

## Should You Move Your SAS Into the Cloud?

Paul Janssen, Janssen Consulting Inc.

### ABSTRACT

Some people should not move their SAS 9.4 into the cloud, and those who do, may run into difficulties. IT departments and consulting firms are learning cloud technology on the fly. Improper installations may cause SAS to run at one third of the speed it could run compared to on-premises. Furthermore, cloud technology can be expensive, and may cost hundreds of thousands of dollars per year more for a very large install, compared to keeping your SAS on-premises.

This paper was written for SAS programmers and IT departments alike. You want to move your SAS 9.4 installation into the cloud, and do it right... the first time! Choosing the most appropriate cloud compute instance type and model, and designing a proper cloud storage configuration can be challenging; done incorrectly, you will have a sub-standard or poorly performing SAS installation, resulting in a competitive disadvantage for your organization and frustrating your SAS users.

This paper defines terms, explains the bread-and-butter components of cloud technology, talks about SAS and Amazon Web Services architecture, covers SAS resource requirements, and presents a methodology for putting your SAS into the cloud, when it makes sense to do so, while being mindful of performance caveats. SAS programmers will gain a better understanding of the IT technologies required to run SAS, and IT professionals will gain better insights regarding SAS performance requirements and how to handle the cloud situation.

SAS has always lived and died on disk I/O throughput<sup>1</sup>, and moving it into the cloud will not change that. If you do not get SAS I/O right, nothing else is going to work properly. This paper will emphasize ways to solve the I/O dilemma, so that a successful SAS cloud deployment is in the cards.

### INTRODUCTION

Cloud is the future, and public cloud adoption is growing at staggering double-digit rates. Gartner calls it the *cloud gold rush*<sup>2</sup>, and organizations that do not join the “cloud first” frenzy risk falling behind. However, cloud adaptation is tricky due to the complexities of cloud technology; it can be more costly (contrary to popular belief); and a lift-and-shift approach to migrating existing applications may backfire due to significant differences between on-premises *systems engineering* and public cloud *software engineering* (hint: you may need to redesign your on-premises application and underlying infrastructure to function properly in the public cloud). Adopting the public cloud, if done incorrectly, can be costly and cause the organization to fall behind as well.

So where does SAS fit into this picture? SAS 9.4 has resource requirements that can be tricky to meet in any environment, especially disk I/O. So how do you properly allocate, manage and monitor CPU, RAM and disk I/O resources in your SAS installation? And how does all that change when you move your SAS installation into the cloud? This paper will address those questions and can help you avoid costly mistakes by covering the following topics:

- **Cloud Technology:** An introduction to cloud technology that covers compute instances and storage options, and defines common cloud terms in the process.
- **SAS and Amazon Web Services Architectures:** An explanation of the essence of SAS 9.4 server

---

<sup>1</sup> Disk input/output (I/O) is simply the act of reading data from disk into computer memory (where the data is processed by the Central Processing Unit or CPU), and writing modified data from computer memory back to disk. Due to the large volume of data that is typical of analytics processing, disk I/O must be sufficiently fast to allow SAS to process jobs that access that data within an acceptable amount of time.

<sup>2</sup> <https://www.gartner.com/smarterwithgartner/7-hidden-cloud-growth-opportunities-for-technology-service-providers/>

topology and Amazon Web Services data center structures such as regions and availability zones.

- **SAS Resource Requirements:** An overview of the resources (CPU, RAM, and disk I/O throughput) that are needed by SAS.
- **Putting SAS in the Cloud – A Methodology:** A methodology to select the most appropriate cloud components to run your SAS installation, and a review of common risks and how to mitigate them.
- **Conclusion.** A summary of what has been discussed in this paper, with final advice.

## CLOUD TECHNOLOGY

A public cloud can be defined as a data center<sup>3</sup> that is managed by a cloud services provider and that makes computing power and data storage available to users over the internet. Larger public clouds typically consist of multiple interconnected data centers, sometimes all over the world. Cloud environments are scalable, meaning that they can be sized up and down manually and automatically to accommodate fluctuating changes in demand, and to accommodate long-term growth (hence the use of the term “elastic” in conjunction with many cloud technology components). Since the cloud provider takes care of the physical equipment that runs the cloud software, customers do not have to worry about hardware maintenance contracts or periodic hardware refresh (replacement) cycles. Through the sharing of resources in a cloud environment, cloud providers can reduce cost due to economies of scale; however, it should be noted that for high-performance applications such as SAS, a cloud deployment may be significantly more expensive than running a comparable SAS deployment on-premises.

The two bread-and-butter components of clouds are elastic compute instances and elastic storage. *The compute instances run SAS software and your SAS code; the elastic storage contains your SAS datasets.*

## COMPUTE INSTANCES

To explain compute instances, let’s compare the various types of machines, from physical machine to virtual machine to instance:

- **Physical Machine:** this is *actual* server, workstation or desktop *hardware* that is purchased through a value-added partner, reseller, or directly from the manufacturer, and which will typically be located on-premises.<sup>4</sup> For example, the HPE ProLiant DL580 Gen10 server is available in six pre-configured models and can be highly customized by an authorized reseller or partner through HPE’s One Config Simple (OCS) website.<sup>5</sup> Available options includes a range of CPUs, various memory configurations, local storage options, various network and storage controller cards, graphics accelerator cards, and so on. If the physical machine will be dedicated, IT staff will install an operating system such as Linux or Windows, and they will install applications as needed.

<\$\$\$INSERT BASIC DIAGRAM HERE>

- **Virtual Machine:** this type of configuration starts with a physical machine on-premises, on which IT staff installs hypervisor software.<sup>6</sup> Hypervisors virtualize and manage physical resources in the machine, such as CPU, RAM, device adapters, the DVD drive, and so on. IT staff then creates virtual machines that are run by the hypervisor. Each virtual machine contains a specific number of virtual

---

<sup>3</sup> A data center is an air-conditioned building or room that contains computer servers, storage devices, security appliances, and telecommunications equipment, often installed in 19-inch wide racks.

<sup>4</sup> On-premises means that the machine is installed on the premises of the organization that uses and manages the machine, for example, a local “server room” or “data center.”

<sup>5</sup> See <https://sce-public.houston.hp.com/SimplifiedConfig/Welcome>. Please note that this author is not affiliated with, or receiving compensation from, Hewlett Packard Enterprise (HPE).

<sup>6</sup> Common hypervisors include VMware vSphere and Microsoft Hyper-V.

VPUs, virtual RAM, virtual peripheral devices and adapters, etc., configured by the system administrator. Finally, inside a virtual machine, IT staff will install an operating system and needed applications. To the operating system, the virtual machine looks and feels the same as a physical machine, but the drivers may be aware of (and communicating with) the hypervisor for optimal performance (these hypervisor-aware operating system drivers are called paravirtual drivers). IT administrators can oversubscribe virtual resources (e.g., create tens or even a few hundred virtual machines that together have more virtual CPUs and virtual RAM than the underlying physical machine). This generally is okay, since not all servers will use all of their allocated resources to the maximum extent at the same time.

<\$\$\$INSERT BASIC DIAGRAM HERE>

- **Compute Instance:** a public cloud instance is similar to a virtual machine, except it lives in a data center managed by a third-party provider rather than on-premises and managed by your own IT department. Examples of public clouds include Amazon Web Services, Google Cloud, and Microsoft Azure. Like virtual machines, cloud instances are managed by a cloud hypervisor and ultimately still run on physical machines. Cloud instances only come in specific pre-defined configurations based on the underlying type of physical host and its storage network connection. \$\$\$ REPLACE THE FOLLOWING TEXT WITH A TABLE). Instances consist of specific combinations of CPU, RAM, local storage devices (for some), and I/O throughput capability, and so on. For example, for a physical host that can make 72 virtual CPUs and 512 GB RAM available through its hypervisor, you may only be able to select instances with 4 cores, 8 cores, 16 cores, 32 cores or 64 cores, but you cannot select your own preferred number of cores (say, 12 cores or 24 cores). Each instance would typically provide a specific amount of RAM and I/O throughput capability (e.g., the 4 core instance might come with 32GB RAM and 100MB/s I/O throughput, the 8 core instance might come with 64GB RAM and 200MB/s I/O throughput, and the 64 core instance might come with 512GB RAM and 1600MB/s I/O throughput). You cannot mix and match resources to meet a specific configuration.

Dedicated physical machines are typically the best performing option (depending on how they have been specified by a designer), but they are also expensive and inflexible. Virtual machines run slower than physical machines (due to hypervisor overhead and due to competing demands for physical resources by other virtual machines), but tend to be less expensive per virtual machine and offer very flexible configuration options. Cloud instances may be very fast and efficient in some ways, but may be slow in other ways (especially with I/O). \$\$\$ IMPROVE THIS COMPARISON

## CLOUD STORAGE

Text here...

## SAS AND AMAZON WEB SERVICES ARCHITECTURES

Text here...

## SAS RESOURCE REQUIREMENTS – HOW MUCH CPU DID YOU SAY YOU NEED?

Resource requirements for SAS, when compared to traditional IT systems, might seem unfathomable to IT administrators who are used to configuring two to four CPU cores to power a medium to large-size file server, or perhaps eight CPU cores to power a small to medium-size SQL database for tens to hundreds of users. By contrast, a single SAS user's server session can fully consume two server CPU cores or more, introducing the potential of having only four power SAS users using as many CPU resources as a few hundred SQL database users!

If that isn't outlandish enough, consider SAS disk input/output (I/O)<sup>7</sup>. The current SAS best practice is to provision systems for a minimum I/O throughput of 100 to 150 megabytes per second per CPU core (100 – 150MB/s/core). SAS typically reads and writes in blocks of 64 kilobytes and since it typically processes large datasets sequentially (oftentimes barreling through entire data sets), SAS qualifies their I/O as “sustained large block sequential I/O”. Traditional storage subsystems are often optimized for much smaller 4 kilobyte or 8 kilobyte blocks that are written more randomly, such as updates to a SQL database. A storage system optimized for small random block Input/Output Operations Per Second (IOPS; pronounced “eye-ops”) may perform poorly with SAS I/O, which behaves more like a barreling stream train than a bee that moves from flower to flower (okay, I exaggerate, but you get the point).

Besides CPU and disk I/O, there is memory (RAM) to consider. At a general requirement of 16GB per CPU core, it is normally not difficult to configure a system that has sufficient RAM for SAS. When in-memory processing is required with SAS LASR or SAS Viya compute servers, RAM becomes a critical resource after all with demands that can reach or exceed 1TB RAM per server.

## PUTTING SAS IN THE CLOUD – A METHODOLOGY

This is the main section of this paper, and focuses on Amazon Web Services. The concepts presented in this section should largely apply to other cloud services platform providers as well, such as Microsoft Azure or Google Cloud.

Here are the steps to follow as part of this methodology with key considerations:

- AWS Elastic Block Store (EBS) storage peaks at approximately 50MB per second per core, with the highest available throughput limited to 1,750MB per second per instance (server). SAS recommends a minimum I/O throughput of at least two to three times this speed (100MB – 150MB per second per core). *AWS storage therefore would run at a speed of up to 33% to 50% of the SAS minimum throughput recommendation.* At a maximum EBS throughput of 1,750 MB per second per instance, at most, 16 CPU cores can be accommodated for a SAS compute server or grid node.
- To provide the best possible I/O performance, a *storage-optimized AWS EC2 instance type* with a local instance store and the fastest available EBS throughput should be selected. The instance store would be used for SAS WORK and SAS UTILLOC to improve overall I/O performance. After the I/O has been optimized this way, processor and memory should be evaluated and selected.
- The AWS i3.16xlarge and i3.metal EC2 instance types each have eight 1.765TB NVMe local instance storage disks which when striped together provide 13.8TB of usable storage. These two instance types can also provide the maximum per-instance EBS throughput of up to 1,750MB/s. I3 is also the only storage-optimized instance type that has an instance store with NVMe disks, rather than HDD's. The i3 instance is currently the SAS recommended instance type for AWS. You'll need a provisioning script that runs on startup to mount and format the drives for use with SAS WORK.
- Disable unused cores at the time the instance is launched. For example, an i3.16xlarge instance can provide up to 1,750MB/s I/O throughput (satisfying 16 cores at up to 109.4MB/s/core or 12 cores at up to 145.8MB/s/core) but actually contains 32 cores (64 vCPUs). The i3.8xlarge instance contains a matching number of cores (16) but can only provide up to 875MB/s throughput (satisfying only 6 to 8 cores). Additional instances may need to be put in the systems design so that the total active core count matches the SAS license. Meanwhile you'll pay for unused CPU power.
- The i3.16xlarge and i3.metal instances have sufficient RAM at 488GB and 512GB, respectively, for 16 core SAS 9.4 compute servers; however, heavy use of SAS Visual Analytics and other in-

---

<sup>7</sup> Disk input/output (I/O) is simply the act of writing data from computer memory to disk, or reading data from disk into computer memory.

memory SAS products in the future may require additional RAM (this would include Viya – although for Viya the need for memory may trump the need for I/O throughput depending on the use case, and you may have to step away from i3 due to its memory limitations).

- Out of the available volume types for EBS volumes, throughput-optimized HDD (st1) is the most appropriate choice for SAS, which requires optimal throughput for its sustained large block sequential I/O. \$\$\$ REPHRASE FOLLOWING if you use io1 – tempting since it is all SSD, and yes, it is faster than st1 -- you'd quickly run into the provisioned IOPS site limit which limits the number of compute servers and/or the number and size of EBS volumes installed at the site – this is one reason why st1 is the SAS and IBM-recommended storage type. And yes AWS customer support can increase the provisioned IOPS site limit but not indefinitely, plus io1 is at least twice as expensive as st1 (which is the other reason why st1 is recommended)
- EBS st1 volumes provide up to approximately 40MB per second I/O throughput *per terabyte of provisioned storage*. Provisioning 12,800GB or greater will enable a single EBS st1 volume to perform up to its full 512MB I/O throughput per second potential while not being dependent on I/O credits. 512MB/s (500MiB/s) is the maximum throughput for any EBS volume regardless of its volume type. \$\$\$ SOMEWHERE PRIOR, EXPLAIN ABOUT ST1, IO1, etc.
- A single EBS volume at 512MB/s is not sufficient to meet the SAS minimum throughput requirement of 100MB – 150MB per second per core if the server contains more than 4 cores. Therefore, multiple EBS volumes must be striped in sufficient numbers to reach the per-instance maximum I/O throughput of 1,750MB/s for 16-core SAS compute servers. *SAS recommends striping a minimum of four to eight EBS st1 volumes.*
- \$\$\$ REPHRASE/REPOSITION There is also the implication here that you will be provisioning anywhere between 51.2TB and 102.4TB per SAS compute server or grid node, just for performance purposes and regardless of whether you actually need the capacity or not. Also, the striping would be done without redundancy built in at the OS-level, relying instead on redundancy supplied by the AWS EBS storage array back-end which is not an AWS best practice, but if you do apply a RAID 5 scheme performance caves significantly below already marginal levels.

\$\$\$ CLEANUP/REORGANIZE THESE PARAGRAPHS The r5 instance type is usually pretty suitable as a high-performance instance type, which focuses on memory and offers good CPU, depending on the exact model of course. The r5 instance models are listed here: <https://aws.amazon.com/ec2/instance-types/r5/>. These instance types may be used for metadata servers, middle tier (web) servers and client machines

The i3 instance type is often used for SAS compute (application servers) since the real performance issues with SAS are usually seen in the compute (application) server. (<https://aws.amazon.com/ec2/instance-types/i3/>) for the compute server in both quotes. Let's compare the i3 instances from both quotes:

\$\$\$ POSITION THIS TABLE ABOVE THE BULLETS ABOVE

Instance Type	vCPUs (Cores)	RAM (GiB)	Instance Store* (SAS Work)	Network Perf.	Max I/O** (MB/s)	Max I/O** per Sec. per Core
i3.large	2 (1)	15.25	1 x 0.5GB NVMe SSD	Up to 10Gb	53.13	53.1MB/s/core
i3.xlarge	4 (2)	30.5	1 x 0.9GB NVMe SSD	Up to 10Gb	106.25	53.1MB/s/core
i3.2xlarge	8 (4)	61	1 x 1.9GB NVMe SSD	Up to 10Gb	212.5	53.1MB/s/core
i3.4xlarge	16 (8)	122	2 x 1.9GB NVMe SSD	Up to 10Gb	437.5	54.7MB/s/core
i3.8xlarge	32 (16)	244	4 x 1.9GB NVMe SSD	10Gb	875	54.7MB/s/core

Instance Type	vCPUs (Cores)	RAM (GiB)	Instance Store* (SAS Work)	Network Perf.	Max I/O** (MB/s)	Max I/O** per Sec. per Core
i3.16xlarge	64 (32)	488	8 x 1.9GB NVMe SSD	25Gb	1,750	54.7MB/s/core
i3.metal	72 (36)	512	8 x 1.9GB NVMe SSD	25Gb	1,750	54.7MB/s/core

\$\$\$ EXPLAIN WHY RED IS RED

\* A single 1.9GB NVMe SSD drive in an i3 instance type can perform up to 1,000MB/s (1GB/s). All eight i3 1.9GB NVMe SSD drives in a striped configuration can hold 12.8TB and can perform up to 6,800MB/s (6.8GB/s). This satisfies the requirements for the file system that will hold SAS WORK \$\$\$ VERIFY NUMBERS AND EXPAND

\*\* Max I/O = Maximum I/O Throughput for the instance (across all EBS volumes).

The assumption for all striped volumes is *Striped (RAID 0)* without OS level redundancy (no RAID 1, 10, 5 or 6).

*Up to 10Gb* means that the storage network connection is shared with other instances, which may significantly impact SAS performance.

Maximum I/O throughput and IOPS reference for AWS EC2 instance types can be found here: <https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/EBSOptimized.html>. The large block throughput I/Os listed for per-instance throughput maximums are twice the size of typical SAS 9.4 I/Os (128KB instead of 64KB) but while testing with an i3 instance-attached EBS storage, the maximum observed throughput was 1,670MB/s with multiple striped 12.8TB EBS volumes, regardless of block size.

As shown in the table, in the EBS-optimized i3 line, the i3.metal and i3.16xlarge can do up to 1,750MB/s (published), but every step below that, maximum I/O throughput is cut in half. So the i3.8xlarge can do 875MB/s max, the i3.4xlarge (quoted) can do 437.5MB/s max, and the i3.2xlarge (quoted) can perform up to 212.5MB/s. \$\$\$ REPHRASE

SAS doubles the number of CPUs for cloud deployments (so for example, a traditional 8 core on-prem compute server license equates to a 16 vCPU cloud license).

\$\$\$ MOVE THIS PARAGRAPH One way to compensate for the approximately 50MB/s/core throughput limitation is to choose an i3 instance type one model higher than the one that matches the SAS license (doubling CPU count but also I/O throughput), and turning off half of the vCPUs when the instance is launched. This way, the CPUs will match the SAS license (so SAS will run), and they are balanced to the SAS minimum I/O throughput requirement. If your SAS 9.4 server license is for 32 cores, you may need to add an additional EC2 instance and a shared file system to allow multiple compute servers access the same file systems. \$\$\$ WORK OUT FURTHER WITH AN EXAMPLE

\$\$\$ MOVE THIS PARAGRAPH The AWS Simple Monthly Calculator can be used to plug in information about your EC2 instances, EBS storage, and other products and services inside AWS: <https://calculator.s3.amazonaws.com/index.html>. This will result in an approximate monthly amount for AWS services. After inputting all of the data, click on the next tab (Estimate of your Monthly Bill) where you can save the information to a unique URL or export it as CSV. Then, if you are using a value-added provider, if you know the cost of the SAS licensing you may be able to make an educated guess about how much your provider profits of the managed AWS arrangement (if applicable).

## CONCLUSION

Text here...

---

If you need to include source code, introduce it with a sentence that ends with a colon:

```
proc ds2;  
data _null_;  
  method init();  
    dcl varchar(16) str;  
    str = 'Hello World!';  
    put str;  
  end;  
enddata;  
run;  
quit;
```

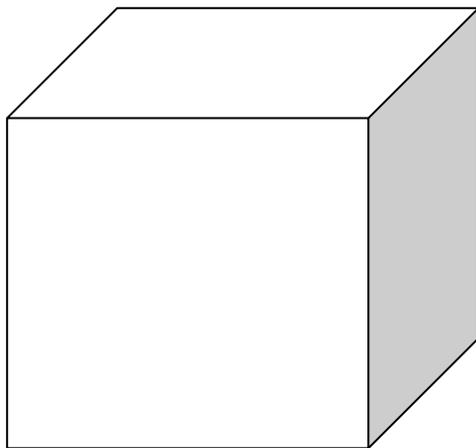
The easiest way to insert code is to copy the preceding code and modify it as needed. Be sure to indent the first line of code three spaces. Other lines can be indented for clarity. Certain words might be in a different color. You can keep the color or change it to black.

When you want to continue text after source code, select the PaperBody paragraph style. (See the section “To select a paragraph style.”)

When you describe the code, use all caps for names of procedures, statements, options, functions, commands, and steps (for example, “the PRINT procedure,” “the DATA step,” and “the PROC PRINT step”). On the first mention of a procedure, use its name, followed by the word “procedure” (for example, “the PRINT procedure”). Subsequently, you can use “PROC PRINT.”

Figure 1 is a sample figure.

Always introduce a figure by inserting a cross-reference. For instructions, see the section “To insert a cross-reference.”



**Figure 1. Caption for Sample Figure**

Always use the Caption paragraph style for figure captions. See the section “To insert a caption.” Following those instructions ensures that the figures are automatically numbered (even if you rearrange them later). Use title style capitalization for captions (as shown in this example). For instructions about inserting a graphic, see the section “To insert a graphic from a file.”

This is a subtopic of a main topic. This paragraph uses the PaperBody style. This paragraph uses the PaperBody style.

Table 1 is a sample table.

Always introduce a table by inserting a cross-reference. For instructions, see the section “To insert a cross-reference.”

SAS Variable Format	DB2 Data Type
\$w. \$CHARw.	CHARACTER
any date format	DATE

**Table 1. DBLOAD Procedure: Default DB2 Data Types for SAS Variable Formats**

Always use the Caption paragraph style for table captions. See the section “To insert a caption.” Following those instructions ensures that the tables are automatically numbered (even if you rearrange them later). Use title style capitalization for captions (as shown in this example).

If you need to include a numbered (ordered) list, copy the following list, paste it, and modify the text. Be sure to introduce your list with a complete sentence that ends with a colon. For example, “Here are the required steps:”.

1. This is a sample numbered or ordered list item. This is list item text. This is list item text. This is list item text.
2. This is a sample numbered or ordered list item. This is list item text.

This is another sample paragraph. This paragraph uses the PaperBody style. This paragraph uses the PaperBody style.

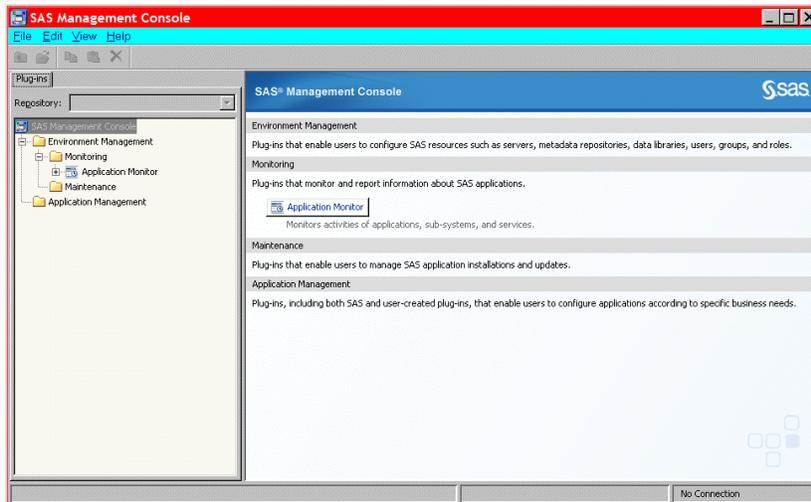
If you need to include a bulleted (unordered) list, copy the following list, paste it, and modify the text. Be sure to introduce your list with a complete sentence that ends with a colon. For example, “Systems that are supported by the product include the following:”.

- This is a sample bulleted list item. This is list item text. This is list item text. This is list item text. This is list item text.
- This is a sample bulleted list item. This is list item text.

This is a continuation of the body of the paper—after an unordered list. This paragraph uses the PaperBody style.

Display 1 is sample display or screen capture.

Always introduce a display by inserting a cross-reference. For instructions, see the section “To insert a cross-reference.”



### Display 1. Former Main Interface for SAS Management Console

Always use the Caption paragraph style for display captions. See the section “To insert a cross-reference.” Following those instructions ensures that the displays are automatically numbered (even if you rearrange them later). Use title style capitalization for captions (as shown in this example).

### SUBHEAD A LEVEL <HEADING 2>

This paragraph uses the PaperBody style.

Output 1 shows an example of how to present output.

Always introduce output by inserting a cross-reference. For instructions, see the section “To insert a cross-reference.”

```
CREATE TABLE ALLACCTX (SourceSystem varchar(4),
cctnum numeric(18,5) CONSTRAINT "ALLACCT_PK" PRIMARY KEY,
ccttype numeric(18,5), balance numeric(18,5), clientid numeric(18,5),
losedate date, opendate date, primary_cd numeric(18,5), status varchar(1))
```

### Output 1. Output from a CREATE TABLE Statement

Always use the Caption paragraph style for output captions. See the section “To insert a caption.” Following those instructions ensures that the outputs are automatically numbered (even if you rearrange them later). Use title style capitalization for captions (as shown in this example). Note that output has the same font as source code, but it is in a box. (This box is not a text box.) The easiest way to properly format output is to copy the preceding output (including the box) and its caption, paste it, and then modify the text.

### SUBHEAD A LEVEL <HEADING 2>

This paragraph uses the PaperBody style.

### Subhead B Level <Heading 3>

This is a subtopic of a subtopic. This paragraph uses the PaperBody style.

### Subhead C Level <Heading 4>

This paragraph uses the PaperBody style.

## CONCLUSION

Do not change the heading style or the text “CONCLUSION” of the preceding heading

This paragraph uses the PaperBody style.

The conclusion summarizes your paper and ties together any loose ends. You can use the conclusion to make any final points such as recommendations, predictions, or judgments.

## REFERENCES

This section is not required. If you include this section, do not change the heading style or the text “REFERENCES” of the preceding heading.

The following references are examples and do not cover the spectrum of references that might be included. The important thing is to be consistent in formatting and organizing your references. If you prefer, you can follow a bibliographic approach such as the American Psychological Association (<http://www.apastyle.org/manual/index.aspx>) or the American Statistical Association (<http://amstat.tjournals.com/asa-style-guide/>).

This is the format for references if you use the author-date format, in which citations are inserted in parentheses in text, such as (Smith 2014) or (Smith, 2014; Doe and Johnson, 2016). If the same author has several references, use a, b, c, and so on. For example, (SAS Institute Inc. 2015a, 2016b, and 2016c).

Book <Author name: last name, first name (or initials)>. <Publication date>. <Book title>. <City, State (abbrev) of publication> : <Publisher name>.

Journal article <Author name: last name, first name>. <Publication date>. “<Article title>.” <Journal title>, <volume no.:page numbers>.

Article in conference proceedings <Author name: last name, first name>. <Publication date>. “<Article title>.” <Title of proceedings such as Proceedings of the SAS Global 2010 Conference>, <City, State (abbrev) of publication> : <Publisher name>. Optional: You can add a URL to access available online proceedings. For example: Available at <http://support.sas.com/resources/papers/proceedings09/TOC.html>.

Website <Author name: last name, first name>. “<Title>.” <Source>. <Date>. Available at <URL>.

Reference examples:

Book Agresti, A. 2013. *Categorical Data Analysis*. 3rd ed. Hoboken, NJ: John Wiley & Sons.

Journal article Akaike, H. 1979. “A Bayesian Extension of the Minimum AIC Procedure of Autoregressive Model Fitting.” *Biometrika*, 66:237–242.

Article in conference proceedings Dorfman, A. H. and R. Valliant. 1993. “Quantile Variance Estimators in Complex Surveys.” *Proceedings of the Survey Research Methods Section*, 866–871. Alexandria, VA: American Statistical Association.

Website Federal Reserve Bank of St. Louis. 2012. “Economic Research.” Accessed November 7, 2012. <http://research.stlouisfed.org>.

## ACKNOWLEDGMENTS

This section is not required. If you include this section, do not change the heading style or the text “ACKNOWLEDGMENTS” of the preceding heading.

This is the text for the acknowledgments. This paragraph uses the PaperBody style.

## RECOMMENDED READING

This section is not required. If you include this section, do not change the heading style or the text “RECOMMENDED READING” of the preceding heading.

This is the format for recommended reading.

- *Base SAS® Procedures Guide*
- *SAS® For Dummies®*

## CONTACT INFORMATION <HEADING 1>

Do not change the heading style or the text “CONTACT INFORMATION” of the preceding heading. Do not change the text of the following paragraph. Replace all fields that are shown in angle brackets.

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

<Name>  
<Enterprise (optional)>  
<Phone (optional)>  
<E-mail>  
<Web (optional)>

The next two paragraphs are **required** and must be in the paper if you are a SAS employee who is authoring a SAS paper. Otherwise, this section is not needed.

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.

Remember to delete the following instructions and all other text that is highlighted in light tan.

## BASIC INSTRUCTIONS

### WRITING GUIDELINES

#### Trademarks and product names

To find correct SAS product names (including use of trademark symbols), if you are a SAS employee, see the [Master Name List](#). Otherwise, see [SAS Trademarks](#).

- Use superscripted trademark symbols in the first use in title, first use in abstract, and in graphics, charts, figures, and slides.
- Do not abbreviate product names. For example, you cannot use “EM” for SAS® Enterprise Miner™. After having introduced a SAS product name, you can occasionally omit “SAS” for certain products, provided that your editor agrees. For example, after you have introduced SAS® Simulation Studio, you can occasionally use “Simulation Studio.”

#### Writing style

- Use active voice. (Use passive voice only if the recipient of the action needs to be emphasized.) For example:

The product creates reports. (active)  
Reports are created by the product. (passive)

- Use second person and present tense as much as possible. For example:

You get accurate results from this product. (second person, present tense)  
The user will get accurate results from this product. (future tense)

- Run spellcheck, and fix errors in grammar and punctuation.

#### Citing references

All published work that is cited in your paper must be listed in the REFERENCES section.

If you include text or visuals that were written or developed by someone other than yourself, you must use the following guidelines to cite the sources:

- If you use material that is copyrighted, you must mention that you have permission from the copyright holder or the publisher, who might also require you to include a copyright notice. For example: “Reprinted with permission of SAS Institute Inc. from *SAS® Risk Dimensions®: Examples and Exercises*. Copyright 2004. SAS Institute Inc.”
- If you use information from a previously printed source from which you haven’t requested copyright permission, you must cite the source in parentheses after the paraphrased text. For example: “The minimum variance defines the distance between cluster (Ward 1984, p. 23)

#### TIPS FOR USING WORD

These instructions are written for MS Word 2007 and MS Word 2010. The steps are similar for MS Word 2003.

#### To select a paragraph style

1. Click the HOME tab. The most common styles in your document are displayed in the top right area of the Microsoft ribbon. If you don’t see a style that you want, click the slanted down arrow at the bottom right corner of the Styles area, and scroll through the list. The main styles for this template are headings 1 through 4, PaperBody, and Caption. Avoid using other styles.
2. To change a paragraph style, click the paragraph to which you want to apply a style, and then click the style that you want in the ribbon.
3. PaperBody (used for most text) is automatically applied when you press Enter at the end of any heading style or the Caption style.

#### To insert a caption

1. Click REFERENCES on the main Word menu.
2. Click Insert Caption.

3. Select the **Label** type that you want.
4. Click **OK**.

### To insert a cross-reference

1. Click **REFERENCES** on the main Word menu.
2. Click **Cross-reference**.
3. In the **Reference type** list box, select **Heading, Figure, Table, Display, or Output**.
4. For a heading:
  - a. In the **For which heading** list, select the heading that you want.
  - b. From the **Insert reference to** list, select **Heading text**.
5. For a figure, table, display, or output:
  - a. In the **For which caption** list, select the caption that you want.
  - b. From the **Insert reference to** list, select **Only label and number**.

### To insert a graphic from a file

1. Click **INSERT** on the main Word menu.
2. Click **Picture**.
3. In the Insert Picture dialog box, navigate to the file that you want to insert.
4. When the name of the file that you want to insert is displayed in the **File name** box, click **Insert**.