloud strategies delivered a set of compelling results, from cost-savings on cloud resources as high as 73%, with minimal runtime overhead, and even in some cases, performance improvements as high as 4.2%. Further, the continuous management and reallocation of containers establishes high reliability for application workloads.

Exotanium demonstrated this technology with the INL MASTODON application, a Multiphysics environment designed to run typical high-performance computing (HPC) simulations for structural dynamics, seismic analysis, and risk assessment. The MASTODON application was packaged into a container using Docker, Deployed on Amazon Web Services (AWS) ECS, and managed through a custom Scale-Out Compute on AWS (SOCA) implementation.

## 2. EXPERIMENT AND RESULTS

The experimental setup utilized a master and worker node configuration on AWS with the m5.large and m5.8xlarge instance types, respectively. This worker node is configured with 32 cores and 128 GiB memory. Both X-Spot and X-Spot Migrate strategies were tested, where the Migrate strategy is distinct for its use of hybrid cloud environments. Additionally, this research has demonstrated the orchestration of containers between multiple cloud providers, at one time beginning a workload on the Google Cloud Platform and ending on AWS.

This experiment produced results in three separate trials, each involving a MASTODON workload of varying intensity. Results are determined by the difference between native and spot-instance runtime and pricing.

First, an X-Spot trial involving a 1-hour workload yielded a 70% cost reduction while forfeiting only a 6% overhead. For X-Spot Migrate, the results were similar with a 69% reduction at only a 7% overhead.

<b>1 Hour Workload - MASTODON</b>		
<b>Instance Strategy</b>	<b>Relative Runtime</b>	<b>Relative Cost</b>
Native	1	1
X-Spot	1.06	0.30
X-Spot Migrate	1.07	0.31



Next, an X-Spot trial involving a 4-hour workload yielded a 73% cost reduction with a 4.2% performance enhancement. For X-Spot Migrate, the results were similar with a 72% reduction and a 3.5% performance enhancement.

<b>4 Hour Workload - MASTODON</b>		
<b>Instance Strategy</b>	<b>Relative Runtime</b>	<b>Relative Cost</b>
Native	1	1
X-Spot	.958	0.27
X-Spot Migrate	.965	0.28

Finally, an X-Spot trial involving a 16-hour workload yielded a 73% cost reduction with a 2.8% performance enhancement. For X-Spot Migrate, the results were similar with a 72% reduction and a 2.7% performance enhancement.

<b>16 Hour Workload - MASTODON</b>		
<b>Instance Strategy</b>	<b>Relative Runtime</b>	<b>Relative Cost</b>
Native	1	1
X-Spot	.962	0.27
X-Spot Migrate	.990	0.28

Further, this research extended the X-Spot platform to support intensive workloads like image processing and machine learning, by creating a peripheral component interconnect (PCI) pass-through to a graphics processing unit (GPU)-based instance. This support enables large, highly compute-intensive Department of Energy applications to reliably run uninterrupted while taking advantage of the spot market's significant discounts.

Exotanium also developed X-Consolidate, which addresses idle resource waste in the cloud by packing idle containers onto a small number of VMs during the idle period, thereby minimizing the number of active machines and reducing the cost of keeping services online. When the workload increases, X-Consolidate relocates containers onto different VMs, without any service interruption. This enables the ability of stateful applications, like databases or streaming processors, to automatically scale with the workload, reducing the need to over-provision resources to accommodate their peak workloads.

### **3. NEXT STEPS**

The next steps include the following. First, expanding the supported cloud environments to include the GovCloud. This will enable more National Laboratory workloads to run using Exotanium's software products. Second, expand to Microsoft Azure, which will allow more public clouds to be used. Third, evaluating support of interactive workloads where optimizations can be significant, such as reducing the resource consumption and costs of servers that are idle. Lastly, evaluating the feasibility of supporting GPUs based migration that will allow accelerators to benefit from the cloud optimizations that Exotanium provides.