Ask the Expert

AI-Driven Transformation – Generating Insights for the Customer

Ankit Dedhia, Senior SAS and Cloud Technical Architect

Ankit Dedhia

Senior SAS and Cloud Architect

Accenture CloudFirst – Data & AI Capability

Ankit is a Senior SAS and Cloud Solution Architect with 14 years of experience. He is a Certified Technology Architect with expertise in analyzing complex on-prem architectures and modernizing them on cloud. He has expertise across SAS technology stacks, Risk and Fraud solutions, driving the SAS modernization journey as a part of Accenture SAS Cloud Factory.

Ask The Expert

AI Driven SAS Transformation: Generating Insights for the Customer

Ankit Dedhia

Senior SAS and Cloud Architect,
Accenture – Data & AI Capability

sas | accenture

Quick Recap

How we accelerate the SAS Journey

Moving SAS technologies from on-premise infrastructure to cloud, quickly and confidently needs an in-depth assessments of multiple parameters across various dimensions such as

Discover

SAS landscape discovery

- Detailed level understanding of the SAS landscapes across its overall utilization across the organization, its computing consumptions, penetration within and across organization and distribution of associated workloads and complexity

Assessment

Shortlisting of Application for migration

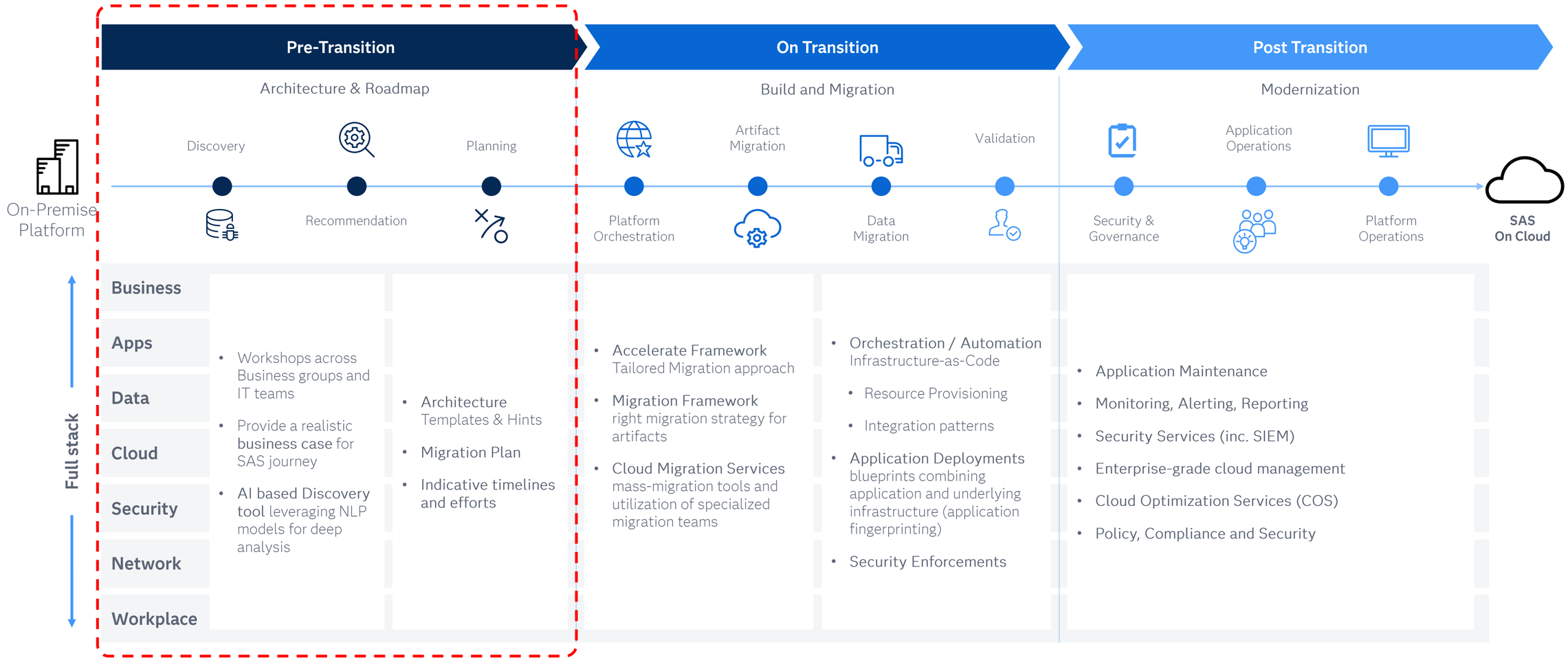
- Conducting a 360-degree technical assessment of multiple metrics across 4 dimensions to identify which SAS application / artifacts should be Rehosted, Replatform, Rearchitect or Retired

Recommendation

Rule based recommendation

- Engage target groups to understand requirements & minimize impact
- Develop training & trainers for new platforms (adoption)
- Provide a high-level set of recommendation of
  - SAS / Non-SAS technologies and tools
  - Cloud Architecture
  - Migration Path

ACN SAS Transformation reduces the time and cost of migrating SAS Estates to the Cloud



CMT Client: Cloud Modernization Initiative

sas | accenture

Use Case

01 Client has a large on-prem SAS Grid platform and is looking to migrate to cloud as a part of cloud modernization initiative

02 They have been exploring SAS Viya as a future state platform

03 The cloud migration needs to be achieved by 2024

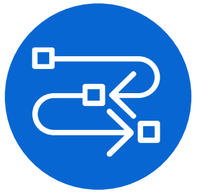
Current State Challenges

Lack of transparency on SAS usage

- There is no centralized logging or oversight of SAS usage
- Half of the workloads are running under local Unix system accounts that are not attributable to a business owner
- Some workloads are running under user-ids that should be inactivated.

Majority of data still resides on-premises

- SAS is being used to access many different on-premise data sources.
- SAS is being used to create department data marts in on-premise databases like SQL Server and Oracle
- Visualization tools are, in some cases, reading directly from SAS datasets.



Long running legacy processes

- There is a high percentage of workloads that are running for at least 1 hour.
- Capacity constraints appear to occur early in the business day.
- Complex legacy processes have been inherited by teams, without full understanding of code.

Lack of cloud skills and knowledge

- Teams that do not rely on Data Lake for most of their processes are not enabled to move to cloud.
- These teams are also not yet aware of how to access data in cloud from on-premises.

Key Observations

Active users

- Most number of active users:
  - Marketing
  - Products & Services
  - Market Intelligence
  - Information Services

Workload patterns

- 22% of workloads analyzed are complex
- No advanced analytics are being used in SAS
- SAS Grid is in use 24x7

Workload Usage

- Teams with the most workloads between Oct to Dec
  - Finance Operations (but cleaned up rogue jobs)
  - Marketing
  - Business Intelligence

Lack of transparency in workloads

- Scheduled jobs are failing over long periods of time
- Workloads are being run under system accounts and user-ids that are not explicitly tied to a business owner
- What data users are accessing is not easy to discern.

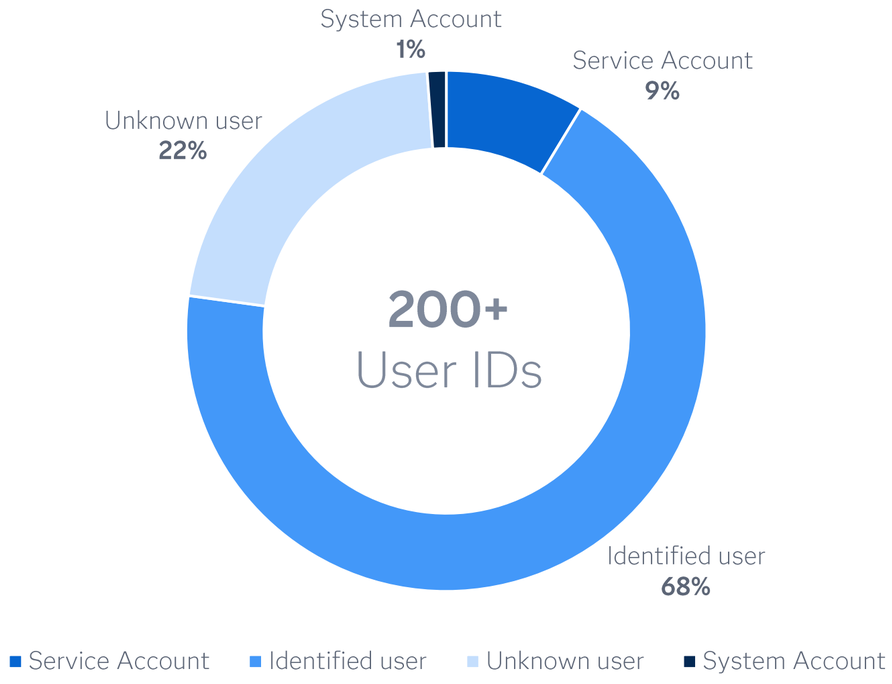
10% of workloads run for >1hr

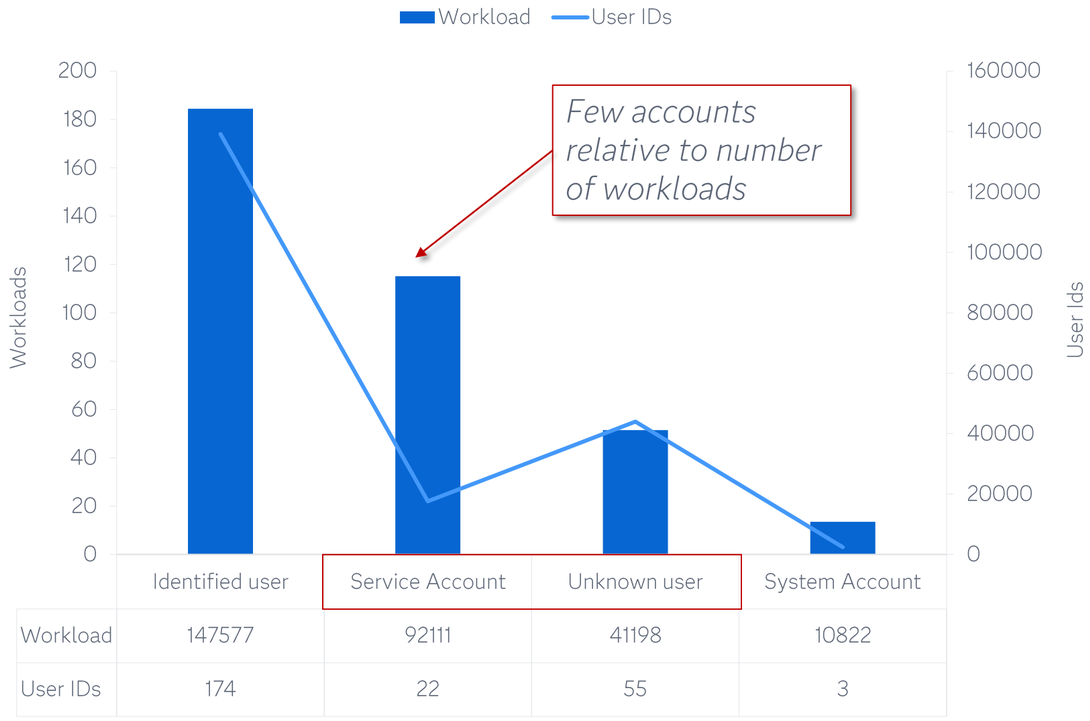
- 6% (over half of long-running workloads) run for >2hrs
- 44% are running under system accounts
- Capacity constraints at 9am on business days are reflected by highest % of long-running jobs

Default tool for reading multiple data sources

- Entrenched SAS users are typically business analysts – creating data marts for downstream consumption by reports or other teams
- 3000+ scheduled flows in SAS MC, with 2000 active

SAS Users





	Identified user	Service Account	Unknown user	System Account
Workload	147577	92111	41198	10822
User IDs	174	22	55	3

- **55** recently active IDs are from unknown users, including original users who no longer work in the organization

- **22** Service Accounts ran **92,111** workloads during the analysis period
- **55** unknown users ran over **41,198** workloads

Key Findings

Not all SAS products in the license are being used: SAS/IML, SAS Visual Analytics, SAS OLAP

Not all SAS capabilities are being leveraged.

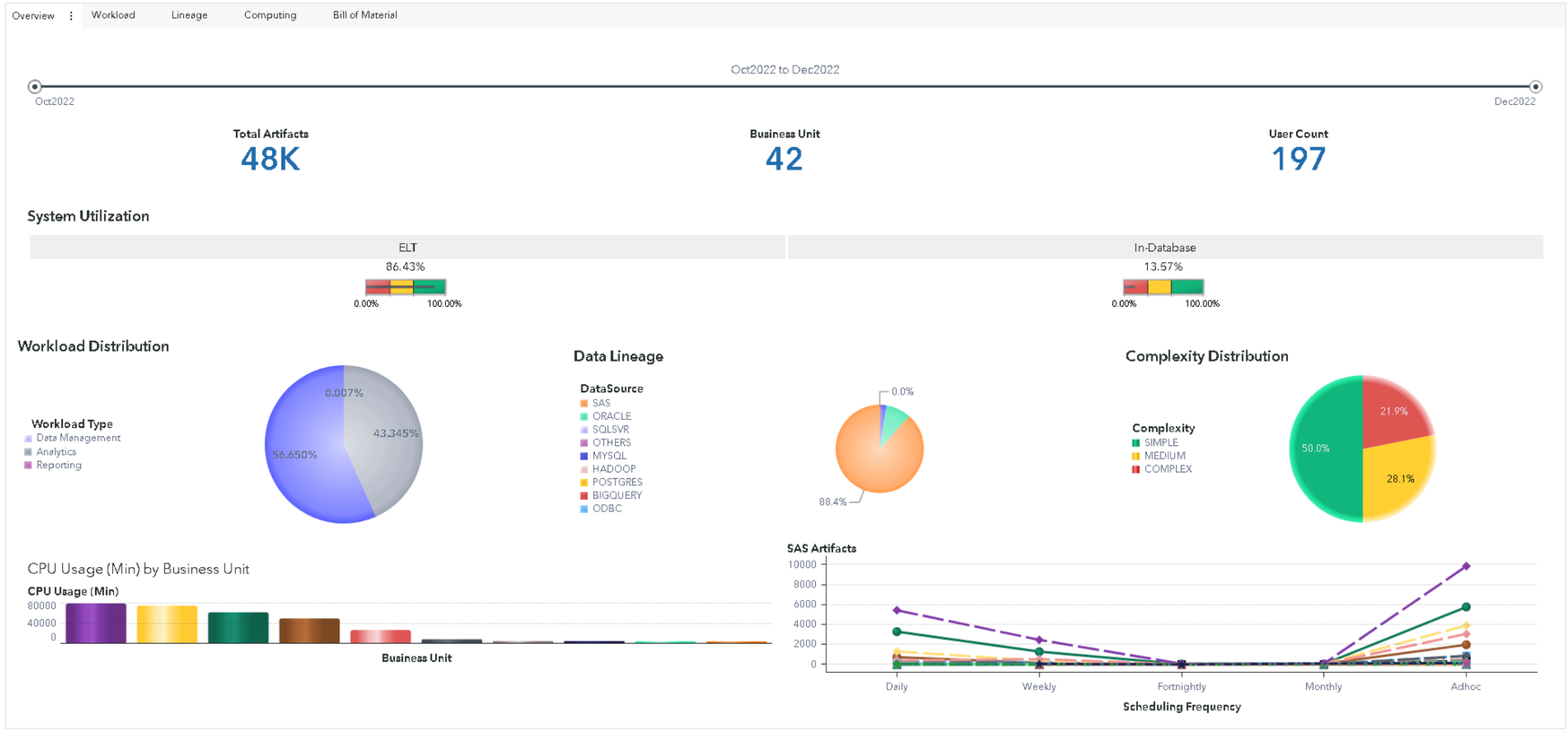
Parallel execution is not being leveraged. SAS Grid has the ability to take individual steps within a SAS program and run them in parallel where possible.

Orchestration of SAS processes within a program to control when a step should be started. For example, it is possible to have one step wait until the completion of a previous step before starting.

SAS is heavily used for Data Management activities. No usage of SAS Advanced Analytics observed.

For external calls from SAS, the majority include X-commands.

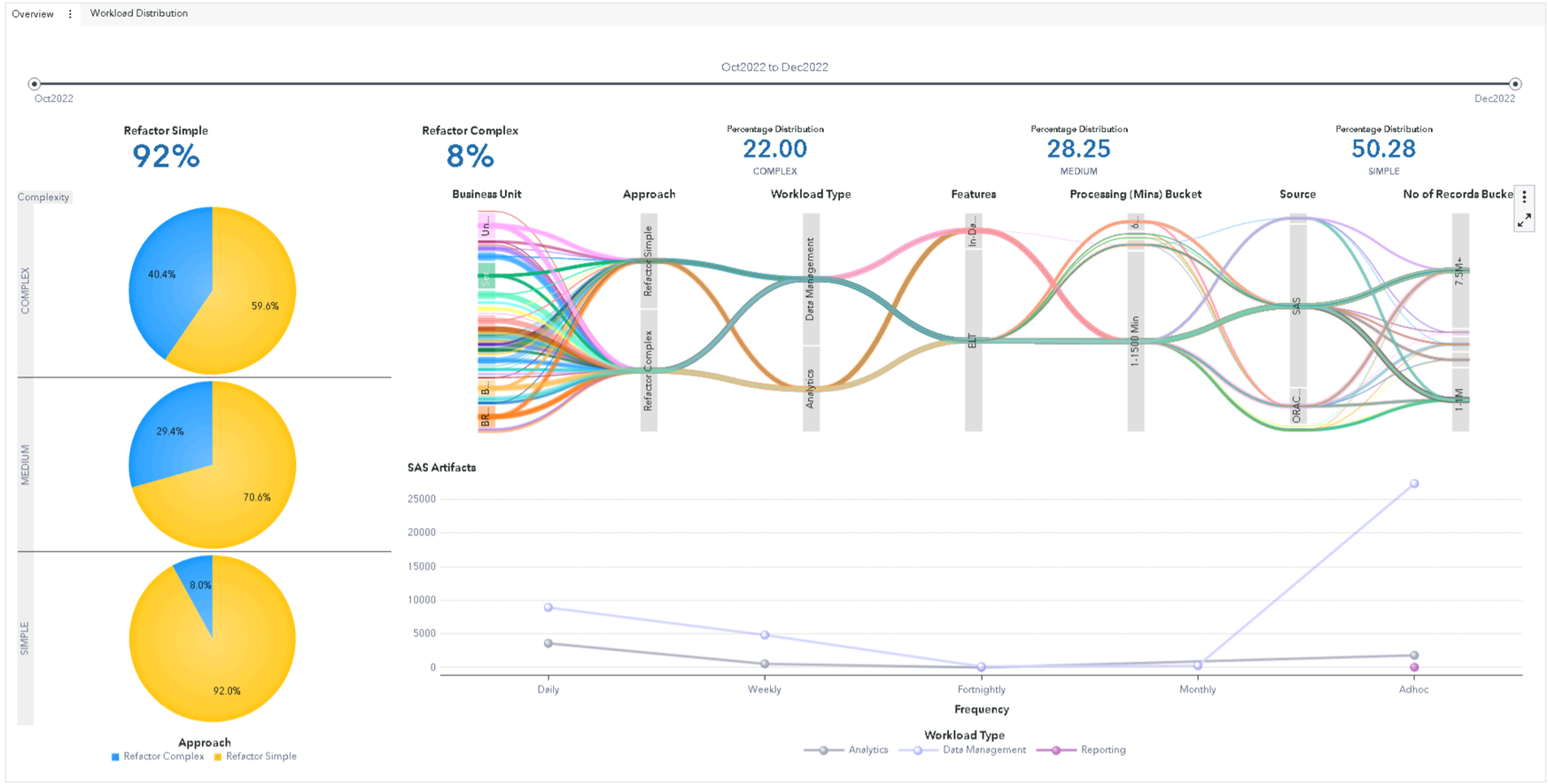
Discovery



- Analysis carried out for the period of Oct – Dec
- 197 Users Identified across 42 Business Units
- Total Artifacts includes SAS Programs and EG Projects

- System utilization in ELT is more compared to In-Database
- Major Workload coming from Unknown, Service-Account
- High percentage of complex workloads, indicating a high number of critical processes

Assessment



- Majority of workloads classified as Simple for migration
- SAS dataset is a major source of data for all BUs

SAS must continue to have a presence

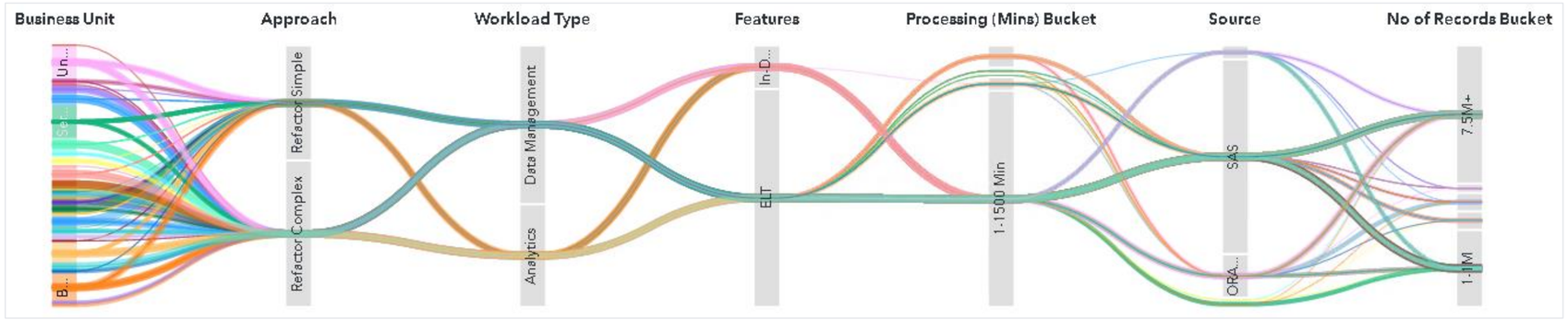
The system is currently used consistently 24x7.

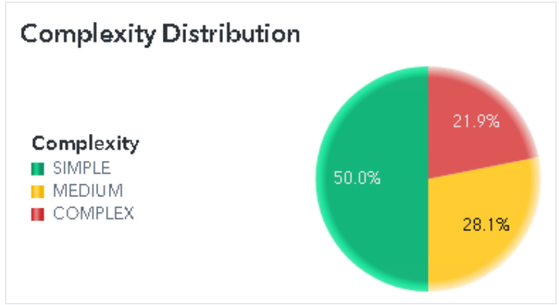
There are many different teams leveraging SAS for data processing large data and performing lightweight analytics.

There is a high percentage of complex workloads (21.9%), indicating that SAS is being used for critical processes, transforming input data sources for further use across the business.

Number of Workloads

Nbr Workl	Hour of Day																							
Row Lab	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	Total
Sun	1,423	1,513	1,509	1,481	1,455	1,516	1,792	2,119	2,095	2,427	2,037	1,834	1,809	1,424	1,376	1,443	1,313	1,157	1,119	989	943	950	1,080	1,046	35,850
Mon	1,046	1,347	1,388	1,497	1,548	1,557	2,018	2,518	2,700	3,355	2,918	2,634	2,498	1,991	1,833	1,821	1,743	1,528	1,360	1,215	1,200	1,259	1,294	1,209	43,477
Tue	1,175	1,382	1,512	1,514	1,542	1,593	2,002	2,508	2,713	3,249	2,838	2,521	2,420	2,072	1,986	1,931	1,786	1,600	1,420	1,239	1,225	1,272	1,270	1,199	43,969
Wed	1,186	1,407	1,508	1,506	1,538	1,580	1,990	2,501	2,654	3,163	2,768	2,536	2,439	1,979	1,894	1,895	1,808	1,581	1,396	1,250	1,219	1,243	1,311	1,218	43,570
Thu	1,180	1,456	1,483	1,509	1,514	1,595	1,990	2,574	2,624	3,206	2,701	2,461	2,386	1,973	1,871	1,866	1,750	1,576	1,427	1,238	1,199	1,264	1,310	1,220	43,373
Fri	1,167	1,393	1,497	1,515	1,505	1,603	2,004	2,538	2,625	2,957	2,682	2,434	2,291	1,930	1,868	1,810	1,748	1,524	1,377	1,218	1,222	1,239	1,276	1,212	42,635
Sat	1,292	1,514	1,651	1,540	1,517	1,623	1,911	2,377	2,365	2,409	2,079	1,866	1,742	1,472	1,461	1,420	1,308	1,202	1,087	193	155	627	1,177	1,346	35,334
Total	8,469	10,012	10,548	10,562	10,619	11,067	13,707	17,135	17,776	20,766	18,023	16,286	15,585	12,841	12,289	12,186	11,456	10,168	9,186	7,342	7,163	7,854	8,718	8,450	288,208





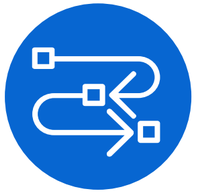
Target Architecture Considerations

Make SAS available in Cloud

- As data is migrated to Cloud, SAS should be migrated as well.
- Capacity vs consumption-based costing
- SAS governance must be built in from beginning
- Move away from EG with SAS Studio for better admin control

Provide access to on-premise databases

- Some databases that are currently being read by SAS are not scheduled to move to Cloud.
- Need to establish data governance policies for accessing and sharing data from SAS processes.



Integration with Cloud Services

- SAS can be configured to take advantage of cloud scalability.
- Job scheduling can be done using Airflow to separate orchestration from the SAS platform.
- May need to redefine op model in cloud for using services, such as email and SFTP.

Clean on-premises environment prior to migration

- Minimize number of assets that need to move to the cloud by removing or archiving inactive assets to reduce cost and effort.
- Review and optimize long-running processes to reduce load on existing Grid and simplify validation when moving to the cloud.
- Establish core governance policies that will continue in the cloud.

Recommendations

Recommendations	Key Elements	Benefits
Optimize on-premises SAS Grid with an eye towards migration to Cloud	• Modify existing Grid configuration to capture SAS log files more consistently • Upgrade to 9.4M7, and add SAS Enterprise Session Monitor (ESM) to the environment • Apply additional governance for use of service accts • Remove or archive inactive assets • Review and optimize long-running SAS processes.	• Provides improved transparency for SAS Grid usage and who is using. • Simplifies ongoing administration of the environment and can support chargeback model • Reduces scope of migration to active assets to minimize effort and future cloud infrastructure costs
Re-host SAS Grid to Cloud as near-term target architecture	• Re-host SAS Grid to Cloud with a scalable architecture. • Consider moving business users to SAS Studio, instead of using Enterprise Guide. • Move SAS workload schedules out of SAS MC into Airflow or other scheduling service.	• Quickest path to cloud for SAS and SAS users • Minimizes changes to SAS processes • Provides environment to standardize needed cloud integrations and operations model.
Once SAS usage is re-baselined after clean-up, code conversions, and cloud re-host, start planning for upgrade to Viya	• Plan for migrating to SAS Viya based on reduced usage of SAS.	• Change management can be focused on interface changes for users. • Streamlines migration to Viya architecturally as patterns and op models have already been established.

Banking Client: Datacenter Exit

sas | accenture

Use Case

01 Client has on-prem based SAS platforms which is heavily used by various Business users

02 One of the datacenter is expected to be decommissioned soon hence the SAS platform needs to be migrated on priority

03 SAS platforms need to be synergized with necessary Governance

04 The Bank has a broader Cloud and Data strategy which needs to be considered in any future transformation of the SAS environments

Key Observations – SAS Platform 1

Environment Insights

- 180+ active users across 20+ Business Units
- ~2000 SAS EGP
- 4500+ SAS Codes
- Total 6900 SAS logs assessed

Workload patterns

- 29.41% of workloads analyzed are complex
- No advanced analytics are being used in SAS

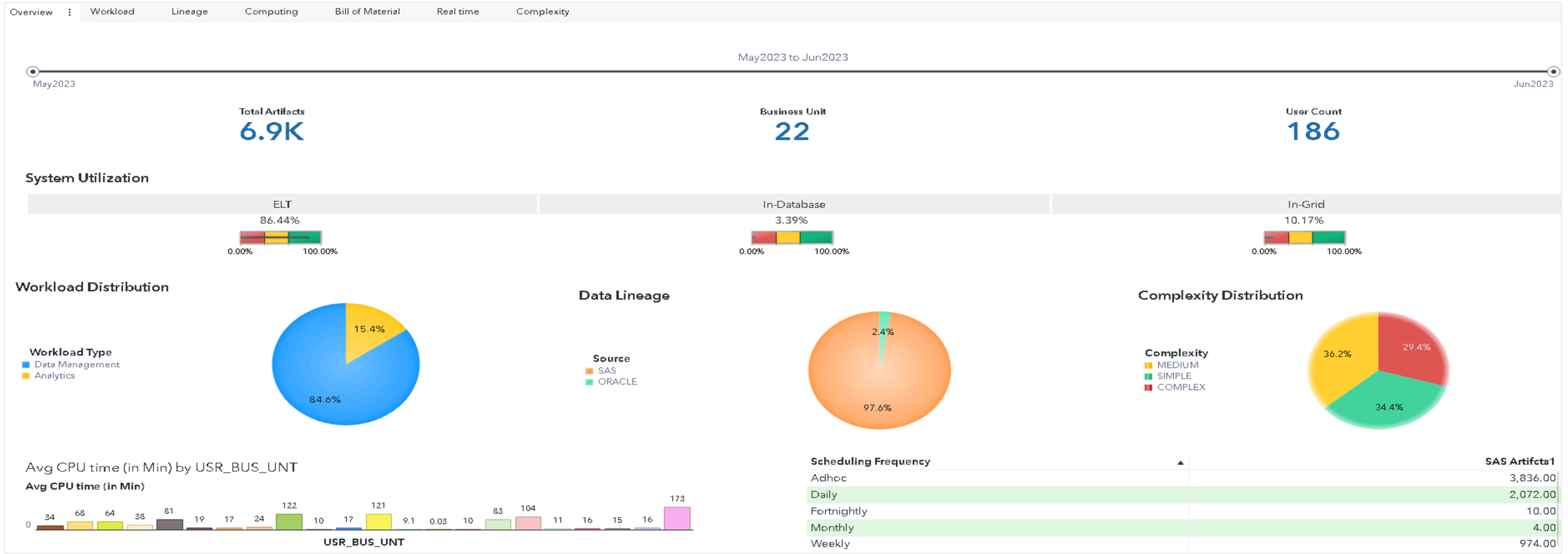
Highest Workloads

- Teams with the most workloads between May and June 2023
  - Marketing
  - Business Intelligence

20% of scheduled workloads

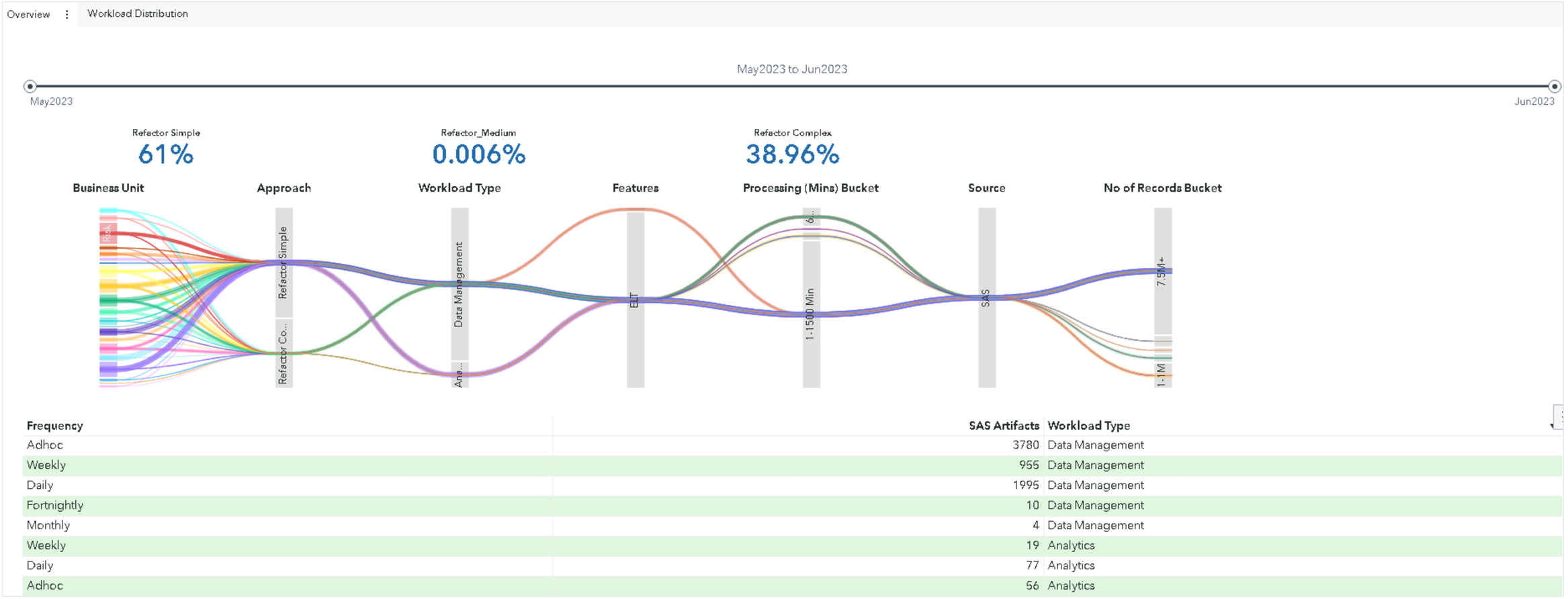
- 50+ scheduled jobs daily basis consisting +300 SAS workloads
- Most of Jobs starts prior 9 am

Discovery – Platform 1



- System utilization in ELT is more compared to In-Database
- 30% Complex indicates critical processes
- SAS is heavily used for Data Management, with 84.6% of usage categorized as Data Management. The remaining usage is 15.4% for Analytics
- SAS is not used for Reporting or Advanced Analytics

Assessment – Environment 1



Frequency	SAS Artifacts	Workload Type
Adhoc	3780	Data Management
Weekly	955	Data Management
Daily	1995	Data Management
Fortnightly	10	Data Management
Monthly	4	Data Management
Weekly	19	Analytics
Daily	77	Analytics
Adhoc	56	Analytics

- For Data Management, majority of workloads runs on Ad-hoc basis.
- Workloads are distributed across Data management and Analytics only. No Advance Analytics workload has been identified.

Key Observations – Platform 2

Environment Insights

- 260+ active users across 7 Business Units
- 2500+ SAS EG Projects
- 7500+ SAS Codes
- Total 11000 Logs assessed

Workload patterns

- 30.51% of workloads analyzed are complex
- No advanced analytics are being used in SAS

Highest Workloads

- Teams with the most workloads between May and June 2023
  - Business Intelligent
  - Finance

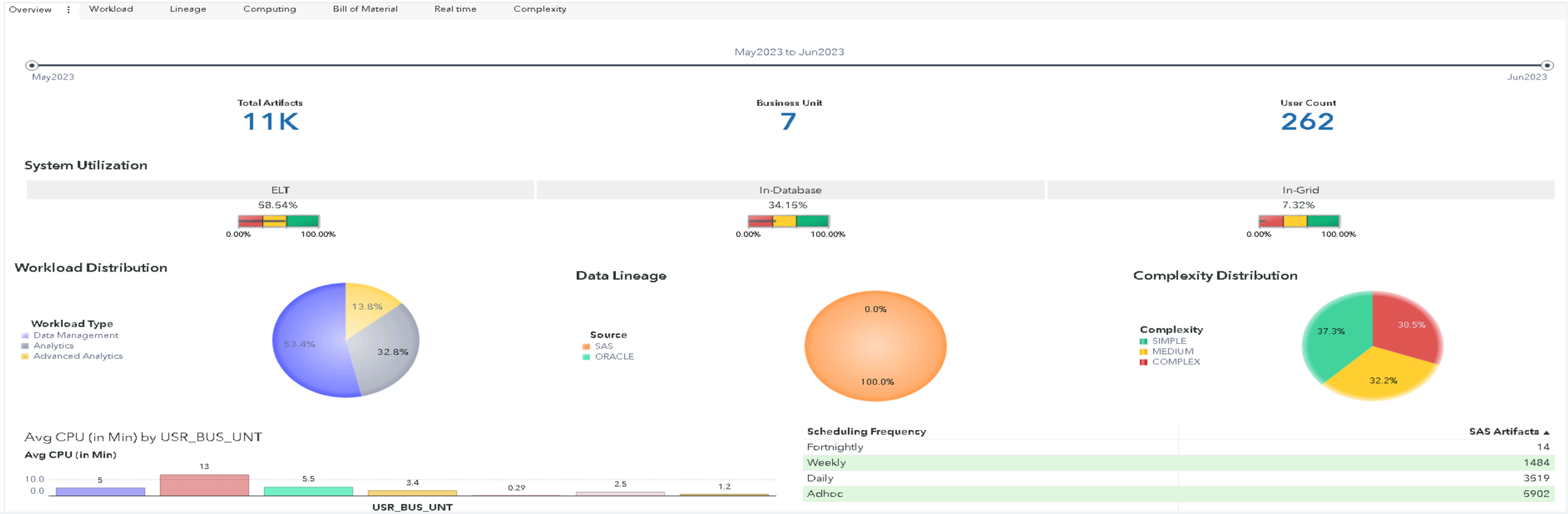
Lack of transparency in workloads

- 9% of SAS workloads are running under User IDs or SAS accounts that are not explicitly tied to a Business Unit (tagged as UNKNOWN). .

48% of scheduled workloads

- 93 scheduled jobs daily basis consisting 684 SAS workloads
- Significant load cause of long-running jobs

Discovery – Environment 2



- 9% of SAS workloads are running under User IDs or SAS accounts that are not explicitly tied to a Business Unit
- SAS, Oracle and flat files are the only data sources
- SAS is used for sending out emails. 4% of the SAS codes sent out emails. SAS is also used to issue OS level X-Commands in 2% of codes. Depending on target server architecture, these code snippets may have to be rewritten.
- Hardcoded paths are used in libname, filename, and other SAS statements in 45% of SAS codes. These hardcoded paths might have to be modified/repointed during migration, depending on the target server architecture.

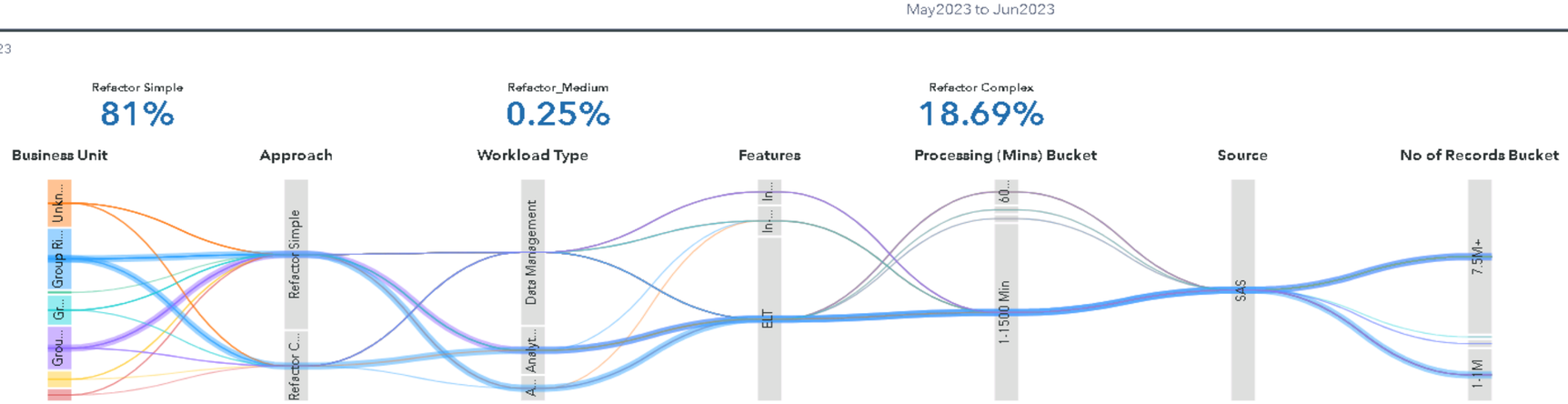
Assessment – Platform 2

Overview | Workload Distribution

May2023 to Jun2023

May2023 — Jun2023

Refactor Simple **81%** | Refactor_Medium **0.25%** | Refactor Complex **18.69%**



Frequency	SAS Artifacts	Workload Type
Daily	3120	Data Management
Daily	46	Advanced Analytics
Daily	353	Analytics
Weekly	1395	Data Management
Weekly	78	Analytics
Weekly	11	Advanced Analytics
Fortnightly	14	Data Management
Adhoc	5504	Data Management
Adhoc	356	Analytics
Adhoc	42	Advanced Analytics

- For Data Management, majority of workloads runs on ad-hoc basis.
- Workloads are distributed across Data management and Analytics only. No Advance Analytics workload has been identified.

Recommendations

With client's current operational criticality of SAS usage coupled with the compelling events, SAS target state implementation will be a multi-phase journey.

Options	Description	Address Compelling Events?	Implementation Efforts	Change Management	Consider?	Rationale
Re-platform directly to SAS Viya	Deploy SAS Viya in cloud and migrate SAS workloads to the cloud					• Move directly to SAS Viya in SAS Cloud, to reduce Overall TCO for SAS, medium Risk is achieving by 2024
Re-host SAS Grid	Upgrade SAS Grid to SAS 9.4M8					• No Change in User Experience • Least Risk option to complete by 2024 • No innovation
Consolidate SAS Platforms	Consolidate onto a single SAS Grid and update OS					• Lower TCO by rationalizing to single governance and policies • Moderate changes to User experience so Risk in Completion by 2024
Re-factor to PySpark	Move SAS workloads to Open Source or cloud native options					• High Risk and High-Cost transformation, and the migration cannot be achieved by 2024

Key Contacts

Amar Bafna

Data & AI Lead

amar.bafna@accenture.com

Jules Sabouret

Global SAS Alliance Lead

jules.sabouret@accenture.com

Ankit Dedhia

SAS and Cloud Architect

ankit.dedhia@accenture.com

Thank you!