00:00:00.000 --> 00:00:24.310
Sprague, John D (HQ-JD010)
All right. Well, why don't we go ahead and get started? I know there'll be more people joining, but looks like we already got a good crowd. Started. So, good morning, West Coast and good afternoon, East Coast. And welcome to the third talk in the AM NCA TI series. This one is aerograph architecture. My name is John Sprague, and I've been on a detail.

00:00:25.210 --> 00:00:56.030
Sprague, John D (HQ-JD010)
Thank you for allowing that person in on a detail for over 2 years from O CIO on this amazing program. Most recently is the acting data services manager and previously as the Data Control Officer. I've been at NASA since 2009 and was previously the Deputy Chief Data officer at NASA. Al Capps is the advanced Air Mobility National campaign deputy program lead and commissioned this whole series, so I wanted to say thanks again and I'll if you're on, if you have anything you'd like to say.

00:01:00.140 --> 00:01:07.340
Sprague, John D (HQ-JD010)
Alright. And uh, Michelle Eshow is also one of the architects helping put this together. Michelle, anything to share?

00:01:08.680 --> 00:01:12.570
Eshow, Michelle M (ARC-AFS)[SimLabs III Contract Management & Technical Services]
I know this is all you, John and all the data services team. Great job.

00:01:13.280 --> 00:01:15.750
Sprague, John D (HQ-JD010)
Yep. Thank you and thank you for all your help in the past too.

00:01:17.080 --> 00:01:45.810
Sprague, John D (HQ-JD010)
So now I'm gonna go over some of the speakers for today and give a little bit of their bio. So first up is Divya Bhadoria and she is the aerospace testing and integration lead. And I see her on the screen. Thank you for jumping on there and at the AMC and she's my detail boss. She is also an aerospace in researcher in the aerospace high Density Operations branch at the Aviation Systems Division.

00:01:46.150 --> 00:02:16.460
Sprague, John D (HQ-JD010)
At NASA Ames Research Center, previously she served as an analytics and benefit lead benefits lead for the Aerospace technology demonstration to the ADT two Subproject. Before joining NASA in 2019. Give you a worked as a computer scientist in Silicon Valley for over 15 years. She was in a variety of industries including biotechnology, semiconductors and surgical robotics.

00:02:16.900 --> 00:02:26.730
Sprague, John D (HQ-JD010)
She received her BA degree in Computer Science and Engineering from India and a master's degree in computer science from the University of South Florida.

00:02:28.190 --> 00:02:47.560
Sprague, John D (HQ-JD010)
Miss Bhadoria is a member of the American Institute for Aeronautics and Astronautics, and her research interests include applications of machine learning techniques towards developing safe and efficient airspace for all those new entrant autonomous flying vehicles, anything to add Divya.

00:02:48.740 --> 00:02:50.370
Bhadoria, Divya (ARC-AFH)
No, thank you for the introduction, John.

00:02:51.020 --> 00:02:51.630
Sprague, John D (HQ-JD010)
Thank you.

00:02:53.390 --> 00:03:01.920
Sprague, John D (HQ-JD010)
Next up is Tim Bagnall. He is a principal analyst in Mosaic at Mosaic ATM. He has over five years.

00:03:02.390 --> 00:03:31.990
Sprague, John D (HQ-JD010)
Uh working here with uh work with the mosaic ATM that is focused on aerospace engineering related projects, including the FA UTM evaluation project, NASA's high Density vertiflex project called HTV, and a NASA management by trajectory research effort. In addition to his current work on the national campaign, Mr Bagnall is also conducting research for the NASA Glenn Research Center regarding detect and avoid solutions.

00:03:32.320 --> 00:04:00.940
Sprague, John D (HQ-JD010)
For general aviation pilots and that's so that they can avoid encounters with drones using the FA's remote ID rule Prior to joining Mosaic, Mr Bagnall provided expert human factors engineering HFV services to the United States Navy for its MQ 4C Triton program, and it's specializing in detecting avoid solutions for remote air vehicle operators. Anything to add, Tim.

00:04:02.070 --> 00:04:03.850
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
No, thank you, John. Pleasure to be here.

00:04:04.520 --> 00:04:27.530
Sprague, John D (HQ-JD010)
Thank you. Next is Amanda Unger. She is. She has worked as a software engineer with Mosaic ATM since 2015. During her time at Mosaic, she has contributed to NASA's a TD two and the DIP projects, as well as joining the national campaign project in spring 2021. Amanda, anything to add?

00:04:30.090 --> 00:04:31.520
Unger, Amanda (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
No, thank you for the introduction.

00:04:32.130 --> 00:04:32.700
Sprague, John D (HQ-JD010)
Thank you.

00:04:33.910 --> 00:05:03.030
Sprague, John D (HQ-JD010)
And our main speaker, Jerry Wilwerding, he has a bad a bachelor's and a master's in mathematics from the University of Nebraska, Lincoln. He is a software engineer since 1996, a software technical team lead since 2004, he has worked. He has been working NASA AMC since the spring of 2020. Previously, you worked at Harris Peraton from 1996 to 2020.

00:05:03.870 --> 00:05:08.610
Sprague, John D (HQ-JD010)
There he worked primarily the weather program with the Air Force and the FAA.

00:05:09.380 --> 00:05:14.530
Sprague, John D (HQ-JD010)
And he has been at Mosaic since February 2020. Anything to add?

00:05:16.430 --> 00:05:17.870
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
No looking forward to the talk.

00:05:16.520 --> 00:05:16.880
Sprague, John D (HQ-JD010)
Turn.

00:05:18.700 --> 00:05:20.970
Sprague, John D (HQ-JD010)
Excellent. Thank you. So.

00:05:22.060 --> 00:05:23.110
Sprague, John D (HQ-JD010)
Let's get started.

00:05:24.390 --> 00:05:26.100
Sprague, John D (HQ-JD010)
If we can go to the next slide, Jerry.

00:05:30.290 --> 00:05:45.360
Sprague, John D (HQ-JD010)
So if you look on there, here's the website for the technical talk series and I don't wanna say say thanks to Rihanna de Los Santos for putting that together and I'm going to go over a little bit about the NASA AMC.

00:05:46.270 --> 00:05:47.500
Sprague, John D (HQ-JD010)
Tech talk series.

00:05:48.580 --> 00:05:58.350
Sprague, John D (HQ-JD010)
What we're trying to do here is and what we want is to engage with the community on types of technologies NASA is designing and developing and support of advanced air mobility.

00:05:59.330 --> 00:06:09.560
Sprague, John D (HQ-JD010)
Deshal talk tech talks are planned like in addition to the one we're doing right now, there's the gonna be an aerograph user interface. We have one for grafana.

00:06:10.640 --> 00:06:42.710
Sprague, John D (HQ-JD010)
Another good one I know a lot of people are gonna want to see is integrated data product and we are looking for more. So if you know of other technologies in support of this program, please let me and Michelle know so that we can get them added to the to the list and you can go through that website and see the other. The other ones that are that are on there because we have a few more ATI and and some others. So some of the ground rules real quick answers to your questions may be in an upcoming slide.

00:06:43.080 --> 00:07:09.440
Sprague, John D (HQ-JD010)
It's OK to ask that question if you you'd like to, but if if you can wait a little bit, please do so please mute your mics unless you need to talk and we'll keep. I'll keep an A parking lot for anything that comes up that that we don't want to like, spend half an hour on, but we definitely want to bring that up and talk about it as much as we can. So thank you. But I might have to use it on a parking lot.

00:07:11.350 --> 00:07:22.060
Sprague, John D (HQ-JD010)
Remember that that this is being recorded for NASA and its partners, and will be posted publicly afterwards after I go through one more approval process and.

00:07:22.930 --> 00:07:27.250
Sprague, John D (HQ-JD010)
And that is it. Uh. Divya, can you kick us off and next slide please.

00:07:28.150 --> 00:07:28.780
Sprague, John D (HQ-JD010)
And thank you.

00:07:29.930 --> 00:08:01.550
Bhadoria, Divya (ARC-AFH)
All right. Thank you, John, and thank you everyone for joining us today for the next chapter of the national campaign, ATI Tech talk series. Today's talk is focused on data processing or data management system that we call error graph. So the agenda for today's talk is First off, Tim will start the talk by providing an overview of national campaigns, data management system or aerograph and he will also talk about the types of data we collect.

00:08:01.640 --> 00:08:07.690
Bhadoria, Divya (ARC-AFH)

The goal and vision for this data and the capabilities and features that are data management system provides.

00:08:08.580 --> 00:08:32.190
Bhadoria, Divya (ARC-AFH)
Next, Jerry will talk, take us through the architecture. He will walk us through the architecture of our data management system and he will also provide details of its various components and how they all come together and work together to make our system. He will also talk about the next steps planned for our data management system and then conclude with a summary of everything that we discussed in this talk.

00:08:32.970 --> 00:09:03.560
Bhadoria, Divya (ARC-AFH)
Today's talk is focused on the back end of a data management system. This data management system produces several data, artifacts and data products, but we are not going to be talking about those in this talk today due to time constraints. However, we do have future talks planned to go into details of some of our data products. So do remember to check out the NC Tech talk schedule on the link that John shared in the previous slide.

00:09:04.790 --> 00:09:05.600
Bhadoria, Divya (ARC-AFH)
Next slide please.

00:09:07.690 --> 00:09:37.200
Bhadoria, Divya (ARC-AFH)
Can you, uh, before I hand it over to Tim, I just want to take a moment to emphasize and talk about the importance of the complexity of data that we we are working with. So this picture that you see here is a conceptual overview of AM ecosystem. And without going into much detail here, I just want you to focus on all the links that you're seeing in this picture and each of the that link is.

00:09:37.700 --> 00:09:48.760
Bhadoria, Divya (ARC-AFH)
Communication, data communication and sound form. This could be communication between different vehicles, what we call vehicle to vehicle communication. It could be vehicle to infrastructure communication.

00:09:49.440 --> 00:10:02.500
Bhadoria, Divya (ARC-AFH)
It could also be a communication in such as operational intent. The precision constraints are messages for contingency management and so on and so forth.

00:10:12.210 --> 00:10:42.440
Bhadoria, Divya (ARC-AFH)
In addition, on national campaign, we have the opportunity to work with vehicle OEM's and through our test activities with them, we get to study the vehicle status under different kinds of maneuvers and we understand that this is a emerging market and very competitive market and our weekly partners have entrusted us with some of their proprietary information. So we take data security and data governance.

00:10:42.540 --> 00:10:57.340
Bhadoria, Divya (ARC-AFH)
Well, very strictly here and we have very stringent data governance and data security policies and procedures in place to make sure that we limit access of data to the right people.

00:10:57.990 --> 00:11:17.350
Bhadoria, Divya (ARC-AFH)
So I hope this gives an appreciation of the complexity of data that we work with and hence the complexity of the data management system that's needed to create a data products, it artifacts out of this. So with that I hand it over to Tim, take it away, Tim.

00:11:19.810 --> 00:11:49.620
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Hey. Thank you, Divya. Like Divya mentioned before Jerry Dells into the technical details of the aircraft architecture, I just have a few slides here that hopefully like the NC One OV that did you just share provide a better context and background for this system that we'll speak of the architecture of. So it is aerograph. And really as you can see with this slide here, it's all about the management of data, so.

00:11:49.750 --> 00:12:17.670
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Paragraph is NASA's data management system for Advanced air mobility, and then you'll see these next two bullets include air graphs, mission and vision statement, and you'll see that they're very closely related. It's mission statement is to support AM research and really to provide a reliable and secure system that collects, stores, protects, and shares am data. So it's quite an end to end solution, really all about managing data.

00:12:18.690 --> 00:12:49.040
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
It's vision statement. It's really just a step further and it's really all about supporting the safety research community, which includes status scientists, aerospace engineers, research analysts as a place that they can trust to obtain and see data to perform those key analysis that they need to do to support the overall and see program. Some of those analysis that you'll see as I talk over the next few slides, we're hoping can be automated within aerograph, as you'll see one of the capabilities.

00:12:50.610 --> 00:13:01.850
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
At the bottom of this slide, here, we just have a few pictures that I hope convey how important data is to not only ATI and data services, but to AM and then the overall national campaign.

00:13:02.620 --> 00:13:09.230
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
And then one of my favorite quotes about data there at the bottom by Moran in terms of you can't have information without data.

00:13:10.440 --> 00:13:11.850
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
OK, next slide please, Jerry.

00:13:13.650 --> 00:13:44.280
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
OK. Like did you mentioned, this slide covers the types of data that error graph is going to help manage. And as you might imagine, it orbits and revolves around flight test events either preparing for flight test events, the actual flight test event itself or artifacts that come out of analysis from those flight test events. We've broken it down into the following five categories. They're more than this and a lot of times what you'll see in my slides here, it's not exhaustive, but it gives you an example of what we're collecting.

00:13:45.100 --> 00:14:14.820
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
So Eric graph does help manage aircraft performance and characterization data can think of that of positions reports telemetry. For instance, an aircraft might be broadcasting ADSB reports. Aircraft is going to help manage that data. Another category is air space, which would include things like operation intents, weight point constraints, all that adaptation that really helps slice and dice the airspace in terms of its three dimensional volume and also a lot of times.

00:14:14.940 --> 00:14:15.530
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Time.

00:14:16.950 --> 00:14:46.660
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Then there's an environmental category, and good examples here would include surface and wind weather, which are really important for the performance of these aircraft, especially as they land or take off from landing areas or verta ports. And then there's an infrastructure category and a good example here is surveillance coverage, so you might imagine that you have a network or satellite of adsb receivers. This would get after how well that network or satellite does in terms of covering certain spots.

00:14:46.820 --> 00:14:54.270
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
To enable the mission, for instance, you might be short a few ADS, B receivers to get the kind of coverage that you need for surveillance.

00:14:55.560 --> 00:15:24.610
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
And then lastly this category, that derivative analytical artifacts, once you have all this great source data, what aircraft can do is invoke a number of algorithms within to produce these key analytical artifacts such as a glide path performance chart, 3D position chart and an integrated data product. And again, this is not an exhaustive list. These three examples are just three examples of many analytical artifacts that error graph can produce.

00:15:25.210 --> 00:15:30.650
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
So we have some examples at the bottom of this chart and the lower left you've got a glide path performance chart.

00:15:31.300 --> 00:15:54.080
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
This was critical during FFT and build 2 and what this gets after is how well an aircraft can conform to different glide angles as it approaches a vertical port. Fotm built two took a look at degrees of 369 and 12 as an aircraft approached, 12 being a relatively steep angle to approach a vertical.

00:15:55.320 --> 00:16:17.590
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
In the center, we've got a 3D track position chart, which aircraft will also produce. This is just a static image here, but within Aerograph, a UI that we'll talk about a bit, it will be dynamic so that the analyst can zoom in, rotate things of that nature. But that will give you the 3D position of the aircraft that allow it to in terms of latitude, longitude and altitude.

00:16:18.940 --> 00:16:49.710
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
And then the lower right we have and this is just a subset, the integrated data product or IDP as you'll hear. And although we were showing just six fields, a typical IDP will include upwards of 200 fields and essentially what the IDP does with all that data, it's collected, it collates, it integrates and it synchronizes time, synchronizes this data into a large table to enable research and a very quick example here is that.

00:16:50.650 --> 00:17:20.370
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
The analysts can tell at a certain time where the aircraft is in terms of its 3D position and say the wind and gust wind that that aircraft is facing. Again, this is just a subset of the type of data that you've seen in IDP. The UDP will most likely include a somewhere around 200 fields, but it's gonna include information such as DPS MU telemetry data, but also other important aircraft performance characteristic data like and sectors that the pilot.

00:17:20.570 --> 00:17:51.490
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Is manipulating the control control surfaces on the aircraft, and in particular those control surface angles, like for instance, let's say a particular aircraft in the cell angle as it rotates throughout flight. So the IDP we're particularly proud of from the data services perspective and you'll notice I've got an asterisk at the bottom right there. John already mentioned this, but the IDP is the topic of a future tech talk and it's slated right now for later this summer. We hope you tune in for that talk.

00:17:51.590 --> 00:17:51.940
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
2.

00:17:53.200 --> 00:17:54.310
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Next slide please, Jerry.

00:17:56.120 --> 00:18:09.390
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Alright. The last slide in my just overview of error graph here. This includes capabilities that aircraft

provides. These are general capabilities. Again, this is not exhaustive, but these five here help provide an overview of what aircraft does.

00:18:10.270 --> 00:18:21.220
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
One of the first things you'll note is the very first capability is shaded in a different color. That's because there's a different team within the ATI team. The data pipeline team that manages it.

00:18:22.000 --> 00:18:38.030
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Nonetheless, the pipelines, all about the ingest of real time data like position, operation, intent, wave points, things of that nature. Nonetheless, we look at it as a sibling capability. And since air graphs all about managing data, we include it here.

00:18:39.330 --> 00:18:51.290
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
About a month ago, John mentioned this. There was a tech talk on the data pipeline given by Irene Smith and her team. If you didn't get a chance to tune into that, you can go to that website there and take a look at that tech talk.

00:18:52.920 --> 00:18:56.690
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
If data cannot be ingested in real time to aircraft.

00:18:57.430 --> 00:19:09.780
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Aircraft does have this post flight ingest, an extraction, transformation, and loading capability. Oftentimes the data instruments that you have available at a field or a test field.

00:19:10.620 --> 00:19:23.110
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
They are independent systems that write files kind of independently, and those files are then given to the data services team or error graph at a later date. Sometimes it's just later that afternoon, sometimes it's the next day.

00:19:23.740 --> 00:19:35.040
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Nonetheless, what aircraft will do is we'll take that CSV file and perform ETL on it to store it into the right databases and secure it so that data is available for analysis.

00:19:36.350 --> 00:20:05.540
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
So between the pipeline, which is real time and the post flight ETL capabilities, you got all this great data and aircraft provides this other capability called what we're calling automated analytics. And you could think of that as a library of algorithms or sometimes we call them computational functions that we can invoke on that data. And you've already seen examples of this that would include the GPA chart or the glide path angle chart. The IDP is also an artifact and a number of other.

00:20:05.600 --> 00:20:12.530
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Artifacts that can be produced and that library is not static. What we're hoping is that as the larger.

00:20:13.320 --> 00:20:25.210
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
At research, community embraces air graph. That library grows. So if you have ideas for algorithms that you'd like to write or or include an error graph, please let us know. We'd be happy to accommodate you.

00:20:26.690 --> 00:20:45.970
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
So between these three capabilities, we have all this great data. The next capability is this access and retrieval capability, right? We'd be remiss if we miss this. The idea is that we want to serve this to the larger NC community so that they could retrieve the raw data or derive analytical artifacts data.

00:20:46.840 --> 00:21:16.770
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
So that they could perform their research as necessary. And there's two ways right now that error graph provides that data. One is through NASA's box interface, which is a secure cloud file and folder sharing system. And then there's a second system called the Error graph front end or UI. And as John mentioned, that's going to be a topic of a future tech talk. Again, probably sometime later this summer, and we hope you tune in for that too. And then lastly, this last capability.

00:20:57.180 --> 00:20:57.410
Bridges, Wayne (ARC-AFS)[SimLabs III Contract Management & Technical Services]
It's.

00:21:16.850 --> 00:21:21.170
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Security is really a cross cutting or overarching capability.

00:21:21.890 --> 00:21:47.220
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
The nature of and see and events. Air mobility is that Nas's established these relationships with industry partners. Those industry partners are entrusting proprietary data to NASA. So this is just gets after the fact that we're applying best practices in security to protect that data. To make sure that only qualified people can access it and see it.

00:21:48.990 --> 00:22:02.330
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
All right, I think that is the summary of these three sites. Hopefully that gives you a good overview of what we see as error graph and what Jerry will now delve into is the technical details of the architecture that makes the aircraft system work.

00:22:07.860 --> 00:22:11.290
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
OK. Thanks, Tim. So at a very high level.

00:22:12.050 --> 00:22:38.390
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Umm this picture kind of depicts the the layout of the architecture. So there was two types of data that we primarily handle. So there's the real time data and the post flight data. So at the bottom is the real time data sets which the pipeline team is responsible for. So that data will come in in real time during the flight test to be routed through the NASA AWS gets stored into an internal data store.

00:22:39.720 --> 00:23:09.230
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Most everything that I'm going to be talking about today is dealing with the post flight data. So on the left hand side we see a nominal representation of what the I guess types of what the post flight data we support more than what's listed here. But generally speaking that data is gathered collected if it needs to be preprocessed or whatever. That's all done when it's ready to be given to us, it's put into NASA's Box, which is basically a secure.

00:23:09.640 --> 00:23:31.080
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Shared folder. From that viewpoint, one of our engineers will take the data, route it through our system where it goes through the ETL which is the extract transform and load that data gets put into a store and then it can be served out through our rest services and consumed either by our custom hair graph user interface.

00:23:32.660 --> 00:23:33.270
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Through.

00:23:33.940 --> 00:23:50.370
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
You know, client produced a web services. I guess clients that can access the web services or the artifacts that we create can be uploaded into box for consumption that way. So this is like a 50,000 foot picture of what the architecture is.

00:23:52.210 --> 00:24:16.380
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Before we dive into some lower level views of it, it's worth mentioning that we make extensive use of free and open source software, which is very, very common with software development processes. Main reasons why we're doing this is proven reliability, faster development times and responsible use of taxpayer dollars because we're not reinventing core infrastructure capabilities. So the types of fasts that we leverage.

00:24:17.350 --> 00:24:39.050
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
We're first going to talk about the data stores, so we'd get lots of data coming into our system. We have to have that data live somewhere. Once it gets into our system. So we've chosen several data stores, the first of which is time scale. So time scale is a Postgres extension which is optimized for time series data. So it's our relational data store.

00:24:41.260 --> 00:25:11.840
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Pretty traditional stuff there. Uh, the Mongo database is what we use for our unstructured data. So it's a document based data storage. I think JSON Data. It's really good with that type of stuff. We also have a graphing database that's called digraph, so that's good for capturing complex relationships which may not lend themselves to the relational data stores quite as well, and it's not really a data store, but we leverage a parquet file format which is very efficient for.

00:25:11.930 --> 00:25:19.400
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Large datasets. It's column oriented, so we use that primarily as our main file store for the IDP.

00:25:22.260 --> 00:25:51.130
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
So as you'll see in later slides, our system has a lot of components and those components have to talk to each other. And so we need to figure out how to do this interprocess communications. We leverage active MQ for most of our messaging that goes between these components. So what's nice about active MQ is that you can send messages to queues or topics and any number of subscribers can be listening to those. And those subscribers can be in multiple languages and on.

00:25:51.220 --> 00:26:00.740
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Multiple platforms, so it gives us a nice kind of way to segregate our pieces and still allow communications to take place without creating a stove pipe kind of system.

00:26:01.560 --> 00:26:30.980
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Umm to integrate these pieces together, we leverage Apache Camel which is just a rich enterprise integration type of the system. So it's has just lots of integration patterns within it. It's a nice way to change the program's behavior through configurations just by kind of wiring in a new piece to do something at a certain point and sending the messages between those pieces. So it's pretty powerful from that viewpoint. It allows us to.

00:26:31.070 --> 00:26:46.530
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Inject new behavior without, you know, affecting the rest of the system and then Apache HTTPD. Access our gateway into our system. Kind of like a reverse proxy. It routes all of our requests between between our components.

00:26:49.370 --> 00:27:21.160
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
So security is a constant theme throughout this talk because of the proprietary nature of the data that we're dealing with. So we leverage Keycloak to give us both the authentication and authorization aspect of things. So key click will integrate in with NASA Launchpad, which allows us to have PIF based access to the system, at least on our production environments. It also provides username and password access, which we typically use in our dev and integration environments. And then we have authorization.

00:27:21.240 --> 00:27:35.240
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Implemented through our back capabilities. So we have a bunch of roles and groups which are associated with users which are mapped down to very granular levels of the data and you have to be in the correct roles and groups to.

00:27:35.850 --> 00:27:37.680
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Basically access that content.

00:27:41.690 --> 00:28:09.780
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
So once we have all this data into the system, we have to have a way of sharing it with others and so the primary ways of sharing data with people is either through how our aerograph user interface, which is just a custom browser application that we have developed which allows individuals to view the analytics that we have created to preview data, or preview the IDP, and also to download those topics. Assuming you have the authorization to do so.

00:28:10.300 --> 00:28:20.870
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Umm, so it's kind of a rich interface which allows access to all the data that we produce and it will be a topic of the future tech talk coming up.

00:28:22.140 --> 00:28:32.800
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
From a machine to machine viewpoint, it's all rest services, so lots of data can be accessed in that manner. Most of the responses are JSON and we'll form schemas.

00:28:34.800 --> 00:28:35.230
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
The.

00:28:36.050 --> 00:28:43.220
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Files that are not Jason would actually be the imagery files that are created which are are metrics or our analytics.

00:28:44.540 --> 00:29:06.310
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
And the final way was share is just by uploading those artifacts into box boxes and secure way of accessing the data. There's an extensive vetting process you got to go through your DUI training, names, approvals and then every folder and file can be locked down to whatever level of controls that we need to make sure only authorized people are accessing that content.

00:29:08.800 --> 00:29:22.890
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
And from a deployment viewpoint, we deploy all of our code currently within Docker, so it gives us a

container based way of deploying the applications based images. Give us a consistent starting point for our various software deployments.

00:29:23.990 --> 00:29:47.180
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
We throw our custom code on top of those base images and then we can tailor that custom code per deployed environment. Dev integration, production. For example, just by changing configuration files, Docker gives you a lot of capabilities out of the box. Things like container, container management and network management ability to scale just through configurations, and the list goes on and on.

00:29:49.570 --> 00:30:18.680
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
OK, so now we can actually talk about components within the system and putting things together. So at this picture I think helps at least see the various parts of the system that we have. So the kind of darker green color in the upper left that's NASA Launchpad, it's not really part of our system, but we integrate with it. So the lighter green colors are all of our various Docker containers. So separate pieces of software that are deployed within the Docker environments.

00:30:19.200 --> 00:30:49.250
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
The purple pieces represent data at rest. Now that data at rest could be a volume that can only be accessed from a single Docker component, or it could be a shared volume such as over here on the left, where we basically write our persistent data that can live outside the lifespan of a particular container. Or we have things like box down at the bottom which lives off in the cloud. But basically what happens when you look at a typical.

00:30:49.540 --> 00:31:19.380
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Kind of interaction going through the system is, you know, some user is going to make a request and let's just pretend they're logging into the UI so that login request is gonna go through our HTTPD component. That component is smart enough to route things through our system, so that request will go to the graph UI. That UI is going to interact with key cloak and our access components to you. Make sure the user is both authorized and authenticated to access data. Once all that's stuff done.

00:31:19.910 --> 00:31:39.300
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
You now can request data by following this path to our rest Docker interface and that rest component is responsible for pulling data out of our Postgres databases, which is a time scale database or our Mongo database and providing that data back up to the individual users.

00:31:40.260 --> 00:31:46.110
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
I'm from an ingest viewpoint. The data comes in through the top through this external volume. It goes into the ingest component.

00:31:46.760 --> 00:32:12.060
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]

Various messages are written to this active MQ component, which then our computational functions, which is our kind of metric per data source generator or however IDP, IDP, Docker component down here we'll both get those messages, do processing, create products and make those products available through various ways. Our static configurations are largely stored up here within digraph.

00:32:13.160 --> 00:32:13.480
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
And.

00:32:15.680 --> 00:32:46.330
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Yeah. So at A at a high level, I guess those are all the pieces that are in play and how they kind of interact with each other. So it's worth calling out this blue box. So the blue box within our architecture allows for these pieces to be to basically deployed, be deployed on a per event basis. So for example, we could deploy like a build up run to set of containers down here and then like an foft.

00:32:46.450 --> 00:32:56.640
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Set of containers a partner, a set of containers. So basically per flight test there could be a separate set of containers that are out there which work well.

00:32:57.300 --> 00:33:03.030
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Together and possibly have different capabilities, different output formats than.

00:33:03.710 --> 00:33:15.480
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
You know the period containers for different flight test event, so this is good for a couple of different reasons. One, it gets rid of backwards compatibility. So if we do something for FFT for example.

00:33:16.220 --> 00:33:17.810
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
And then want to completely change everything.

00:33:19.070 --> 00:33:48.830
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
We can we have that freedom to completely change everything because the FFT software is kind of isolated from, you know, future flight test events. So that's, you know, a nice component. It allows for each flight test to be tailored independently. If you had multiple flight tests going on at the same time, they'd have separate systems. So the architecture allows for that, you know, truth and advertising, it's not fully developed yet, but it's very, very close.

00:33:49.150 --> 00:34:08.340
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
So, but it's a key part. A key portion of the architecture just to allow that type of capability and that's where the reverse proxy coming in here is pretty important because it has to be smart enough to route any request to the correct event instance of the rest containers.

00:34:12.110 --> 00:34:14.910
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
OK. So if we wanted to talk nominally about how?

00:34:15.750 --> 00:34:20.240
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
You know, our automated processing happens so.

00:34:21.500 --> 00:34:44.400
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
I'm going to walk you through the ingest kind of process, so generally speaking, a set of files or dropped into a directory and there's some process that's monitoring that directory. So when those files show up, we recognize that we send those products to you, hey, validation queue and some type of validation takes place. These files would we expect them to be.

00:34:45.060 --> 00:34:58.520
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Uh, do we have all the required metadata about those files? If that answers yes, we can send that content to an extraction queue in which we actually extract that content or parse them out once that data is parsed out.

00:34:53.560 --> 00:34:54.070
Bridges, Wayne (ARC-AFS)[SimLabs III Contract Management & Technical Services]
Tough love.

00:34:57.750 --> 00:34:58.430
Bridges, Wayne (ARC-AFS)[SimLabs III Contract Management & Technical Services]
The parking.

00:34:59.270 --> 00:35:10.290
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
We can transform that data and transforming that data includes things like unit conversions, creating drived parameters. Basically whatever needs to be done to that data.

00:35:11.510 --> 00:35:40.740
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Prior to storing it into a data store, and once that's done, we can send another message to a queue to actually do the persisting or the loading of that data into our data stores. Once we have the data within our system, we then can kind of broadcast out several messages at once, which is what this horizontal line is trying to represent, and so we send messages to our computational functions, sorting decompositions at the same time as we send messages to our IDP.

00:35:41.260 --> 00:36:07.340
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Processing queues. So within the sorting processing workflow will actually query some of the real time data, the event markers to determine all the parts of the particular sorting that were relevant for that flight test, and then we wanna create metrics on a per source basis. For example a DSP, DPS, the weather sources, whatever they may be. We wanna create our our analytics.

00:36:09.160 --> 00:36:13.820
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
On a per maneuver per sorty basis. So we create a whole bunch of products from that.

00:36:14.710 --> 00:36:44.460
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
I'm concurrently in parallel. We do the same thing with the ADT, where we'll generate the IDP itself for the product that was just ingested, and then we will go through and create the glide path analysis and the track overlay products again in a broadcast kind of manner. So these are in parallel. So I was talking earlier about Camel and how you can inject behaviors. So essentially through configuration changes we can inject as many other queues as we want within this processing chain and.

00:36:44.610 --> 00:36:45.790
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Create new products.

00:36:46.480 --> 00:37:14.540
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Whatever we need to do, Umm, same kind of thing on the left hand side where we're talking about doing the product decompositions and the ETL's of it. You know, we can add new queues here that are specific to a particular data source and handle those appropriately. And that also led illustrates the language independence that we have. So generally speaking, everything on the left side is written in Java, everything on the right side is written in Python, but they can be any languages that.

00:37:15.200 --> 00:37:30.450
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
You know, people are familiar with, you know, the core requirements are that it has to integrate with active MQ because that's where the messages are being sent to you. And then you have to be able to read the data from the data stores or from the IDP itself. So there's lots of.

00:37:31.680 --> 00:37:39.940
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
I guess ways to extend the software just through changing the configurations and the queues and who's listening to those queues.

00:37:42.570 --> 00:37:46.240
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
OK. So if we wanted to dive into the weeds a little bit.

00:37:47.470 --> 00:37:57.120
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
And we can talk about how ingest actually works at a more granular level. So I've set that there's these set of files that need to be.

00:37:57.950 --> 00:38:31.040
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Ingested. So there's actually the raw data file and the metadata file, so the metadata file contains things that the raw data file doesn't necessarily contain. Things like the partner who is associated with this

flight test event, the actual flight test event name, the sorty of the given day, things like that are in the metadata file. And so that metadata file and the raw data file is kind of go together. And the first thing our ingest processes are going to do is verify that that stuff is correct.

00:38:31.300 --> 00:38:39.580
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Assuming everything's correct, we store the metadata information into our database and then we send our ETL messages to you active MQ.

00:38:40.700 --> 00:39:02.850
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
And once that's done, our ingest data processing code is going to recognize that message. It's gonna read the message. It's gonna perform the extract, extract transforms and loads. So in our current Java implementation, this is all kind of done within one key step. Even though anomaly, it could be 3, so we could easily split that out if we needed to.

00:39:04.050 --> 00:39:20.510
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Most everything we've seen so far, it's just easier to do it all within one logical operation. Once the data is stored off to the data store, we send off those two messages in parallel that I just talked about going off to our computational functions and going off to our IDP.

00:39:22.460 --> 00:39:41.190
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
So once those messages are sent to the broker, we essentially now have parallel processing going on because the computational function generator is going to read those messages at the same time as the IDP generators going to be reading those messages. So we get, you know the.

00:39:41.910 --> 00:39:47.040
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
I guess parallel processing making things go faster, so let me get back here.

00:39:54.000 --> 00:40:25.630
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Within the computational functions, we're going to do that query for the maneuvers within the sorty determine all the products that we have to create based upon the number of maneuvers and the number of data sources that we have for each one of those products, we're going to send a separate message to activemq and sending the separate message again allows more stuff to be done in parallel, because if you had more than one consumer to that queue, you could generate these products in parallel all the way through, because each one of these consumers will be reading the products.

00:40:25.920 --> 00:40:32.300
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Generating the products and persisting that information and the same kind of thing is happening down here with the IDP processing.

00:40:39.370 --> 00:40:45.420
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
So this is a slide with a lot of words just describing what I said. This slide is mainly here for those who.

00:40:46.230 --> 00:41:10.020
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Are not listening to this presentation, so they can kind of make sense of that previous slide, important things to pull out of these words are design patterns that are in play, which make it very easy to ingest new data sources and new, I guess analytics or metrics that can be algorithms that can be ran within our function.

00:41:10.980 --> 00:41:12.370
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
I'm using the.

00:41:14.080 --> 00:41:42.150
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Active in Q uh queues and topics can allow for vertical scaling just by adding consumers to those queues. So assuming your server has enough resources, it's a very easy way to scale the application and one thing that I didn't call out specifically is that when we store these analytic products that we create, all of that metadata goes within Mongo, which gives us an easy way to query that data and make it available to.

00:41:43.760 --> 00:41:44.570
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
I guess clients.

00:41:48.150 --> 00:42:08.370
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
OK, son, data is in the system now. We've got to make it come out right. So we're we leverage a rest services for that kind of interaction post and get methods or what we're leveraging the more complex kind of request or leveraging the post request with a JSON object within the body that has I guess.

00:42:09.070 --> 00:42:32.000
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Content conforming to a schema that allows us to filter objects based upon whatever the users requesting. The more straightforward, simple request just leverage Git following the KVP pattern, so that's the key value pair patterns. As I mentioned before, the responses are mostly JSON coming back out, and the types of things that you can request from our system would be the metadata about.

00:42:33.250 --> 00:42:40.790
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
The data or the analytics that we created, so that's an important thing. So we tell you what we have. So then you can actually query the underlying data.

00:42:41.910 --> 00:42:47.950
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
You can retrieve the actual data that was stored, or you can retrieve the analytics that we have created.

00:42:50.380 --> 00:42:53.010
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
What is the basic retrieval look like?

00:42:56.190 --> 00:43:14.100
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
So your basic retrieval is going to start off with a clients, whether that's our browser or you know the custom client that somebody creates, you're going to have to request a token from keycloak so this access token will give you, I guess it serves as a.

00:43:15.150 --> 00:43:24.360
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Well, yeah and authorization token. So you make this request, we get routed through all the way to keycloak. Keycloak will then interact with Launchpad to.

00:43:25.530 --> 00:43:25.940
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Umm.

00:43:26.900 --> 00:43:53.230
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Authenticate the user, make sure they're credentials are correct and that they're valid from anatha system. Once that's done, we're going to do some interactions to get the roles and groups associated with that particular user and all that information kind of comes back within the token. So any request going forward must include that token in it, and the token is valid for a period of time. I think it's 5 minutes by default.

00:43:54.450 --> 00:44:18.020
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
And you can renew that token or if it expires, you can request a new one. But with that token coming in you now can ask for data. So I request for data will come in with that token. The rest services first thing it's gonna do is validate that token, make sure everything is good, it's gonna validate your access permissions and your authorization for whatever data sets you're trying to retrieve.

00:44:36.820 --> 00:44:49.060
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
OK, so the preview of our user interface, so this is currently undergoing a redesign. So this is one of the first I guess, sneak peeks of it it allows.

00:44:49.710 --> 00:44:53.920
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Basically, filtering of a bunch of metadata here on the left hand side.

00:44:54.860 --> 00:44:58.160
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
And then a bunch of products that you can choose.

00:44:59.050 --> 00:45:09.470
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Any products you choose, you basically can see within this grillable window pane, and then you'll also have the ability to download the analytic products that are here.

00:45:10.670 --> 00:45:14.730
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
The raw data behind those products and also the ID key.

00:45:15.830 --> 00:45:20.620
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
So basically it can be a way for anybody who doesn't want to write custom code to.

00:45:21.250 --> 00:45:39.340
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Grab the OR deleverage the capabilities of air graph. The UI again makes extensive use of the RBAC to make sure only authorized authenticated people have access to the data, and again this will be a feature tech talk where you can see the.

00:45:40.250 --> 00:45:43.200
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Latest implementation of the user interface at that time.

00:45:47.010 --> 00:45:47.820
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
So as far as.

00:45:48.630 --> 00:46:07.400
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Next steps go UM, we're continually valving of the paragraph system, so there's always new data sources with every single flight test that we need to support. There's always a need for new analytics that need to be supported. We're currently in the middle of automating.

00:46:08.540 --> 00:46:17.430
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
The entire set of generation of these products on a per maneuver basis within a sorting, so that should be done here in the very, very near future.

00:46:20.480 --> 00:46:48.950
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
We need to scale the application up to handle the concurrent events going on at the same time. Given the schedules that are being laid out within the AEM program and obviously the AEM program is very dynamic and evolving, right? So we're we're evolving with the program to meet all of those needs. So it's pretty exciting looking forward to those changes. So as far as future tech talks go.

00:46:49.030 --> 00:47:02.270
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]

Related to everything I talked about within this particular presentation, there's going to be something on the user interface and the integrated data product, so look forward to those presentations.

00:47:03.300 --> 00:47:04.520
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Within the next few months.

00:47:06.670 --> 00:47:12.140
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
So what we talked about today was just a high level view of the air graph architecture.

00:47:14.330 --> 00:47:27.340
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Emphasize the fact that error graph is NASA's A AM data management system and we provide, you know, a secure way of collecting, storing and providing that data to.

00:47:27.980 --> 00:47:48.830
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Clients uh data government governance is baked into the system to make sure only authenticated and authorized people have access to the data. Future tech talks are in the Lync, so please see that and if you have questions outside of this presentation, you can direct those to either Divya or John. Their contact information is listed below.

00:47:51.080 --> 00:47:51.580
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Chats.

00:47:54.900 --> 00:47:55.460
Sprague, John D (HQ-JD010)
Alright.

00:47:56.340 --> 00:47:58.850
Sprague, John D (HQ-JD010)
Thank you, Jerry. Appreciate it very much.

00:48:00.030 --> 00:48:13.280
Sprague, John D (HQ-JD010)
I wanna thank Divya, Tim, Amanda and Jerry for very informative brief. We do have a hard cut off on the top of the hour, so I, but I think we have planned time for questions. First, I want to make.

00:48:14.350 --> 00:48:44.630
Sprague, John D (HQ-JD010)
Some of the comments that showed up in the chat, so Yuri mentioned a really good one about the flight data portal and would this ability here this capability have be one of the feeds to it or from it. And I don't know the answer on that yet, but hopefully we can explore that. I think that that's a great and I've heard of that portal. I've I've been to it a couple times and I I know I can see its value if anybody else has a comment on that please let me know.

00:48:41.280 --> 00:48:41.880
Capps, Al (ARC-AFO)
I I don't.

00:48:45.420 --> 00:49:09.810
Capps, Al (ARC-AFO)
That John, this is Al. Sorry I joined late, but I have a comment on that and thanks for bringing that up on the flight data portal. I did wanna mention and pass along to you and Divya and others that we have had a couple of out reaches with the FDP folks. But as as some may know, they were really kind of focusing on other projects and programs to start it off like the X 57.

00:49:10.500 --> 00:49:13.510
Capps, Al (ARC-AFO)
And that was a flat file format.

00:49:14.610 --> 00:49:40.440
Capps, Al (ARC-AFO)
You know HDF 5 or whatever and you know what we're we're really trying to do here was leverage, you know, relational data stores and structures. There's actually information in the structure of the data itself for cross validation and things like that. So in the comments, I believe it was Yuri mentioned that that might be an input into error graph and I and I do agree with that. I don't think it would be the other way around where these data would feed into.

00:50:09.200 --> 00:50:27.680
Mederos, Luis (ARC-AT)
Hey, Alan. John, when back in 2019 when we first started AEROGRAPH, we also did reach out to the flight data portal people and as it's been posted, it's more of a flight, a file and folder storage. It does not leverage any relational or graph features. That error graph has. So yeah, I concur with you Al that.

00:50:29.090 --> 00:50:32.200
Mederos, Luis (ARC-AT)
AFTP could be an input to aerograph, not the other way around.

00:50:34.350 --> 00:50:37.040
Sprague, John D (HQ-JD010)
Love it, we answered the question here on the.

00:50:37.960 --> 00:50:43.270
Sprague, John D (HQ-JD010)
On the uh talk. So thank you. That's very good. Any other comments or questions, please let us know.

00:50:43.910 --> 00:51:08.240
Sprague, John D (HQ-JD010)
UM also saw uh. Some mentions about the pointer and yeah, I I saw at the pointer was really small and and I might work on that for the next one. So thank you for bringing that up. And then Joseph Rios wanted to see functional functions mapped to the prior noted dockers. And he was wondering if it was a mini to one relationship. So Jerry or or somebody, if that wants to jump on and and maybe go through that.

00:51:08.790 --> 00:51:28.730
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Sure. So generally speaking, our analytic functions are either going to be mapped to the error graph CF component, or the error graph IDP component. Technically, they could go anywhere, but that's just kind of where we have them right now and the differentiation is based upon.

00:51:29.370 --> 00:51:31.060
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Umm, what their?

00:51:31.860 --> 00:52:01.410
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
I guess how many data sources are involved? So if it's just a single data source right now that is going into our aerograph CF. So all of the adsb metrics or all of the sodar metrics would be calculated within this particular component based upon just the ingest data flow happening. If a product or a query, I guess an analytics capability requires multiple functions. Right now we have those within the IDP.

00:52:02.010 --> 00:52:07.850
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Mainly because the IT and we really didn't talk. I guess Tim talked about it. The IDP is a.

00:52:08.560 --> 00:52:40.000
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Data set that fuses everything together so all the data that would come in is on different time stamps, some different frequencies and so the IDP basically takes those data, gets them all into similar time steps which makes it easier to create analytics off of those. So those type of things are done within this container. So things like the glide path analysis for example, which is showing the rate of descent, it also contains some weather information within it or the track overlays which.

00:52:40.080 --> 00:52:53.780
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Show the positions of the various telemetry sources we have all on one graph that you can toggle those products on and off. Those are all created within this particular container. So did that answer your question?

00:52:55.170 --> 00:53:24.960
Rios, Joseph (ARC-AF)
Uh, not, not quite. I I think that this is a great presentation by the way. I just I get very interested in the how you guys are doing this and the architecture. So like this and if you Click to the next slide that was my main question is mapping these dockers to the functions you outline on the bottom here like I couldn't do that mapping in my mind and that's just where my mind went. So and then again the next step after you put the sequence diagrams up, I didn't wanna go there, but I also was very interested in which of the dockers are the actors in the sequence diagram. I couldn't make that mapping.

00:52:56.790 --> 00:52:57.020
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
OK.

00:53:10.320 --> 00:53:11.560
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Ah, OK.

00:53:25.260 --> 00:53:29.470
Rios, Joseph (ARC-AF)
I'm just very interested in how you guys put this together, so I just may follow up later and I appreciate it.

00:53:29.920 --> 00:54:00.610
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Yeah, absolutely. But to try to answer that relatively quickly, everything to the left of the load data is being done within our ingest Docker component. The these two CF ones are within the paragraph CF Docker component and the IDP one followed with these GPA and tracks are in our IDP component. So this particular sequence diagram is covered basically within this top ingest this middle CF.

00:54:00.910 --> 00:54:06.270
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
And then this lower IDP of course with support from our data stores as well.

00:54:09.850 --> 00:54:10.590
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
And yeah, there's.

00:54:11.510 --> 00:54:19.340
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
A lot of lower level information that you know if you're curious about by all means you can write those questions through John and we can get you more detailed answers.

00:54:23.190 --> 00:54:25.100
Sprague, John D (HQ-JD010)
Yeah. Good question. Thank you.

00:54:25.560 --> 00:54:34.630
Sprague, John D (HQ-JD010)
Umm so I I saw and by the way, the IDP brief that you keep talking about will be coming up in the next few months. So thank you.

00:54:35.930 --> 00:54:39.500
Sprague, John D (HQ-JD010)
Jay Will Young had a question too. I saw your hand up. Go ahead.

00:54:40.490 --> 00:54:43.980
Jung, Jaewoo (ARC-AFH)
Yeah, yeah, this is jeru not a TMX project.

00:54:45.490 --> 00:55:14.970
Jung, Jaewoo (ARC-AFH)
In the introduction, it was mentioned that the no data don't information but quality of data matters too.

So I'm curious how you check the quality of data like a format you know if you're expecting decimal degree for latitude but the the the the degree second comes in integrity like ohh everything after the decimal is 0 for latitude and then continuity things like I'm expecting one or something but it's coming in every other minute.

00:55:15.420 --> 00:55:18.230
Jung, Jaewoo (ARC-AFH)
So those type of things or does that get checked?

00:55:19.420 --> 00:55:27.210
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Yeah, so data validation can be an extremely complex and deep rabbit hole. Once you start jumping into it.

00:55:28.790 --> 00:55:29.550
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
So the.

00:55:30.690 --> 00:55:47.050
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Simple answer, which is not necessarily a great answer, is that the extract process is expecting certain formats of data, right? So you have that check where the extraction will fail if the formats are not.

00:55:48.230 --> 00:55:49.800
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
You know of the expected.

00:55:50.660 --> 00:55:51.030
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Uh.

00:55:52.570 --> 00:55:53.990
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Yeah, of the expected formats.

00:55:55.230 --> 00:56:24.790
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Things like all of the latitudes not having decimals on them when they're decimal degrees for example really is not cut within our system yet, except it will be painfully obvious in the output that we create. So from an soft viewpoint, what we were basically doing was creating these products as both a quick look and then a final viewpoint and that quick look gives the initial indication.

00:56:24.880 --> 00:56:26.040
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Of if something was.

00:56:27.010 --> 00:56:28.580
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Wrong within those data sets.

00:56:29.680 --> 00:56:35.010
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
So there there's no doubt that you know the actual data validation.

00:56:36.260 --> 00:56:46.410
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Definitely could be improved upon because it's not really there outside of you know, we expect you know 10 columns of floats kind of a deal in this order.

00:56:47.910 --> 00:56:48.610
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
But yeah, so.

00:56:50.630 --> 00:56:51.730
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
That's what's currently done.

00:56:53.510 --> 00:57:09.480
Li, Qiang (ARC-TI)[SimLabs III Contract Management & Technical Services]
Hey, Joe, this is Sean. From this Sean from data pipeline, just make comments so they data pipeline collecting data inject to the error graphic. So we defined all the endpoint with the swagger file. We have a data validation.

00:56:53.870 --> 00:56:54.780
Jung, Jaewoo (ARC-AFH)
Ohh OK so.

00:57:09.850 --> 00:57:28.500
Li, Qiang (ARC-TI)[SimLabs III Contract Management & Technical Services]
Uh, I just swagger for level. Also don't make a database a database. Also do the data validation before we inject data to the error graph. So just let you know the backward so data pipeline has really close with the error graph together for the data flow.

00:57:29.640 --> 00:57:35.100
Li, Qiang (ARC-TI)[SimLabs III Contract Management & Technical Services]
That's at the back wiring for that in the in the data pipeline and the database level.

00:57:29.950 --> 00:57:30.150
Wilwerding, Jerome C. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Yep.

00:57:37.650 --> 00:57:55.630
Jung, Jaewoo (ARC-AFH)
OK, I I mentioned this as we found the need to the check the data quality quickly, especially when there's my flight going on so that we can decide, OK, you gotta have to do it again versus now. Let's move on to the next thing. But anyway you have to follow up if you have questions about this.

00:57:40.150 --> 00:57:42.280
Sprague, John D (HQ-JD010)
As in the.

00:57:57.390 --> 00:57:57.870
Sprague, John D (HQ-JD010)
Great.

00:57:57.730 --> 00:58:27.820
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
Jay, I'll just. I'll just interrupt say that I, I do think that there's some opportunities to improve the data fabrication like you're suggesting in terms of when the data is coming in, do we have the right amount of columns or the in the right units, somewhat similar to a JSON schema. If you have a model that way and to know that there's errors right away validation, I look at as say at the next step that looks at the quality of the data and sometimes that's much more difficult to inspect. There's a great example in.

00:57:57.980 --> 00:57:58.230
Bhadoria, Divya (ARC-AFH)
Hey.

00:57:59.780 --> 00:58:00.080
Bhadoria, Divya (ARC-AFH)
Hey.

00:58:27.910 --> 00:58:59.460
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
FFT the follow on flight test where a new DGPS system was added. So we had these two DGPS systems capturing telemetry. The helicopter that was flying and it wasn't until that data was charted say just a 2D Birds Eye view of that data and Sarah Egan was the one. So she gets credit for this. That noticed that this new system placed the aircraft in say slightly different place. If I remember looking top down I was kind of to the South.

00:58:59.550 --> 00:59:23.530
Bagnall, Timothy M. (ARC-AFO)[UNIVERSITIES SPACE RESEARCH ASSN]
West of where it should have been a DSB and the previous DPS system were tracking. I'm just adding this that data verification and validation depending on the level of fidelity that you mean can kind of go from very simple to more advanced. And I think there are ways to validate it, but oftentimes it's human eyes on it that will capture things that are really anomalous.

00:59:25.380 --> 00:59:27.630
Sprague, John D (HQ-JD010)
Hey everybody. I'm go ahead, Divya.

00:59:25.480 --> 00:59:56.070
Bhadoria, Divya (ARC-AFH)
Check what I'll just quickly. Sorry, John. I'll just quickly add to that. So Jerry alluded to this quick look at a product that we create. So it's it's sort of a data verification of validation mechanism we have in place wherein we produce, we process all the data and produce quick some quick look results within a few

hours after each Shorty and that that the the that analysis or those products or the validation there helps us determine and it feeds into the next days.

00:59:56.160 --> 01:00:26.370
Bhadoria, Divya (ARC-AFH)
Or next sorties, maneuvers or data collection, so that sort of provides the initial feedback whether certain data quality is not being is not as good as expected, and if we need to redo some data collection or renewed need to do redo some maneuvers. So I think that's the whole idea behind the quick look that we produce. And of course, yeah, I agree there's more we can do more we can add for data validation. But you already have some initial mechanism in place.

01:00:26.440 --> 01:00:28.010
Bhadoria, Divya (ARC-AFH)
You look that kind of validation.

01:00:29.930 --> 01:00:45.390
Sprague, John D (HQ-JD010)
Thank you. Thank you, Divya very much. Hey, everybody. Unfortunately, we've gone over time and I know folks have to get to other meetings. I will say I did see Raj, Brad and Dan, your questions. I copied them down. So we'll get answers back to you.

01:00:46.660 --> 01:00:50.870
Sprague, John D (HQ-JD010)
Thank you everybody for taking the time to see this. If you have other.

01:00:51.920 --> 01:01:00.000
Sprague, John D (HQ-JD010)
Uh technologies that you'd like to share, like this one that we're sharing right now. Please let let us know so that we can get that scheduled.

01:01:00.690 --> 01:01:05.060
Sprague, John D (HQ-JD010)
I thank you everybody and thank you to our speakers. Phenomenal job. Appreciate it.

01:01:10.220 --> 01:01:10.760
Bhadoria, Divya (ARC-AFH)
Thank you.