The Advantages and Pitfalls of Implementing SAS® in an Amazon Web Services Cloud Instance

Jeff Lehmann, Slalom Consulting, LLC

ABSTRACT

The implementation of the best in breed advanced analytics software, SAS®, on a cloud instance managed by the largest cloud service competitor, Amazon Web Services (AWS), is likely to create a compelling solution. This paper nonetheless will describe not just the positive aspects of such a combination but will delineate any potential issues and concerns.

INTRODUCTION

As the analytics function using SAS® grows and matures in many enterprises around the world, many of these same enterprises are seeking relevant analytics competitive advantages. One SAS® administration topic worthy of investigation and suitable for new analytics investments that outperform would be implementing and migrating to an instance of SAS® using a cloud service provider such as AWS.

This paper analyzes both the benefits and the downsides that accrue to such a cloud implementation scenario. The objective is to determine the likelihood of a high return on investment leading to a competitive advantage in analytics. This paper is divided very simply into two sections:

1. Advantages to implementing SAS® in an AWS instance

2. Pitfalls and concerns for implementing SAS® in an AWS instance

In order to emphasize the prospective positive returns over any possible negative, undesirable effects, IT Managers, SAS administrators and SAS users need to reach a consensus on the suitability of an investment of time and labor effort in a new SAS® cloud instance implementation. This paper intends to be useful in such a decision-making process.

ADVANTAGES TO IMPLEMENTING SAS® IN AN AWS INSTANCE

Cloud service providers such as Amazon continue to "commodify" the process to create a server instance. Choosing a desirable OS and applying sizing requirements for a SAS implementation has been simplified. As a result, a definite advantage is the speed and simplicity of getting started with a SAS installation. Planning horizons are non-existent and initial financial outlay is economized since no server hardware procurement occurs, no data center space reserved nor any hardware/OS engineers assigned to participate in the initial server instance creation. The cloud infrastructure seems to make the OS irrelevant, an afterthought and even just an extension of SAS software. In addition, if the initial sizing, memory allocation or disk space selection results later in some deficiency or errors in SAS processing, the flexibility of the virtual server instance allows for the instance image to be saved and restored to a new larger or performance-enhanced instance at relatively low cost and minor inconvenience for production users.

Once logged on with an authenticated ID, internet connectivity established, a sas installer ID created and a web browser started, it's just a matter of downloading the SAS download manager to begin the creation of the SAS software depot. Many Amazon cloud instances have download speeds that tend to be greater than data center-based solutions and this results in shorter processing time to create the depot. Installing SAS via the Deployment Wizard is not dissimilar on a cloud instance versus a server instance and all the same challenges (e.g. SSL, authentication and single-sign on, repository migration) apply. Overall, SAS administrators have an optimal, straight-forward and low-cost

opportunity to deploy additional SAS instances running different versions or more complex configurations (e.g. SAS Grid, resource-based load balancing, SAS jobs split and run parallel across multiple nodes).

In essence, opting for an AWS solution means that the enterprise is opting to trade capital expense for variable expense. With traditional infrastructure, you have to guess how much capacity you will need over the next 3-7 years, and pay for most or all of it on day one. This ties up budget resources that could be allocated for other things. Moreover, it's very difficult to forecast capacity requirements accurately over such a long period of time. With AWS, you pay only for what you need, when you need it. With AWS, you don't need to guess future capacity needs, since you can utilize additional capacity during busy periods and relinquish it when it is no longer needed.

What follows is a compiled list [1] of cloud advantages to be considered for a SAS® implementation:
.
1. Commodity

2. Massive economies of scale

3. Globally accessible

4. Appliance aspect

5. Limited planning horizon required

6. Small initial financial outlay

7. No required data center space

8. No hardware/OS expert staffed

9. OS as an extension of the application

10. Effortless sizing and resizing

11. Download speeds

12. Short implementation period

13. Straight-forward

14. Low cost and industry competition has led to price reductions

15. Momentum for additional deployments

16. Cost efficiency

17. Convenience and continuous availability

18. Availability

19. Backup and recovery

20. Cloud is environmentally friendly

21. Resiliency and redundancy

22. Scalability and performance

23. Quick deployment and ease of integration

24. Increased storage capacity

25. Device diversity and location independence

26. Smaller learning curve

27. Speedy and agile implementation

28. Hybrid solutions of on-premises resources integrating with cloud resources

29. Flexible purchasing options

PITFALLS AND CONCERNS FOR IMPLEMENTING SAS® IN AN AWS INSTANCE

While the main advantages of using a cloud instance to deploy a new SAS implementation tend to revolve around efficiency, speed and affordability, its prospective pitfalls have to do with the customer's shared responsibility regarding cloud instance security. The same easy, low-cost server instance launch process also has a potential negative flip side: a customer's OS inexperience and lack of oversight can lead to vulnerable security schemes. Because you're building systems on top of the AWS cloud infrastructure, the security responsibilities will be shared: AWS has secured the underlying infrastructure and you must secure anything you put on the infrastructure or connect to the infrastructure. The amount of security configuration work you have to do varies depending on how sensitive your data is and which services you select.

Just like any other data center-based infrastructure, AWS cloud infrastructure faces similar vulnerability to external attacks, thereby highlighting the importance of "shared" security implementations. Only very experienced or trained Linux admins can implement the appropriate security precautions to prevent, for example, the spread of the Shellshock vulnerability for web servers. As well, these same engineers typically are the only knowledgeable staff that know how to patch their physical and virtual systems to remediate against a Shellshock vulnerability intrusion. To avoid the ruinous, targeted attacks where even credentials can be compromised, potential SAS customers must staff their SAS 9.4 AWS implementations with precisely the same experienced engineers that they would staff similar data-center based initiatives. Only under this staffing arrangement will SAS AWS customers meet their security shared responsibilities.

Extra steps have to be taken to prevent the aforementioned attacks and fortunately, there are cloud-based methods available. By creating a Virtual Private Cloud (VPC) instance, AWS users can restrict access by originating IP address while also requiring additional administration including have to create entries for application ports that require external access. Moreover, with each step towards more secure cloud implementations, there are additional complexities that arise, including having to make additional changes/compromises with corporate firewall policy and user authentication methods. Fortunately, enterprises can utilize AWS virtual private gateways to create a VPN connection between the VPC and an external corporate network to further enhance authentication.

In spite of the reported vulnerabilities, an indication that AWS is up to the security challenge is that both the US federal government and the CIA have adopted and implemented AWS instances in what AWS terms "government cloud" and "CIA cloud".

Below appears a catalogued list of some concerns about an AWS implementation:

1. Sensible security precautions are required to limit vulnerability to intrusions and external attacks

2. Billing is based upon past use which requires tracking and oversight

3. Added corporate firewall and authentication complexity

CONCLUSION

Due to the prospective cost savings and speed of implementation of SAS® on an AWS cloud instance, enterprises using SAS® should immediately begin budgeting for proof of concept (POC) trial uses of AWS, such as spinning up an AWS test instance of SAS®. On the other hand, migrating a successful production data center-based instance of SAS® to an AWS instance is a more complex decision and will require further independent study to review each production scenario on a case-by-base basis.

REFERENCES

[1] ”Advantages and Disadvantages of Cloud Computing – Cloud computing pros and cons” Available at: http://www.javacodegeeks.com/2013/04/advantages-and-disadvantages-of-cloud-computing-cloud-computing-pros-and-cons.html

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name:	Jeff Lehmann
Enterprise:	Slalom Consulting
E-mail:	jeffl@slalom.com
Twitter:	@jefflehmann

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.