The STAC Big Data Special Interest Group
Questions and Answers

Updated 7 May 2014

Q: What is STAC?

A: STAC (the Securities Technology Analysis Center) facilitates the STAC Benchmark Council (www.STACresearch.com/council), which is a group of financial organizations and vendors with two missions: 1) to define and deploy standard technology benchmarks that are meaningful to financial firms, 2) to facilitate substantive dialog among customers and vendors with respect to technology topics. The Council started several years ago by focusing on low-latency, real-time workloads in automated trading. In recent years, it has expanded into tick data analytics, derivatives risk management, and other areas where the size or complexity of the workload is a bigger challenge than the speed of the data. It is expanding further by casting the net to encompass workloads that the Council considers to be "big data".

Q: How does the Council define "Big Data"?

A: The STAC Benchmark Council uses the term "big data" to describe problems rather than solutions. For the purposes of organizing Council activities and delineating benchmarks, the Council's definition of big data workloads are those which are difficult or expensive to handle using traditional technologies, largely due to data scale or complexity. This broad definition is being made more concrete as the STAC Big Data SIG develops a set of reference use cases.

Q: What is the STAC Big Data SIG?

A: The STAC Big Data SIG is a group of user firms and vendors that meets periodically (sometimes physically, sometimes virtually) to discuss key issues related to "big data" challenges, as defined above. The group is charged with defining benchmark specifications based on these challenges and eventually to establish a research agenda to address shared interests. It is also a forum to discuss new technologies with potential to solve these challenges (NoSQL, NewSQL, new compute, storage, and networks, etc.).

Q: What workloads are in scope?

A: The SIG maintains an inventory of big data workloads for the benefit of SIG members. The workloads range from fraud detection and risk management to operational intelligence and social analytics for trading. This inventory serves to focus technology discussions and provides the starting point for benchmark specifications. As the SIG decides to develop benchmarks around specific workloads, it may spin off dedicated working groups.

That has happened in one case so far. Many industry observers consider backtesting of trading strategies to be one of the cononical "big data" workloads in finance, given the large volumes of data involved and the need to move compute to the data in a highly parallel fashion. Because this workload is well understood and the economics surrounding backtesting tend to be quite different from those of other big data workloads, the Council quickly spun out a separate working group to develop technology benchmarks for backtesting (see www.STACresearch.com/a3).

For more information or to indicate your interest, visit STACresearch.com/bigsig

Q: Is this SIG related just to investment banking?

A: No. Any type of financial institution can get involved. STAC started out in the securities industry, but this special interest group is broader. The main representation is currently from investment and retail banking, but other financial institutions (e.g., asset management, insurance) are also welcome. Some of the workloads in the current inventory are specific to retail banking or capital markets, while some of them (e.g., operational intelligence) are probably relevant to all large, IT-intensive organizations, particularly heavily regulated ones.

Q: Is this just about performance benchmarks?

A: No. A range of benchmarks are valuable in the rapidly evolving big data space. SIG members have expressed a desire for benchmarks that cover security & entitlements, resilience, and other functionality, in addition to performance. Moreover, the SIG also has information meetings that are not specific to benchmarks.

Q: How often does the SIG meet?

A: For the last several months, STAC has been working with SIG members 1:1 to develop the initial inventory of use cases. How the SIG decides to proceed with proposed benchmarks and what topics SIG members want to cover in information sessions will determine the frequency of meetings. Most of these will be by teleconference, though there are occasional in-person meetings. Big data topics are also covered at STAC Summits in cities around the world (see www.STACresearch.com/meetings).

Q: What are the benefits of participation for end user firms?

A: Several:
- Learn how other end users and vendors are approaching big data problems.
- Help ensure that the industry dialog around big data serves your firm's interests.
- Once SIG-directed testing programs begin, access benchmark information not available to the public.
- Influence and understand benchmark standards for big data solutions.

Q: What are the benefits of participation for vendor firms?

A: Several:
- Understand the big data challenges that this market faces.
- Convey your company's big data message to key customers in various ways.
- Influence and understand benchmark standards for big data solutions.
- Shorten the sales cycle for new products.

Q: What firms can participate?

A: The SIG is currently open to Contributing and Recipient Members of the STAC Benchmark Council anywhere in the world, including vendor members. To determine what kind of membership your firm has, see www.STACresearch.com/MyMemberInfo. For membership information, see www.STACresearch.com/council.

For more information or to indicate your interest, visit STACresearch.com/bigsig

Q: What does it cost to participate?

A: There is no additional fee for the members above to participate in the SIG at this time. Fees may be required at some point in the future, to access test harness software or fund research projects directed by the SIG. Other interested organizations (e.g., academia, consultants, etc.) should discuss with STAC.

Q: What kind of individuals should be on this SIG?

A: Depending on the meeting topic, CTOs, data architects, data scientists, infrastructure architects, and performance engineers may be interested. On the vendor side, both business and technical functionaries may be interested, such as product managers, heads of engineering, solution architects, business development managers, pre-sales engineers, and performance engineers.

Q: Can multiple people represent a single firm on the SIG?

A: Yes. However, in-person meetings (and perhaps even some teleconferences) may have limits on the number of reps per firm.

Q: Where do I go to find out what's happening with the SIG?

A: www.STACresearch.com/bigsig. Note that most materials will only be accessible to the member types above.

For more information or to indicate your interest, visit `STACresearch.com/bigsig`